Evolution and Cellular Resistance Mechanisms of the Immunity-Related GTPases

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Julia Hunn
aus Singen (Htwl.)

Köln 2007

Berichterstatter:	Prof. Dr. Jonathan C. Howard
	Prof. Dr. Thomas Langer
	Prof. Dr. Otto Haller
Tag der mündlichen Prüfung:	7. February 2008

Success is nothing more than going from failure to failure with undiminished optimism.
Winston Churchill

Table of contents

Abbreviations

AD	activation domain
ADAR	adenosine deaminase acting on RNA
ADRP	adipocyte differentiation-related protein
APOBEC3G	apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G
APS	ammonium peroxide sulfate
ATP	adenosine triphosphate
ATPase	adenosine triphosphatase
BAC	bacterial artificial chromosome
BD	DNA binding domain
BLAST	basic local alignment search tool
BSA	bovine serum albumin
CARD	caspase recruitment domain
CID	central interacting domain
CIITA	MHC class II transactivator (CIITA)
DAPI	4',6-Diamidine-2'-phenylindole dihydrochloride
DC	dentritic cell
DBD	DNA binding domain
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethylsulfoxid
Ds	double stranded
DTT	dithiothreitol
ECMV	encephalomyocarditis virus
EDTA	ethylen diamine tetraacetate
EF-G	elongation factor G
EF-Tu	elongation factor thermal unstable
EGFP	enhanced green fluorescent protein
ER	endoplasmic reticulum
eRF1	elongation release factor 1
ERV9	endogenous retrovirus 9
EST	expressed sequence tag
EtOH	ethanol
FAT10	HLA-F adjacent transcript 10
FCS	fetal calf serum
FLUAV	influenza A virus
FTase	farnesyl transferase
FPLC	fast protein liquid chromatography
G domain	GTPase domain
GAF	gamma activated factor
GAL4	galactose induced gene 4
GAS	gamma-activated sequence
GAP	GTPase activating protein
GBP	guanylate binding protein
GDI	guanine nucleotide dissociation inhibitor
GDP	guanosine diphosphate
GDPβS	guanosine-[β-thio]-diphosphate
GED	GTPase effector domain
GEF	guanin nucleotide exchange factor
GFP	green fluorescent protein
GGTase	geranylgeranyl transferase
GPI	glycosyl-phosphatidylinositol
GppNHp	5'-guanylylimidodiphosphate
GRA	dense granule protein
GST	glutathione-S-transferase
GTP	guanosine triphosphate
GTPase	guanosine triphosphatase
GTPγS	guanosine 5'-(3-O-thio)triphosphate
GTPI	GTPase IFN-induced
HRP	horse radisch peroxidase
HSV-1	herpes simplex virus type 1
IBDV	infectious bursal disease virus
IDO	indoleamine 2,3-dioxygenase
IF	immunofluorescence
IFN	interferon
IFNAR	IFNα receptor
IFNGR	IFNγ receptor
IFNLR	IFNλ receptor
IGTP	inducibly expressed GTPase
IIGP	IFN-inducible GTPase
IL	interleukin
IL10R2	interleukin 10 receptor 2
iNOS	inducible nitric oxide synthase
IPTG	isopropyl-β-D thiogalactoside
IRF	IFN regulatory factor
IRG	immunity-related GTPases
IRG-47	IFN-regulated GTPase 47 kDa
ISGF3	IFN-stimulated gene factor 3
ISG15	IFN-stimulated gene 15
ISG20	IFN-stimulated 3'-5' exonuclease gene 20
ISRE	IFN-stimulated response element
Jak	Janus kinase
LD	lethal dose
LPS	lipopolysaccharide
LRG-47	LPS-regulated GTPase 47 kDa
LTR	long terminal repeat
LZ	leucine zipper
MEC	mouse oviduct epithelial cell
MEF	mouse embryonic fibroblast
mGDP	mant guanosine diphosphat
mGTPγS	mant GTPγS
MHC	major histocompatibility complex
MIC	microneme protein
Mif	Mifepristone
MOI	multiplicity of infection
MW	molecular weight
MxA	Myxovirus resistance protein A
NEDD8	neural precursor cell expressed, developmentally down-regulated 8
NK cell	natural killer cell
NMT	*N*-myristoyl transferase
NOS2	nitric oxide synthase 2

NRAMP1	natural resistance-associated macrophage protein 1	Sox	Sry-type high-mobility-group domain box
2'-5' OAS	2'-5'-oligo-adenylate synthetase	SRP	signal recognition particle
o.n.	over night	SSAHA	sequence search and alignment by Hashing algorithm
OD	optical density	STAT	signal transducer and activator of transcription
ORF	open reading frame	SUMO	small ubiquitin-related modifier
PAGE	polyacrylamind gel electrophoresis	TGTP	T cell specific GTPase
PBS	phosphate buffered saline	TIM	triose-phosphate isomerase
PH	pleckstrin homology	TK	HSV thymidine kinase
phox	phagocyte oxidase	TNFα	tumor necrosis factor α
PKR	dsRNA-dependent protein kinase	TRIM5α	tripartite motif protein 5α
PML	promyelocytic leukemia	tsvGORASP	testis specific splice variant Golgi reassembly and stacking protein 2
PMSF	phenylmethylsulphonyl fluoride	Tyk2	tyrosine kinase 2
PRD	proline-rich domain	U	units
PV	parasitophorous vacuole	UAS	upstream activating sequence
PVM	parasitophorous vacuolar membrane	URG4	up-regulated gene 5
Rab	Ras genes from rat brain	UTR	untranslated region
Ran	Ras related nuclear protein	VLIG	very large inducible GTPase
Ras	rat sarcoma	VSV	vesicular stomatitis virus
Rho	rat sarcoma homologue	WB	Western blot
RNAi	RNA interference	Y2H	yeast two-hybrid
RON	rhoptry neck protein	YFP	yellow fluorescent protein
ROP	rhoptry protein	YPD	yeast-peptone-dextrose medium
ROS	reactive oxygen sspecies	ZAP	zinc-finger antiviral protein
RT	room temperature		
SD	synthetic defined		
SDS	sodium dodecylsulfate		

I. Introduction

I.1. Infection, immunity and cytokines

Multicellular organisms are constantly threatened by the invasion of microorganisms. The arms race between the host trying to prevent infection and eliminate the invaders and the pathogen seeking to exploit the host and to counteract its defence mechanisms led to the evolution of a complex, multilayered immune defence. The immune system has traditionally been divided into innate and adaptive components. The innate immune system predates the adaptive immune response evolutionary (Medzhitov 1997). Even single-cell organisms have heritable defence mechanisms, and every multicellular organism appears to have a complex innate immune system, whereas adaptive immunity is found only in vertebrates (Beutler 2004; Medzhitov 1997). The innate response includes both constitutive and inducible mediators and produces generic receptors that recognise conserved pathogen-associated molecular patterns (PAMPs) to trigger an inflammatory response that limits pathogen invasion (Janeway 2002). Specific adaptive immunity, by contrast, depends upon somatic diversification of antigen-receptor genes to generate a vast repertoire of lymphocytes, each expressing a different antigen receptor. Recognition of specific antigenic pathogen compounds by these cell-surface receptors triggers clonal amplification, cellular differentiation and production of secreted receptors with the same antigen binding specificity (Burnet 1959). In the vertebrate immune system innate and adaptive components synergise in the clearance of pathogens. The innate response is crucial in limiting the early replication and spread of infectious agents. By contrast, the generation of an adaptive immune response involves considerable lag time but culminates in the production of specialised effector mechanisms that are highly efficient in eliminating the pathogen and in the formation of an immunological memory (Le Bon 2002). In addition to these complementary activities, there is a fundamental connection and extensive crosstalk between innate and adaptive immunity. Namely, the magnitude and quality of the adaptive immune response is dependent on signals derived from the innate response to infection (Medzhitov 1997). Besides cell-cell interactions a central mode of communication within the immune system is via cytokines produced by hematopoietic and non-hematopoietic cells. Among the cytokines, interferons play a complex and central immunomodulatory role mediating host resistance to pathogens (reviewed in (Boehm 1997; Stark 1998; Young 2007)).

I.2. Interferons

Interferons (IFNs) were originally discovered as agents that interfere with viral replication (Isaacs 1957a; Isaacs 1957b). Beside their antiviral activity, IFNs exhibit growth-inhibitory effects and have important roles in immunosurveillance for malignant cells and in immunomodulation. The importance of interferons is apparent from mice that lack expression of IFNγ or its receptor due to targeted deletions and consequently are susceptible to many infectious agents (Dalton 1993; Huang 1993; Janssen 2002; Jouanguy 1999a; Jouanguy 1999b; Orange 1995; Ottenhoff 2002; Scharton-Kersten 1996) (see Table 1). Humans with naturally occurring mutations in their IFNγ receptor genes, however, exhibit a severe, profound and selective susceptibility to weakly virulent

mycobacteria only (Doffinger 2002; Jouanguy 1997). IFNs are induced either directly by infection and tissue damage or by immune and inflammatory stimuli. IFNs are classified into type I, type II and type III according to sequence homology, cellular sources and receptor specificity.

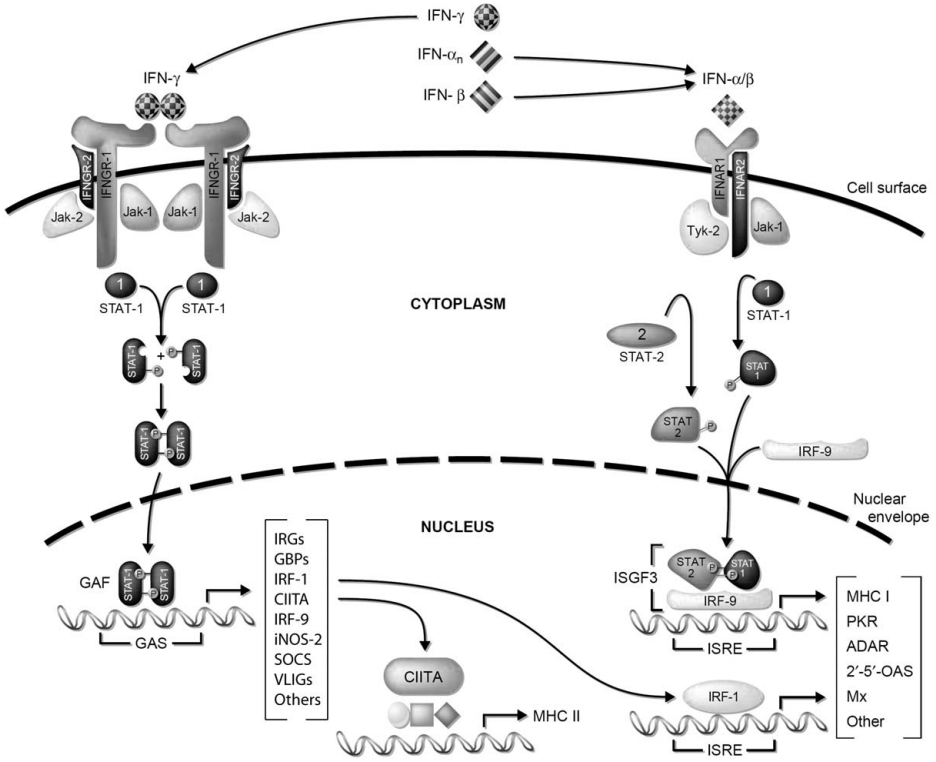


Figure 1 Interferon signalling pathways. Type I and type II interferons bind to their respective receptors activating related, partially overlapping but distinct signalling pathways of the Janus kinase family (Jak)–signal transducer and activator of transcription (STAT) type. These signalling cascades activate the expression of genes containing GAS (gamma-activated sequence) and ISRE (IFN-stimulated response element) elements in their promoter. Type III IFNs (λ-IFNs) signal through the same Jak/STAT pathway as type I IFNs driving expression of a common set of genes but engage a distinct heterodimeric receptor (IFNLR1/IL10R2) (not shown). The transcriptional response to IFNs includes numerous factors participating in the host immune response to viral and microbial pathogens. Protein kinase R (PKR), 2‘-5‘-oligoadenylate synthetase (2‘-5‘ OAS), adenosine deaminase acting on RNA (ADAR), Mx and guanylate-binding proteins (GBPs) have been implicated in viral resistance, whereas indoleamine 2,3-dioxygenase (IDO), nitric oxide synthase 2 (NOS2), phagocyte oxidase (phox), natural resistance-associated macrophage protein 1 (NRAMP1) and immunity-related GTPases (IRGs, p47 GTPases) all have been reported to inhibit the replication of bacteria, protozoa and viruses. Other proteins that are inducible by IFNs include MHC proteins and transcription factors IRF-1, IRF-9 and CIITA (modified from (Shtrichman 2001)).

Type I interferons are comprised of multiple IFNα subtypes (14-20 depending on species) (Pestka 1987; van Pesch 2004), IFNβ (Mogensen 1999), IFNδ (Lefevre 1998), IFNε (Conklin 2002; Pestka 2004), IFNκ (LaFleur 2001), IFNτ (Bazer 1997; Martal 1998), IFNυ (Samarajiwa 2006; Uze 2007), IFNω (Hauptmann 1985) and limitin/IFNζ (Oritani 2000). IFNα and IFNβ are secreted by almost all cell types upon virus infection and by activated immune cells such as dendritic cells and macrophages. The other type I IFNs are not necessarily present in all mammalian species or expressed upon viral infection. For example, mice have no orthologues of IFNυ, IFNκ is constitutively expressed by keratinocytes, secretion of IFNτ has only been reported in ungulate ruminants and has a specific function in maternal recognition of pregnancy (reviewed in (Schroder 2004)). Type I IFN family members bind a common heterodimeric cell-surface

receptor, the IFNα receptor (IFNRA), which consists of IFNAR1 and IFNAR2 chains and is activated by ligand-induced dimerisation (Darnell 1994; Stark 1998) (Figure 1). Each receptor subunit binds constitutively to a single specific member of the Janus kinase (Jak) family: IFNAR1 to tyrosine kinase 2 (Tyk2) and IFNAR2 to Jak1. Ligand binding induces the phosphorylation of Jak1, Tyk2, intracellular tyrosine residues of the IFNA receptor chains and the signal transducers and activators of transcription 1 and 2 (STAT1 and STAT2). Together with IFN-regulatory factor 9 (IRF9) phosphorylated STAT1 and STAT2 form a transcriptional complex called IFN-stimulated gene factor 3 (ISGF3), which translocates into the nucleus and binds to IFN-stimulated response elements (ISREs) present in the promoters of many IFN-regulated genes (Stark 1998).

Although the three type III IFNs or λ-IFNs (IFNλ1-3 or IL-28A/IL-28A/IL-29) differ genetically from type I IFNs, they exhibit similar biological antiviral, antitumour and antiproliferative activity and their expression is regulated in a similar fashion (Kotenko 2003; Osterlund 2007; Sheppard 2003; Uze 2007). Furthermore, they signal through the same Jak/STAT signalling pathway driving expression of a common set of genes. Importantly, however, λ-IFNs bind to a distinct membrane receptor composed of the IFNLR1 and IL10R2 receptor chains. This specific receptor usage suggests that this cytokine family does not merely replicate the type I IFN system and justifies its designation as type III IFN.

Interferon gamma (IFNγ) is the sole type II IFN and is synthesised only by certain activated effector cells of the innate and adaptive immune systems including natural killer (NK) cells (Bancroft 1993), T lymphocytes (Mosmann 1989), macrophages (Gessani 1998; Munder 1998) and dendritic cells (DCs) (Ohteki 1999). Although originally defined as an agent with direct antiviral activity, the properties of IFNγ include regulation of several aspects of the immune response, stimulation of bactericidal activity of phagocytes, stimulation of antigen presentation through class I and class II major histocompatibility complex (MHC) molecules, orchestration of leukocyte-endothelium interactions, effects on cell proliferation and apoptosis, as well as the stimulation and repression of a variety of genes. The two chains of the IFNγ receptor, IFNGR1 and IFNGR2, heterodimerise upon binding of an IFNγ homodimer leading to the activation of the respective receptor associated Janus kinases, Jak1 and Jak2 (Figure 1) (Darnell 1994; Stark 1998). The subsequent tyrosine phosphorylation cascade results in phosphorylated and activated STAT1. Upon phosphorylation STAT1 forms a homodimer termed gamma activated factor (GAF), translocates into the nucleus and initiates transcription by binding to gamma activated sequences (GAS) in the promoters of IFNγ inducible genes (Stark 1998).

Thus, binding of IFNs to their specific cell surface receptors leads to the activation of distinct but related components of the signal transduction and transcriptional activation machinery, resulting in the stimulation of the transcription of more than thousand genes belonging to partially overlapping sets (Boehm 1997; Darnell 1994; Der 1998; Ehrt 2001; Stark 1998; Takaoka 2000; Valente 1992) (Figure 1). Among these are a wide range of mediators that undermine the ability of pathogens to survive in host cells, contributing to organismal and cellular resistance involving both the innate and adaptive immune system (Boehm 1997).

I.3. IFN-mediated cell autonomous resistance

It has become clear over the last decades that the resistance against pathogens does not solely rely on specialised immune cells like lymphocytes or macrophages that patrol through the body. To counteract intracellular pathogens every single body cell is equipped with multiple defence mechanisms that are independent of specialised immune cells. Many of the molecular players of this so-called cell autonomous resistance are induced by interferons and therefore strictly speaking not completely independent of other cells. Still, apart from the need of a 'danger signal' to induce cell autonomous immunity, the process of pathogen counteraction and clearance is independent from the world outside the cell boundaries.

The list of proteins implicated in cell autonomous immunity is long and still growing and includes factors active against all kinds of pathogens. The antiviral factors include double stranded (ds)RNA-dependent protein kinase (PKR) (Garcia 2007; Lee 1993; Meurs 1992; Stojdl 2000; Yang 1995), 2'-5' oligoadenylate synthetase (2'-5' OAS)/RNase L (Chebath 1987; Chebath 1983; Samuel 2001), dsRNA adenosine deaminase (ADAR1) (Jayan 2002; Wong 2002), ISG20 (interferon stimulated 3'-5' exonuclease gene 20kDa) (Degols 2007; Espert 2003), promyelocytic leukemia protein (PML/TRIM19) (Chee 2003; Chelbi-Alix 1998; Everett 2007; Regad 2001), zinc-finger antiviral protein (ZAP) (Gao 2002; Guo 2007), APOBEC3G/CEM15 (apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G) (Harris 2003; Harris 2004; Mangeat 2003; Sheehy 2002; Turelli 2004) and tripartite motif protein 5 alpha (TRIM5α) (Nisole 2005; Stremlau 2004; Stremlau 2006). Natural resistance-associated macrophage protein 1 (NRAMP1) (Atkinson 1997; Barton 1999; Bradley 1979; Gros 1981; Hackam 1998; Plant 1976; Vidal 1993) mediates resistance against bacterial and protozoal pathogens. A number of resistance molecules have been reported to counteract bacterial and protozoal as well as viral pathogens. Indoleamine 2,3-dioxygenase (IDO) (Bodaghi 1999; Murray 1989; Pantoja 2000; Pfefferkorn 1984; Pfefferkorn 1986) and the phagocyte oxidase (phox) complex (Jackson 1995; Vazquez-Torres 2001) belong to this group. Inducible nitric oxide synthase (iNOS/NOS2) has an even broader field of action, displaying cytostatic or cytotoxic activity against viruses, bacteria, fungi, protozoa, helminths and tumour cells (Bogdan 2001; Kapur 1999; MacMicking 1997; Turco 1986; Vazquez-Torres 2001). The underlying resistance mechanisms range from inhibition of mRNA translation (PKR), degradation of RNA (OAS, ISG20, ZAP), RNA deamination (ADAR1, APOBEC3G) and binding and premature disassembly of viral capsids (TRIM5α) to depletion of divalent cations, tryptophan and arginine (NRAMP1, IDO and iNOS respectively) and cytostatic or cytotoxic effects mediated by nitric oxide (iNOS) and reactive oxygen species (ROS) (phox). Among this ever growing group of cell autonomous resistance factors are numerous members of the superfamily of large, interferon-inducible GTPases.

I.4. Guanosine triphosphatases (GTPases)

Guanosine triphosphatases (GTPases) are a diverse family of proteins that carry out various cellular functions, including membrane trafficking (Rab, dynamin), cell signalling and migration (Ras, Rho, Gα), nuclear transport (Ran), translation and protein

translocation (EF-Tu, EF-G, signal recognition particle (SRP)) (Leipe 2002) and cell autonomous resistance against intracellular pathogens (Mx, IRGs) (Martens 2006). Guanine nucleotide binding is essential for protein function and is mediated by five motifs termed G1–G5, of which the G1 (GX_4GKS/T), G3 (DXXG), and G4 (N/TQ/KXD) are more or less universally conserved (Bourne 1991; Dever 1987). The G1 motif or P-loop interacts with the phosphate groups of the nucleotide, the G3 motif binds the magnesium ion and makes contact to the γ-phosphate and the G4 motif confers specificity by contacting the base of the guanine nucleotide (Bourne 1990; Bourne 1991).

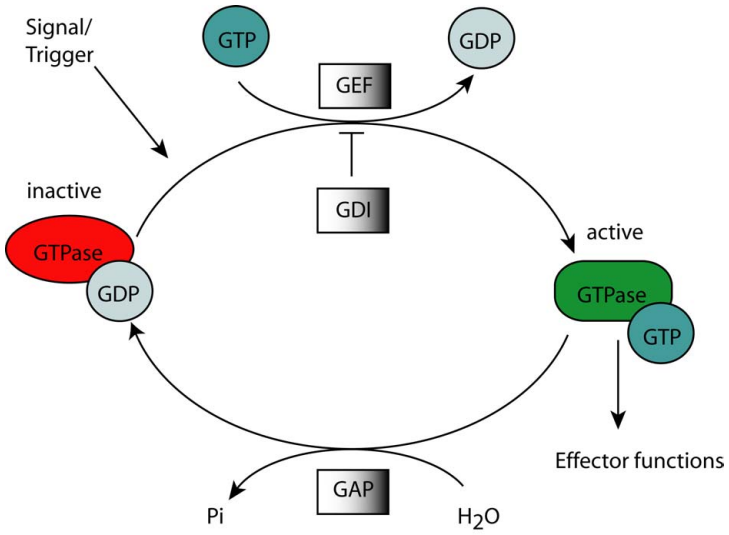


Figure 2 The GTPase cycle. Simplified depiction of a model GTPase cycle. The GDP-bound form of the GTP-binding protein is considered inactive, whereas the GTP-bound form represents the active form mediating effector functions. For many GTPases the transition between the GDP- and GTP-bound states is regulated by other proteins. Guanine nucleotide dissociation inhibitors (GDIs) prevent dissociation of GDP, keeping the GTPase in the inactive form. Guanine nucleotide exchange factors (GEFs) release the bound GDP from the GTPase enabling GTPase reactivation by GTP-binding. GTPase activating proteins (GAPs), in contrast, trigger GTP hydrolysis, restoring the inactive GDP-bound form.

All GTPases analysed to date share a nucleotide-binding domain with a common structural fold (Leipe 2002) and cycle between two alternative conformations induced by binding of guanosine diphosphate (GDP) and triphosphate (GTP), respectively, often functioning as molecular switches (Bourne 1990; Bourne 1991) (Figure 2). The GDP-bound form of the GTP-binding protein is considered inactive, whereas the GTP-bound form represents the active form mediating effector functions. GTPases associate with different regulator and effector molecules depending on their position in the nucleotide hydrolysis and exchange cycle. For many GTPases the transition between the GDP- and GTP-bound states is regulated by other proteins. Three classes of regulatory proteins are distinguished. Guanine nucleotide dissociation inhibitors (GDIs) prevent dissociation of GDP, keeping the GTPase in the inactive form. Guanine nucleotide exchange factors (GEFs) release the bound GDP from the GTPase. This enables GTPase reactivation by GTP-binding due to the higher intracellular concentrations of GTP than GDP and the higher affinity for GTP of most GTPases. GTPase activating proteins (GAPs), in contrast, trigger GTP hydrolysis, restoring the inactive GDP-bound form (Vetter 2001). The GAP activity essential for GTP-hydrolysis in many GTPases can be provided from a separate GAP protein, as known from the GAP proteins of H-Ras that insert a catalytic arginine side chain (the so called arginine finger) into the active site of the GTPase

(Sprang 1997) and from the exceptional Rap1Gap that provides a catalytic asparagine (Daumke 2004). Alternatively, GTPases can provide their own GAP activity, either in *cis* from a separate protein domain as it is the case with Gα proteins (Sprang 1997) or in *trans* by self-association as seen in dynamins (Tuma 1994) and large IFN-inducible GTPases (Irga6 (Uthaiah 2003) and Pawlowski unpublished results; hGBP1 (Ghosh 2006)).

I.4.1. Dynamin

Dynamins are GTPases of about 100 kDa molecular weight found in animals, plants and yeast that exert various functions including vesicle formation, vesicle transport, organelle division, and cytokinesis (Praefcke 2004b). Dynamins also regulate membrane dynamics in the context of cell motility (Kruchten 2006) and have been shown to interact with the actin cytoskeleton via a large number of actin binding proteins (Schafer 2002; Schafer 2004). Dynamin-related proteins have also been reported in bacteria (Low 2006; van der Bliek 1999). Mammals possess three dynamin genes each coding for various alternative splice forms (Cao 1998; Urrutia 1997). Dynamin I is neuronal specific, dynamin II ubiquitously expressed and dynamin III is restricted to lung, brain and testis.

Dynamins contain a large N-terminal nucleotide binding domain (~300 aa) followed by the middle domain, a pleckstrin homology (PH) domain, a GTPase effector domain (GED) and a C-terminal proline-rich domain (PRD) mediating self-assembly, membrane targeting, GAP activity and interaction with other proteins, respectively (Praefcke 2004b; Urrutia 1997). In addition to these structural features, dynamin and dynamin-related proteins are distinguished from small GTPases by their oligomerisation-dependent GTPase activation and cooperative GTP hydrolysis, their low nucleotide-binding affinities in the micromolar range and the ability of many family members to interact with and tubulate lipid membranes (reviewed in (Hinshaw 2000; Praefcke 2004b; Song 2003).

Under physiological salt conditions in absence of nucleotide dynamin is in a monomer-tetramer equilibrium *in vitro* (Binns 1999) showing rather high basal GTP hydrolysis (Praefcke 2004b) that is further stimulated by self-assembly into ring-like oligomers in solution (Hinshaw 1995; Tuma 1994; Tuma 1993) and into ring- and spiral-like structures on lipid membranes (Marks 2001; Stowell 1999). Within these oligomers, dynamin functions as its own GAP (Muhlberg 1997). The presence of lipids massively accelerates GTP hydrolysis and enhances nucleotide-dependent oligomerisation and self-assembly (Song 2003; Tuma 1994).

Dynamin tubulates membranes in the presence of GTP-analogues *in vitro* (Hinshaw 1995) and *in vivo* (Marks 2001) whereby vesicle scission requires GTP hydrolysis-dependent conformational changes (Marks 2001). The mechanism by which dynamin distorts membranes is under debate (Praefcke 2004b; Song 2003). Dynamin could either act as a mechanochemical enzyme pinching off vesicles from donor membranes (Hinshaw 1995; Marks 2001) or, alternatively, mediating vesicle scission by GTP-dependent recruitment of effector proteins to the neck of budding vesicles (Sever 1999; Stowell 1999; Sweitzer 1998) where dynamin assembles in spiral-like structures *in*

vivo (Hinshaw 1995). Although the classical dynamins do not seem to play roles in IFN-mediated immune resistance, all known IFN-inducible GTPase families share characteristics with the dynamins.

I.4.2. Large, interferon-inducible GTPases

The expression of four described families of large GTPases is regulated by IFNs: The Mx GTPases (Haller 2002), the 65 kDa (p65) guanylate binding proteins (GBPs) (Cheng 1986; Cheng 1983), the very large inducible GTPases (VLIGs) (Klamp 2003) and the immunity-related or 47 kDa (p47) GTPases (IRGs) (Bekpen 2005b; Boehm 1998). The large, IFN-inducible GTPases – despite negligible conservation outside the GTPase domain - share numerous dynamin-like features that clearly distinguish them from small, Ras-like GTPases (Martens 2006; Praefcke 2004b). Among these features are the possession of one or more domains in addition to the conserved G-domain, low micromolar nucleotide affinities, GTP-dependent oligomerisation leading to cooperative GTP hydrolysis, GAP activity provided by self-interaction, binding – and in some cases also tubulation – of lipid vesicles *in vitro*. In addition, these GTPases are all strongly induced by interferons and many have been shown to mediate cell autonomous resistance against intracellular pathogens. Two families of large, IFN-inducible GTPases, Mx proteins and the guanylate binding proteins (GBPs), share conserved domains with and belong to the dynamin superfamily of GTPases (Praefcke 2004b).

I.4.2.1. The Mx GTPases

The 70–80-kD Mx GTPases are strongly induced exclusively by type I (IFNα/β) and type III IFNs (IFNλ) (Aebi 1989; Haller 1980; Holzinger 2007; Kotenko 2003; Simon 1991; Staeheli 1986a). They were initially discovered in an inbred mouse strain (A2G) that showed an exceptionally high degree of resistance against infection with influenza A viruses (FLUAV) (Lindenmann 1962; Lindenmann 1963). The resistance phenotype was inherited as a single autosomal dominant trait, was specific for members of the orthomyxovirus family and was dependent on type I IFN (Haller 1979; Lindenmann 1963; Lindenmann 1964). The single gene responsible for the resistance phenotype was termed orthomyxovirus resistance gene 1 (Mx1) (Lindenmann 1963). The Mx1 antiviral effect was shown to be cell autonomous as macrophages from A2G mice were resistant to FLUAV infection *in vitro* (Lindenmann 1978) and to be independent of other IFN-induced factors (Arnheiter 1990; Staeheli 1986b). In contrast to wild mouse species, most laboratory inbred mouse strains carry non-functional Mx1 alleles and are highly susceptible to mouse adapted FLUAV strains (Haller 1987; Jin 1998; Staeheli 1988a). The second IFN-regulated mouse Mx gene, Mx2, (Reeves 1988; Staeheli 1986c; Staeheli 1988b) was found to be disrupted in all laboratory strains examined (Staeheli 1988b). Functional Mx2 was discovered later in *Mus musculus musculus* and *Mus spretus* derived mouse strains of feral origin (Jin 1999). The widespread absence of functional Mx genes in inbred mice strains kept under pathogen free conditions and the fact that $Mx1^{-}$ alleles are frequent in wild mice may indicate that possession of functional Mx alleles is linked to high evolutionary costs. However, the absence of functional Mx genes from inbred mice might be due to a founder effect (Haller 2007b).

The mouse Mx1 and Mx2 proteins localise to the nucleus (Dreiding 1985) partially colocalising with PML bodies (Engelhardt 2004; Engelhardt 2001) and the cytoplasm (Meier 1988) respectively, and display antiviral activities against RNA viruses that replicate in these particular subcellular compartments. Thus, Mx1 confers resistance to orthomyxoviruses (e.g. FLUAV) and Thogoto viruses, whereas Mx2 provides resistance to bunyaviruses (e.g. La Crosse virus) and rhabdoviruses (e.g. vesicular stomatitis virus (VSV)) (Haller 1998; Jin 1999; Zurcher 1992b). Humans possess two Mx homologues, MxA and MxB (Aebi 1989; Staeheli 1985). MxA was shown to be partially soluble and partially associated with the smooth ER in punctate granula (Accola 2002; Staeheli 1985; Stertz 2006). It displays antiviral activity against a large range of RNA viruses *in vivo* and *in vitro*, including bunyaviruses, orthomyxoviruses, paramyxoviruses, rhabdoviruses, togaviruses, picornaviruses, and hepatitis B virus, a DNA virus with an RNA intermediate (Chieux 2001; Haller 1998; Hefti 1999; Landis 1998; Melen 1996). MxA can function in mice independent of other IFN-inducible factors (Hefti 1999) and is active even in mosquito cells (Miura 2001). In contrast, no antiviral function could be detected for MxB (Melen 1996; Pavlovic 1990). MxB contains a functional NLS at the N-terminus and mainly localises to the nuclear envelope in a nucleotide-dependent manner (King 2004; Melen 1996). Recently, MxB has been implicated in the regulation of nuclear transport and cell-cycle progression (King 2004).

As other large GTPases, Mx proteins have a relatively high molecular mass, a low affinity for nucleotides, and a high intrinsic rate of GTP hydrolysis (Haller 2007b; Richter 1995; Staeheli 1993). Furthermore they show cooperativity in GTP hydrolysis and self-assemble into highly ordered homo-oligomers forming ring-like and helical structures *in vitro* in presence of non-hydrolysable GDP and GTP analogues respectively (Accola 2002; Kochs 2002a; Melen 1992; Nakayama 1993). Self-assembly seems to be critical for GTPase activity, protein stability (Janzen 2000; Schumacher 1998), and recognition of viral target structures (Johannes 1997; Kochs 2002b; Zurcher 1992a). MxA was shown to bind and tubulate phosphatidyl serine vesicles *in vitro* in a nucleotide-independent manner (Accola 2002). Mx proteins contain an N-terminal GTPase (G) domain, a central-interacting domain (CID) in the middle and a C-terminal GTPase effector domain (GED) with leucine-zipper (LZ) motifs (Haller 2002; Haller 2007b; Melen 1992). Interaction of the LZ region of MxA with the G-domain increases the GTPase activity (Schwemmle 1995). The GED can function even when supplied in *trans* but the catalytic mechanism remains elusive (Schwemmle 1995). Further intra- or intermolecular interactions of Mx domains including interaction of the GED with the CID have been proposed (Di Paolo 1999; Janzen 2000; Ponten 1997; Schumacher 1998; Schwemmle 1995). Homo-oligomerisation was proposed to result from binding of the LZ region of one molecule to the CID of a second neighbouring molecule leading to an enhancement of GTPase activity (Janzen 2000; Schumacher 1998). Homomeric MxA and MxB interactions were shown to occur *in vivo* (Melen 1997; Ponten 1997).

Mx proteins were found to bind to essential viral components and to interfere with viral trafficking, assembly, and replication. A physical interaction of MxA with viral nucleocapsid proteins was demonstrated for Thogoto virus, La Crosse virus and – weakly – for influenza viruses (Kochs 1999a; Kochs 1999b; Kochs 2002b; Kochs 1998; Turan 2004). MxA blocked the transport of viral nucleocapsids into the nucleus in case

of Thogoto virus and depleted nucleocapsid proteins from the viral replication sites in the case of La Crosse virus (Kochs 1999b; Kochs 2002b; Reichelt 2004). Models of MxA action propose that the protein forms two types of assemblies in the cell: the first is a resting oligomeric pool of MxA occurring in the absence of viral infection, and the second is a co-polymer formed from monomeric MxA that breaks away from the oligomeric pools to complex with target viral proteins, effectively sequestering and inactivating them (Haller 2002; Haller 2007b). Mouse Mx1 inhibits nuclear primary transcription of the influenza virus RNA (Pavlovic 1992), possibly via the viral RNA-dependent RNA polymerase (Huang 1992; Stranden 1993). Cytoplasmic human MxA normally inhibits influenza virus replication at a later step, while primary viral transcription is unaffected (Pavlovic 1992). However, when artificially targeted to the nucleus, MxA also inhibits primary viral transcription (Zurcher 1992b), whereas mouse Mx1 has no activity against influenza virus in the cytoplasm (Zurcher 1992c). GTP binding is necessary for the antiviral effect of Mx proteins but, surprisingly, oligomerisation and GTP hydrolysis seem to be dispensable (Janzen 2000; Pitossi 1993; Ponten 1997). The C-terminal domain shows the highest sequence divergence between different Mx proteins and is required for virus inhibition, thus, it may determine the specificity of the antiviral activity (Johannes 1997; Ko 2002; Kochs 2002b; Ponten 1997; Zurcher 1992a).

IFN-inducible, antivirally active Mx proteins are found in most vertebrate species, including pigs, cows, birds and fish (Leong 1998; Watanabe 2007). However, several Mx proteins, such as human MxB and rat Mx3, appear devoid of antiviral activity (Meier 1990; Pavlovic 1990). Mx genes are polymorphic in most species and the resulting amino acid differences were show to account for variations in the antiviral activities of the allelic gene products in many cases (Haller 1987; Jin 1999; Nakajima 2007; Palm 2007; Seyama 2006). Recently, the first invertebrate Mx protein was identified in disk abalone *(Haliotis discus discus)*. It shares certain features with fish and mammalian Mx proteins and is induced by PolyI:C injection at the message level (De Zoysa 2007).

I.4.2.2. The Guanylate Binding Proteins (GBPs)

The 65-67 kDa guanylate-binding proteins (GBPs) are among the most abundant proteins that accumulate in response to IFNγ stimulation (Boehm 1998; Cheng 1985; Cheng 1986; Cheng 1983; Nguyen 2002). Induction of GBPs, though somewhat weaker, has also been reported by type I IFN, interleukin-1 alpha and beta (IL1α/β), tumour necrosis factor alpha (TNFα), and lipopolysaccharide (LPS) (Cheng 1986; Guenzi 2001; Guenzi 2003; Lubeseder-Martellato 2002; Nantais 1996; Tripal 2007; Vestal 1996). Furthermore, *Listeria monocytogenes* infection resulted in IFNγ-dependent induction of mGBP2 and mGBP4 expression in mouse liver (Boehm 1998). The GBP family is represented by seven genes in human (hGBP1-7) and ten genes in mouse (mGBP1-10) (Boehm 1998; Cheng 1991; Degrandi 2007; Fellenberg 2004; Han 1998; Luan 2002; Olszewski 2006; Wynn 1991), and is well conserved throughout the vertebrates (Robertsen 2006). ISRE and GAS elements have been detected in the putative promoters of most GBP genes (Olszewski 2006).

Most family members harbour an unusual G4 motif, T(L/V)RD, instead of the canonical (N/T)(K/Q)XD (Praefcke 1999), a property shared with the IFN-inducible giant GTPase, VLIG-1 (see below) (Klamp 2003). GBPs bind nucleotides with low micromolar affinity (Praefcke 1999; Praefcke 2004a; Schwemmle 1994), display GTP-dependent oligomerisation (Praefcke 1999; Prakash 2000a) and hydrolyse GTP cooperatively (Praefcke 2004a; Prakash 2000a). All GBPs analysed to date bind GTP, GDP and GMP with similar micromolar affinity (Cheng 1985; Cheng 1983; Praefcke 1999; Praefcke 2004a; Staeheli 1984). GBPs, in contrast to other GTPases, hydrolyse GTP to GDP and GMP in two successive cleavages of orthophosphate rather than by pyrophosphate generation (Neun 1996; Praefcke 1999; Schwemmle 1994). Hydrolysis of GDP to GMP involves the same catalytic machinery as GTP hydrolysis, whereby the β-phosphate of the GDP product of the first round of hydrolysis is brought into the position of the former γ-phosphate by movement of the phosphate cap (see below; (Ghosh 2006)). GMP is the predominant product for hGBP1 at any time of the reaction, whereby the product ratio of GDP and GMP for hGBP1 is temperature dependent (Kunzelmann 2006; Schwemmle 1994). GDP can be bound from solution but cannot serve as substrate for hydrolysis (Neun 1996; Schwemmle 1994). For hGBP2, the main product of hydrolysis is GDP (Neun 1996).

The structure of human GBP1 has been determined by X-ray crystallography (Prakash 2000a; Prakash 2000b). The N-terminal GTP-binding domain (G domain), associated with an elongated C-terminal helical domain, has several insertions relative to the canonical Ras domain (Pai 1989). The insertions include the unique guanine and phosphate caps that shield the nucleotide from the solvent (Prakash 2000b). The helical domain consists of 5 α-helices in two three helix bundles, representing the middle and the GED domain respectively, followed by a long penultimate helix that reaches back to and contacts the G-domain and a short C-terminal helix (Prakash 2000a). Recombinant human GBP1 is monomeric in the nucleotide free, GMP- and GDP-bound state but forms dimers in the GTP-bound state and tetramers in the GDP+AlFx-stabilised transition state (Ghosh 2006; Praefcke 2004a; Prakash 2000a; Prakash 2000b). The isolated nucleotide-binding domain of hGBP1 crystallised as G-domain–G domain dimer in the GTP- and GDP-bound forms (Ghosh 2006). No multimers containing more than four GBP molecules have been reported. GTP-dependent dimerisation increases the GTP hydrolysis rate probably by providing a catalytic arginine, R48, in *cis* (Ghosh 2006). The arginine residue in the P loop (R48) is conserved in most GBPs (GxxRxGKS) and mutation inhibits GTP-hydrolysis but not nucleotide binding (Praefcke 2004a).

Human and mouse GBP1, -2 and -5 contain a C-terminal CaaX motif (Olszewski 2006), defined by a cysteine (C) residue, followed by two small, generally aliphatic (a) residues, and the X residue, which can be many amino acids (Clarke 1988). This sequence enables posttranslational modification by covalent, irreversible attachment of an isoprenoid lipid catalysed by one of three protein prenyltransferases: protein farnesyltransferase (FTase), protein geranylgeranyltransferase type I (GGTase-I) and type II (GGTase-II or Rab GGTase) (Lane 2006). FTase and GGTase-I transfer a farnesyl (C15) or geranylgeranyl (C20) isoprenoid, respectively, to the cysteine of a C-terminal CaaX motif. The specificity is determined by the X residue with a general preference of the FTase for methionine, serine, glutamine, or alanine and of the GGTase-

I for leucine or phenylalanine in that position (Lane 2006). After covalent attachment of the isoprenoid in the cytoplasm, most CaaX proteins undergo two further prenylation-dependent processing steps at the endoplasmic reticulum: proteolytic removal of the aaX tripeptide by endopeptidases and carboxymethylation of the prenylcysteine residue by carboxyl methyltransferases (Clarke 1992; Schafer 1992). These modifications are thought to facilitate membrane binding and certain protein-protein interactions. Judging from the sequence of their CaaX motifs, hGBP1 and mGBP5 should be farnesylated whereas hGBP2, hGBP5, mGBP1 and mGBP2 are predicted to be geranylgeranylated (Olszewski 2006). Prenylation *in vivo* has been demonstrated for rat GBP (Vestal 1996), mGBP1 (Stickney 2000), mGBP2 (Vestal 1998) and hGBP1 (Nantais 1996).

In subcellular fractionations, hGBP1, mGBP1 and mGBP2 were largely cytosolic (Modiano 2005; Nantais 1996; Stickney 2000; Vestal 1998; Vestal 2000). A punctate cytoplasmic localisation in vesicle-like structures was observed for endogenous mGBP2 not colocalising with a variety of organellar markers (Gorbacheva 2002; Vestal 2000). Neither GTP binding nor other IFNγ-induced factors but an intact CaaX box were required for this localisation (Gorbacheva 2002; Vestal 2000). In contrast, N-terminally tagged mGBP1 was diffusely distributed throughout the cytoplasm and not in puncta (Vestal 2000). GFP-tagged hGBP1, hGBP3, and hGBP5 were exclusively detected in the cytoplasm, whereas hGBP-2 and hGBP4 were also detected in the nucleus of endothelial cells (Tripal 2007). Treatment with aluminium fluoride, which can trap GTPases in the transition state (Combeau 1988), led to relocalisation of hGBP1 and hGBP2 but not of hGBP3 and hGBP4 to the Golgi apparatus (Modiano 2005; Tripal 2007). This required IFN induction of the cells and for hGBP1 also GTP binding and a functional isoprenylation motif. Human GBP5 was detected at the Golgi independent of IFN and AlFx treatment (Tripal 2007).

Several distinct, apparently unrelated functions have been reported for GBPs. For hGBP1 and -2 a weak antiviral effect against VSV and encephalomyocarditis virus (ECMV) has been reported in cultured cells and hGBP1 also mediated a similar effect on hepatitis C virus (HCV) (Anderson 1999; Carter 2005; Itsui 2006). Furthermore, mGBP1 and hGBP1 have been implicated in the regulation of angiogenesis by proinflammatory cytokines (Guenzi 2001; Guenzi 2003). Hereby hGBP1 is thought to inhibit angiogenesis by downregulating the expression of the matrix metalloproteinase-1 in vascular endothelia cells (Guenzi 2003). In addition, the C-terminal helical domain of hGBP1 was reported to mediate anti-proliferative effects of proinflammatory cytokines on vascular endothelial cells independent of the GTPase activity and isoprenylation of the molecule (Guenzi 2001). In contrast, mGBP2 was reported to mediate IFNγ-induced enhanced proliferation of murine fibroblasts in a nucleotide-binding dependent manner even in absence of IFNγ (Gorbacheva 2002). Thus, the current picture of GBP function is incomplete and controversial. Furthermore, none of the reported effects provides a conclusive explanation for the striking dependence of the GBPs on IFNγ induction in all cell types, their high intracellular concentrations, and their unique mechanism of GTP hydrolysis. Only very recently a study reported the involvment of several murine GBPs in the response to *Toxoplasma gondii* (Degrandi 2007) (see discussion for details).

I.4.2.3. Very large inducible GTPases (VLIGs)

The very large inducible GTPases (VLIGs) are a family of giant GTP binding proteins with a molecular weight of approximately 280 kDa. The family is represented in mouse by 6 members on chromosome 7 (Klamp 2003) and in humans by a single conserved homologue encoded on chromosome 11. Mouse VLIG-1 was massively induced by IFNγ and less strongly by IFNβ in cultured cells and by infection with *Listeria monocytogenes* in different mouse strains. VLIG-1 is a soluble protein found in the cytosol and nucleus. The open reading frame of VLIG-1 is encoded on a single very large exon and there is evidence that VLIG-1 is polymorphic in mice of different genetic backgrounds. The greatest part of the protein sequence does not show significant relationship to other known protein families. However, VLIG-1 possesses canonical G1 and G3 GTP-binding motifs embedded in a local sequence environment that resembles the nucleotide binding domains of the other IFN-inducible GTPase families, especially of the GBPs. The relation of VLIG-1 to other GTPase superfamily members is more distant. VLIG-1 is indeed a guanine nucleotide-binding protein as it was shown to bind strongly to GDP-agarose and very weakly, if at all, to GTP- and GMP-agarose. Recently a VLIG-1 related sequence was reported in a proteomic analysis of rat microglia cells (Zhou 2005) and infectious bursal disease virus (IBDV) was shown to induce a VLIG-1 homologue in chicken fibroblasts (Wong 2007). Sequence similarity of the central part of VLIG-1 with both caspase recruitment domain protein 6 (CARD6) and up-regulated gene 4 (URG4) has been reported in a region that is predicted to be an inosine 5'-monophosphate dehydrogenase/GMP reductase domain in those proteins (Dufner 2006). This domain family forms a triose-phosphate isomerase (TIM) barrel structure and is involved in biosynthesis of guanosine nucleotide (Andrews 1988; Collart 1988; Sintchak 2000). CARD6 has been reported to associate with microtubules and to modulate NFκB activity while URG4 enhances cell proliferation and both proteins contain a caspase recruitment domain (Dufner 2006). The relevance of these observations remains to be determined experimentally. VLIG function remains elusive but the IFN-inducibility of VLIG-1 in mice and the sequence relationship to other large, IFN-inducible GTPases suggest a role in cell autonomous resistance to intracellular pathogens.

I.4.2.4. The immunity-related GTPases (IRGs)

The immunity-related GTPases (IRGs) or p47 GTPases are guanine nucleotide binding proteins with a molecular weight of approximately 47 kDa but no sequence homology to any other known GTPase family outside the G-domain (Boehm 1998) (Figure 3). At the beginning of this study six mouse IRG genes had been described, namely *Irgd (IRG-47)* (Gilly 1992), *Irgm1 (LRG-47)* (Sorace 1995), *Irgb6 (TGTP/Mg21)* (Carlow 1995; Carlow 1998; Lafuse 1995), *Irgm3 (IGTP)* (Taylor 1996; Taylor 1997), *Irga6 (IIGP)* and *Irgm2 (GTPI)* (Boehm 1998) and no homologues from other species had been reported (for details about the nomenclature of IRG genes and the newly identified genes see results section (III.1.1) and (Bekpen 2005b)). Three of the mouse IRG proteins, Irgm1-3, have the non-canonical sequence GX_4GMS in place of the otherwise universally conserved GX_4GKS in the first nucleotide-binding motif (G1) correlating with other sequence features to define the GMS (IRGM) and GKS subfamilies respectively (Bekpen 2005b;

Boehm 1998) (Figure 3). In the course of this study the number of IRG genes in *Mus musculus domesticus* was extended to a total of 21 genes falling into 25 coding units (see results section (III.1.2) and (Bekpen 2005b).

Figure 3 Amino acid alignment of six IRG proteins. Manual alignment of amino acid sequences of Irgd (IRG-47), Irgb6 (Mg21/TGTP), Irga6 (IIGP), Irgm1 (LRG-47), Irgm3 (IGTP) and Irgm2 (GTPI). Positions where at least four sequences share the same residue are shaded. The N- and C-terminal domains show low sequence homology between the family members whereas the core G-domain of the proteins is conserved. The consensus sequences of the conserved GTP binding motifs are shown below the sequence. Three of the proteins contain an unusual methionine instead of the otherwise universally conserved lysine in the first nucleotide binding motif (G1) (from (Boehm 1998).

I.4.2.4.1. Expression of the IRG proteins

The IRG proteins are typically encoded on a single long exon and the promoters contain GAS and ISRE motifs mediating IFN-responsiveness but no other recurrent promoter motifs (Bekpen 2005b; Gilly 1996). In the mouse, 14 IRG genes have been shown to be inducible by IFNγ in fibroblasts. The mouse Irgc is a notable exception in that it lacks IFN regulation but is expressed exclusively in haploid spermatids in the mature testis (Bekpen 2005b; Rhode 2007). IRG protein expression is rapidly and strongly induced by IFNγ in a STAT1-dependent manner *in vivo* and *in vitro* in all cell types analysed (Boehm 1998; Collazo 2002; Gavrilescu 2004; Gilly 1992; Lafuse 1995; MacMicking 2003; Sorace 1995; Taylor 1996). IFNα/β also triggers expression, albeit to a lesser extent, as does lipopolysaccharide indirectly by stimulating type I IFN (Bafica 2007; Carlow 1998; Lafuse 1995; Lapaque 2006; Sorace 1995; Taylor 1996; Zerrahn 2002). Other cytokines tested have little effect on expression (Boehm 1998; Lafuse 1995; Sorace 1995). In mice, IRG proteins are expressed at modest levels in absence of infection in tissues with a significant leukocyte contribution such as thymus and spleen (Collazo 2001; Taylor 1996) and in bone marrow derived haematopoietic stem cells (Advani 2004; Terskikh 2001; Venezia 2004). A developmentally regulated burst of IFN synthesis may be responsible for this expression. Infection induces strong IRG expression in most tissues by triggering IFN production (Boehm 1998; Collazo 2001; Feng 2004; MacMicking 2003; Taylor 2000; Zerrahn 2002). IRG protein expression in

the documented absence of IFN signalling has only been reported for Irga6, which shows a high constitutive expression in hepatic parenchymal cells (Zeng 2007, Parvanova 2005). This expression persists in mice deficient in STAT1 and both the type I and II IFN receptors and is driven by a second independent promoter containing numerous binding sites for liver-enriched transcription factors of the hepatocyte nuclear factor group (HNFs) (Bekpen 2005b, Zeng 2007, Parvanova 2005) known to mediate liver specific expression (Costa 2003; Schrem 2002).

I.4.2.4.2. Biochemical properties of the IRG GTPases

GTPase activity was first documented *in vitro* for Irgm3 immunoprecitated from IFNγ-induced cells and for recombinant Irgm3 (Taylor 1996). Furthermore, radiolabelled GTP rather than GDP co-immunoprecipitated with Irgm3 from cells (Taylor 1997). As neither the affinities of the protein for the different nucleotides nor the hydrolysis rate are known and the assay does not detect nucleotide free protein, one can conclude only that Irgm3 is capable of binding GTP in IFNγ-induced cells. Subsequently, also recombinant Irgb6 was shown to hydrolyse GTP to GDP *in vitro* (Carlow 1998). To date, only Irga6 has been systematically characterised biochemically (Ghosh 2004; Uthaiah 2003). Irga6 hydrolyses GTP to GDP cooperatively with a maximum rate of 2 per minute, oligomerises in the presence of GTP, and the oligomers resolve upon GTP hydrolysis (Uthaiah 2003). The oligomer is probably the site of rapid hydrolysis as the addition of aluminium fluoride to trap the transition state stabilises the oligomers (Ghosh 2004; Uthaiah 2003). Thus, Irga6 is a self-activating GTPase, and to date no exogenous regulators of its activity have been described. Irga6 has a low nucleotide affinity, which is higher for GDP (1 μM) than for GTP (15 μM). Therefore, Irga6 should be predominantly GDP bound at cellular concentrations of 300 μM GTP and 100 μM GDP (Kleineke 1979) though GTP may be captured in oligomeric complexes *in vivo* (Pawlowski, unpublished data). Irga6 apoprotein and GDP-bound forms both crystallised as rotationally symmetrical dimers with identical contact interfaces (Ghosh 2004) (Figure 4). The G-domain is essentially Ras-like (Pai 1989), and lies between a three-helix bundle in the N-terminus, and a complex series of helices and loops in the C-terminus (Ghosh 2004). A destabilising mutation in the crystal dimer interface (M173A) yielded a monomeric Irga6 crystal with the non-hydrolysable GTP analogue GppNHp (Ghosh 2004) displaying only small structural changes relative to the GDP-bound structure. It is not clear whether the monomeric Irga6(M173A):GppNHp structure accurately reflects the actual GTP structure. Mutations in the dimer interface reduced oligomerisation and cooperative GTP hydrolysis but none eliminate the activity completely ((Ghosh 2004); Pawlowski unpublished data). This and the lack of strong conservation of the dimer contact residues between IRG family members (Bekpen 2005b; Ghosh 2004) suggests that the crystal dimer interface is not identical to one of the two interfaces needed for oligomerisation. Recent mutational analysis defined the primary interface as a G-domain-G domain interaction involving bound nucleotides (Pawlowski unpublished data). Formation of this contact is thought to induce large conformational changes exposing the putative second interface.

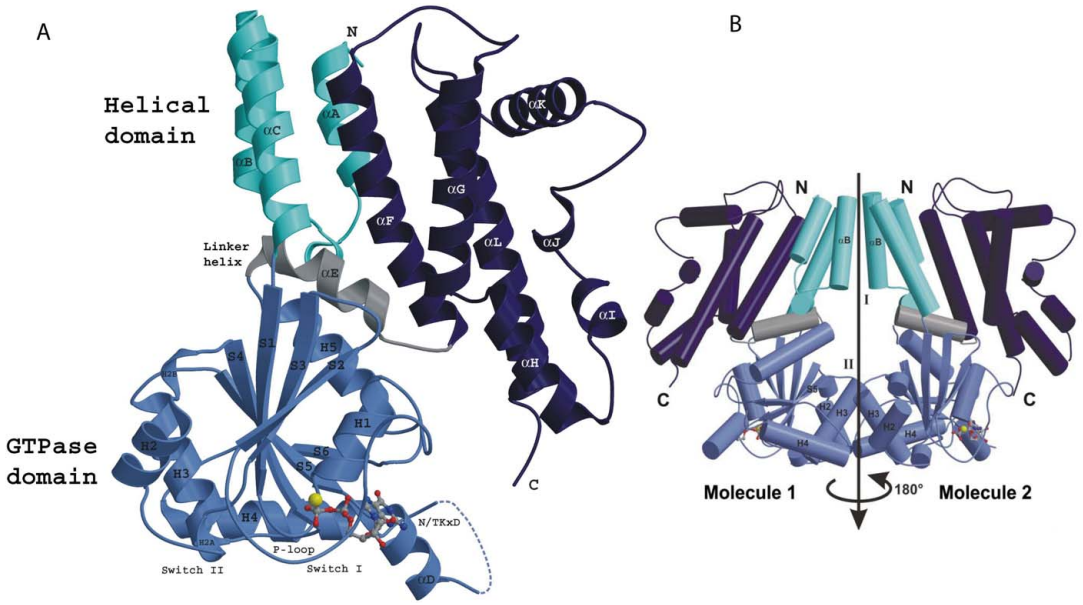


Figure 4 Crystal structure of Irga6 in the GDP bound form. (A) One molecule of the Irga6-GDP dimer is shown in ribbon presentation. The N-terminal domain (cyan) is composed of three α-helices and is followed by the Ras-like GTP binding domain (light blue). The helical C-terminal domain (dark blue) is connected to the G-domain by the linker helix αE (grey). The helices in Irgm1, Irgm2 and Irgm3 correlating to αK mediate membrane targeting of these proteins (see main text). GDP and Mg^{2+} are shown as atomic stick figure and yellow sphere respectively. The first 13 amino acids of Irga6 are not resolved in the structure. (**B)** Structure of the Irga6-GDP dimer. The subdomains are colour coded as in A. Secondary structure elements involved in the dimer interfaces I and II are labelled. The two-fold noncrystallographic symmetry axis is shown (from (Ghosh 2004).

I.4.2.4.3. Intracellular localisation of IRG proteins

IRG proteins associate with distinct subcellular membrane compartments by different targeting mechanisms. Irga6 concentrates at the endoplasmic reticulum (ER) (Martens 2004a) and at least in some cells also at the Golgi apparatus (Kaiser 2004; Zerrahn 2002) and contains a myristoylation sequence at the N-terminus (MGQLFSS) (Martens 2004a; Uthaiah 2003). The consensus sequence for this lipid modification is MG{EDRHPFYW}X_2[STAGCNDEF]{P} (Maurer-Stroh 2002), whereby the curly brackets indicate non-permissive amino acids, the square brackets essential amino acids, X any amino acid. Serine is favoured at position five of the motif, thus, most myristoylation motifs can be simplified as MGXXXS. In the myristoylation process a methionyl aminopeptidase first removes the initiator methionine (Farazi 2001) before the *N*-myristoyl transferase (NMT) catalyses the covalent attachment of the 14-carbon fatty acid myristate to the now N-terminal glycine residue (Casey 1995; Farazi 2001; Johnson 1994; Rajala 2000). This cotranslational modification is generally regarded as a constitutive process, has only been observed in eukaryotes and appears to be irreversible (Johnson 1994) though removal of the lipid from mature protein has been reported in *Dictyostelium* (da Silva 1990). Mutations that convert the glycine at position two to alanine (G2A) completely abolish myristoylation (Rajala 2000). Myristoylation is critical for mediating protein-protein and/or protein-membrane interactions for some proteins (Johnson 1994). For others, interactions with cell membranes require accessory factors or other covalent modification. In addition, an attached myristoyl can also serve a structural role in proteins (Ames 1994; Zheng 1993). An N-terminal peptide of mouse Irga2 (in that

study misleadingly called human homolog of IIGP) was myristoylated in an *in vitro* assay (Maurer-Stroh 2004) and Irga6 is efficiently N-terminally myristoylated *in vivo* (Martens 2004b). Furthermore, the N-terminal 68 amino acids of Irga6 targeted EGFP to endomembranes in a myristoylation dependent manner (Martens 2004b). In contrast, lipid modification is largely dispensable for the membrane association of full length Irga6. Thus, Irga6 must employ other, yet unclarified, mechanisms for membrane association in uninfected cells. Ten other mouse IRG proteins also contain myristoylation sequences (see Table 7 in result section) but their subcellular localisation and myristoylation *in vivo* have not yet been explored. Irgm1 localises to the Golgi as a consequence of an amphipathic helix in its C-terminus (αK in the Irga6 crystal structure, Figure 4) (Martens 2004b) and Irgm2 and Irgm3 to the Golgi and ER, respectively, through the analogous helix, which in these cases lacks amphipathic character (Martens 2006; Martens 2004b; Taylor 1997). Irgm3 further localised to circular, vesicle-like cytoplasmic structures of unknown origin (Taylor 1997). The degree of membrane association varies among the IRG proteins. Irgm1 and Irgm3 are both >90% membrane bound, while Irga6, Irgb6 and Irgd are 60–70%, 20–30% and <10% membrane bound, respectively (Martens 2004b). The precise subcellular localisation of the non-cytosolic fraction of Irgb6 and Irgd remains to be determined.

I.4.2.4.4. Function of the IRG proteins

Only two interaction partners of IRG proteins have been reported to date. First, Irga6 interaction with the microtubule motor–binding protein, Hook-3, was detected in a yeast two hybrid (Y2H) screen and in cell lysates (Kaiser 2004). The interaction was abolished with the Irga6(S83N) mutant, which was assumed to be nucleotide-binding deficient by analogy to Irgm3(S98N) (Taylor 1997). Thus, the Irga6-Hook3 interaction is nucleotide dependent. The abrogation of the co-immunoprecipitation of Irga6 with Hook3 from IFNγ-induced cells by GDPβS (Kaiser 2004) strongly argues for GTP as the relevant nucleotide. It was proposed that Irga6 might influence cytoskeleton-based membrane trafficking via this route. Second, the rat homologue of Irgd was shown to interact with adipocyte differentiation-related protein (ADRP) in Y2H, a fatty acid binding protein that coats lipid droplets and is involved in their formation and in lipid uptake (Yamaguchi 2006). The interaction was confirmed in glutathione-S-transferase (GST)-pulldown and mapped to the C-terminal half of Irgd (aa 246-420). Thus, Irgd might be involved in cellular lipid homeostasis, though the interaction remains to be confirmed in mammalian cells.

Over the last years a number of IRG proteins have been implicated in resistance to intracellular protozoal and bacterial pathogens mainly by studying mice with targeted deletions of single family members (Table 1). Irgm1-deficient mice were highly susceptible to infection with *Toxoplasma gondii* (Collazo 2001), *Leishmania major* (Taylor 2004), *Trypanosomas cruzi* (Santiago 2005), *Chlamydia trachomatis* (Coers manuscript in preperation), *Listeria monocytogenes* (Collazo 2001), *Salmonella typhimurium* (Taylor 2007; Taylor 2004), *Mycobacterium tuberculosis* (MacMicking 2003) and *M. avium* (Feng 2004). In most cases, Irgm1$^{-/-}$ mice succumbed to infection with kinetics similar to that of IFNγ-deficient mice. The other IRG proteins tested so far

have a narrower pathogen specificity. Irgm3-deficient mice lost resistance against *Leishmania major* (Taylor 2004), *Toxoplasma gondii* (Taylor 2000), *Chlamydia trachomatis* (Coers manuscript in preparation) and *C. psittaci* (Miyairi 2007). Irgm3 was shown to be required in hematopoietic as well as non-hematopoietic cells for resistance to *Toxoplasma in vivo* (Collazo 2002). Irgd, in contrast, displayed only a partial loss of resistance to *T. gondii* in the chronic phase of the infection (Collazo 2001). IRG function in pathogen resistance is non redundant since Irgm3, Irgd and Irgm1 all independently contribute to the resistance of mice to the protozoan parasite *Toxoplasma gondii*.

	wt	IFNγ	Irgm1	Irgm3	Irgd	Irga6	Irgm2	Irgb10
Intracellular protozoa								
Leishmania major	**R**	**S**	**S**	**S**	ND	**R**	ND	ND
Toxoplasma gondii	**R**	**S**	**S** $S_{ko\text{-}M\Phi}$	**S#** $S_{ko\text{-}astroc/M\Phi}$	**S*** $R_{ko\text{-}M\Phi}$	**R** $S_{ko\text{-}astrocy}$	ND	ND
Trypansomoma cruzi	**R**	**S**	**S** $S_{ko\text{-}M\Phi}$	**R**	S_{RNAi}	ND	ND	ND
Plasmodium chabaudi	**R**	**S**	ND	**R**	ND	ND	ND	ND
Intracellular bacteria								
Chlamydia trachomatis	**R** R_{MEFs}	**S**	**S** $S_{ko\text{-}MEFs}$	**S** $S_{ko\text{-}MEFs}$	ND	**R** $R_{ko\text{-}MEFs}$ $S_{RNAi\text{-}MECs}$	R_{RNAi}	$S_{RNAi\text{-}MEFs}$
Chlamydia psittaci	**R**	**S**	ND	**S**	ND	ND	**S_{strain}** S_{RNAi}	**S_{strain}**
Listeria monocytogenes	**R**	**S**	**S**	**R**	**R**	**R**	ND	ND
Mycobacterium tuberculosis	**R**	**S**	**S** $S_{ko\text{-}M\Phi}$	**R**	**R**	ND	ND	ND
Mycobacterium avium	**R**	**S**	**S** $R_{ko\text{-}M\Phi}$	**R**	ND	ND	ND	ND
Salmonella typhimurium	**R**	**S**	**S** $S_{ko\text{-}cells}$	**R**	**R**	ND	ND	ND

Table 1 IFNγ and IRG proteins in pathogen resistance. Summary of the susceptibility phenotypes of mice or cells (indicated in subscript) genetically (ko) or experimentally deficient for the indicated genes. ND: not determined; S: susceptible; R: resistant; strain: susceptible mouse strains; MΦ: macrophages; MEFs: mouse embryonic fibroblasts; astroc: astrocytes; MECs: mouse oviduct epithelial cells; RNAi: expression suppressed by RNA interference; * in the chronic phase only; # in hematopoietic and non-hematopoietic cells *in vivo*: Note: weak antiviral effects observed in cellular overexpression studies for Irgm3 and Irgb6 in cellular are not listed here. See main text for more information and for references.

Recently, two studies linked mouse strain specific differences in resistance to *Chlamydia trachomatis* and *C. psittaci* to differential expression of Irgb10 (Bernstein-Hanley 2006; Miyairi 2007). In the case of *C. psittaci* Irgm2 was also involved in determining the resistance differences (Miyairi 2007). Irga6-deficient mice developed cerebral malaria with higher incidence than wild type mice following infection with *Plasmodium berghei* but displayed normal resistance to *Chlamydia trachomatis* and *Toxoplasma gondii* infection as well as to various other pathogens tested (Coers manuscript in preparation)(Martens 2006; Martens 2005; Parvanova 2005). In most cases, there is a strong correlation between loss of resistance in IRG-deficient mice and loss of IFNγ mediated pathogen control in cultured host cells, indicating that regulation of innate pathogen killing is a cell autonomous function of the IRG proteins. Thus, Irgm1 and Irgm3 are required for control of *T. gondii, Salmonella typhimurium* and *Legionella* growth in macrophages (Butcher 2005; Ling 2006; Taylor 2007; Taylor 2004)(Coers and Dietrich personal communication), Irgm3 and Irga6 for control of *T. gondii* in astrocytes (Halonen 2001; Martens 2005), Irgm1 for control of *M. tuberculosis* and *T. cruzi* in macrophages (MacMicking 2003; Santiago 2005), Irgm3 for control of *Chlamydia trachomatis* in fibroblasts (Bernstein-Hanley 2006) and Irgd for control of *T. cruzi* in

macrophages (Koga 2006) (Table 1). Thus, Irga6 contributes to the cell autonomous resistance against *T. gondii* although Irga6-deficient mice were not significantly more susceptible than wild type mice (Martens 2005). Irga6 was reported to participate in *C. trachomatis* resistance in mouse oviduct epithelial cells (Nelson 2005) but was dispensable in mouse embryonic fibroblasts (MEFs) (Coers manuscript in preparation).

Few studies have been performed trying to elucidate the mechanisms by which IRGs mediate their non-redundant function in pathogen resistance on a cellular level. Irgm1 and Irgm3 have been shown to localise rapidly to phagosomes containing latex beads in activated macrophages as the vacuoles form and remain associated as the vacuoles mature (Butcher 2005; Martens 2005). The mechanism by which the proteins move from their resting localisation at the ER/Golgi or in the cytoplasm to the vacuole is not clear but was shown to require a functional GTP binding domain in the case of Irgm1 (Martens 2004b). Furthermore, defective killing of *Mycobacterium tuberculosis* in IFNγ-activated Irgm1-deficient macrophages was associated with impaired maturation of bacteria-containing phagosomes that otherwise recruited Irgm1 in wild-type cells (MacMicking 2003). Autophagy has subsequently been implicated in the Irgm1-mediated enhanced maturation of *M. tuberculosis* containing phagosomes (Gutierrez 2004). Five IRG proteins (Irga6, Irgb6, Irgm2, Irgm3 and Irgd) have been shown to concentrate at the parasitophorous vacuolar membrane in IFN-induced, *T. gondii*–infected primary astrocytes and macrophages (Ling 2006; Martens 2005). Irgm1, however, does not seem to localise to the *T. gondii* vacuole (Butcher 2005; Martens 2005), yet it is essential to eradicate the pathogen. The molecular mechanisms by which IRG proteins promote *T. gondii* destruction are not clear though both Irga6 and Irgm3 have been shown to contribute to the vesiculation and disruption of the parasitophorous vacuolar membrane (PVM) and the parasite itself (Ling 2006; Martens 2005). Subsequent engulfment of the parasite into an autophagosome-like vacuole ultimately fusing with lysosomes seems to be involved in either the killing or the disposal of the dead parasite (Ling 2006). It is not clear yet whether IRG proteins play a direct role in the disruption and the associated membrane vesiculation. In addition to controlling the fate of pathogens in IFNγ activated hosts cells, Irgm1 has also been implicated in the regulation of haematopoiesis following chronic infection. Irgm1-deficient mice infected with *M. tuberculosis* developed profound lymphopenia, anaemia and thrombocytopenia that coincided with increased bacterial burdens and ultimately decreased survival (Feng 2004). Similarly, infection with *M. avium* resulted in a decreased expansion of haematopoietic progenitors (Taylor 2007) and *T. cruzi* infection led to a profound lymphopenia and bone marrow failure in Irgm1-deficient mice (Santiago 2005). The mechanism by which the absence of Irgm1 affects haematopoiesis has not been defined and the molecular basis for the heightened susceptibility has not been disentangled.

At present, the role of the IRG GTPases in viral resistance is not as clear as their roles in protozoan and bacterial resistance. Two *in vitro* studies suggest that Irgb6 and Irgm3 may have antiviral properties: Overexpression of Irgb6 in mouse fibroblasts decreased vesicular stomatitis virus- but not herpes simplex virus-induced cell lysis (Carlow 1998), while overexpression of Irgm3 in HeLa cells diminished coxsackie virus-induced lysis (Zhang 2003). In the latter study, it was suggested that Irgm3 mediates its antiviral effect by both directly blocking viral induced lysis, and modulating viral-

induced apoptosis (Zhang 2003). Currently, there is no data supporting a function of IRG proteins in viral resistance *in vivo.* Resistance of Irgm1-, Irgm3- and Irgd-deficient mice to murine cytomegalovirus (Collazo 2001; Taylor 2000) and of Irgm3-deficinet mice to Ebola virus was normal (Taylor 2000).

In summary, multiple IRG proteins have been implicated in the cell autonomous resistance to a wide variety of intracellular bacterial and protozoal pathogens. Yet, the resistance to only one pathogen, *Toxoplasma gondii,* involved all IRG family members tested thus far. Hence, it is of central importance to understand the biology of this protozoan to gain further insight into the antimircobial effects mediated by the IRGs.

I.5. Toxoplasma gondii

Suppression of both the acute and chronic phase of *Toxoplasma gondii* infection *in vivo*, as well as in cultured cells, requires IFNγ-induced responses (Yap 1999b) and all IRG proteins analysed so far contribute to pathogen resistance (Martens 2006). Hence, this is an excellent system to assess IRG protein function. *Toxoplasma gondii* is a unicellular, obligate intracellular protozoan parasite belonging to the phylum *Apicomplexa*, many members of which are human or animal pathogens (Levine 1988). *T. gondii* exhibits an extremely broad host range, capable of infecting most nucleated cells of virtually any warm blooded animal (Black 2000; Sibley 2003). Apicomplexan parasites are bounded by the pellicle, a composite structure consisting of the plasma membrane and the closely apposed inner membrane complex, which is comprised of two layers of flattened vesicles (Figure 5). They contain, in addition to the nucleus, mitochondria, Golgi and ER, a unique set of apical organelles called the apical complex and a chloroplast-like organelle, the apicoplast (reviewed in (Binder 2004; Black 2000; Vaishnava 2006). The apical complex consists of conoid, rhoptries and micronemes and is essential for invasion and proliferation. The role of the apicoplast is uncertain but recent results indicate synthesis of lipids, heme and isoprenoids as possibilities.

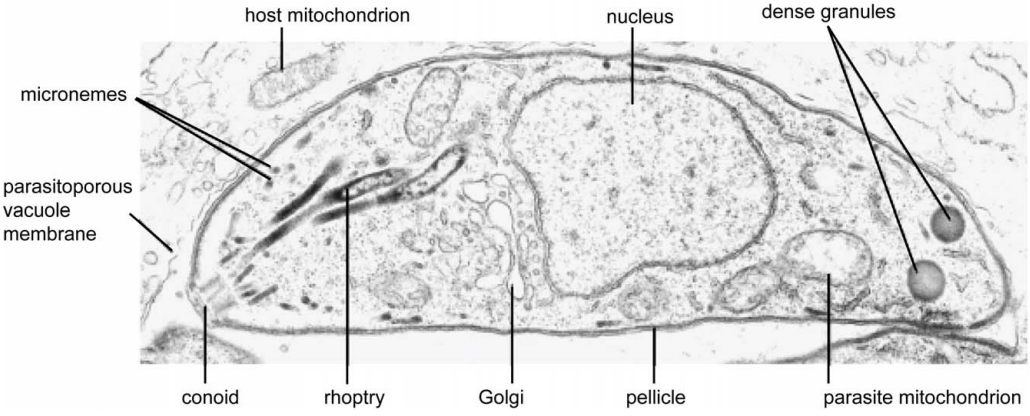


Figure 5 Transmission electron micrograph of a *T. gondii* tachyzoite (RH strain) within a host cell. *T. gondii* tachyzoites are polarised cells with organelles common to all eukaryotes (mitochondria, nucleus and Golgi complex). The micronemes, rhoptries and dense granules are secretory organelles specific to the phylum *Apicomplexa*. Sequential release of organellar contents is thought to be essential for host cell invasion. The triple membrane surrounding the parasite is composed of a plasma membrane and two membranes that form vesicular structures called the inner membrane complex. The parasite lives within a parasitophorous vacuole that is lined with host mitochondria and ER. The vacuole is modified by secreted contents of rhoptries and dense granules. The conoid is a cytoskeletal structure at the apex of the tachyzoite (from (Binder 2004)).

The life cycle of *Toxoplasma* is unusual in that the organism is capable of indefinite replication using either sexual or asexual subcycles (Boothroyd 2002). The asexual cycle can occur in any infected animal and consists of two stages, the rapidly multiplying tachyzoites and the more slowly dividing bradyzoites, which can encyst in brain, heart and other tissues. Acute infection, associated with the tachyzoite stage, is normally controlled in the immunocompetent host (Lieberman 2002; Suzuki 1988). However, in the chronic phase of infection the cysts in which the bradyzoite state can persist are apparently refractory to the robust immune response induced by tachyzoites (Sims 1989; Weiss 2000). Persistent bradyzoites can initiate a new infection upon ingestion of bradyzoite-infected tissue by a predator or scavenger. Suppression of both acute and chronic phase of infection *in vivo,* as well as in cultured cells, requires IFNγ-induced responses (Yap 1999a).

Sexual reproduction of *Toxoplasma* occurs only in feline species. It includes full gametogenesis and mating within the intestinal epithelium and culminates in the generation of millions of extremely stable and highly infectious oocysts containing haploid sporozoites that are shed in the faeces (Boothroyd 2002). Genetic exchange between different strains can occur only in the rare event of a feline becoming infected simultaneously with more than one strain. There is no predetermined mating type and a single organism can complete the entire life cycle (Pfefferkorn 1977).

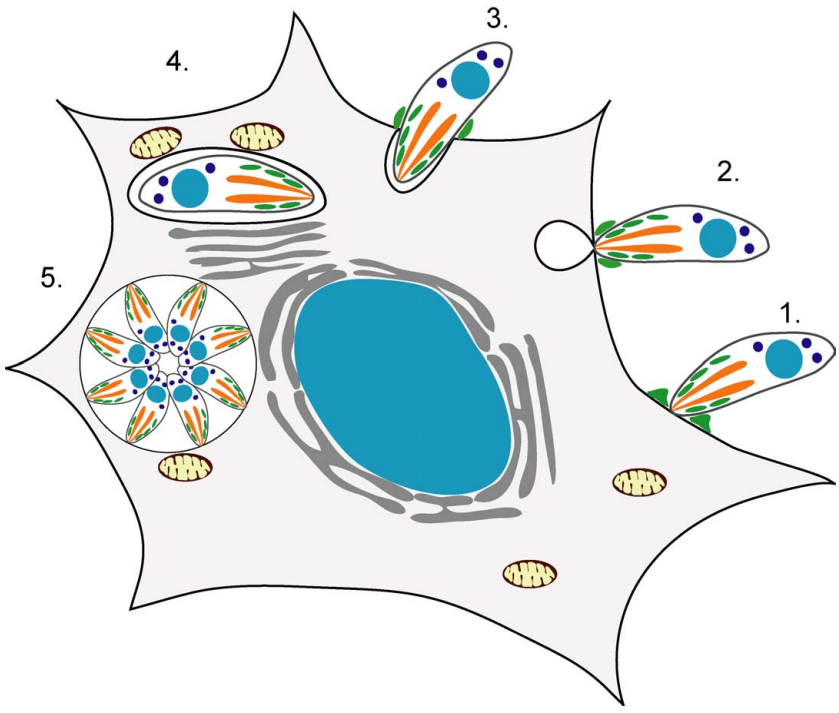


Figure 6 *Toxoplasma gondii* host cell invasion and intracellular replication. Invasion is a rapid process that is accompanied by sequential secretion of micronemes (green), rhoptries (orange) and dense granules (dark blue). (1) Recognition of and attachment to the host cell surface by the parasite is mediated by secreted microneme proteins. (2) Rhoptry neck and rhoptry protein secretion. MIC and RON proteins form a moving junction, seen as a constriction in the parasite as it enters host cells, that is thought to exclude host proteins from the nascent parasitophorous vacuole. (3) Entry of the parasite using its actin-myosin motor invaginating the host cell plasma membrane to form the PV. Microneme proteins are excluded form the PV. (4) Dense granule secretion occurs after parasite entry is complete. The parasite remodels the intracellular vacuole inhibiting lysosomal fusion and recruiting host mitochondria and portions of the host ER probably to secure energy and nutrition supply. (5) The parasites undergo several rounds of replication by endodyogeny inside the PV before lytic egress from the host cell (not shown). The released parasites quickly invade neighbouring cells completing the cycle. Note that the figure is not to scale.

Several studies have concluded that, in Europe and North America, the majority of *T. gondii* strains comprises three distinct clonal lineages now known as types I, II and III, respectively (Grigg 2001; Howe 1995; Sibley 1992) that differ genetically by 1% or less (Saeij 2005; Su 2003). Detailed genetic analysis showed that within type variation is extremely rare (Fazaeli 2000; Grigg 2001; Lehmann 2000), except at the highly polymorphic microsatellite markers (Ajzenberg 2002; Blackston 2001). Hence, it is clear that these three lines have emerged relatively recently without much opportunity for genetic drift. Population genetic analyses suggest that recombination(s) between two discrete ancestral gene pools produced recombinant progeny, of which a small number have recently come to dominate over most other strains in many parts of the world. As a result, only two alleles exist for a majority of genes in a majority of strains.

In mice, *T. gondii* strains differ enormously in their virulence, by up to 4 orders of magnitude, with striking mouse strain specific differences (Araujo 1976; Boothroyd 2002; Howe 1995). Almost all of the virulent strains typified as RH on the basis of the well-characterised RH strain originally isolated by Sabin in 1939 (Sabin 1941) are a single clonal genotype (type I), whereas the vast majority of the non-virulent lines are type II or type III (reviewed in (Howe 1995; Saeij 2005)). Even though association of certain *T. gondii* strains with specific manifestations of toxoplasmosis (like ocular toxoplasmosis) has been documented, no equivalent virulence difference of strains can be found in human (Boothroyd 2002).

The process of host cell invasion is an active parasite-mediated process involving the parasite's cytoskeleton and differs from the facilitated endocytosis seen in bacterial or viral infections (Dobrowolski 1996). Entry into the host cell is a rapid process completed in about 15 to 20 s. During invasion of host cells, *T. gondii* directs sequential secretion of proteins from the three specific types of secretory organelles of the apical complex: the micronemes, rhoptries and dense granules (Figure 5, Figure 6) (Binder 2004; Carruthers 1999). During attachment to the host plasma membrane, microneme proteins (MICs) are secreted and mediate recognition of, and attachment to, the host cell and provide the force for host plasma membrane invagination (Carruthers 1997; Dobrowolski 1996). During invasion, host proteins are largely excluded from the nascent vacuole by establishment of a moving junction containing microneme and rhoptry neck proteins (RONs) (Alexander 2005; Mordue 1999a). These results in a dramatic remodelling of the intracellular vacuole the parasite resides in, termed parasitophorous vacuole (PV), which is primarily composed of host cell lipids (Suss-Toby 1996). Upon contact with the host cell rhoptry proteins (ROPs) involved in the formation and modification of the nascent parasitophorous vacuole are secreted from the club-shaped rhoptries at the apical tip of the parasite (Carruthers 1997; Hakansson 2001). ROPs either associate with the PVM or are transported across into the host cell cytoplasm (Bradley 2005; Gilbert 2007). Recently two rhoptry kinases, ROP16 and ROP18, have been identified as important virulence factors accounting at least for part of the tremendous virulence differences of *Toxoplasma* strains, whereby ROP18 seems to function at the PV membrane and ROP16 in the host cell nucleus (El Hajj 2006; Saeij 2006; Saeij 2007; Taylor 2006). The mechanisms by which these kinases enhance pathogen virulence remain to be clarified. Once safely inside the host cell, the parasite secretes proteins from dense granules (GRAs) that also modify the vacuolar membrane and associate with a vesicular network

within the PV (Carruthers 1997; Dubremetz 1993; Sibley 1995). The modification of the vacuole by rhoptry contents (e.g. ROP2; (Carey 2004; Nakaar 2003; Sinai 2001) and dense granule proteins (e.g. Gra3; (Henriquez 2005)) results in a vacuole that is nonfusogenic, not acidified, and associated with host mitochondria and endoplasmic reticulum (Jones 1972; Sinai 1997). This prevents degradation of the parasite by host cell lysosomes (Mordue 1999b; Mordue 1997) and secures access to host cell nutrients, facilitating parasite replication within this protected intracellular niche (Carruthers 1997; Hakansson 2001; Sibley 2004; Sinai 2001). Inside the parasitophorous vacuole *T. gondii* replicates by endodyogeny, a unique form of budding, in which daughter cells are assembled within a mother cell (Figure 6) (Goldman 1958; Wildfuhr 1966). When the daughter cells are fully mature, they bud from the mother, adopting her plasma membrane. The cycle is repeated inside the PV until the infected cell eventually bursts and tachyzoites are released.

I.6. The aim of this study

It has been shown that the IRG resistance system is essential for survival in a world rife with pathogens. Furthermore, profound host species-specific and pathogen strain-specific differences in susceptibility and employed resistance mechanisms against certain pathogens like *Toxoplasma gondii* have been reported (MacKenzie 1999). Thus, a systematic phylogenetic analysis was undertaken to clarify the age and distribution of IRG genes in the animal kingdom. It was also analysed whether the IRG family displayed signs of recent expansion and diversification in any of the analysed species as well as contraction in other species – features that are considered to be characteristic for multigene families associated with pathogen resistance due to host-pathogen coevolution (Angata 2004; Borghans 2004; Delarbre 1992; Hood 1975; Kumar 2000; Leister 2004; Mashimo 2003; Noel 1999; Trowsdale 2001)}.

Though there is much important data underscoring the essential, cell autonomous, non-redundant function of mouse immunity-related GTPases in host resistance to intracellular pathogens, at the beginning of this study no investigations of the molecular bases of the resistance mechanism had been performed. Only one study had analysed the effect of a single IRG protein, Irgb6, on pathogen replication, in this case VSV and HSV-1 (herpes simplex virus type 1), in absence of IFNγ-induction (Carlow 1998). In contrast to the complete inhibition of viral replication by MxA in a similar overexpression setting (Haller 2007a), the documented effect was small. Furthermore, the observed antiviral effect was marginal compared to the resistance phenotypes documented in bacterial and protozoal infections (compare (Carlow 1998) to (Collazo 2001)). Most of the more recent functional studies were performed with cells derived from IRG-deficient mice, where only the influence of the absence of single IRGs can be assessed but not the influence of individual family members on pathogens, cellular functions and other IRG members. Moreover, no link between the biochemical properties of the IRG proteins and their resistance function *in vivo* has been drawn to date and nothing is known about the regulation of IRG GTPase activity and function in cells.

To elucidate the cell biology and molecular mechanisms of host resistance mediated by IRG proteins on a single cell level and to disentangle the effect mediated by

single IRG proteins from the rest of the IFNγ response murine fibroblasts, stably and hormone-inducibly expressing single IRGs were established. Furthermore, co-expression studies were used to reveal possible interplay between family members. Mutants with impaired biochemical properties were characterised and used to link GTPase activity to protein function *in vivo*. *Toxoplasma gondii* was used as a model for infection as all IRG proteins analysed so far contributed to resistance against this pathogen (reviewed in (Martens 2006)).

II. Material and Methods

II.1. Material

II.1.1. Mammalian cells and media

Gs3T3 mouse fibroblasts (Invitrogen; derived from NIH/3T3 cells ((Jainchill 1969) ATCC CRL-1658) by stable transfection with pSwitch vector (Invitrogen)) and murine embryonic fibroblasts (MEFs) derived from C57/BL6 mice were cultured in high glucose DMEM (Gibco BRL). Both media were supplemented with 10% FCS (Sigma-Aldrich), 2 mM L-glutamine, 1 mM sodium pyruvate, 1x MEM non-essential amino acids, 100 U/ml penicillin and 100 µg/ml streptomycin (all Gibco BRL). Human foreskin fibroblasts (Hs27, ATCC CRL-1634) were cultured in IMDM supplemented with 5% FCS and 2 mM L-glutamine. Sterile trypsin/EDTA solution in PBS (10x trypsin/EDTA solution: 0.05% (w/v) trypsin (1:250, Gibco BRL)/17 mM EDTA/145 mM NaCl)) was used to detach adherent cells from culture dishes (TRP) or flasks (Sarstedt).

II.1.2. Bacterial strains and media

E. coli strain	Genotype	Reference
K-12 DH5α	*80dlacZ*ΔM15, recA1, endA1, gyrA96, thi-1, hsdR17 (r_B^-, m_B^+), supE44, relA1, deoR, Δ*(lacZYA-argF)*U169)	(Bachmann 1983; Bachmann 1990)
BL21 (DE3)	B, F^-, *hsdS* (r_B^-, m_B^-), *gal*, *dcm*, *ompT*, λ(DE3)	(Studier 1986)

For bacterial culture LB (Luria-Bertani) medium (1 % bacto tryptone (w/v), 0.5 % yeast extract (v/w), 1 % NaCl (w/v) (Bertani 1951), for plates including 1.5 % (w/v) bacto agar) was used. For selection ampicillin or kanamycin were added to the medium at a final concentration of 100 µg/ml and 50 µg/ml respectively (all reagents from Roth).

II.1.3. Yeast strains and media

Strain	Genotype	Reference
PJ69-4A	*MATa* trpl-901 leu2-3,112 ura3-52 *his3-200* ga14Δ ga18O Δ LYSZ::GALl-HIS3 GAL2-ADE2 met2::GAL7-lacZ	(James 1996)
PJ69-4a	*MATα* trpl-901 leu2-3,112 ura3-52 *his3-200* ga14Δ ga18O Δ LYSZ::GALl-HIS3 GAL2-ADE2 met2::GAL7-lacZ	(James 1996)

Yeast cells were either cultured in complete medium (YPD: 1% (w/v) Yeast extract (MP Biochemicals), 2% (w/v) Peptone (Roth), 2% (w/v) Glucose (Roth), for plates including 2% (w/v) Bacto agar (Roth)) or in synthetic defined (SD) medium (Wickerham 1951) (2% (w/v) α-D-glucose (Roth), 0.67% (w/v) yeast nitrogen base including ammonium sulfate (Difco), 0.11 mM arginine, 0.08 mM isoleucine, 0.27 mM lysine, 0.07 mM methionine, 0.36 mM phenylalanine, 0.42 mM L-threonine, 0.36 mM uracil, 0.46 mM L-leucine, 0.20 mM tryptophan, 0.15 mM adenine and 0.06 mM histidine (all Merck), for plates including 2% (w/v) Bacto agar (Roth)). Depending on the required selection leucine, tryptophan, adenine and/or histidine were omitted from the medium. SD plates lacking histidine were supplemented with 1 mM 3-amino-1,2,4-triazole (Sigma-Aldrich) to decrease the leakiness of the *HIS3* reporter (Durfee 1993).

II.1.4. *Toxoplasma gondii* strains

ME49: type II strain, avirulent, originally isolated from sheep muscle (California, USA) in 1965 (Guo 1997).

RH: type I strain, virulent, originally isolated in 1939 by Albert B. Sabin from a child in Ohio (Sabin 1941), stable YFP expressing transgenic RH strain (2F-1 YFP_2, termed RH-YFP in this study) generated by Gubbels (Gubbels 2003).

II.1.5. Vectors

The following expression vectors were used: pSwitch, pGene/V5-HisA (both Invitrogen), pGW1H (British Biotech, Oxford, UK) and pEGFP-C3 (Clontech) for mammalian cells, pGEX-4T-2 (GE Healthcare Life-Sciences) for bacteria and pGAD-C(x) and pGBD-C(x) (James 1996) for yeast cells (see appendix V.2 for vector maps).

II.1.6. Generation of expression constructs

The coding regions of Irgm1, -m2, -m3, -a6, -d, and Irgb6 were derived from PCR amplification of full-length cDNAs from IFN-γ-stimulated mouse embryonic fibroblasts according to standard procedures (Boehm 1998). The open reading frame (ORF) of Irgc was obtained by PCR from a bacterial artificial chromosome (BAC) (RP23-57J6; Rohde diploma thesis) derived genomic fragment. The PCR fragments with added SalI restriction sites on both ends and a Kozak consensus sequence (Kozak 1987) upstream of the start codon were cloned with restriction enzymes from New England Biolabs into the SalI site of the following expression vectors: pGW1H (British Biotech, Oxford, UK), pGEX-4T-2 (GE Healthcare Life-Sciences), pEGFP-C3 (BD Biosciences), pGAD and pGBD (James 1996) (Table 2-4; for vector maps see appendix V.2). The native open reading frames of the IRG genes (appendix V.3) were cloned into pGene/V5-HisA as follows: Irgm1 and Irga6 by BamHI-NotI digestion from the pGEX-4T-2-IRG constructs and Irgm2, Irgm3, Irgd and Irgb6 by EcoRI-ApaI restriction from the respective pEGFP-C3 constructs.

Mutations in the first nucleotide binding motif (G1) and epitope tags were introduced by site directed mutagenesis (see II.3.9). The ctag1 C-terminal epitope tag replaces the last two residues (RN) of Irga6 with the sequence KLGRLERPHRD (formerly IIGP1-m, (Uthaiah 2003). In case of Irgd-ctag1 the same 13 residues were added after the last amino acid of the protein. The FLAG-tagged constructs carry the sequence DYKDDDDK at the C-terminus. The bacterial expressed Irga6 proteins all carry the extension GSPGIPGSTT at the N terminus due to cleavage of the GST fusion by thrombin. All constructs were verified by sequencing.

vector	insert	N-terminus	C-terminus	source
pGex-4T-2	Irga6	*GST*-LVPRGSPGIPGSTT**MGQL**	**CLRN**	R Uthaiah
	Irga6(K82A)	*GST*-LVPRGSPGIPGSTT**MGQL**	CLRN	R Uthaiah
	Irga6(S83N)	*GST*-LVPRGSPGIPGSTT**MGQL**	**CLRN**	This study

Table 2 Bacterial expression constructs. Amino acid sequences derived from the respective ORFs are depicted in bold, the fused glutathione S-transferase (GST) protein is indicated in italics and the thrombin cleavage site is highlighted in blue. The enzyme cuts between the arginine and the glycine residue.

vector	insert	N-term..	C-terminus	source
pGW1H	Irgc	**MATS**	**SEDK**	C. Rhode
	Irgm2	**MEEA**	**IPHP**	U. Böhm
	Irgm2(S78N)	**MEEA**	**IPHP**	N. Schröder
	Irgm3	**MDLV**	**PEIH**	S. Martens
	Irgm3(S98N)	**MDLV**	**PEIH**	N. Schröder
	Irgm1	**MKPS**	**PPQI**	S. Martens
	Irgm1(S90N)	**MKPS**	**PPQI**	S. Martens
	Irga6	**MGQ**	**CLRN**	U. Böhm
	Irga6-Ctag1	**MGQL**	**CL**KLGRLERPHRD	N. Papic
	Irga6(K82A)	**MGQL**	**CLRN**	R. Uthaiah
	Irga6(K82A)-Ctag1	**MGQL**	**CL**KLGRLERPHRD	N. Papic
	Irga6(S83N)	**MGQL**	**CLRN**	This study
	Irga6(S83N)-Ctag1	**MGQL**	**CL**KLGRLERPHRD	This study
	Irga6(E106R)-Ctag1	**MGQL**	**CL**KLGRLERPHRD	N Pawlowski, N Papic
	Irga6(173A)-Ctag1	**MGQL**	**CL**KLGRLERPHRD	R. Uthaiah, N. Papic
	Irga6(K196D)-Ctag1	**MGQL**	**CL**KLGRLERPHRD	N Pawlowski, N Papic
	Irgd	**MDQF**	**VNVA**	Martens, Luetke
	Irgd-Ctag1	**MDQF**	**VNVA**KLGRLERPHRD	This study
	Irgb6	**MAWA**	**YWEA**	U. Böhm
	Irgb6-FLAG	**MAWA**	**YWEA**	S. Martens
	Irgb6(K69A)-FLAG	**MAWA**	**YWEA**	S. Martens
	Irgb6(S70N)-FLAG	**MAWA**	**YWEA**	This study
pGene/V5-HisA	Irgc	**MATS**	**SEDK**	This study
	Irgm2	**MEEA**	**IPHP**	This study
	Irgm3	**MDLV**	**PEIH**	This study
	Irga6	**MGQL**	**CLRN**	This study
	Irgd	**MDQF**	**VNVA**	This study
	Irgm1	**MKPS-**	**PPQI**	This study
	Irgb6	**MAWA**	**YWEA**	This study

Table 3 Mammalian expression constructs. Amino acid sequences derived from the respective ORFs are depicted in bold, epitope tags in regular letters and alterations from the native sequence are highlighted in red.

vector	insert	N-terminus	C-terminus	source
pGAD-C1	Irgc	*GAD*-IEFPGGSIDVD**MATS**	**SEDK**	C Rhode
	Irgc (K65A)	*GAD*-IEFPGGSIDVD**MATS**	**SEDK**	This study
	Irgc (S66N)	*GAD*-IEFPGGSIDVD**MATS**	**SEDK**	This study
pGAD-C3	Irgm2	*GAD*-IEFPGDPSMSTT**MEEA**	**IPHP**	This study
	Irgm2(S78N)	*GAD*-IEFPGDPSMSTT**MEEA**	**IPHP**	This study
	Irgm3	*GAD*-IEFPGDPSMSTT**MDLV**	**PEIH**	This study
	Irgm3(S98N)	*GAD*-IEFPGDPSMSTT**MDLV**	**PEIH**	This study
	Irgm1	*GAD*-IEFPGDPSMSTT**MKPS**	**PPQI**	This study
	Irgm1(S90N)	*GAD*-IEFPGDPSMSTT**MKPS**	**PPQI**	This study
	Irga6	*GAD*-IEFPGDPSMSTT**MGQL**	**CLRN**	This study
	Irga6(K82A)	*GAD*-IEFPGDPSMSTT**MGQL**	**CLRN**	This study
	Irga6(S83N)	*GAD*-IEFPGDPSMSTT**MGQL**	**CLRN**	This study
	Irgd	*GAD*-IEFPGDPSMSTT**MDQF**	**VNVA**	This study
	Irgb6	*GAD*-IEFPGDPSMSTT**MAWA**	**YWEA**	This study
	Irgb6(K69A)	*GAD*-IEFPGDPSMSTT**MAWA**	**YWEA**	This study
	Irgb6(S70N)	*GAD*-IEFPGDPSMSTT**MAWA**	**YWEA**	This study
pGBD-C1	Irgc	*GBD*-PEFPGGSIDVD**MATS**	**SEDK**	C Rhode
	Irgc (K65A)	*GBD*-PEFPGGSIDVD**MATS**	**SEDK**	This study
	Irgc (S66N)	*GBD*-PEFPGGSIDVD**MATS**	**SEDK**	This study
pGBD-C3	Irgm2	*GBD*-PEFPGDPSMSTT**MEEA**	**IPHP**	This study
	Irgm2(S78N)	*GBD*-PEFPGDPSMSTT**MEEA**	**IPHP**	This study
	Irgm3	*GBD*-PEFPGDPSMSTT**MDLV**	**PEIH**	This study
	Irgm3(S98N)	*GBD*-PEFPGDPSMSTT**MDLV**	**PEIH**	This study

vector	insert	N-terminus	C-terminus	source
	Irgm1	*GBD*-PEFPGDPSMSTT**MKPS**-	**-PPQI**	This study
	Irgm1(S90N)	*GBD*-PEFPGDPSMSTT**MKPS**-	**-PPQI**	This study
	Irga6	*GBD*-PEFPGDPSMSTT**MGQL**-	**-CLRN**	This study
	Irga6(K82A)	*GBD*-PEFPGDPSMSTT**MGQL**-	**-CLRN**	This study
	Irga6(S83N)	*GBD*-PEFPGDPSMSTT**MGQL**-	**-CLRN**	This study
	Irgd	*GBD*-PEFPGDPSMSTT**MDQF**-	**-VNVA**	This study
	Irgb6	*GBD*-PEFPGDPSMSTT**MAWA**-	**-YWEA**	This study
	Irgb6(K69A)	*GBD*-PEFPGDPSMSTT**MAWA**-	**-YWEA**	This study
	Irgb6(S70N)	*GBD*-PEFPGDPSMSTT**MAWA**-	**-YWEA**	This study

Table 4 Yeast expression constructs. Amino acid sequences derived from the respective ORFs are depicted in bold, the fused Gal4 activation (AD) and DNA binding domains (BD) are indicated in italics.

II.1.7. Primers

side directed mutagenesis (SDM) primers (name/ 5'-3' sequence)	
Irga6(S83N) SDM fwd	ggagacgggatcagggaagAAcagcttcatcaataccctg
Irga6(S83N) SDM rev	cagggtattgatgaagctgTTcttccctgatcccgtctcc
Irgc(K65A) SDM fwd	gagtcgggagccggcGCgtcctccctcatcaatgc
Irgc(K65A) SDM rev	gcattgatgagggaggacGCgccggctcccgactc
Irgc(S66N) SDM fwd	gagtcgggagccggcaagAActccctcatcaatgc
Irgc(S66N) SDM rev	gcattgatgagggagTTcttgccggctcccgactc
Irgb6(K69A) SDM fwd	ggaaacaggcgcagggGCgtccactttcatcaatacc
Irgb6(K69A) SDM rev	ggtattgatgaaagtggacGCccctgcgcctgtttcc
Irgb6(S70N) SDM fwd	ggaaacaggcgcagggaagAAcactttcatcaatacc
Irgb6(S70N) SDM rev	ggtattgatgaaagtgTTcttccctgcgcctgtttcc
Irgd-ctag1 SDM fwd	gaaacagtaaatgttgccaaactaggccgactcgagcggccgcatcgtgactgagtggtcgacctgcagg
Irgd-ctag1 SDM rev	cctgcaggtcgaccactcagtcacgatgcggccgctcgagtcggcctagtttggcaacatttactgtttc
Irgm1(S90N) SDM fwd	gggactctggcaatggcatgAATtctttcatcaatgcacttcg
Irgm1(S90N) SDM rev	cgaagtgcattgatgaaagaATTcatgccattgccagagtccc
Irgm2(S78N)/ Irgm3(S98N) SDM fwd	gggactctggcaatggcatgAATtctttcatcaatgcccttagg
Irgm2(S78N)/ Irgm3(S98N) SDM	gggactctggcaatggcatgAATtctttcatcaatgcccttagg
sequencing primers (name/ 5'-3' sequence)	
pGW1H-5' (British Biotech)	ctttccatgggtcttttctg
pGW1H-3' (British Biotech)	tcaggggaggtgtgggagg
pGene Forward Primer (Invitrogen)	ctgctattctgctcaacct
pcDNA3.1/BGH Reverse Primer (Invitrogen)	tagaaggcacagtcgagg
IS837 (pGAD fwd primer)	ccactgtcacctggttggacgg
IS838 (pGAD/BD rev primer)	cacagttgaagtgaacttgcggg
KU1012 (pGBD fwd primer)	gtgcgacatcatcatcggaag

Table 5 List of primers. Primers were from Operon. Fwd: forward primer, rev: reverse primer. The sequences of the ctag1 epitope tag is highlighted in red and mutations introduced in the G1 motive of the IRGs are depicted in capital letters.

II.1.8. Primary immunoreagents

name	recognised antigen	type	dilution	source
39/3°	mouse Irgc peptides LVEKRSTGE-GTSEEA; YILDSWKRRD LSEDK	rabbit polyclonal	WB 1:10000 IF 1:500	Eurogentec
αIGTP clone 7	mouse Irgm3 (aa283-423)	mouse monoclonal	WB 1:2000 IF 1:250	BD Transduction Laboratories
165	recombinant mouse Irga6	rabbit polyclonal	WB 1:25000 IF 1:8000	(Martens 2004a; Uthaiah 2003)
10D7	recombinant mouse Irga6	mouse	WB 1:2000	Jens Zerrahn, Berlin

name	recognised antigen	type	dilution	source
	(4.3 mg/ml)	monoclonal	IF 1:500	(Martens 2005; Zerrahn 2002)
10E7	recombinant mouse Irga6 (2.6 mg/ml)	mouse monoclonal	WB 1:1000 IF 1:200	J. Zerrahn, Berlin (Martens 2005; Zerrahn 2002)
2078	mouse Irgd peptides CKTPYQHPK-YPKVIF; CDAKHLLRKIETVNVA	rabbit polyclonal	WB 1:1000 IF 1:500	Eurogentec
L115 B0	mouse Irgm1 peptides QTGSSRLP-EVSRSTE, NESLKNSLGVRDDD	rabbit polyclonal	WB 1:2000	Eurogentec
A19	mouse Irgm1 N-terminal peptide	goat polyclonal	WB 1:200 IF 1:100	Santa Cruz Biotechnology
A20	mouse Irgb6 N-terminal peptide	goat polyclonal	WB 1:500 IF: 1:200	Santa Cruz Biotechnology
H53	mouse Irgm2 N-term. peptide MEEAVESPEVKEFEY	rabbit polyclonal	WB 1:1000 IF 1:1000	Eurogentec
αIrgb10	mouse Irgb10 C-terminal peptide LKKKVFLQDSVDSE	rabbit polyclonal	WB 1:4000 IF 1:2000	(Miyairi 2007); Coers, Harvard, Boston
2600	ctag1 peptide KLGRLERPHRD	rabbit polyclonal	WB 1:12000 IF 1:5000	Eurogentec
SPA-865	canine Calnexin N-terminal peptide	rabbit polyclonal	WB 1:5000 IF 1:250	StressGene
G65120	C-terminus of rat GM130	mouse monoclonal	IF 1:1000	BD Transduction Laboratories
αGiantin	human Giantin	mouse monoclonal	IF 1:1000	H. P. Hauri, Basel
FK2	Ubiquitin	mouse monoclonal	WB 1:1000	BioTrend
αToxo-plasma	*Toxoplasma gondii* (strain C56)	rabbit polyclonal	IF 1:1000	BioGenex
αGRA7 5–241–178	*Toxoplasma gondii* protein GRA7	mouse monoclonal	IF 1:3000	R. Ziemann, Abbott Laboratories, USA
αSAG1	*Toxoplasma gondii* p30 protein (SAG1)	mouse monoclonal	IF 1:4000	Biodesign
M2	FLAG-epitope (DTKDDDDK)	mouse monoclonal	IF 1:4000	Sigma Aldrich
GAL-TA (C-10)	GAL4 activation domain (aa768-881)	mouse monoclonal	WB 1:500	Santa Cruz Biotechnology
GAL4 DBD (RK5C1)	GAL4 DNA binding domain (aa94-147)	mouse monoclonal	WB 1:1000	Santa Cruz Biotechnology

Table 6 Primary immunoreagents. (WB: Western blot; IF: immunofluorescence)

II.1.9. Secondary immunoreagents

The following secondary immunoreagents were used: goat anti-mouse Alexa 488 and 546, goat anti-rabbit Alexa 488 and -546, donkey anti-goat Alexa 350, -488 and -546, donkey anti-mouse Alexa 488 and -555, donkey anti-rabbit Alexa 488 and -555 (all Molecular Probes; all used 1:1000 for immunofluorescence), donkey anti-rabbit HRP (Amersham Bioscience), donkey anti-goat HRP (Santa Cruz) and goat anti-mouse HRP (PIERCE) (all horse radish peroxidase (HRP)-coupled sera were used 1:5000 for immunodetection of Western blots). 4',6-Diamidine-2'-phenylindole dihydrochloride (DAPI, Roche) was used for nuclear counterstaining at a final concentration of 0.5 µg/ml.

II.2. Methods: Phylogenetic analysis

II.2.1. Use of database resources

All available public databases (Ensembl (http://www.ensembl.org); NCBI (National Centre for Biotechnology Information; http://www.ncbi.nlm.nih.gov)) were extensively screened by BLAST (basic local alignment search tool) and related searches for sequences belonging to the IRG family. In the case of the mouse, transcript sequences derived from the C57BL/6 strain were given preference over sequences of other and undefined strain origin, and compared with genomic sequence available (Ensembl mouse genome v28.33d.1, February 2005). Newly identified IRG sequences were confirmed wherever possible by multiple sequence comparisons at transcriptional and genomic levels. The obtained sequences were analysed using DNA Strider v1.3 alignment and matrix functions (Marck 1988). Chromosomal locations and synteny analyses of chromosomes was initiated through the Ensembl genome browser. The translated open reading frames of the IRG family members identified and used in this paper are given in the appendix (chapter V.13).

II.2.2. Generation of multisequence alignment and phylogenetic trees

Multisequence alignments were performed with ClustalW via the EBI server (http://www.ebi.ac.uk/clustalw) using the default settings and edited with SeaView (Galtier 1996). Shading of alignments was performed with Genedoc (Nicholas 1997). Phylogenetic trees were generated with PhyML (Guindon 2003) and edited with MEGA v3.1 (Kumar 2004). The appropriate model of protein evolution was determined with ProtTest (Abascal 2005) and is indicated in the respective figure legends.

II.3. Methods: Mutagenesis and cloning

II.3.1. Preparation of chemical competent bacteria

A single colony of *E. coli* strain DH5α was grown over night in LB medium with 20 mM $MgSO_4$ and 10 mM KCl at 37°C. The culture was diluted 1:100 in fresh aforementioned medium and grown at 37°C to an OD_{600} of 0.45. Cultures were incubated on ice for 10 min, cells were pelleted (2800 *g,* 4°C, 5 min), resuspended in cold TFB I (30 mM KOAc/50 mM $MnCl_2$/100 mM RbCl/10 mM $CaCl_2$/15% (v/v) Glycerol (pH 5.8); 30 ml/100 ml culture) and incubated on ice for 5 min. The bacteria were pelleted (2800 *g,* 4°C, 5 min), resuspended in cold TBF II (10 mM NaMOPS pH 7, 75 mM $CaCl_2$, 10 mM RbCl, 15% (v/v) Glycerol (pH 5.8); 4 ml/100 ml culture), shock frozen and stored at –80°C.

II.3.2. Transformation of competent bacteria

Competent cells were incubated with plasmid DNA or ligation mix (see below) for 30 min on ice. After a heat shock at 42°C for 30 sec followed by a 2 min incubation on ice, 400 µl of LB medium were added and the mixture was shaken for 30 min at 37°C. Different amounts of the transformation were plated on agar dishes containing antibiotics (ampicillin 100 µg/ml or kanamycin 50µg/ml) and grown overnight at 37°C.

II.3.3. Plasmid DNA isolation

For small-scale isolation of plasmid DNA from bacteria, overnight cultures were pelleted by centrifugation at room temperature (RT) with 23000 *g* for 1 min. The pelleted bacteria were resuspended in 100 µl P1 (50 mM Tris pH 8, 10 mM EDTA, 100 µg/ml RNase A). After addition of 100 µl P2 (200 mM NaOH, 1% SDS), samples were gently mixed and incubated for 5 min at RT. 140 µl of P3 (3 M KAc pH 5.5) were added and the reaction was spun for 15 min at 23000 *g*. The supernatant was transferred into a new tube, the DNA was precipitated by addition of two volumes of 100% ethanol (EtOH) and pelleted for 15 min at 23000 *g*. The DNA pellet was washed with 70% EtOH, air-dried and resuspended in 10 mM Tris pH 8. For preparation of large amounts of plasmid DNA the Qiagen Midi Plasmid Preparation Kit were used according to the manufacture's instructions.

II.3.4. Agarose gel electrophoresis

DNA was analysed by agarose gel electrophoresis in 40 mM Tris/0.5 mM EDTA (pH 7.5). Migration of the samples in the electric field was visualised by bromphenol blue. Fragment size was determined in comparison to the 1 kb ladder for agarose gels (Gibco). The DNA was visualised by use of ethidium bromide (0.3 µg/ml gel) and UV light.

II.3.5. Purification of DNA fragments from agarose gel

DNA fragments obtained by restriction endonuclease digest and separated on agarose gels were cut out of the gel. The DNA fragments were eluted from the gel blocks using the Rapid PCR Product Purification Kit (Boehringer) according to the manufacturer's protocol. Purity and yield of DNA was determined using the spectrophotometer measuring the optical density at 260 and 280 nm (concentration = OD_{260nm}* 50 µg/ml*dilution factor).

II.3.6. Restriction digest

All restriction endonucleases were purchased from New England Biolabs and used according to the manufacturer's advice. Double digests were either performed simultaneously in the advised NEB buffer or subsequently with adjustment of the salt concentration after the first digest.

II.3.7. Ligation

Purified insert DNA was ligated into pre-cut, de-phosphorylated (with shrimp alkaline phosphatase (Amersham Bioscience)), purified vector in a molar ratio of 3:1 using T4 DNA ligase (NEB) according to the manufacturer's instructions at 16°C overnight.

II.3.8. Sequencing

All constructs were verified by sequencing using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems), which is based on the didesoxy-chain termination method of sequencing developed by Sanger (Sanger 1977). In a total volume of 10 µl, 0.5

µg of template DNA, 10 pmol of the respective primer and 0.5 µl BigDye Terminator Ready Reaction Mix were mixed and PCR was carried out with the following program: 5 min 96°C, 25 cycles of 30 sec 96°C/15 sec 50°C/4 min 60°C. The DNA was precipitated with 70% EtOH/300 mM sodium acetate and pelleted with 23000 *g* at 4°C for 15 min. The pellet was washed with 70% ethanol, air-dried and rehydrated for sequencing on the ABI 3730 sequencer at the Cologne Centre of Genomics.

II.3.9. Site directed mutagenesis

Site directed mutagenesis was carried out using a modification of the protocol supplied with "QuikChangeTM XL Site-Directed Mutagenesis" Kit from Stratagene. The amplification was carried out using 20 ng plasmid as template, 125 ng of the sense and antisense primers and 2.5 U of Pfu-polymerase (Promega) in a total volume of 50 µl. The following PCR program was used: 1x 30 sec 95°C; 25x (30 sec 95°C/1 min 55°C/2 min/kb of plasmid length 72°C); 1x 15 min, 72°C. The template DNA of bacterial origin was digested with the dam methylation sensitive restriction enzyme DpnI (NEB) for 1 hrs at 37°C. 5 µl of the reaction were used to transform competent DH5α cells.

II.4. Methods: Protein Biochemistry

II.4.1. Expression and purification of recombinant protein

pGEX-4T-2-Irga6 (Uthaiah 2003), -Irga6(K82A), and –Irga6(S83N) were transformed into *E. coli* strain BL-21. Cells were grown at 37 °C in LB medium (Bertani 1951) to an OD_{600nm} of 0.8. Irga6 proteins were expressed as N-terminal GST fusions upon overnight induction with 0.1 mM isopropyl-β-D thiogalactoside (IPTG) at 18 °C. The cells were harvested (5000 *g*, 15 min, 4°C), frozen at –20°C, resuspended in PBS/2 mM DTT containing "complete mini protease inhibitors, EDTA free" (Roche) and lysed using a microfluidiser (EmulsiFlex-C5, Avestin) at a pressure of 150000 kilopascals. The lysate was cleared by sequential centrifugation at 4°C with 50000 *g* for 15 min and 75600 *g* for 30 min. The soluble fraction was purified on a glutathione Sepharose affinity column (GSTrap FF 5ml, GE Healthcare Life-Sciences) equilibrated with PBS/2 mM DTT. GST was cleaved off by overnight incubation of the resin with 5 units/ml thrombin (Serva) at 4 °C. Free Irga6 was eluted with PBS/2 mM DTT. Protein containing fractions were analysed by SDS-PAGE and visualised by Coomassie staining (de St. Groth 1963). Irga6 containing fractions were pooled and subjected to size exclusion chromatography on a Superdex 75 column (GE Healthcare Life-Sciences) equilibrated in 50 mM Tris/HCl pH 7.4/5 mM $MgCl_2$/2 mM DTT. The fractions were analysed by SDS-PAGE, and those containing Irga6 were concentrated by a centrifugal concentrator (Vivaspin 20, Sartorius, 10 kDa cut-off). Aliquots were shock-frozen in liquid nitrogen and stored at -80 °C. The concentration of Irga6 was determined by UV spectrophotometry at 280 nm (Gill 1989). The purified proteins carry the extension GSPGIPGSTT at the N terminus due to digestion of the GST fusion with thrombin (see Table 2).

II.4.2. Guanine nucleotide binding parameters

The nucleotide binding affinities of Irga6, Irga6(K82A) and Irga6(S83N) for 2'(3')-O-methylanthraniloyl (mant) GDP (mGDP) and mGTPγS, a nonhydrolysable analogue of GTP (both from Jena Bioscience) were determined by equilibrium titration as described in Uthaiah *et al* 2003. In short Irga6 proteins were titrated in a range from 0 to 100 µM against 0.5 µM mant nucleotides in 50 mM Tris/HCl pH7.4/5 mM $MgCl_2$/2 mM DTT. The mant nucleotides were excited at 355 nm, and the fluorescence was monitored at 448 nm (Aminco-Bowman 2 Luminescence Spectrometer; SLM Instruments). The increase in fluorescence upon the stepwise addition of the protein was measured, and each value was averaged over 30 sec. The equilibrium dissociation constants, Kd, were obtained by fitting a quadratic function to the data as described in (Herrmann 1996) by using the Sigma Plot program (Systat). To determine whether the nucleotide binding of Irga6 depended on the presence of magnesium, the equilibrium titration was also carried out in absence of $MgCl_2$ and presence of 10 mM EDTA.

II.4.3. GTP hydrolysis assay

Irga6 wild type and mutant recombinant proteins (80 µM) were incubated with 10 mM GTP (Sigma-Aldrich) containing traces of αP^{32}-labeled GTP (GE Healthcare Life-Sciences) at 37°C in 50 mM Tris/HCl pH7.4/5 mM $MgCl_2$/2 mM DTT. At the indicated time points aliquots of the reaction were spotted onto PEI Cellulose F thin layer chromatography (TCL) plates (Merck). Dried plates were run in 1 M acetic acid/0.8 M LiCl. Signals were detected by use of the BAS 1000 phosphoimager analysis system (Fujifilm) and quantified with the AIDA Image Analyser v3 software (Raytest).

II.4.4. Analysis of protein oligomerisation by light scattering

Oligomerisation behaviour of Irga6 and the Irga6 mutants K82A and S83N was determined by conventional and dynamic light scattering. Light scattering of 80 µM protein in the presence of 10 mM GDP and GTP respectively (Sigma-Aldrich) in 50 mM Tris/HCl pH7.4/5 mM $MgCl_2$/2 mM DTT was analysed at 37°C. Samples were cleared by ultracentrifugation prior to addition of nucleotides. In conventional light scattering, samples were excited at 350 nm and scattered light was detected at the same wavelength in an Olis DM45 Spectrofluorimeter (Olis) at a fixed angle of 90°. Dynamic light scattering (DLS) was performed using a DynaPro molecular sizing instrument equipped with a MircoSampler temperature control unit (MSTC800, Protein Solutions, Wyatt Technologies). The scattering of light with 650 nm wavelength by the sample was measured at 37°C for 30 min in a Quartz cuvette (acquisition time 10 sec). Data were obtained and analysed using the DYNAMICS software (v.5). The hydrodynamic radius (R_H) was calculated from the translational diffusion coefficient (D_T), obtained by autocorrelation of the data, using the Stokes-Einstein Equation ($R_H = k_b\ T/6\pi\eta D_T$; k_b: Boltzmann constant, T: absolute temperature in Kelvin, η: solvent viscosity). The molecular weight (MW) was estimated from the hydrodynamic radius R_H using the standard curve of MW versus R_H for globular proteins ($MW = (R_H \text{ factor} * R_H)^{Power}$, for globular proteins R_H factor = 1.68, Power = 2.3398).

II.5. Methods: mammalian cells

II.5.1. Freezing and thawing of mammalian cells

Cells were harvested, pelleted at 400 *g* for 5 min, resuspended in ice-cold sterile FCS/10% DMSO (v/v) ($4*10^6$ cells/ml), frozen slowly in cryotubes (Greiner) and transferred to liquid nitrogen for long-term storage. Cells were thawn at 37°C, immediately transferred into a large volume of medium, pelleted and plated in fresh medium.

II.5.2. Transfection of mammalian cells

FuGENE6 transfection reagent (Roche) was used for transient transfections of the indicated amounts of DNA according to the manufacturer's instruction using a FuGENE to DNA ration of 3:1. For stable transfection calcium phosphate co-precipitation was used (Graham 1973). Cells were plated on 10 cm dishes to be 50-80% confluent on the day of transfection. The medium was changed at least 2 hrs prior to transfection (9 ml/plate). DNA solutions (0.1 µg - 5 µg) were prepared in 500 µl of 250 mM $CaCl_2$, pipetted drop wise into 500 µl of 2xHBS (280 mM NaCl, 50 mM HEPES, 1.5 mM Na_2HPO_4/H_2O, pH 7.1) while shaking vigorously. The mix was incubated for 20 min at RT and added drop wise to the cells while shaking the plates cautiously. The medium was exchanged 24 hrs later with fresh medium containing selection agents.

II.5.3. Hormone-inducible mammalian expression system (GeneSwitch)

The hormone-inducible mammalian expression system used in this study (GeneSwitch, Invitrogen) (Wang 1994)) is based on two plasmids. The pSwitch vector expresses a hybrid regulatory protein, the so called GeneSwitch protein, containing the yeast GAL4 DNA binding domain (GAL4-DBD), a truncated human progesterone receptor ligand binding domain (hPR-LBD) and the p65 activation domain (p65-AD) from human NF-κB under the control of a hybrid promoter consisting of yeast GAL4 upstream activating sequences (UAS) (Giniger 1985; Wang 1994) linked to a minimal promoter from the Herpes Simplex Virus thymidine kinase (TK) gene. The pGene/V5-His plasmid inducibly expressing the gene of interest, is controlled by a hybrid promoter consisting of yeast GAL4 upstream activating sequences linked to the TATA box sequence from the adenovirus major late E1b gene (Lillie 1989).

In the absence of the synthetic progesterone antagonist Mifepristone, low basal transcription of the GeneSwitch gene from pSwitch occurs from the minimal TK promoter. Once translated, the GeneSwitch protein is predominantly localised in the nucleus in an inactive form. Upon addition, Mifepristone binds with high affinity to the truncated hPR-LBD (Vegeto 1992) and causes a conformational change of this domain resulting in dimerisation and activation of the GeneSwitch protein. The ligand-bound homodimer then interacts with the GAL4-UASs of both pGene/V5-His and pSwitch and activates transcription of both the gene of interest and the GeneSwitch gene, resulting in a positive feedback loop (Figure 7) (Marmorstein 1992; Wang 1994) (for vector maps see appendix V.2).

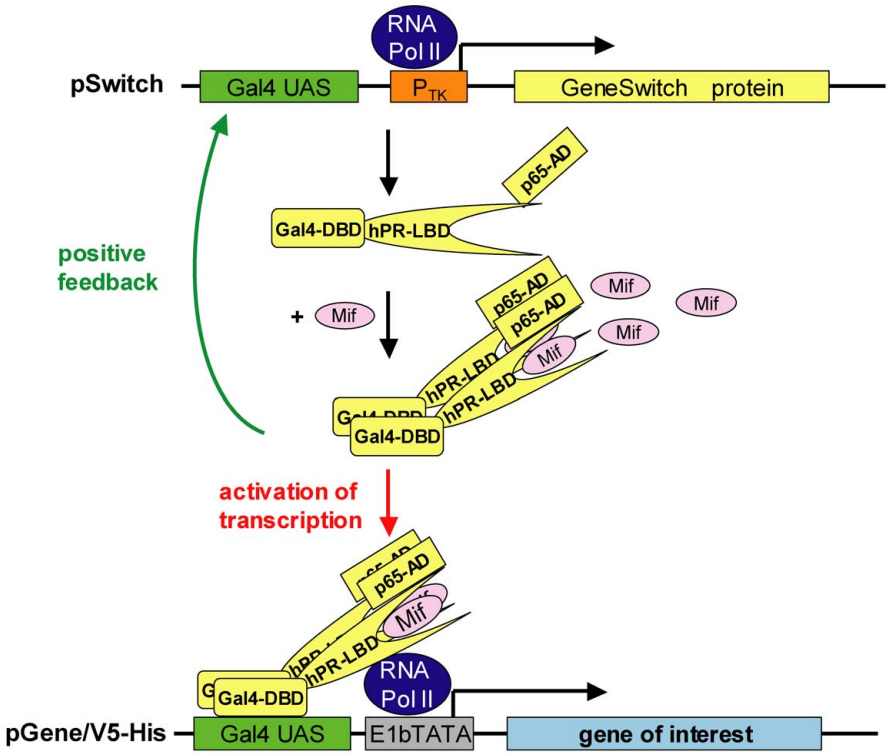


Figure 7 GeneSwitch system for inducible expression. The GeneSwitch hybrid regulatory protein is expressed constitutively at low level from the minimal thymidine kinase promoter. Upon binding of the synthetic steroid Mifepristone the GeneSwitch protein undergoes a conformational change resulting in dimerisation and activation. The ligand-bound homodimer then interacts with the GAL4-binding sites in the GAL4 UAS of both pGene/V5-His and pSwitch activating the transcription of both the gene of interest from the E1b TATA box and the regulatory fusion gene itself. (GAL4-UAS: yeast upstream activating sequences with binding sites for GAL4 transcription factor; E1bTATA: TATA box sequence from adenovirus major late E1b gene; P_{Tk}: herpes simplex virus thymidine kinase minimal promoter; GAL4-DBD: yeast GAL4 DNA binding domain; hPR-LBD: truncated human progesterone receptor ligand binding domain; p65-AD: activation domain of human NFκB; GeneSwitch protein: GAL4-DBD/hPR-LBD/p65-AD regulatory fusion protein; Mif: Mifepristone).

II.5.4. Generation of stable inducible cell lines

Following transfection by calcium phosphate precipitation with single pGeneA-IRG constructs, cells were kept under selection (50µg/ml Hygromycin and 200 µg/ml Zeocin, both Cayla) on 10 cm culture dishes until larger clones were visible. Cells were trypsinised, counted in a Neubauer chamber and cloned by limiting dilution. Single clones were expanded and expression of the respective gene after 24 hrs induction with Mifepristone was determined by immunofluorescence and Western Blot. Clones with Mifepristone-induced IRG expression levels comparable to IFN-induced one and no detectable background expression were subcloned and used for further experiments.

II.5.5. Induction with IFNγ and Mifepristone

Cells were supplemented with fresh medium containing the desired amount of inducer (200 U/ml recombinant murine IFNγ (Cell Concepts) and 10^{-7} to 10^{-10} M Mifepristone (Invitrogen)) and incubated for 24 hrs unless noted otherwise.

II.5.6. Immunofluorescence

Cells were grown on heat-sterilised cover slips, induced with IFNγ (Cell Concepts) or Mifepristone (Invitrogen) for 24 hrs, transfected or left untreated and fixed with PBS/3% (w/v) paraformaldehyde for 20 min at RT. Subsequently cells were washed four times with PBS, permeabilised with washing buffer (PBS/0,1% (w/v) Saponin) for 10 min at RT and blocked with PBS/0,1% (w/v) Saponin/3% (w/v) BSA (blocking buffer) for 1 hr at RT. Incubation with the primary antibodies diluted in blocking buffer was performed in a humid chamber for 1 hr at RT or over night (o.n.) at 4°C. Cover slips were washed with washing buffer (3x5 min) followed by incubation with the secondary antibodies and DAPI (0.5 µg/ml) diluted in blocking buffer for 30 min at RT. After washing 3x5 min with washing buffer, cover slips were mounted on slides with ProLong Gold antifade reagent (Invitrogen) and sealed with nail polish. Images were made on a Zeiss Axioplan II fluorescence microscope equipped with an AxioCam MRm camera (Zeiss) using the Metamorph (Universal Imaging Corporation) and Axiovision (v4; Zeiss) software.

II.5.7. Cell proliferation assay

For analysis of proliferation 1000 cells/well were plated onto 96well plates and kept under induction with either IFNγ (200 U/ml) or Mifepristone (10^{-9} M) for the indicated length of time. Each day cell proliferation was evaluated using the CellTiter 96 AQueous non-radioactive cell proliferation assay (Promega) according to the manufacturer's instructions. The absorption of the bioreduced form (formazan) of a substrate (MTS) generated by metabolically active cells during incubation at 37°C for 4 hrs was measured in an ELISA reader (Vmax, Molecular Devices) at 490 nm (Barltrop 1991; Cory 1991; Mosmann 1983). The quantity of formazan product is directly proportional to the number of living cells in the culture.

II.5.8. Cell cycle assay

Single cell suspensions in PBS ($1*10^6$ cells/ml) were fixed with 70% (v/v) ethanol for 30 min on ice. Cells harvested by centrifugation were resuspended in PBS and passed through a 25-gauge syringe needle. Cells were incubated with 100 µg/ml RNase A and 40 µg/ml propidium iodide for 30 min at 37°C and analysed using a FACSCan flow cytometer and the CellQuest Pro 4.0.2 software (both BD Bioscience).

II.6. Methods: Analysis of cellular protein

II.6.1. Generation of cell lysates for SDS-PAGE

If not mentioned otherwise, cells were harvested by scraping in cold PBS (Cellscraper; Sarstedt), pelleted at 500 *g* and 4°C for 5 min and resuspended in PBS/1% Triton X-100 (Sigma-Aldrich)/"Complete Mini Protease inhibitor Cocktail" (Roche) (200µl/$3*10^5$ cells). Samples were incubated 30 min on ice and nuclei were pelleted for 30 min at 23.000 *g* and 4°C. The supernatant was analysed directly in immunoblot while the pellet was boiled in protein loading buffer (50 mM Tris/HCl pH 6.8/0.7 % β-ME/1% SDS/5 % glycerol/0.0025 % (w/v) bromephenol blue) to detect aggregated protein.

II.6.2. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

For the analysis of proteins, discontinuous one-dimensional gel electrophoresis (Laemmli 1970) was carried out under denaturing conditions in the presence of 1% SDS on 7.5-12% polyacrylamide gels (separating gel; stacking gel: 5%). The protein samples were boiled for 10 min in protein loading buffer (50 mM Tris/HCl pH 6.8/0.7 % β-ME/1% SDS/5 % glycerol/0.0025 % (w/v) bromephenol blue (Sigma-Aldrich)) prior to loading. Small gels were run for several hours at RT and large gels over night at 4°C in electrophoresis buffer (25 mM Tris/190 mM glycine/0.1 % SDS). Wide range protein standard marker (Sigma-Aldrich) was used for size determination of protein bands.

II.6.3. Western Blotting and Ponceau S staining

After SDS-PAGE proteins were transferred to nitro-cellulose membranes (Protan 0.45 µm, Schleicher & Schuell) by electro-blotting in transfer buffer (15 mM Tris/190 mM glycine) for 1 hr at RT and 0.5 A in a blotting chamber (Ideal Scientific Company). Ponceau S staining (0.1% Ponceau S (w/v) in 5% acetic acid) was used to locate proteins on the membrane after Western blotting. Membranes were blocked with PBS/5% skim milk powder/0.1% Tween 20 for 1 hr at RT, washed twice for 20 min with PBS/0.1% Tween 20 and incubated with the appropriate primary antibodies diluted in PBS/1% FCS (1.5 hrs at RT or o.n. at 4°C). After washing the blots 3x10 min with PBS/0.1% Tween 20, they were incubated for 30 min at RT with the respective secondary HRP-coupled antisera and washed with PBS/0.3% Tween 20 (3x10 min). Bound primary antibodies were detected using HRP-coupled secondary reagents and visualised by chemiluminescence (100 mM Tris/HCl pH 8.5, 12.5 mM 3-aminophthalhydrazide, 200 µM p-coumaric acid, 0.01% H_2O_2) using a Super RX film (Fujifilm).

II.6.4. Sequential Triton X-114 partitioning assay

24 hrs after transient transfection with 2 µg pGW1H-IIGP and/ or induction with 200U/ml IFNγ and/or 10^{-10} M Mifepristone, $3*10^5$ gs3T3 cells/sample were harvested by scarping and lysed in 400 µl ice cold PBS/1% Triton X-114 (Sigma)/"CompleteMini protease inhibitor cocktail, EDTA free" (Roche) for 1 hr on ice. Lysates were centrifuged for 15 min at 3000 *g* and 4°C to remove nuclei. Supernatants were incubated for 5 min at RT to induce phase separation seen as cloudiness of the solution and centrifuged for 1 min at 23.000 *g* at RT. The detergent phase was adjusted to the same volume as the aqueous phase with PBS and stored on ice. The aqueous phase was subjected to a subsequent round of partitioning by adding ice cold 10% Triton X-114 to a final concentration of 1% Triton X-114. The samples were incubated on ice until the solution became clear again and were then shifted to RT to induce phase separation. The samples were further treated as described above. In total four subsequent partitionings were performed. Equal volumes (40 µl) of all collected fractions were used for SDS-PAGE and immunoblot with 165 anti-Irga6 serum.

II.6.5. Co-Immunoprecipitation

IFNγ-induced (24 h, 200U/ml) and untreated gs3T3 cells were lysed in lysis buffer (PBS/0.1% Thesit (Sigma-Aldrich)/3 mM $MgCl_2$/"CompleteMini protease inhibitor cocktail, EDTA free" (Roche)) either in absence of nucleotide or in presence of 0.5 mM GDP and 0.5 mM GTPγS respectively (both Sigma-Aldrich). Lysates were cleared by high-speed centrifugation (30 min, 23000 *g*, 4°C). The Irga6 specific 165 serum (Uthaiah 2003) was coupled to protein A Sepharose CL-4B (GE Healthcare Life-Sciences) (14 µl antibody/50 µl dry beads). Beads were washed extensively and bound immunglobulins were crosslinked to the protein A Sepharose using 20 mM dimethyl pimelimidate (DMP) (Sigma-Aldrich) (Harlow 1988). Coupled beads were washed with PBS and lysis buffer and incubated with the cell lysates for 2 hrs at 4°C. Beads were washed twice with lysis buffer and twice with PBS/3 mM $MgCl_2$. Bound proteins were eluted by incubation with 100 mM Tris pH 8.5/0.5% SDS for 30 min at room temperature (RT). Residual beads were removed by centrifugation and filtration (filter tubes, pore size 0.45 µM; Millipore). Eluates were supplemented with SDS-PAGE protein loading buffer, boiled for 5 min and subjected to SDS PAGE. For the co-immunoprecipitation of Irgm3 with Irga6 one quarter of the eluate was used for detection with anti-Irga6 antibody 10D7 and three quarters were used for detection with the anti-IGTP (Irgm3) antibody clone 7. For the immunoprecipitation performed to detect ubiquitinated Irga6 equal amounts of precipitate were analysed by SDS-PAGE and detected with the anti-Irga6 antibody 10D7 and the anti-ubiquitin antibody FK2 respectively.

II.6.6. Pull down

Irga6 and Irga6(S83N) GST fusion proteins were expressed and harvested as described above (chapter II.4.1). Cleared bacterial lysates were incubated with glutathione Sepharose (High Performance, GE Healthcare Life-Sciences) for 2 hrs at 4°C. Beads were washed 10 times with PBS/2 mM DTT, pelleted by centrifugation (2 min, 2000 *g*) and resuspended in the same buffer. To control protein binding, beads were boiled in protein loading buffer and the supernatant was subjected to SDS-PAGE followed by Coomassie staining (de St. Groth 1963). The protein bound glutathione Sepharose was washed once with PBS/5 mM $MgCl_2$/1 mM DTT and preincubated with or without 1 mM nucleotide (GDP, GTPγS or mGDP (both Sigma-Aldrich) or a combination of 100 µM GDP plus 300 µM GTPγS) or 10 mM EDTA in PBS/5 mM $MgCl_2$/1 mM DTT buffer for 1 h at RT. Beads were washed with lysis buffer (PBS/0.1% Thesit/3 mM $MgCl_2$/"CompleteMini protease inhibitor cocktail, EDTA free") containing 10 µM of the respective nucleotide. Gs3T3 cells were induced for 24 hrs with 200 U/ml IFNγ and 10^{-9} M Mifepristone respectively, harvested and lysed in lysis buffer containing 0.5 mM of the respective nucleotides or 100 µM GDP plus 300 µM GTPγS or 10 mM EDTA for 1 hr at 4°C. The lysates were cleared by centrifugation (23000 *g* 30 min 4°C), mixed with the nucleotide pre-treated protein-glutathione Sepharose and incubated at 4°C o.n. Beads were washed twice with lysis buffer and twice with PBS/3 mM $MgCl_2$. Cellular proteins bound to the beads were eluted with 100 mM Tris pH 8.5/0.5% SDS for 30 min at RT, supplemented with protein loading buffer and boiled for 5 min. The samples were separated by SDS-PAGE and subjected to Western blot. Input of recombinant protein

was monitored by Ponceau S staining. Immunodetection of Irgm3 was performed with αIGTP monoclonal antibody (clone 7) over night at 4°C.

II.6.7. Coomassie staining

For unspecific detection of all proteins, SDS-PAGE gels were stained with (0.1% (w/v) Coomassie Brilliant Blue R-250 (Serva)/40% (v/v) ethanol/10% (v/v) acetic acid (de St. Groth 1963) shaking 20 min at RT. Excess staining was removed by incubation in 40% (v/v) ethanol/10% (v/v acetic acid) at RT for 1 hr. Gels were washed o.n. with H_2O and dried under vacuum pressure at 80°C for 2 hours.

II.6.8. Analytical size exclusion chromatography of cellular IRG GTPases

Gs3T3 cells ($1x10^6$ cells/sample) were induced with 200 U/ml IFNγ of 10^{-9} M Mifepristone respectively for 24 hrs and lysed for 1 hr at 4°C in 500 µl PBS/0.1% Thesit/"CompleteMini protease inhibitor cocktail, EDTA free" (Roche). Postnuclear supernatants cleared by ultracentrifugation (45000 *g*, 30 min, 4°C) were separated on a Superose 6 HR 10/30 size exclusion column (Amersham Bioscience) in PBS/0.1% Thesit with a flow rate of 0.2 ml/min on an ÄKTA fast protein liquid chromatography (FPLC) (monitor UPC-900, pump P-920, fraction collector Frac-950, valves INV-907; Amersham Bioscience) using the Unicorn 5.01 software. The column was calibrated with carbonic anhydrase (29 kDa), bovine serum albumin (BSA; 66 kDa), alcoholic dehydrogenease (150 kDa), β-amylase (200 kDa). 0.2 ml fractions were collected and analysed for their IRG protein content in SDS-PAGE followed by immunoblot with specific immunoreagents. Thesit was used for cell lysis, as the low UV absorption of this non-ionic detergent did not interfere with the monitoring of protein elution during chromatography at 280 nm.

II.7. Methods: Infection assays

II.7.1. *In vitro* passage of *Toxoplasma gondii*

Tachyzoites from *T. gondii* strains ME49 and RH-YFP (Gubbels 2003) were maintained by serial passage in confluent monolayers of human foreskin fibroblasts (Hs27, ATCC number CRL-1634) cultured in 25 cm^2 flasks with IMDM/5% FCS/2 mM L-glutamine. Following inoculation of fibroblasts with $2x10^6$ and $0.5x10^6$ parasites respectively, parasites actively invade the host cells, replicate intracellularly and egress from their host cells approximately 3 days later. Extracellular parasites were harvested from the supernatant and purified from host cell debris by differential centrifugation (5 min at 100 *g*, 15 min at 500 *g*). Parasites were resuspended in medium, counted using a Neubauer chamber and immediately used for inoculation of host cells. RH-YFP parasites were propagated in presence of Chloramphenicol (3.2 µg/ml, Sigma-Aldrich) to counter select for loss of the stable integrated YFP expression plasmid containing a Chloramphenicol acetyltransferase selectable marker (Gubbels 2003).

II.7.2. Infection of murine cells with *Toxoplasma gondii*

Murine fibroblasts seeded onto cover slips were stimulated with 200 U/ml IFNγ (Cell Concepts) and/or 10^{-9} M Mifepristone (Invitrogen) and/or transfected with the indicated expression constructs using FuGENE6 reagent (Roche) for 24 hrs or were left untreated. The cells were then inoculated with *T. gondii* ME49 or RH-YFP tachyzoites for 2 hrs at 37°C at a multiplicity of infection (MOI) of 4 to 6. Infection was synchronised by centrifugation. Extracellular parasites were removed by washing 3x with medium and once with PBS. The infected cells were fixed for 20 min with PBS/3% paraformaldehyde at RT and processed further for immunofluorescence staining as described above. Parasites were visualised either by *T. gondii* specific immunoreagents (see above, Table 6) or by detection of YFP fluorescence (RH-YFP). Intracellular parasites were either identified in phase contrast or by immunostaining for the *T. gondii* dense granule protein GRA7, a 29 kDa-predicted transmembrane protein that is released into the parasitophorous vacuole by intracellular parasites shortly after invasion and associates with the PV membrane (Bonhomme 1998; Fischer 1998).

II.7.3. Quantification of IRG signals on *T. gondii* parasitophorous vacuoles

The intensities of IRG specific signals on *T. gondii* parasitophorous vacuoles were quantified using the ImageJ software (Wayne Rasband, NIH). Two to three intensity profiles orthogonal to each other were generated per vacuole followed by subtraction of the background fluorescence and averaging of the 4 (6) peak values obtained for each vacuole. The average pixel intensity of each vacuole was plotted for each IRG protein in a scatter plot using Excel (Microsoft).

II.8. Methods: Yeast 2 Hybrid

II.8.1. Lithium acetate transformation of *Saccharomyces cerevisiae*

Fresh yeast colonies (strain PJ69-4a-a or -α) were washed once with 100 mM LiAc. The pellet was resuspended in 100 mM LiAc/33% (w/v) polyethylenglycol 3350 (Sigma-Aldrich)/25 µg calf thymus DNA (Sigma-Aldrich)/1 µg of pGAD/BD expression construct), incubated 30 min at 37°C and heat shocked for 20 min at 42°C. The lithium acetate transformed yeast cells (Gietz 1995) were pelleted and plated on selection medium (SD lacking leucine (SD-L) and tryptophan (SD-T) respectively). After 2-3 days incubation at 30°C colonies were picked and replated on SD-L or –T. Transformed strains were frozen in 15% glycerol at –80°C for storage.

II.8.2. Analysis of protein interaction by yeast two-hybrid (Y2H)

The complete coding regions of the IRG proteins studied were expressed constitutively at low level as N-terminal fusions with the Gal4 DNA-binding (BD) and Gal4 activation domain (AD) in haploid yeast reporter strains of opposite mating types (PJ69-4a-α and PJ69-4a-a) using tryptophan and leucine selection respectively (SD-T, SD-L) (James 1996) (Figure 8). PJ69-4a-a-pGAD-IRG and PJ69-4a-α-pGBD-IRG yeast cells were mated with each other on YPD plates and selected for diploid cells on SD-L-T medium. Single colonies were assayed for protein-protein interaction driving expression of the

pyrophosphate-aminoimidazole carboxylase and the imidazoleglycerolphosphate dehydratase reporter genes involved in adenine and histidine biosynthesis (Stotz 1993), respectively, by growth on SD-L-T also lacking adenine and/or histidine. 1 mM 3-amino-1,2,4-triazole (Sigma-Aldrich) was used to inhibit autoactivation of the *HIS3* reporter (Durfee 1993). At least two independent crossings were performed for each interaction.

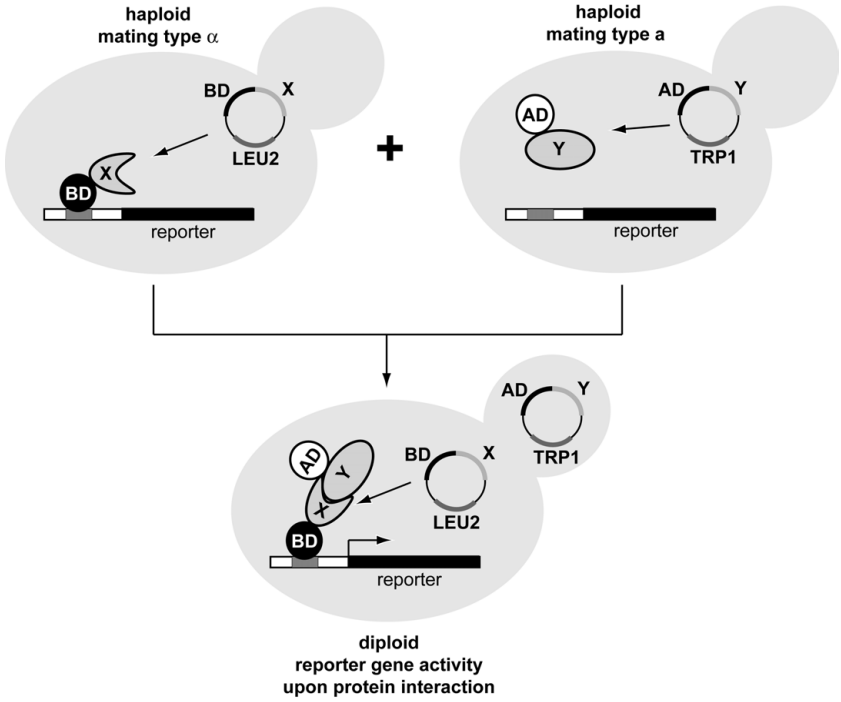


Figure 8 Classical yeast-two hybrid assays using yeast mating. Classical yeast-two hybrid assays using yeast mating**.** Constructs encoding putative interacting proteins (X and Y) fused to the DNA-binding (BD) and transcription activation (AD) domains, respectively, are transformed separately into yeast strains of opposite mating type (a and α). After mating, both constructs are present in the same yeast cell and, if proteins X and Y interact, reporter gene activity is detected (modified from (Causier 2002)).

III. Results

III.1. Genomic organisation, synteny and phylogenetic relationship of IRG GTPases

At the beginning of this study only six *IRG* genes from mouse, namely Irgd (IRG-47) (Gilly 1992), Irgm1 (LRG-47) (Sorace 1995), Irgb6 (TGTP/Mg21) (Carlow 1995; Carlow 1998; Lafuse 1995), Irgm3 (IGTP) (Taylor 1996; Taylor 1997), Irga6 (IIGP) and Irgm2 (GTPI) (Boehm 1998), were published but evidence for the presence of additional family members was accumulating from analyses of genomic bacterial artificial chromosome (BAC) clones and expressed sequence tags (ESTs) available in the public databases (identification of several Irga6-like genes, (Parvanova 2005)). Upon completion of the mouse genome project , a thorough analysis of all p47 GTPase genes in C57BL/6 strain was feasible. Hence, BLAST (Basic Local Alignment Search Tool; (Altschul 1990; Altschul 1997)) and SSAHA (Sequence Search and Alignment by Hashing Algorithm; (Ning 2001)) searches with the known IRG genes were performed. The relevant parts of the genome containing homologues were extracted and analysed further for gene order, orientation and homology to the published IRGs using DNA Strider v1.3 (Marck 1988). Where ambiguities persisted in the mouse genomic map, especially on chromosome 18 in the region of *Irga6-Irga8* (Mb 60.878-60.958) and on chromosome 11 in the region from *PA28*βψ to *Irgb7*ψ (Mb 57.570-57.700), primary BAC and cosmid sequences were used to reach a consensus view (positions refer to the Ensembl mouse genome release v28.33d.1, February 2005). Additionally, transcript sequences, preferentially from C57BL/6 strain, were compared with the genomic sequences available via Ensembl. Subsequently the search for and analysis of IRG homologues was extended to other species in order to shed light onto the phylogenetic relationship, age and evolutionary history of this family of large GTPases.

III.1.1. Nomenclature

Altogether over 180 novel p47 GTPases from 33 different species were identified during this study (summarised in Table 7). The protein sequences of all the genes are listed in the appendix (chapter V.13). To assure a consistent naming of new family members, a general nomenclature based on phylogenetic principles using the stem IRG (immunity-related GTPase) was introduced. This terminology not only allows the priority in nomenclature to IRG-47 as the first *IRG* gene identified (Gilly 1992). It also intentionally does not refer to IFN-inducibility since the highly conserved *IRGC* genes of mammals are clearly related phylogenetically to the IFN-inducible immunity genes but are definitely not IFN-inducible (see below). The stem IRG is followed by a single-letter suffix each identifying an individual deep monophyletic clade in the phylogenetic trees of IRGs (e.g. IRGM or IRGA, see Figure 14 and Figure 13). Individual genes in one clade are distinguished by consecutive numbering (e.g. Irga6). To date nine IRG suffixes based on the phylogenetic analyses were defined (A-H and Q) (Figure 14, Figure 22). It has to be noted that some of the suffix defining clades are less deep than others in a broad phylogenetic comparison, especially in case of IRGA, IRGB and IRGD (Figure 22). Still these clades are very well defined in the mammals, thus justifying the distinction (Figure

14, Figure 18). The analysis of more fish sequences is essential to clarify whether the distinction between *IRGG* and *IRGE* genes is justified or if those genes rather belong to one joint clade (Figure 22).

(Sub/Infra)CLASS/order	Species	GMS genes	Classical GKS genes	IRGC genes	Quasi IRGs	Total #
MAMMALIA (Eutheria)						
Rodentia	*Mus musculus domesticus*	3	17 (D, 8 A, 8 B)	1 (C)	1 (Q)	22
	Rattus norvegicus	3	11 (1 D, 7 A, 3 B)	1 (C)	1 (Q)	16
	Cavia porcellus	3	7 (3 B, 4 D)	1 (C)	1 (Q)	12
	Spermophilus tridecemlineatus	1	1 (B)	1 (C)	1 (Q)	4
Lagomorpha	*Oryctolagus cuniculus*	1 (G1: GIS)	1 (A)			2
Insectivora	*Erinaceus europaeus*		1 B, 1 BQ		1 (Q)	3
	Sorex araneus			2 (C, C-like)	1 (Q)	2
	Echinops telfairi	2 (1 GLS in G1)	4 (1 D, 1 A, 2 B)	1 (C)	1 (Q)	8
Chiroptera	*Myotis lucifugus*			2 (C, C-like)	1 (Q)	3
Carnivora	*Felis catus*			2 (C, Clike)	1 (Q)	3
	Canis familiaris	3	6 (4 B, 1 D, 1 DQ)	2 (C, C-like)	1 (Q)	12
Artiodactyla	*Bos taurus*		1 (D)	2 (C, C-like)	1 (Q)	4
	Sus scrofa			2 (C, C-like)		2
Perissodactyla	*Equus caballus*			1 (C)	1 (Q)	2
Proboscidea	*Loxodonta africana*	2	3 (2 A, 1 D)	1 (C)	1 (Q)	7
Cingulata	*Dasypus novemcinctus*	1	4 (2 B, 1 BQ, 1 D)	1 (C)	1 (Q)	7
Primates (Haplorrhini)	*Homo sapiens*	1*	---	1 (C)	1 (Q)	3
	Pan troglodytes	1*	---	1 (C)	1 (Q)	3
	Pongo pygmaeus	9*				2
	Macaca mulatta	1*(G1: GMN)		1 (C)	1 (Q)	3
	Callithrix jacchus	1* (G1: SIS)		1 (C)	1 (Q)	3
Primates (Strepsirrhini)	*Otolemur garnetti*		1 (A)	1 (C)	1 (Q)	2
	Microcebus murinus	≥ 4	≥ 3 (2 A, 1 D)	1 (C)	1 (Q)	≥ 9
MAMMALIA (Metatheria)						
Didelphimorphia	*Monodelphis domestica*		2 (2 D)	2 (C,C-like)		4
MAMMALIA (Prototheria)						
Monotremata	*Ornithorhynchus anatinus*		1 (B)	5 (1 C, 4 C-like)		6
AMPHIBIA						
Anura	*Xenopus tropicalis*		1		1 (XQ)	2
Caudata	*Ambystoma tigrinum tigrinum*			1 (C-like)		1
REPTILIA						
Squamata	*Anolis carolinensis*		5 (3 E)	5 (C-like)		10
ACTINOPTERYGII						
Cypriniformes	*Danio rerio*	---	11 (6 E,4 F 1 G)		3 (XQ)	11
Tetraodoniformes	*Takifugu rubripes*		≥ 3 (3 F)			≥ 3
	Tetraodon nigroviridis		≥ 2 (2 F)			≥ 2
CHONDRICHTHYES						
Rajiformes	*Leucoraja erinacea*		1			1
Squaliformes	*Squalus acanthias*		2			2
Cephalochordata						
Amphioxi	*Branchiostoma floridae*		≥ 14 (14 H)			≥ 14

Table 7 IRG GTPases in the animal kingdom. List of all the IRG genes identified in the public databases during this study. The number of genes belonging to certain IRG clades is given in brackets for the GKS genes. Conservative substitutions present in the G1 motif of two GMS GTPases are specified in brackets. Some of the listed genes are represented only by fragments due to the premature state of many of the databases used (for references of the used databases see main text and material and methods section; for sequences see appendix V.13). Note that there are 12 *IRGB* sequences in the mouse genome, that represent 8 genes four of which are tandem genes with two IRG sequences as coding exons. The GMS genes of the higher primates (*) are considerably shorter than all other IRG genes and shorter than H-Ras-1 terminating shortly before the G5 motif. Grey highlighting marks genes with a single intron situated between the G4 and G5 motif.

As the nomenclature of IRG genes is based on their phylogenetic relationship, an adequate naming of newly identified sequences requires a thorough phylogenetic analysis. Novel genes were provisionally named applying the following principles: Firstly each gene name starts with a two-letter prefix for the species of origin (e.g. RN for *Rattus norvegicus,* see appendix V.4 for details); secondly the name of the IRG subfamily the respective sequence belongs to (GKS or GMS, see below) was added; thirdly the genes were numbered in order of discovery (e.g. LA_GMS1 for the first elephant *(Loxodonta africana)* GMS sequence identified). In order to avoid repeated renaming, the provisional names were given up in favour of the general nomenclature only upon complete and unambiguous phylogenetic characterisation. The IRG

nomenclature using different forms for mouse (Irg), human (IRG), dog (IRG) and zebrafish (irg) (see below) has been accepted by the gene nomenclature committees of human and mouse, and by the zebrafish sequencing project. As is customary, the same nomenclature rules were used for mice and rat (Irg) (http://www.informatics.jax.org/mgihome/nomen/gene.shtml). For all other mammals the capitalised form (IRG) was used. The principles of zebrafish gene nomenclature were applied to all fish, amphibian and reptile sequences as suggested by others (http://genome.jgi-psf.org/Xentr4/docs/Xenopus_gene_naming.pdf).

The nomenclature of the 'quasi IRG' proteins (*IRGQ)* departs from the phylogenetic principle. The original IRGQ nomenclature published in Bekpen *et al.* 2005 (Bekpen 2005b) grouped all IRG homologous sequences that are devoid of GTPase function due to modifications in the GTP-binding site together, though they clearly represent a polyphyletic group. However, further analysis revealed that the mammalian genes highly conserved with the radically modified murine *Irgq* (also known as FKSG27) form a monophyletic group distant from the other IRG clades (Figure 20). It was decided to keep the Irgq nomenclature for these genes. Other mammalian IRG genes with disrupted GTPase motifs possess otherwise conserved clade specific features (Figure 18-19). Therefore, it was decided to allocate them to the respective clade adding the suffix “q” to the name (e.g. EE_ GKS2/IRGBQ1) thereby departing from the original principles (Bekpen 2005b) to accommodate more information about the phylogenetic origin of the respective sequence. As it is rather difficult to predict whether a distinct substitution in the nucleotide-binding site renders a GTPases non-functional, only a clear disruption of the universally conserved G1 motif/P-loop was considered sufficient to qualify an IRG member for the suffix “q”. The GMS GTPases are excluded from this rule despite the unusual methionine in the G1 motif since they have been shown to be functional *in vivo* (Collazo 2001; Taylor 2000; Taylor 1997). All IRG sequences with non-canonical substitutions in the G1, G2 and G3 motifs are listed in Table 8. At least some of the IRGs with modified G3 or G4 motifs only will surely also prove to lack GTPase function upon experimental analysis. The quasi IRGs, which are neither orthologous to murine Irgq nor can be assigned to a defined phylogenetic clade of IRG genes, will be preliminary named *IRGXQ* (e.g. DR_*irgxq1*, formally termed *irgq1*). The entire IRGQ nomenclature should be considered provisional.

III.1.2. IRG genes of the C57BL/6 mouse

The C57BL/6 mouse (*Mus musculus domesticus)* possesses 26 p47 GTPase (IRG) related sequences, including the six previously described members of the family (Boehm 1998), localised in four clusters on chromosomes 7 (band A2), chromosome 11 (bands B1.2 and B2) and chromosome 18 (band D2) (mouse genome release v28.33d.1, February 2005) (Figure 9, Figure 10). Twenty-five of the sequences contain all the features typical for the IRG family while one sequence is more distantly related and contains drastic alterations in the GTPase domain (Irgq, see below, Table 8). The 26 sequences represent 21 genes, as four of the IRG coding units on chromosome 11 are apparently only transcribed in tandem with a second coding unit, thus giving rise to a protein twice the size of a typical immunity-related GTPase (see below). Chromosome 7

contains two IRG genes in close proximity, termed *Irgc* or CINEMA and *Irgq* (also known as FKSG27) (Figure 9, Figure 15), while chromosome 18 contains 8 closely related genes in a stretch of 220 Mb, (Figure 9, Figure 10). Chromosome 11 carries 16 IRG-related sequences in two clusters of 230 and 32 Mb separated by 65 Mb (Figure 9, Figure 10).

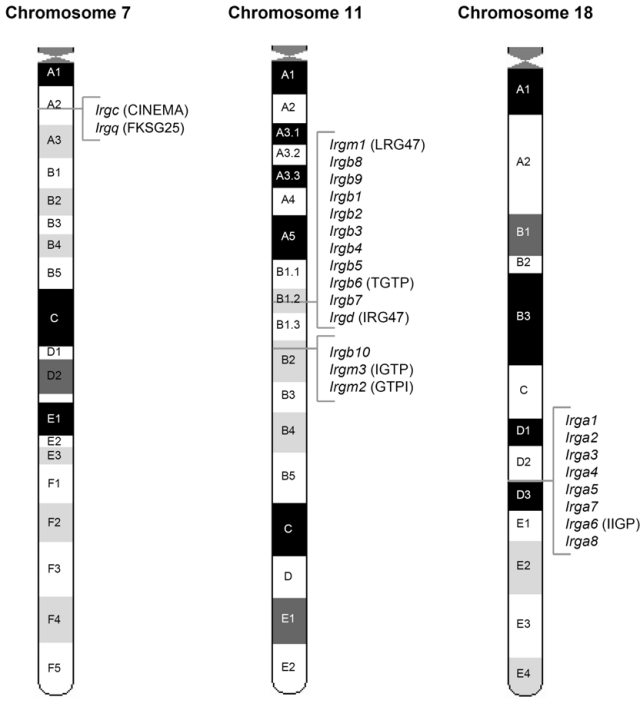


Figure 9 Chromosomal location of the IRG GTPase in Mus musculus domesticus. Disposition of the IRG sequences on the mouse karyotype (mouse genome Ensembl release v28.33d.1, February 2005). Individual IRG genes are listed in correct gene order in each cluster. In the most recent Ensembl database version, the cluster on chromosome 7 lies in band A3 (v43.36d, Feb. 2007). The location of the other sequences is unaltered.

The protein sequences encoded by the identified IRG genes were aligned based on the crystal structures of Irga6 (IIGP) (Ghosh 2004) using ClustalW ((Higgins 1994), http://www.ebi.ac.uk/Tools/clustalw) (Figure 11). The G-domains were further aligned on the structure of H-Ras-1 (Pai 1990; Wittinghofer 1991). As in Irga6, Irgb6 and Irgd (Boehm 1998; Ghosh 2004), four classical GTPase motifs (Bourne 1991) are conserved in most newly identified murine IRG proteins: the P-loop or G1 motif (Gx_4GKS/T), the G3 motif (DxxG), the G4 motif (N/T/SxxD), and the G5 motif (SAK/L in small GTPases, most commonly SNF in the IRGs). The IRG proteins are typically 400-450 amino acids in length. In contrast, the murine 'quasi' IRG protein Irgq is extended amino-terminally relative to other IRG GTPases by about 180 residues encoded on a single exon preceded by two untranslated upstream exons. The remaining 407 residues, encoded on a single long exon, are clearly homologous to and collinear with the other IRG proteins (Figure 13, see also (Bekpen 2005b)), especially in the amino- and carboxyl-terminal parts of the main exon. The region of lowest similarity is in the G-domain, and conserved GTP-binding motifs are lacking (Table 8, Figure 13, Figure 14; see also DNA matrix *Irgq* vs. *Irgc* in (Bekpen 2005b)). Thus, *Irgq* is not a GTPase despite its phylogenetic relationship to the other *IRG* proteins. *Irgq* is closely linked to *Irgc* in mouse, humans, chimpanzee, rat and dog and probably also in other mammals

(Figure 15-17). As the *IRGQ* genes diverge so much from the classical IRG genes and form an isolated, conserved group in the phylogenetic analysis (Figure 20), it can be questioned whether they truly belong to the family of IRG GTPases. Contrariwise, *IRGQ* genes clearly carry sequence features that are characteristic for the IRG family, especially outside of the GTPase domain where IRGs are most divergent from other GTPases (Figure 13, Figure 14). Furthermore, *IRGQ* genes are easily recovered from public databases by classical TBLASTN searches with the sequences of classical IRG family members. The expectation values for the resulting alignments are low, indicating significance (e.g. 1.1^{-36} for the alignment of mouse Irgc with mouse Irgq). The expectation values for pairwise alignments of divergent members of the classical IRG proteins are comparable (e.g. 7.5^{-41} for Irgc with Irgm1) though the values are significantly lower for more closely related family members (already 1.9^{-85} for Irgc with Irgd). Members of other GTPase families do not constitute significant hits in these BLAST searches. Together with the close spatial linkage to *IRGC* genes, this argues for the *IRGQ* genes as legitimate, though distantly related, members of the IRG family. Nevertheless, on grounds of the profound differences to the classical IRGs, it is unlikely that IRGQ proteins are functionally related to the other IRG proteins.

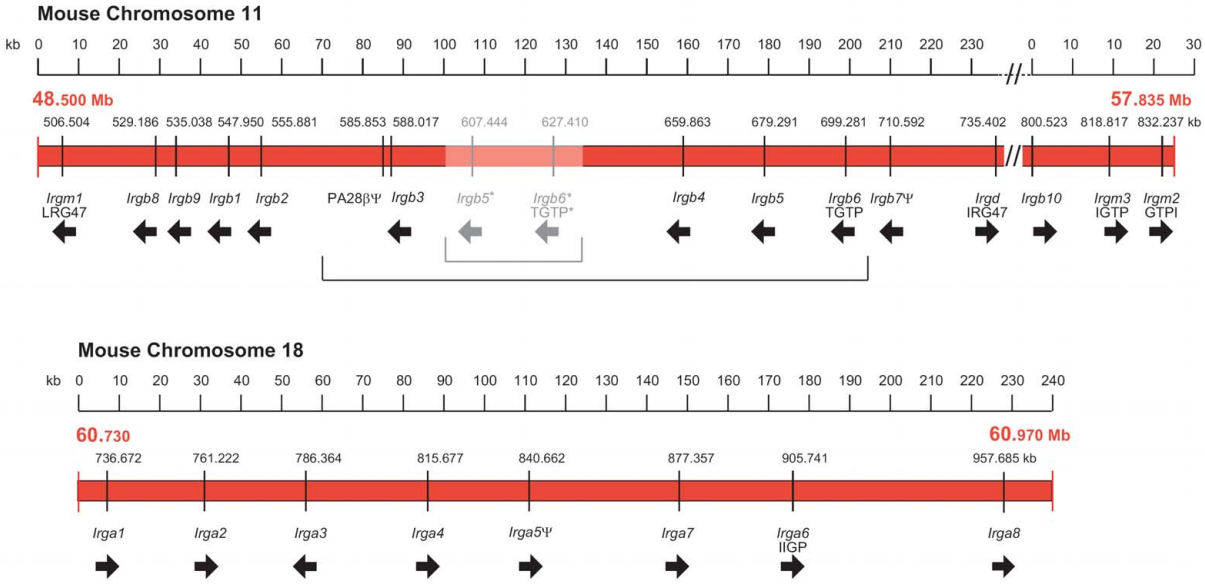


Figure 10 Genomic positioning of mouse IRG GTPases *(Mus musculus domesticus).* Positioning and orientation of the murine IRG genes in the chromosome 11 and 18 clusters. Positions of genes refer to the location in mouse genome release v28.33d.1 (February 2005) of the first G of the second glycine codon of the G1 motif (Gx_4**G**KS or **G**MS) of each gene. The segments of the chromosome 11 cluster indicated with square brackets are regions of uncertain structure. Gene orientation is given by black arrows. The shaded region of the chromosome 11 maps might represent a duplication introduced in the Ensemble mouse genome in an attempt to resolve a region of high ambiguity indicated by the longer square bracket. Though BAC sequence analysis indicated a misalignment of genomic sequences, the existence of independent transcripts reflecting the two distinct copies of *Irgb6* suggests a recent genomic duplication with little time to diverge by genetic drift. The independent existence of the two sibling sequences *Irgb3* and *Irgb4* is proved by the proximity of the PA28βΨ retropositioned pseudogene to *Irgb3* but not to *Irgb4*, in addition to consistent sequence differences. (*) Indicates recently duplicated genes with not or only minor sequence differences to their sibling genes.

In Irgm1, Irgm2 and Irgm3, the otherwise universally conserved Gx_4GKS/T sequence in the P loop is replaced by Gx_4GMS (Figure 11). This striking difference correlates with other sequence features to define the GMS subgroup of the IRG GTPases,

which form a distinct clade in the phylogenetic analysis of both nucleotide and amino acid sequences (Figure 11, Figure 12; (Boehm 1998)). From alignment of Irga6, Irgb6, Irgd, Irgm1, Irgm2 and Irgm3 alone it was not predictable which of the (partially) conserved threonines corresponding to positions 102, 107 or 108 in the switch I region of Irga6 functions as G2 motif (Ghosh 2004). The alignment in Figure 11 reveals that only Thr108 is conserved in all murine IRG proteins. Furthermore, the threonine residue at this position is highly conserved in IRG proteins throughout the euchordates (Figure 13, Figure 19, Figure 23) making it a good candidate as potential equivalent of the G2 motif Thr35 in H-Ras (Wittinghofer 1991). However, in the GppNHp-bound crystal structure of Irga6 neither Thr102 nor Thr108 are involved in Mg^{2+}- or γ-phosphate coordination (Ghosh 2004). IRG specific sequence features that are absent from H-Ras are apparent over the complete length of the proteins, though the N- and C-terminal parts are clearly more divergent than the GTPase domain (Figure 11). For eleven members of the IRGA and IRGB group a potential myristoylation motif was identified at the N-terminus (Table 9; MGQLFSS in Irga6, for consensus sequence see figure legend).

IRG	IRGQ name	G1 motif	G3 motif	G4 motif
MM_Irga5Ψ		**G**x$_4$**GK**Y		
MM_Irgb2			**D**xxS	
MM_Irgb7Ψ			ExxA	
EE_GKS1				Mxx**D**
EE_GKS2	EE_Irgbq1	**G**x$_4$**G**NF		
EE_GKS3			PxxN	
CF_GKS2			PxxH	
CF_GKS3	CF_IRGDQ1	Qx$_4$RMY		
DN_GKS1				Axx**D**
DN_GKS3	DN_IRGBQ2	**G**x$_4$K--		
DN_GKS4				DxxH
DN_GMS1			Yxx**G**	
OA_GKS6				**T**xxV
OG_GKS1			**D**xxA	
MN_IRGM		**G**x$_4$**GM**N		
CJ_IRGM		**G**x$_4$**SIS**		
MU_GMS3		**G**x$_4$**DMS**		
MU_GMS4			T**xx**A	
Mammalian Irgq		**G**x$_4$**G**F/LV	T/S/A/HxxL	I/**T**xxQ/E/K
DR_irgxq1	DR_irgxq1	**G**x$_4$**GVS**		
DR_irgxq2	DR_irgxq2	**G**x$_4$E**K**A	**D**xxP	LxxS
DR_irgxq3	DR_irgxq3	**G**x$_4$**GSS**		
XT_GKS1	XT_irgxq4	**G**x$_4$VRH		AxxE
XT_GKS2			ExxL	
Consensus motif		**G**x$_4$**GK/MS/T**	**D**xx**G**	**N/T/S**xx**D**

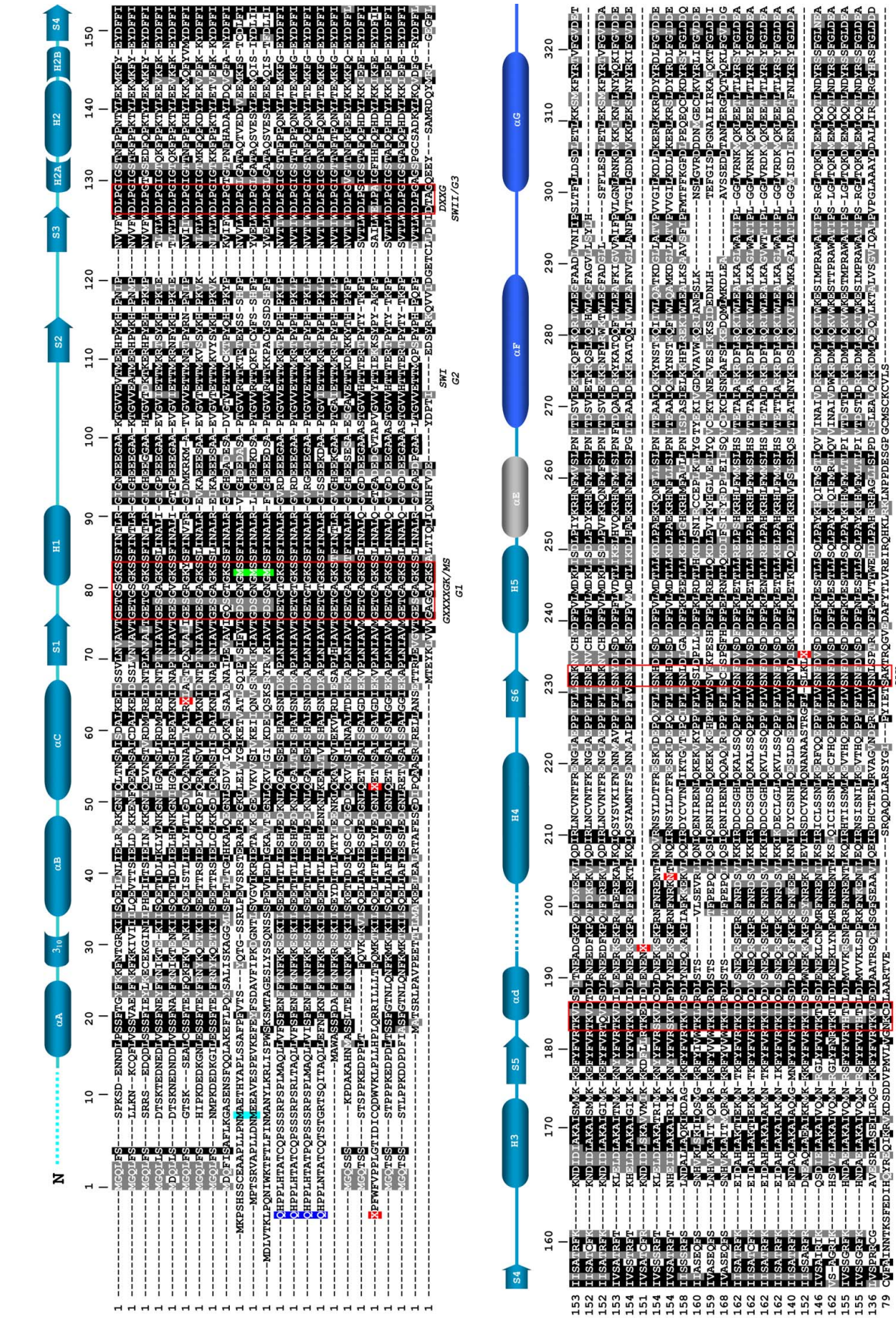
Table 8 IRG GTPases with modified GTPase motifs. List of all IRG proteins with non-canonical substitutions in one or more of the GTPase motifs. Only the altered motifs are given. Consensus residues are printed in bold. X: irrelevant residues. As MM_Irga5Ψ and MM_Irgb7Ψ terminate before the G1 motif due to premature stop codons they are not considered to be 'quasi' IRGs but pseudogenes.

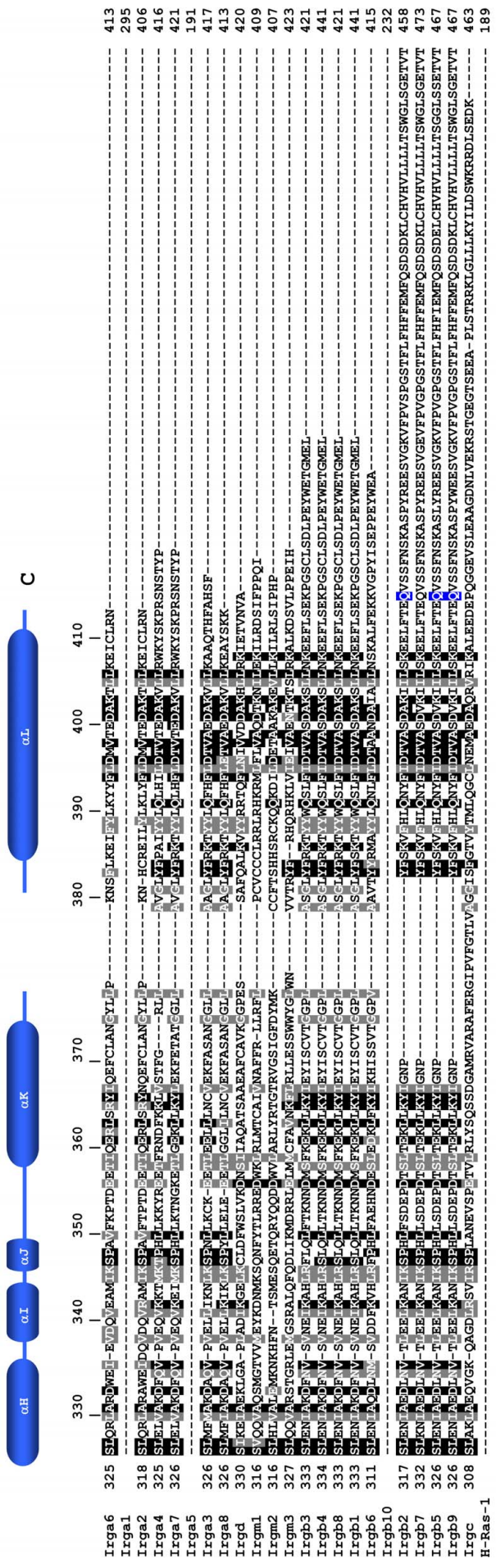
All the IRGs on chromosome 18 show closest sequence homology to Irga6 (IIGP) and cluster together in phylogenetic analysis (Figure 11, Figure 12). The members of this subfamily (IRGA) were subsequently termed Irga1 to -a8 (Figure 10). The two clusters of IRG genes on chromosome 11 include *Irgd* (IRG-47) and the 3 GMS GTPases *Irgm1* (LRG-47), *Irgm2* (GTPI) and *Irgm3* (IGTP) (Figure 10, Figure 11, Figure 12). The remaining 12 IRG sequences belong to the IRGB group and are most similar to *Irgb6* (TGTP) (Figure 12). Three of these sequences, *Irgb4, Irgb5* and *Irgb6*, seem to have been duplicated in conjunction giving rise to a second, closely related block of three IRGB sequences: *Irgb3* differing by only 11 nucleotides (resulting in six amino acid exchanges) from *Irgb4,* a second identical copy of *Irgb5 (Irgb5*)* and *Irgb6**, which

differs from *Irgb6* by four nucleotides but is unaltered at the amino acid level (square brackets in Figure 10) (see appendix V.6 and V.7). As these two regions are also highly similar in the intragenic space, two explanations are conceivable for their origin. This stretch of DNA might either represent a misalignment of sequences during the assembly of the mouse genome in a region of high ambiguity and therefore an artificial duplication *in silico* or a very recent genomic duplication with little time to diverge. BAC sequence analysis made the first hypothesis favourable, especially as the region is rich in repetitive elements and contains a large number of closely homologous IRG genes. In earlier versions of the mouse genome assembly, this region and the area of chromosome 18 carrying the *IRGA* genes were subjected to extended changes and rearrangements, frequently leading to the loss or duplication of one or the other *IRG* gene from the database for months. The independent existence of *Irgb3* and *Irgb4,* however, is proved by the proximity of the PA28βΨ retropositioned pseudogene to *Irgb3* but not to *Irgb4*, in addition to consistent sequence differences (Figure 10, appendix V.7). Furthermore, recent detailed analysis of novel ESTs and cDNA sequences available in the public databases confirms the existence of the two independent versions of *Irgb6* distinguishable by 4 non-coding substitutions in the ORF and a few additional exchanges in the untranslated regions (UTRs) (*Irga6** (AK163978, BE632518), *Irga6* (L38444); Table 10, appendix V.6). Taken together, this supports the presence of both sequence blocks in the mouse genome.

Our understanding of the situation will grow as new versions of the mouse genome, transcript databases and detailed experimental data become available as nicely exemplified by the history of *Irgb10*. This gene was thought to be truncated at the C-terminus due to a premature stop codon right after the G5 motif based on the version of the Ensembl mouse genome release used for analysis (v28.33d.1, February 2005), despite its IFNγ-inducibility in murine fibroblasts (Bekpen 2005a; Bekpen 2005b). However, Frank Kaiser identified full length Irgb10 ("cIGP_9", XM_137576) as an interaction partner of Irga6 (IIGP) in a yeast two hybrid screen (Kaiser 2005). Furthermore Bernstein-Hanley and colleagues revealed that *Irgb10* is an intact gene functional in resistance against *Chlamydia trachomatis* (Bernstein-Hanley 2006). Recently, *Irgb10* was also implicated in resistance to *Chlamydia psittaci* (Miyairi 2007). In the latest version of the mouse genome, *Irgb10* is represented as an intact gene on chromosome 11 at bp 58,004,193 (first base of the glycine codon of the G1 motif; mouse genome Ensembl release v43.36d, Feb 2007).

Figure 11 Amino acid alignment of the mouse IRG GTPases (see next pages). Amino acid sequences of 23 mouse IRG sequences showing the close homology extending to the carboxyl-terminus, aligned on the known secondary structure of Irga6 (indicated in blue above sequence alignment). The sequences of notional products of the two pseudogenes Irga5 and Irgb7 have been partially reconstructed; premature terminations are indicated by red highlighting. In the C57BL/6 mouse the sequence of the *Irga8* gene is damaged by an adenine insertion, indicated by the red highlighted N at position 204. The sequence given after this point is derived by correcting the frame shift. *Irga8* from *Mus musculus musculus* (CZECHII) lacks the extra adenine (BC023105). The turquoise-highlighted M in Irgm1 and Irgm2 are initiation codons that are dependent on alternative splicing; the unusual methionine residues in the G1 motif of GMS proteins are highlighted in green. The blue background Q residue of Irgb5, Irgb2 and Irgb8 indicate the point where tandem splicing occurs to Irgb4/b3, Irgb1 and Irgb9, respectively. Canonical GTPase motifs are indicated by red boxes. Note that the C-terminal part of Irgb10 is missing due to a premature stop codon in previous versions of the Ensembl mouse genome (v28.33d, Feb 2005) (see main text below; for alignment of the full length Irgb10 protein see Figure 13). (The figure was kindly provided by Cemalettin Bekpen and modified slightly; (Bekpen 2005b))




αH
αI
αJ
αK
αL
C
Irga6
Irga1
Irga2
Irga4
Irga7
Irga5
Irga3
Irga8
Irgd
Irgm1
Irgm2
Irgm3
Irgb3
Irgb4
Irgb8
Irgb1
Irgb6
Irgb10
Irgb2
Irgb7
Irgb5
Irgb9
Irgc
H-Ras-1

The open reading frame of an *IRG* gene is typically encoded on a single long 3' exon behind one or more 5'-untranslated exons with the exception of *Irgq* (see above) and of alternative splice forms of *Irgm1* and *Irgm2* where the initial methionine is encoded on the penultimate exon (Bekpen 2005b). Furthermore, four of the closely related IRGB sequences on chromosome 11 *(Irgb1, Irgb3, Irgb4* and *Irgb8)* apparently occur only as tandem transcripts in-frame with the respective closely linked upstream IRGB sequences (*Irgb2, Irgb5*, Irgb5* and *Irgb9)* (for accession numbers see appendix V.5). To date, neither dedicated 5'-untranslated exons nor putative promoters containing IFN-inducible elements could be identified for *Irgb1, Irgb3, Irgb4* and *Irgb8*. Furthermore, no independent transcripts were detected for these downstream tandem domains. Therefore, the downstream parts of the tandems are strictly speaking not genes but exons of genes twice the size of typical IRG GTPases. Three of the tandems are supported by ESTs in the databases (*Irgb2-b1, Irgb5-b4, Irgb9-b8*; for accession numbers see appendix V.5), though further information is needed to define the 5' prime region of the *Irgb9-Irgb8* transcript (AK144287, AK165747, and Li unpublished data). The donor and acceptor sites mediating the splicing from the upstream *IRGB* unit into the downstream one are conserved between the three distinct tandems and in different mouse subspecies (C57Bl/6 (*Mus musculus musculus*), Jyg (unidentified Chinese subspecies, (Imai 1994)), CZECHII (*Mus musculus domesticus*) (appendix V.5). The fourth tandem, *Irgb5*-Irgb3*, was postulated in analogy to the other tandem pairs. Indeed, the 60 nucleotides upstream of the start codon of *Irgb3* including the splice acceptor site are conserved with the other downstream units of the tandem genes (Figure 13, appendix V.13). To date, no ESTs could be identified in the databases representing the *Irgb5*-Irgb3* tandem, but recent experimental analysis confirmed the presence of all four tandem transcripts in IFNγ-induced mouse embryonic fibroblasts (Li, unpublished data). Furthermore, the upstream tandem elements were detected as independent transcripts, indicating that two alternative splice variants exist for the tandem genes – a tandem version and a normal sized one. Hence, chromosome 11 contains 12 *IRG* genes, four tandem genes and eight classical, single exon genes. Since the majority of IRG GTPases is encoded on one long exon and for simplicity reasons, the individual exons of the tandem genes will nevertheless be treated as single units for the phylogenetic analysis.

Of all CB57BL/6 mouse IRGs, only *Irga5* and *Irgb7* are unambiguously pseudogenes due to premature stop codons before the first GTP-binding motif (Figure 11). The remaining 21 classical *IRG* genes are intact across the GTP-binding domain, although *Irga1* and *Irga8* are carboxyl-terminally truncated relative to the majority, and no transcripts of *Irga7* have yet been found. *In silico* analysis revealed that most mouse p47 GTPases, except for *Irgc* (CINEMA), *Irga7* and those genes transcribed only as the 3' half of a tandem, contain IFN-inducible elements (GAS and ISRE) in their putative promoters, and 14 have been shown experimentally to be IFN-inducible (namely *Irga2, -a3, -a4, -a6, -a8, Irgb1, -b2, -b5, -b6, -b10, Irgd, Irgm1, -m2* and *-m3)* (Bekpen 2005a; Bekpen 2005b; Boehm 1998; Gilly 1996). The proximal promoter region of *Irgc*, which was neither infection- nor IFN-induced, is devoid of ISRE or GAS elements but contains putative Sox5, Sox17 and NF-Y transcription factor binding sites (Sox: Sry-type high-mobility-group domain box; (Bekpen 2005b; Rhode 2007)). Sox5 and Sox17 are expressed in haploid spermatids but are not testis specific (Wegner 1999) while NF-Y is

a universal transcription factor (Mantovani 1999). *Irgc* is expressed exclusively and constitutively in the testis (Bekpen 2005b; Rhode 2007). While their IFN-inducibility suggests a function related to immunity for most of the IRG GTPases, the developmental regulation paralleling sexual maturity and the expression limited to haploid spermatids implicate *Irgc* in sperm development (Rhode 2007). However, no infertility or other phenotype has yet been detected in Irgc-deficient mice (Rhode 2007).

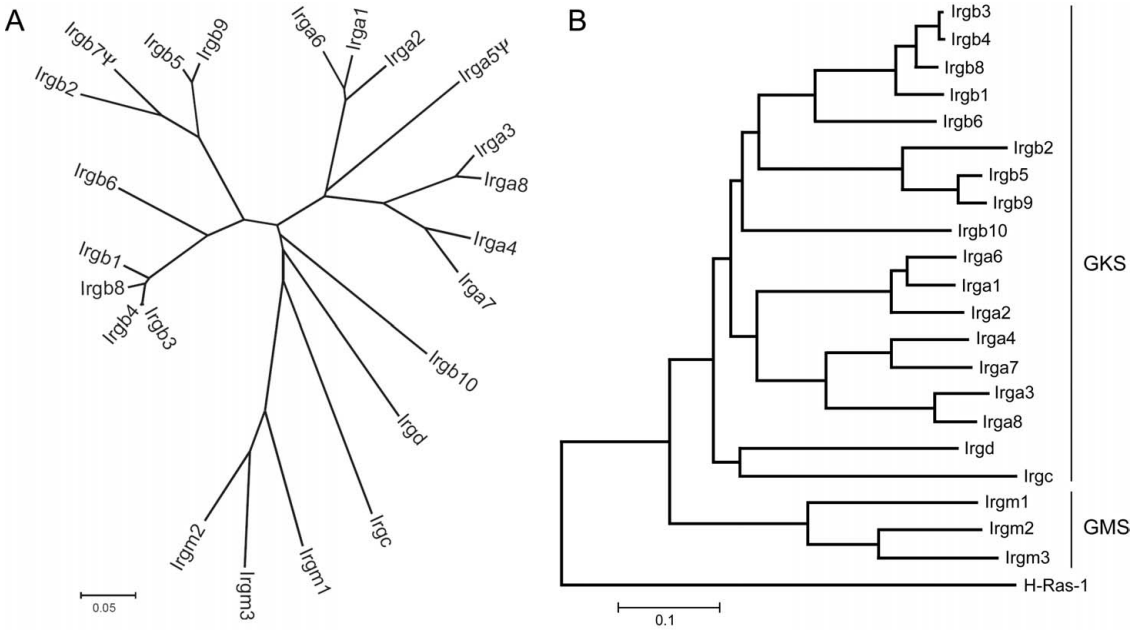


Figure 12 Phylogenetic relationship of mouse IRG GTPases. (A) Unrooted phylogenetic tree (p-distance based on neighbour-joining method) of nucleotide sequences of the G-domains of 23 mouse IRG GTPases, including the two presumed pseudo-genes *Irga5* and *Irgb7*. The amino acid sequences of all IRG sequences are given in the appendix V.13. (B) Phylogenetic tree of the amino acid sequences of the G-domains of 21 mouse IRG GTPases rooted on the G-domain of H-Ras-1 (accession number: P01112). The products of the two presumed pseudo-genes *Irga5* and *Irgb7* are excluded from the analysis. (The figure was kindly provided by Cemalettin Bekpen and modified slightly (Bekpen 2005b))

Chromosome 11 contains the most divergent IRG sequences (Figure 10, Figure 11, Figure 12), including all three GMS (IRGM) GTPases as well as *Irgd* and the IRGB genes (Boehm 1998), suggesting that this cluster is relatively ancient. In contrast, the eight IRGA genes on chromosome 18 cluster phylogenetically, suggesting more recent divergence, probably from a translocated member of the IRGB cluster on chromosome 11. The isolated IRG gene on chromosome 7, *Irgc*, is an ancient root with no obvious systematic relationship to the other subfamilies. Within the chromosomal clusters, more recent duplication events are apparent (see also above). The downstream elements of the tandem genes, *Irgb1*, *Irgb3*, *Irgb4*, and *Irgb8,* cluster phylogenetically and appear to have been duplicated in tandem with the likewise clustered upstream units, *Irgb2*, *Irgb5*, Irgb5* and *Irgb9*, respectively (Figure 10, Figure 12). The pattern of divergence in the mouse IRG tree suggests an old gene family that has undergone a succession of duplication-divergence cycles over time, a pattern of evolution that is still ongoing in several of the subfamilies.

A systematic study of polymorphism has not yet been completed but it is already clear that nearly all IRG sequences derived from the CZECHII cDNA libraries (*Mus musculus musculus*) differ from C57BL/6 sequences. These differences make allocation of many CZECHII sequences to individual clade members of the C57BL/6 mouse

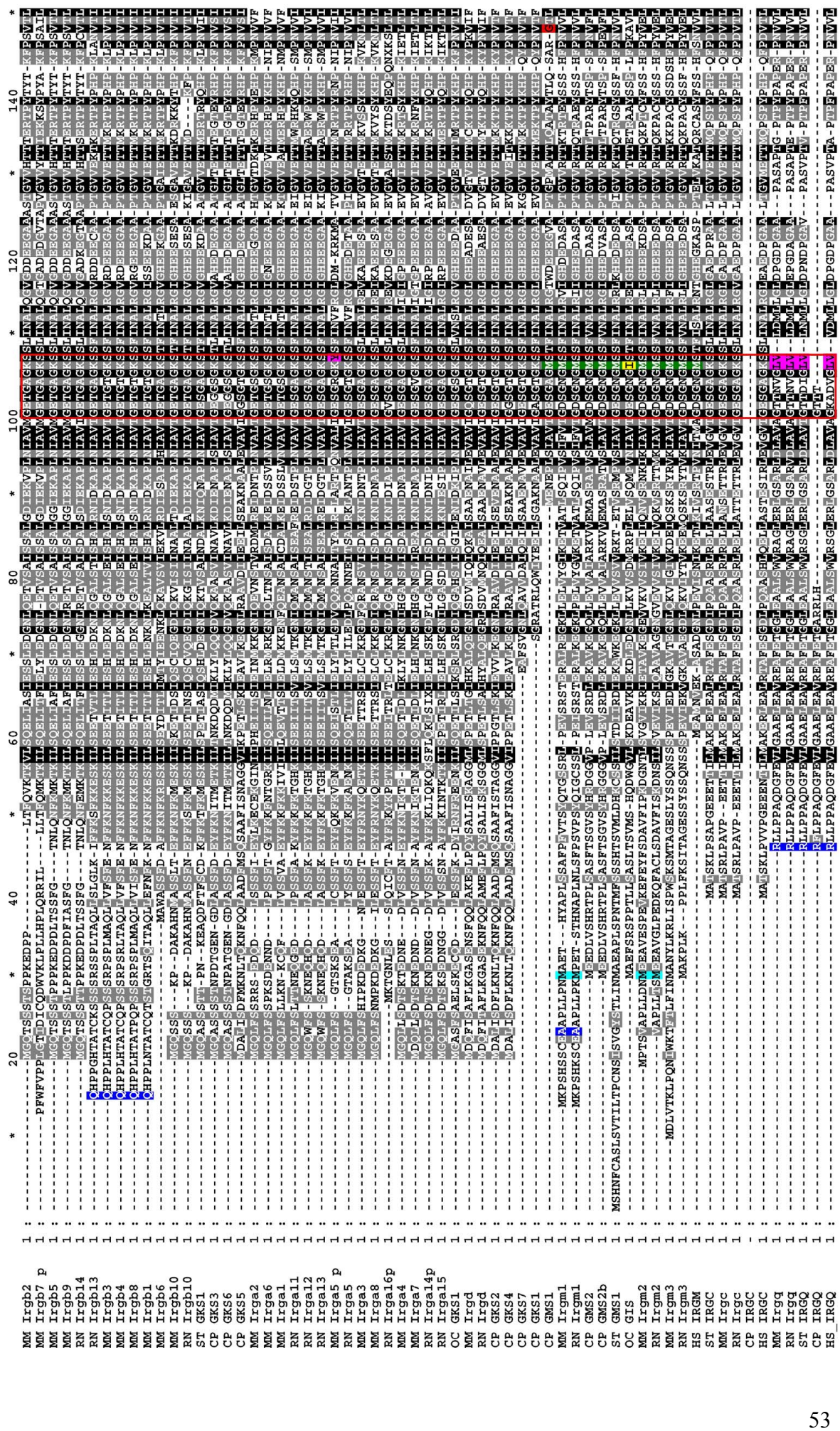
problematic. A thorough analysis of the IRG gene content of the south-eastern Asian mouse *(Mus musculus castaneus)* is currently under way (Li unpublished data). Homology searches revealed that the rat *(Rattus norvegicus)* genome contains a comparable complex set of IRG genes, whereas the representation of the IRG family is radically different in human.

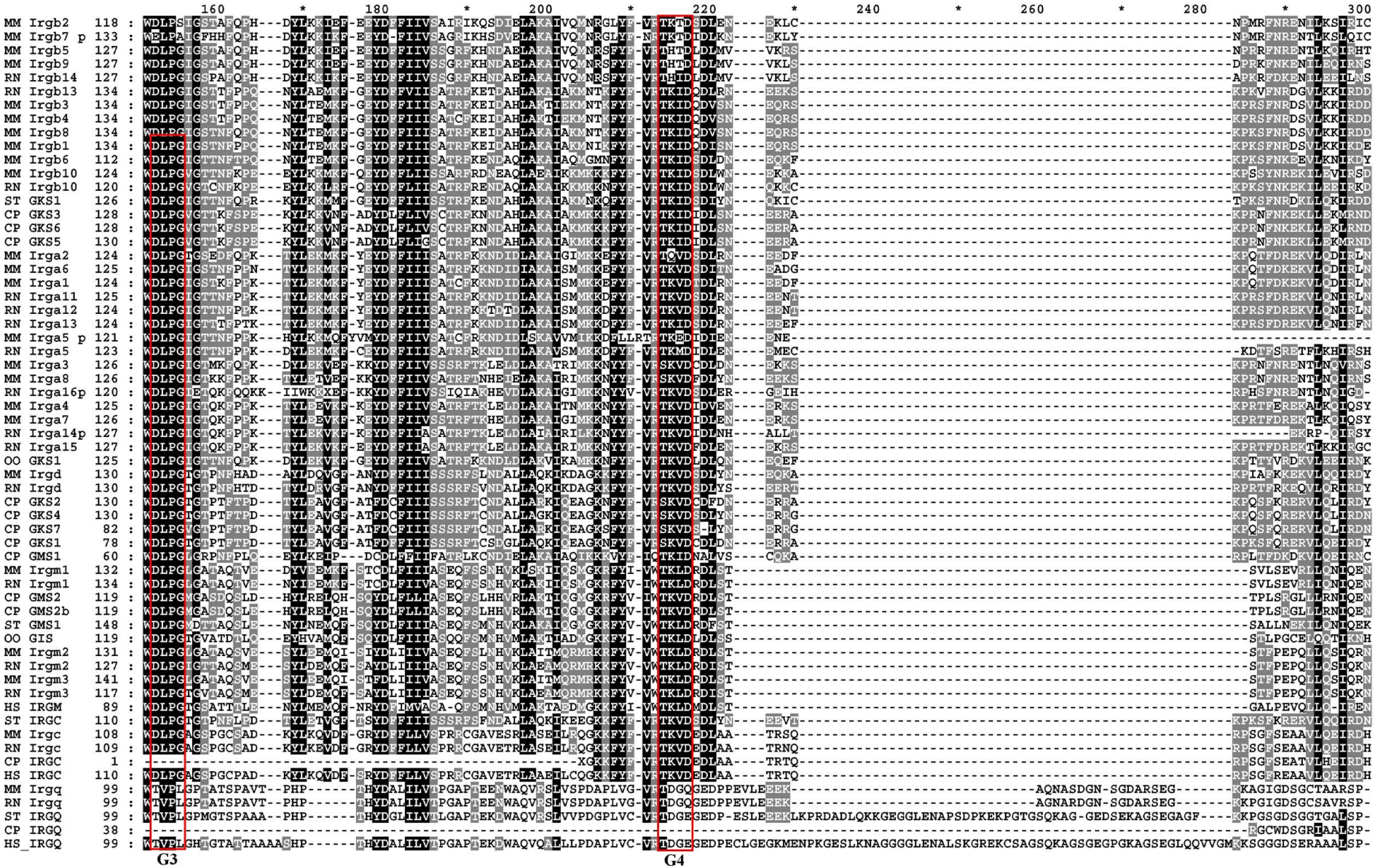
III.1.3. Human IRG genes and their synteny relationship to mouse IRGs

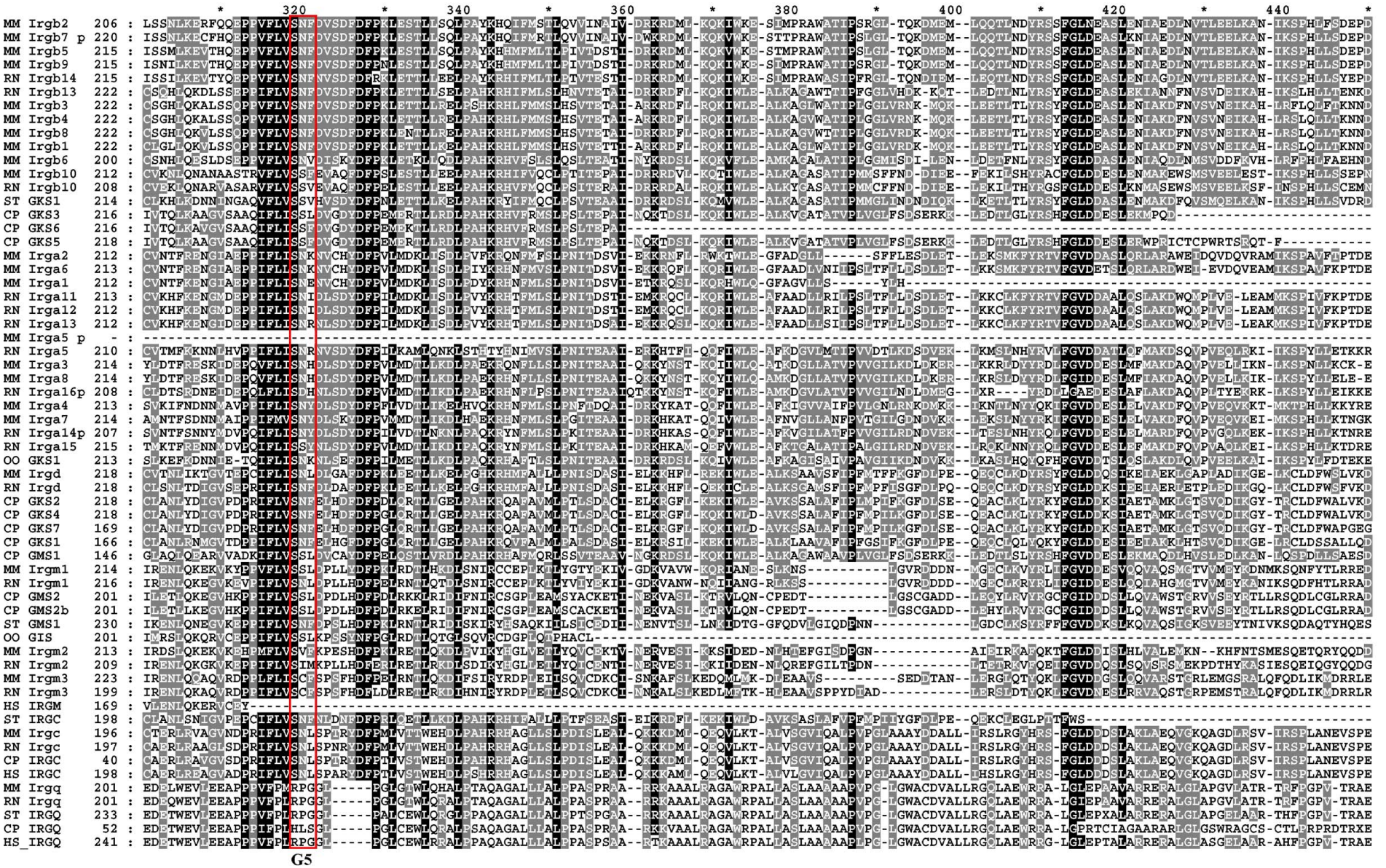
Only three IRG sequences, all transcribed, are present in human, two (*IRGC* and *IRGQ*) on chromosome 19 (19q13.31) and one (*IRGM*) on chromosome 5 (5q33.1). Human IRGC is more than 85% identical at the nucleotide level and 90% at the amino acid level to the isolated mouse gene *Irgc* on chromosome 7 (Figure 13). Human *IRGQ* is over 80% identical to mouse *Irgq* on the amino acid level and is likewise radically modified in all GTPase motifs but contains an insertion absent from the mouse protein (Table 6, Figure 13. appendix V.9). *IRGM* encodes an amino- and carboxyl-terminally truncated G-domain homologous to the IRGM (GMS) subfamily of mouse p47 GTPases (Figure 13, Figure 14). Human IRGM protein is 183 amino acids in length, as opposed to around 400 residues for the mouse IRGM proteins, and is 54%, 53% and 51% identical to mouse Irgm1, Irgm2 and Irgm3, respectively. Hence, human IRGM is not significantly closer related to Irgm1 in sequence than to the other two mouse GMS proteins, though this has occasionally been implied in the literature (Deretic 2006; Miyairi 2007; Singh 2006). Phylogenetic analysis confirmed that human IRGM is equally distant to all three mouse GMS proteins (Figure 14, Figure 21).

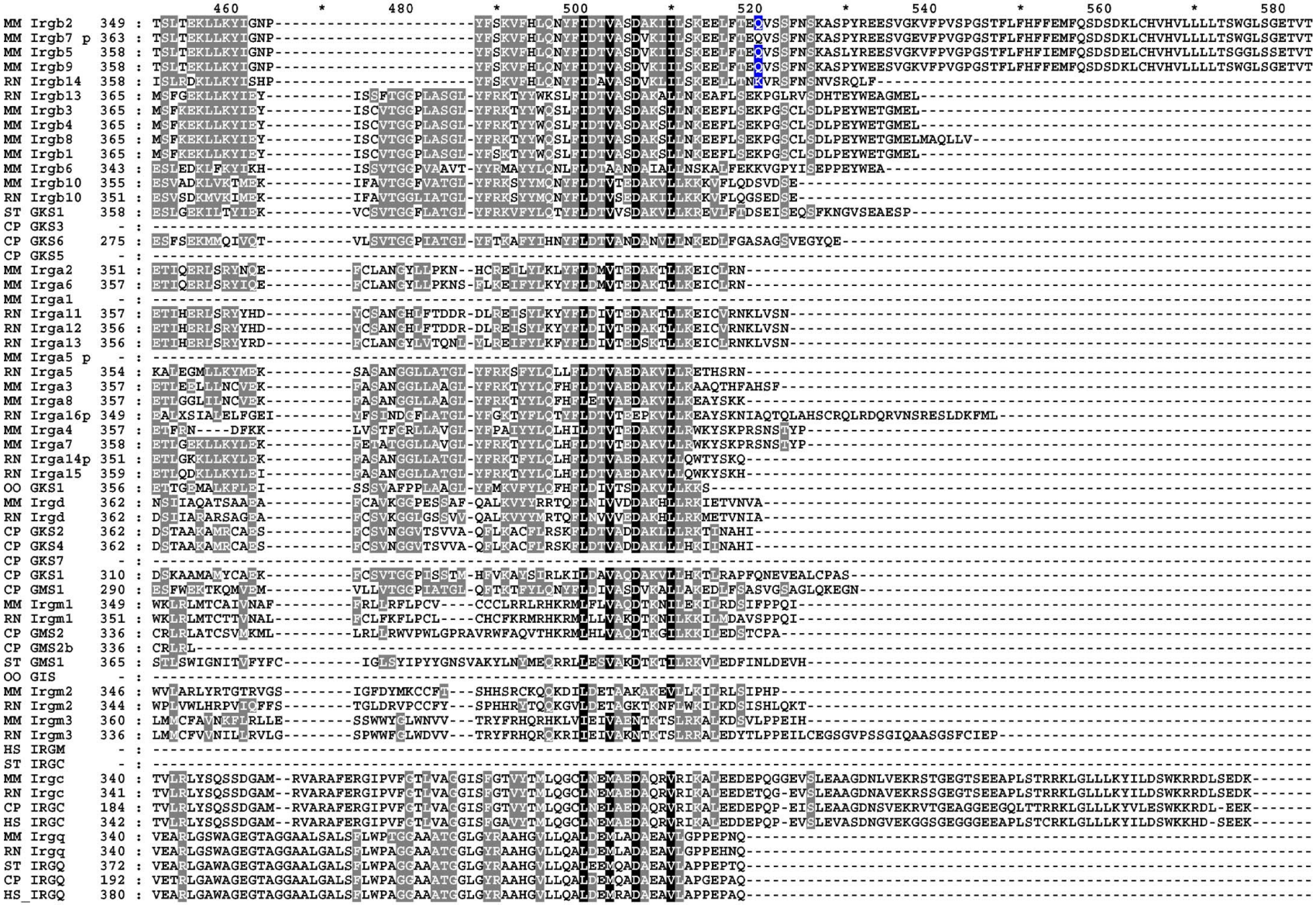
The mouse and human *IRGC* and *IRGQ* genes sit in chromosomal regions syntenic between chromosomes 7 and 19, respectively (Figure 15 A) and are clearly orthologous. The proximal promoter region of human *IRGC* is largely conserved with that of mouse *Irgc* and contains the same transcription factor binding sites (Sox5, Sox17 and NF-Y) (Rhode 2007). As in the mouse, no interferon response elements are found either in the proximal conserved region or in divergent regions up to 10 kb upstream of the transcriptional start (Bekpen 2005b; Rhode 2007). Human *IRGC*, like mouse *Irgc*, is not inducible *in vitro* by interferons and is strongly expressed exclusively in adult testis (Bekpen 2005b; Rhode 2007). The human genomic segments syntenic to the IRG gene containing regions on mouse chromosome 11 and 18 both mapped to human 5q33.1, suggesting that the interferon-inducible IRG proteins were once encoded in a single block ancestral to the human chromosome 5 region (Figure 15 B). *IRGM* (150.2 Mb) maps only 180 kb upstream and 90 kb downstream of the closest syntenic markers *TNIP1* and *DCTN4,* respectively (human genome Ensembl release v43.36d, Feb. 2007).

Figure 13 Alignment Glires and human IRG proteins (see next pages). The sequences of Irga8 and the two pseudogenes Irga5 and Irgb7 have been partially reconstructed (see Figure 11). The unusual methionine residue in the G1 motif of GMS proteins is highlighted in dark green, conservative exchange to isoleucine in yellow. The turquoise-highlighted M in Irgm1 and Irgm2 are initiation codons that are dependent on alternative splicing (dark blue). The dark blue shaded residues indicate the positions at which splicing from upstream exons or tandem splicing occurs form Irgb5, Irgb2, Irgb9 and Irgb14 to Irgb4/b3, Irgb1, Irgb8 and Irgb13, respectively. Disruptive mutations in the G1 motif are marked in pink. CP_GMS1 represents a hybrid sequence combining IRGM- and IRGB10-like features separated by a frame shift causing a premature stop codon (red highlighting). GTPase motifs are indicated by red boxes. Black shading: 80%, grey shading: 25% conserved. CP: *Cavia porcellus*, HS: *Homo sapiens,* MM: *Mus musculus*, OC: *Oryctolagus cuniculus,* RN: *Rattus norvegicus*, ST: *Spermophilus tridecemlineatus.*


G1
G2


160 * 180 * 200 * 220 * 240 * 260 * 280 * 300
MM Irgb2 118 :
MM Irgb7 p 133 :
MM Irgb5 127 :
MM Irgb9 127 :
RN Irgb14 127 :
RN Irgb13 134 :
MM Irgb3 134 :
MM Irgb4 134 :
MM Irgb8 134 :
MM Irgb1 134 :
MM Irgb6 112 :
MM Irgb10 124 :
RN Irgb10 120 :
ST GKS1 126 :
CP GKS3 128 :
CP GKS6 128 :
CP GKS5 130 :
MM Irga2 124 :
MM Irga6 125 :
MM Irga1 124 :
RN Irga11 125 :
RN Irga12 124 :
RN Irga13 124 :
MM Irga5 p 121 :
RN Irga5 123 :
MM Irga3 126 :
MM Irga8 126 :
RN Irga16p 120 :
MM Irga4 125 :
MM Irga7 126 :
RN Irga14p 127 :
RN Irga15 127 :
OO GKS1 125 :
MM Irgd 130 :
RN Irgd 130 :
CP GKS2 130 :
CP GKS4 130 :
CP GKS7 82 :
CP GKS1 78 :
CP GMS1 60 :
MM Irgm1 132 :
RN Irgm1 134 :
CP GMS2 119 :
CP GMS2b 119 :
ST GMS1 148 :
OO GIS 119 :
MM Irgm2 131 :
RN Irgm2 127 :
MM Irgm3 141 :
RN Irgm3 117 :
HS IRGM 89 :
ST IRGC 110 :
MM Irgc 108 :
RN Irgc 109 :
CP IRGC 1 :
HS IRGC 110 :
MM Irgq 99 :
RN Irgq 99 :
ST IRGQ 99 :
CP IRGQ 38 :
HS_IRGQ 99 :
G3
G4


G5



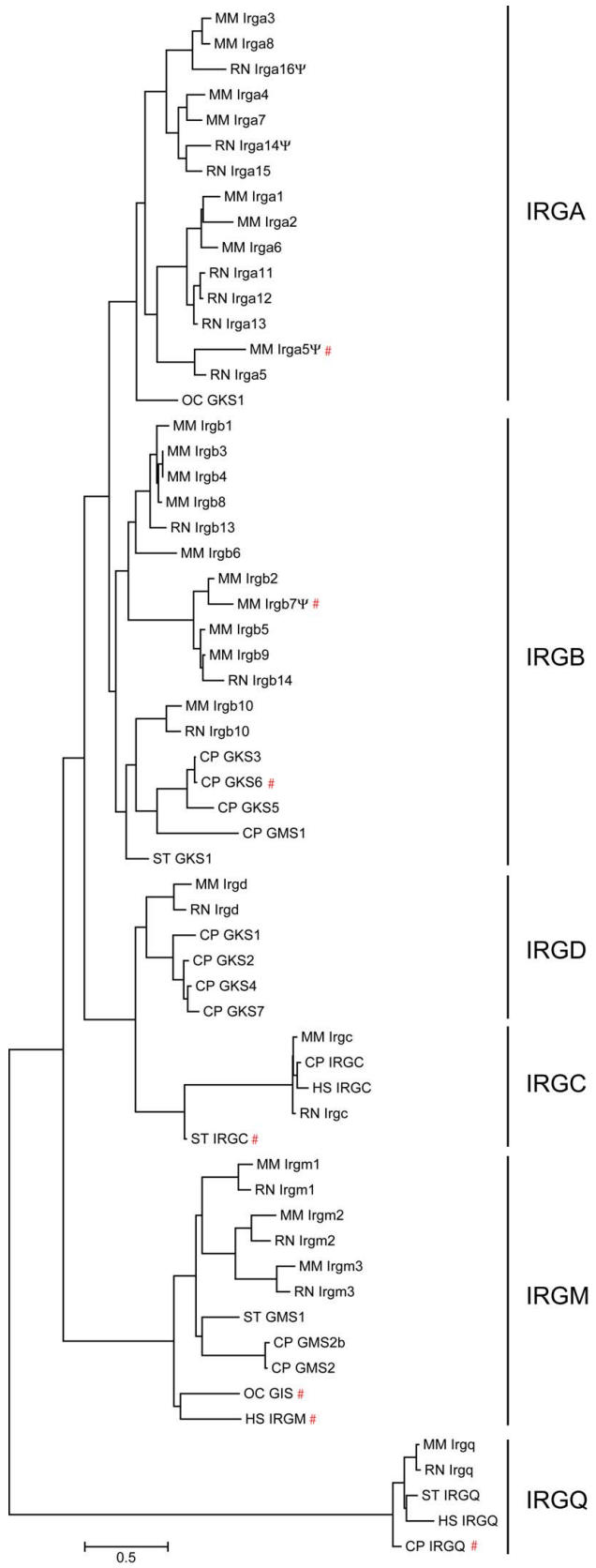


Figure 14 Phylogenetic relationship of Glires and human IRG proteins. Maximum likelihood consensus bootstrapped tree based on the alignment of Glires and human IRG proteins (see Figure 13) rooted on the IRGQ proteins. Only the G-domains were used for the phylogenetic analysis applying the JTT+I+G amino acid replacement matrix (Jones 1992). CP_GMS1 represents a hybrid sequence combining IRGM- and IRGB10-like features separated by a frame shift causing a premature stop codon right before the G3 motif (red highlighting in Figure 13). The red # indicates sequences that are shorter than the classical *IRG* genes. CP: *Cavia porcellus*, HS: *Homo sapiens,* MM: *Mus musculus*, OC: *Oryctolagus cuniculus,* RN: *Rattus norvegicus*, ST: *Spermophilus tridecemlineatus.*

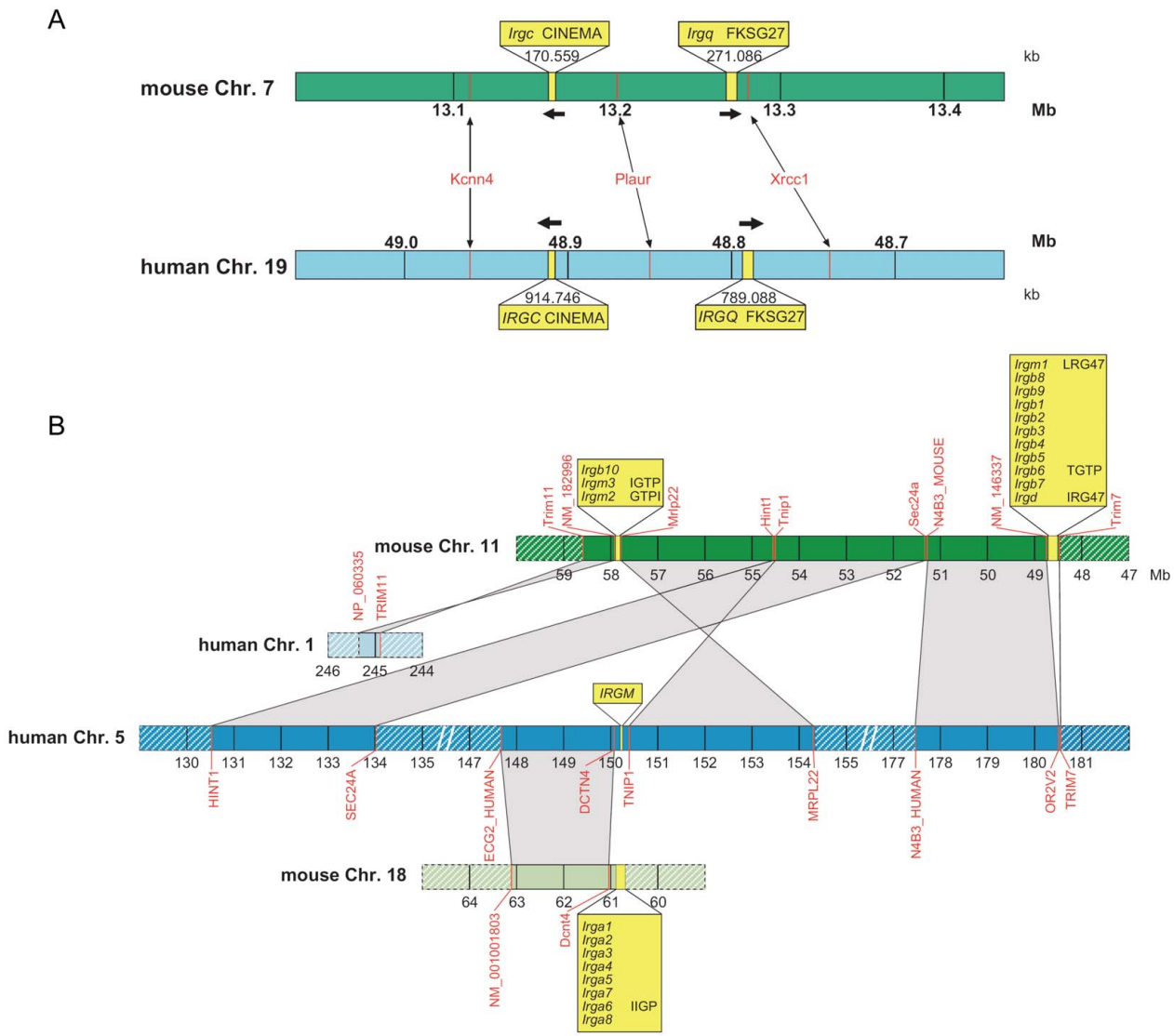


Figure 15 Synteny relationship between the human and mouse IRG genes. Figures indicate distances from the centromere in megabases. The locations of *IRG* genes are shown in the yellow panels. Positions of diagnostic syntenic markers are also indicated. Syntenic blocks are given in full colour, and the rest is shaded. A. Synteny between mouse chromosome 7 and human chromosome 19 in the region of the *IRGC* and *IRGQ* genes. Gene orientation is given by black arrows. B. Complex synteny relationship between human chromosome 5 and mouse chromosomes 11 and 18 in the regions containing the mouse IRG genes.

IRGM is transcribed in unstimulated human tissue culture lines with no increase after interferon induction (Bekpen 2005a; Bekpen 2005b). The promoter region corresponds to the ERV9 (endogenous retrovirus 9) U3 long terminal repeat (LTR) and does not contain any interferon response elements (Bekpen 2005a; Bekpen 2005b; Ling 2002). Polyadenylated transcripts of *IRGM* occur with five 3' splicing isoforms, three of which are probably subject to rapid RNA degradation by nonsense mediated decay (NMD) due to the presence of a stop codon more than 50 bp upstream of an exon-intron boundary (Singh 2003). IRGM protein was detected in several human cell lines in absence of IFN (Singh 2006). At the protein level the shortest isoform of IRGM is shorter than a canonical G-domain, being truncated in the middle of β-strand 6 just before the G5 sequence motif, which interacts with the guanine base of the bound nucleotide (Figure 19, see Figure 11 for alignment on Irga6 crystal structure, (Bekpen 2005b; Bourne 1991; Ghosh 2004)). The longer isoforms are terminated by short

sequence extensions that are unrelated to known GTPase domains. Hence, all residual IRG genes of human lack the character of functional resistance genes.

In order to clarify whether the situation in the mouse or the human represents the exception, the analysis was extended to other species including different orders of vertebrates.

III.1.4. IRG homologues in rodents and lagomorphs (Glires)

All rodents analysed so far contain numerous divergent IRG GTPases including representatives of the GKS as well as GMS subtype (Table 7; mouse: 19 GKS/3 GMS, rat: 13 GKS/3 GMS, guinea pig 9 GKS/3 GMS, thirteen-lined ground squirrel 3 GKS/1 GMS). While the GMS genes as well as *Irgd, Irgc, Irgq, Irga5* and *Irgb10* are reasonably conserved within the *Muridae*, other *IRGA* and *IRGB* genes have diverged to an extent that makes clear assignment of rat (*Rattus norvegicus*) sequences to specific mouse genes impossible (*Irga11-a16, Irgb13 and Irgb14;* Figure 14, see also (Bekpen 2005a); rat genome Ensembl release RGSC 3.4 v45.34o, Feb 2006). Analysis of the genomic positioning of the rat *IRG* genes revealed a situation very similar to the mouse. Rat IRG genes are found in four clusters on three chromosomes (18, 1 and 10) syntenic to the respective regions of mouse chromosome 18, 7 and 11 (Figure 16). Rat chromosome 18 (18q12.1) carries 7 *IRGA* genes clustered in a stretch of 230 kb. Rat *Irgc* and *Irgq* are situated in the band q21 of chromosome 1. Rat chromosome 10 contains two clusters of IRG genes separated by 9 Mb, one consisting of *Irgm1, Irgb13, Irgb14* and *Irgd* and the other of *Irgb10, Irgm3* and *Irgm2.*

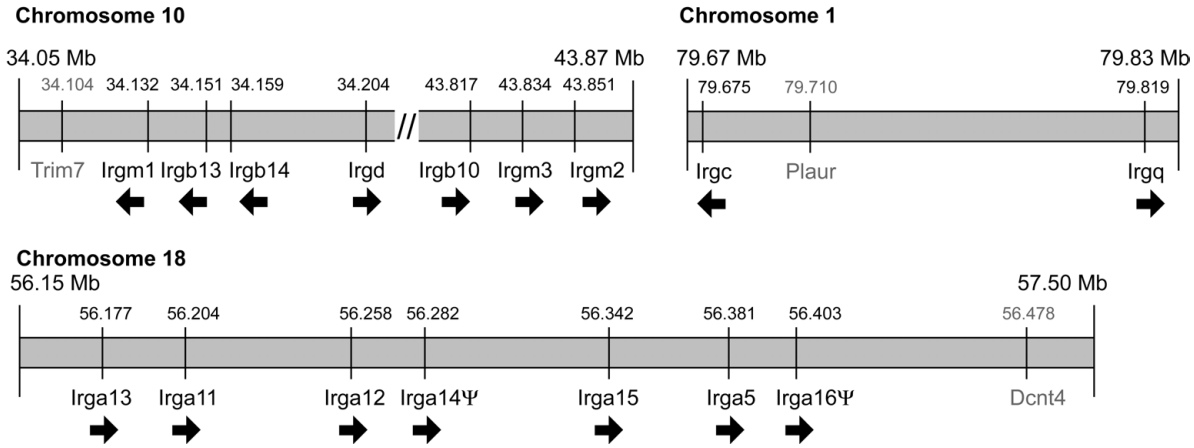


Figure 16 Genomic positioning of rat *(Rattus norvegicus)* IRG GTPases. Positioning and orientation of IRG genes in the rat chromosome 10, 1 and 18 clusters. Positions of genes refer to the location in rat genome Ensembl release v45.34o (Aug. 2007); Gene orientation is given by black arrows. *Irgb14* and *Irgb13* are probably expressed as tandem transcript. The closest syntenic markers are indicated in grey.

Interestingly, one of these IRGB genes, *Irgb13,* lacks a start codon and is closest related to mouse *Irgb1, -b3, -b4* and *-b8*, all of which have been shown to be expressed as the downstream half of a tandem transcript (see above and Figure 14, Figure 13). Sequence analysis revealed that the 60 nucleotides upstream of the mutated start codon including the (putative) splice acceptor site are conserved between rat *Irgb13* and the downstream parts of the mouse tandems (Figure 13, appendix V.13). Furthermore, a second IRGB gene, *Irgb14,* is situated only 8 kb upstream of *Irgb13* in the same

orientation (Figure 16). *Irgb14* clusters phylogenetically with the upstream units of the mouse IRGB tandems (*Irgb2, Irgb5*/Irgb5* and *Irgb9)* (Figure 14) and the splice donor site used by the mouse tandems is conserved (see appendix V.13). Thus, *Irgb14* and *Irgb13* are probably expressed as a tandem, though to date no confirming transcripts could be identified in the databases. Interestingly, only one IRG gene tandem pair is present in the rat genome as opposed to four pairs in the mouse indicating recent successive duplication events in the murine line (compare Figure 10 and Figure 16). Thus, the only independently transcribed *IRGB* gene in rat is *Irgb10.* If one reconsiders the mouse situation on this basis, it becomes likely that *Irgb6,* which is transcribed as a single unit, was also derived from a tandem gene. Interestingly, *Irgb6* clusters phylogenetically with the downstream halves of the known tandem genes (*Irgb1, -b8, -b3 and -b4).* Furthermore, the pseudogene *Irgb7* is situated only 11 kb upstream of *Irgb6* in the same orientation. *Irgb7Ψ* clusters phylogenetically with the upstream halves of the IRGB-tandems (*Irgb2, -b5/b5* and -b9)* and possesses the conserved splice donor site used in the other transcripts. This makes it likely that *Irgb7* and *Irgb6* were originally duplicated together from an ancestral tandem gene. While *Irgb7* became a pseudogene due to insertion of premature stop codons, *Irgb6* apparently managed to acquire a functional promoter and became an independent gene.

Between *Muridae* and *Caviidae* the sequence divergence of all members besides *Irgc* and *Irgq* is either too pronounced to identify individual homologous pairs (GMS genes) or different genes of the family have duplicated and diversified further (Figure 14). *Irgd* and *Irgb10,* which are single genes in mouse and rat, are represented by three and four closely related genes in the guinea pig *(Cavia porcellus),* respectively (guinea pig genome Ensembl release cavPor2 v44.1a, April 2007). Notably, one of the guinea pig GMS genes (CP_GMS1) represents a hybrid sequence combining *IRGM*- and *IRGB*-like features separated by a frame shift causing a premature stop codon right before the G3 motif Figure 13, appendix V.13). It remains to be analysed whether these features are due to incorrect genome assembly or not. To date, no *IRGA* genes could be detected in the guinea pig genome.

In the thirteen-lined ground squirrel, *Spermophilus tridecemlineatus,* four GKS IRG genes were identified, representing *Irgc, Irgq* and *Irgb10,* plus one member of the GMS subfamily (Figure 14) (squirrel genome Ensembl release speTri1 v44.1, April 2007). In the rabbit *(Oryctolagus cuniculus)*, a representative of the *Lagomorpha,* two IRG sequences were identified (rabbit genome Ensembl release RABBIT v44.1b, Apr 2007). One represents an *IRGA* gene, the other clearly belongs to the GMS subfamily of IRG GTPases (Figure 14) though it has an isoleucine instead of a methionine in the G1 motif (Figure 13 and see below). Thus, all Glires analysed so far possess multiple and divergent full-length IRG genes.

III.1.5. IRG homologues in the carnivore *Canis familiaris*

For the domestic dog (*Canis familiaris),* as a representative of the *Carnivora*, eleven *IRG* genes plus the partial sequence of a 12th genes could be recovered from the public genome database (Table 7, appendix V.13; dog genome Ensembl release CanFam2.0 v44.2b, April 2007). Of these, three are GMS genes (*IRGM4, -M5* and *-M6*)

and appear to have diversified independently from the mouse GMS genes (Figure 18, Figure 20). One gene clearly represents dog *IRGD* and four are closely related to murine *Irgb10* (*IRGB11, –B12*, CF_*GKS1*, and CF_*GKS2ψ*). One of these IRGB10 like sequences contains a single frame shift (CF_*GKS1*) and another one is a pseudogene by a number of criteria (CF_*GKS2ψ).* Two genes are clearly dog *IRGC* and *IRGQ,* respectively, whereas the partial sequence is novel but most closely related to *IRGC* (CF_*IRGC-like)* (Figure 18-20). The dog possesses an additional 'quasi' IRG gene, *IRGDQ1,* which is clearly related to *Irgd* and contains several premature stop codons as well as a disrupted G1 motif (Qx_4RMY) (Figure 18, Figure 19, Table 8). Thus, both GMS and GKS genes are represented and representatives of both subgroups are inducible by interferon in dog MDCK epithelial cells (*IRGM4-6* and *IRGB12*; (Bekpen 2005a; Bekpen 2005b)). As in human and mouse, dog *IRGC* was not induced by IFNγ *in vitro* (Bekpen 2005a; Bekpen 2005b). The canine IRG genes are situated in separate clusters on chromosome 1 (*IRGC, IRGQ*, *IRGC-like*), chromosome 10 (*IRGM4-6, IRGB11-12*, CF_*GKS1, IRGD, IRGDQ1*) and chromosome 29 (CF_*GKS2ψ*) (Figure 17). The IRG genes containing region of dog chromosome 10 is syntenic to the respective regions of mouse chromosome 11 and human chromosome 5. The region of canine chromosome 1 containing the *IRGQ* and *-C* genes is syntenic to the respective regions of murine chromosome 7 and human chromosome 19. For the isolated pseudogene on chromosome 29, CF_*GKS2ψ*, no syntenic linkage to IRG gene containing regions of the human or mouse genomes has been reported. Overall, the IRG gene status of the dog clearly resembles that of mouse rather than that of humans.

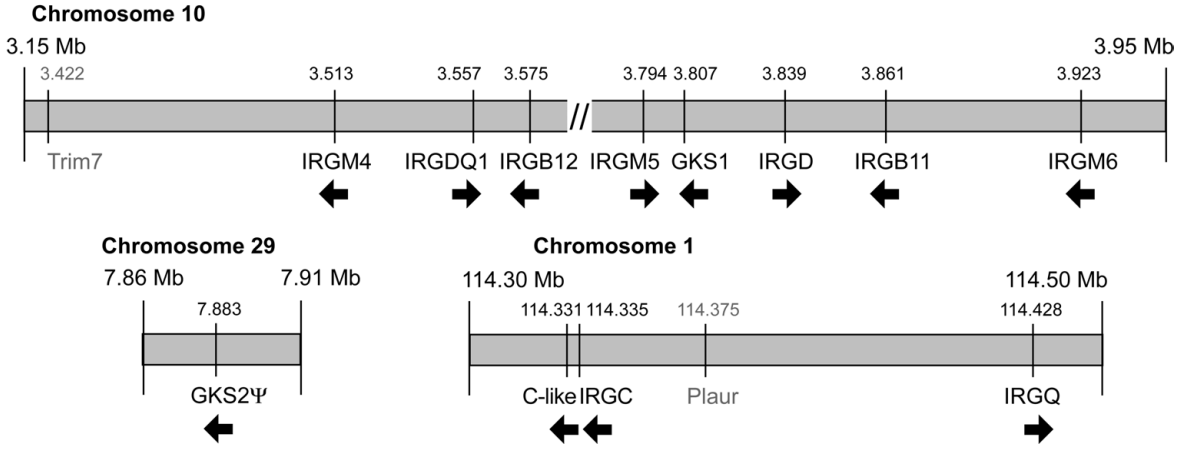


Figure 17 Genomic positioning of dog *(Canis familiaris)* IRG GTPases. Positioning and orientation of IRG genes in the dog chromosome 10, 1 and 29 clusters. Positions of genes refer to the location in dog genome Ensembl release v44.2b (CanFam 2.0 v44.2b April 2007). Gene orientation is given by black arrows. The closest syntenic markers are indicated in grey.

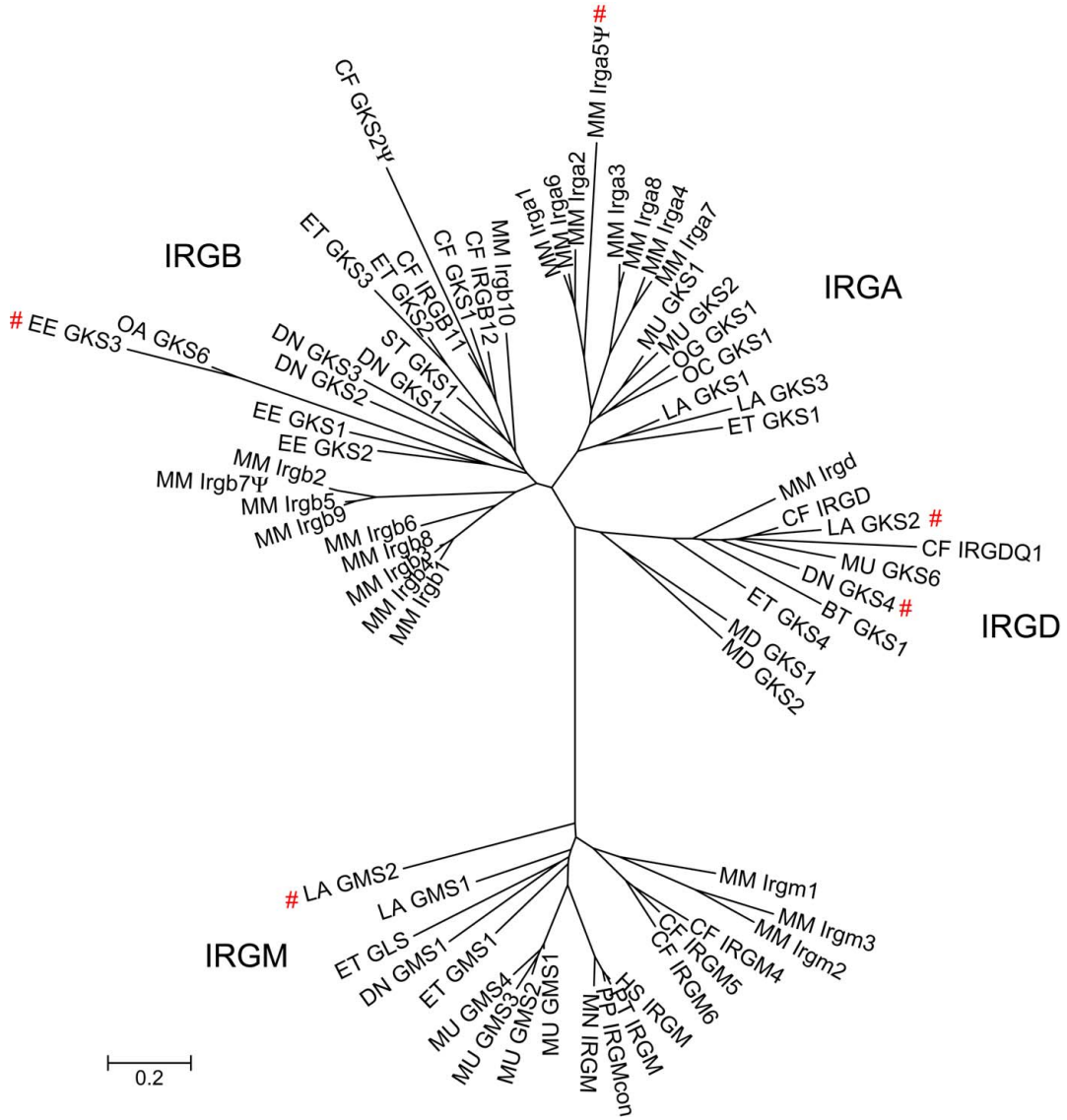
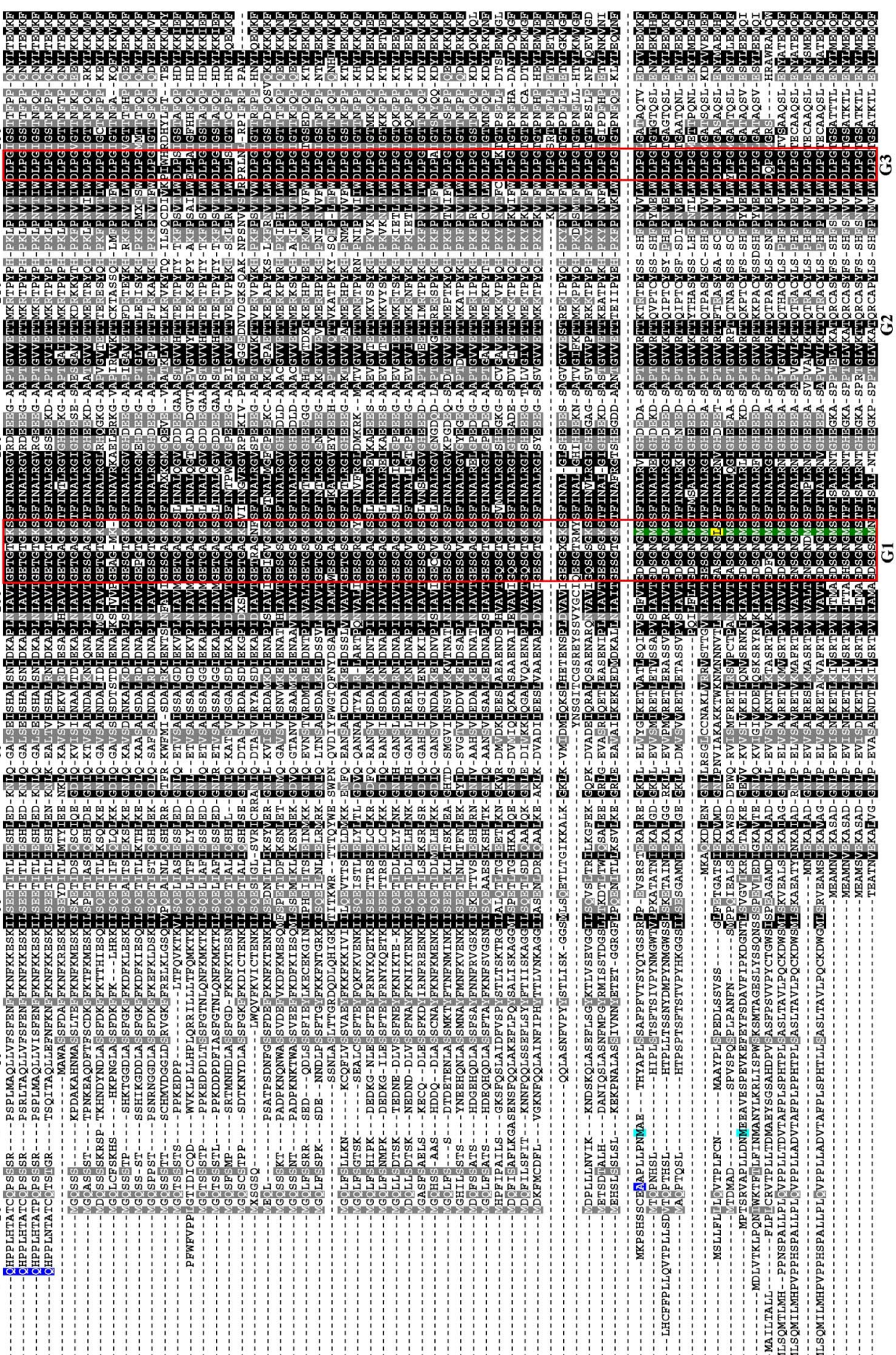
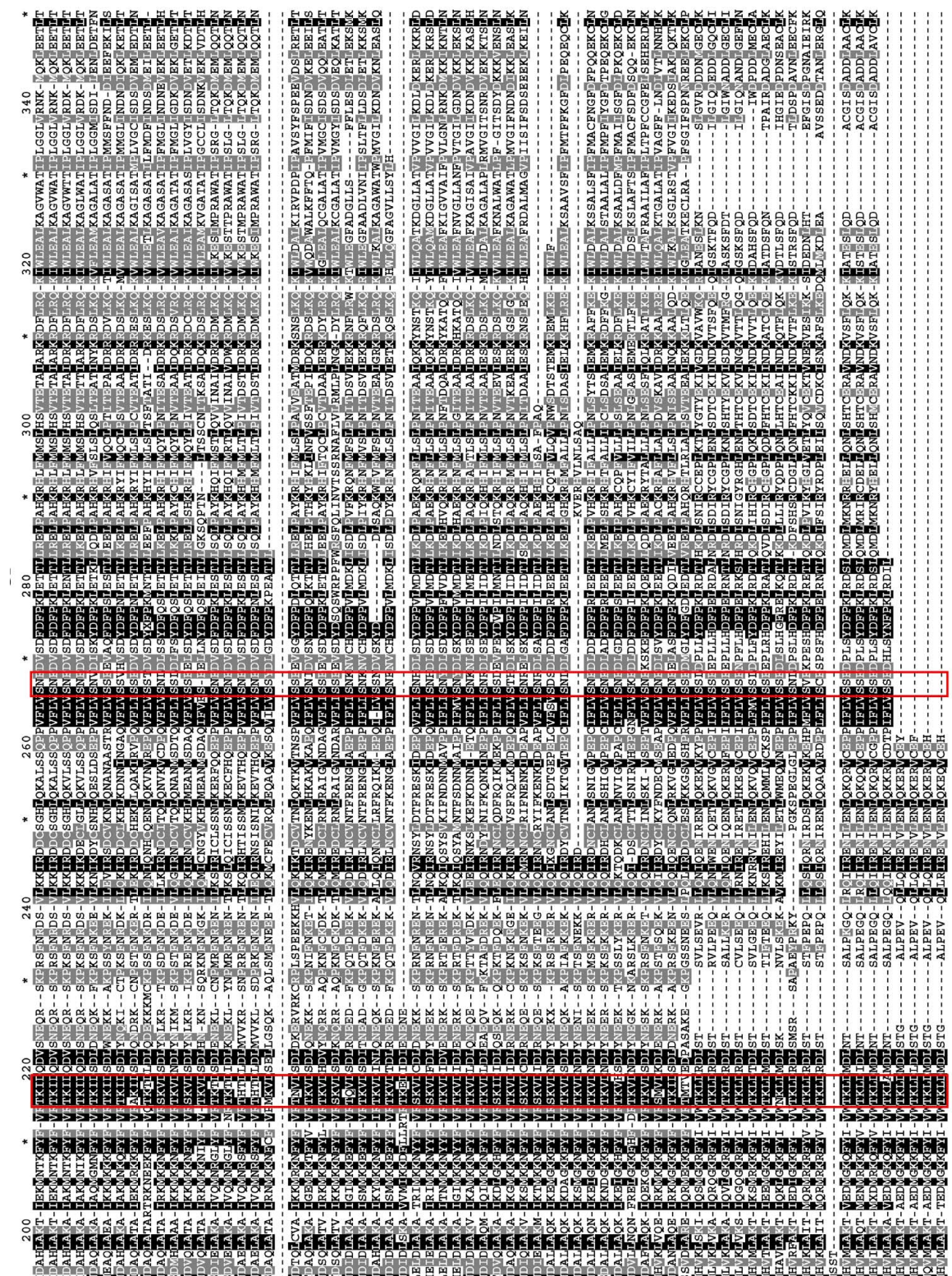


Figure 18 Phylogenetic relationship of mammalian IRG GTPases. Unrooted maximum likelihood consensus bootstrapped tree based on the alignment of the G-domains of mammalian IRG proteins (see Figure 19) using the JTT+I+G amino acid replacement matrix (Jones 1992). IRGC proteins and the IRGs from non-murine Glires were analysed separately (Figure 14, Figure 20). The nucleotide and amino acid sequences of all IRG sequences are given in the appendix (chapter V.13). The red # indicates sequences that are shorter than the classical IRG genes. Note that DN_GKS3 and EE_GKS2 are 'quasi' IRG proteins (DN_IRGBQ2 and EE_IRGBQ1 respectively). BT: *Bos taurus*, CF: *Canis familiaris*, DN: *Dasypus novemcinctus*, EC: *Equus caballus*, EE: *Erinaceus europaeus*, ET: *Echinops telfairi*, FC: *Felis catus*, HS: *Homo sapiens,* LA: *Loxodonta africana,* MD: *Monodelphis domestica,* ML: *Myotis lucifugus,* MM: *Mus musculus,* MN: *Macaca mulatta,* MU: *Microcebus murinus*, OA: *Ornithorhynchus anatinus*, OG: *Otolemur garnettii*, PP: *Pongo pygmaeus*, PT: *Pan troglodytes,* SA: *Sorex araneus*, SS: *Sus scrofa.*

Figure 19 Alignment of mammalian IRG proteins (see following pages). Amino acid alignment of all mammalian IRG GTPases identified in this study besides the IRGs from the non-murine Glires, the IRGC and the IRGQ proteins (analysed separately, Figure 13, Figure 20). The unusual methionine residue in the G1 motif of GMS proteins is highlighted in dark green, the conservative exchange to leucine in yellow. The dark blue background Q residues indicate the point where tandem splicing occurs from Irgb5/b5*, Irgb2 and Itgb9 to Irgb4/b3, Irgb1 and Irgb8, respectively. Canonical GTPase motifs are indicated by red boxes. Black shading: 70% conserved; grey shading: 50% conserved


G1
G2
G3


G4
G5

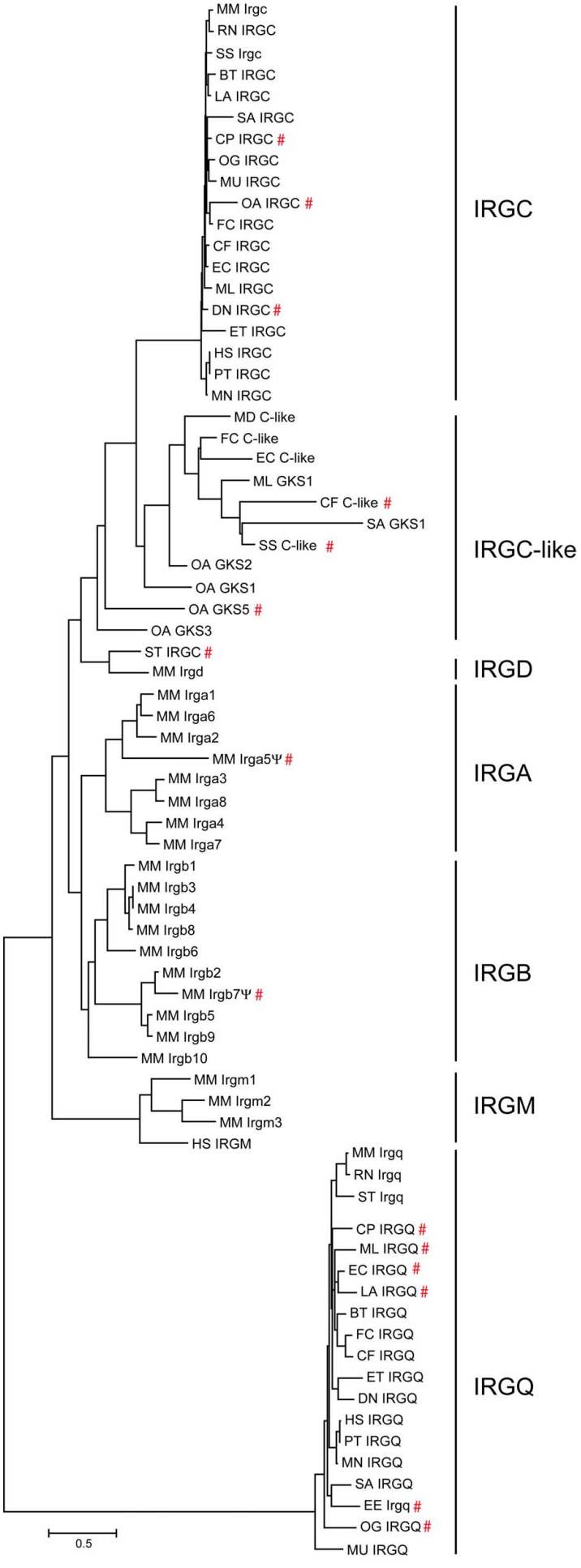

MM Irgc
RN IRGC
SS Irgc
BT IRGC
LA IRGC
SA IRGC
CP IRGC #
OG IRGC
MU IRGC
OA IRGC #
FC IRGC
CF IRGC
EC IRGC
ML IRGC
DN IRGC #
ET IRGC
HS IRGC
PT IRGC
MN IRGC
IRGC
MD C-like
FC C-like
EC C-like
ML GKS1
CF C-like #
SA GKS1
SS C-like #
OA GKS2
OA GKS1
OA GKS5 #
OA GKS3
IRGC-like
ST IRGC #
MM Irgd
IRGD
MM Irga1
MM Irga6
MM Irga2
MM Irga5Ψ #
MM Irga3
MM Irga8
MM Irga4
MM Irga7
IRGA
MM Irgb1
MM Irgb3
MM Irgb4
MM Irgb8
MM Irgb6
MM Irgb2
MM Irgb7Ψ #
MM Irgb5
MM Irgb9
MM Irgb10
IRGB
MM Irgm1
MM Irgm2
MM Irgm3
HS IRGM
IRGM
MM Irgq
RN Irgq
ST Irgq
CP IRGQ #
ML IRGQ #
EC IRGQ #
LA IRGQ #
BT IRGQ
FC IRGQ
CF IRGQ
ET IRGQ
DN IRGQ
HS IRGQ
PT IRGQ
MN IRGQ
SA IRGQ
EE Irgq #
OG IRGQ #
MU IRGQ
IRGQ
0.5

Figure 20 Phylogenetic relationship of mammalian IRGC and IRGQ GTPases. Maximum likelihood consensus bootstrapped tree based on the alignment of murine IRG proteins, human IRGM, mammalian IRGC and IRGQ proteins rooted on the IRGQ proteins. Only the G-domains were used for the phylogenetic analysis applying the JTT+I+G amino acid replacement matrix (Jones 1992). The red # indicates sequences that are shorter than the classical IRG genes. Note that the available fragment of *Spermophilus* IRGC is so short that it forms a branch with murine Irgd in the phylogenetic analysis though it is clearly an IRGC protein (see alignment appendix V.8). Note that the macaque IRGC sequence was reconstructed from a fragment derived from the *Macaca mulatta* genome and several ESTs from *Macaca nemestrina* (DY750065, EB520540). BT: *Bos Taurus*, CF: *Canis familiaris*, CP: *Cavia porcellus*, DN: *Dasypus novemcinctus*, EC: *Equus caballus*, EE: *Erinaceus europaeus*, ET: *Echinops telfairi*, FC: *Felis catus*, HS: *Homo sapiens*, LA: *Loxodonta africana,* MD: *Monodelphis domestica,* ML: *Myotis lucifugus,* MM: *Mus musculus,* MN: *Macaca mulatta*, MU: *Microcebus murinus*, OA: *Ornithorhynchus anatinus*, OC: *Oryctolagus cuniculus,* OG: *Otolemur garnettii*, PP: *Pongo pygmaeus*, PT: *Pan troglodytes,* RN: *Rattus norvegicus*, SA: *Sorex araneus*, SS: *Sus scrofa,* ST: *Spermophilus tridecemlineatus.*

III.1.6. IRG homologues in other mammalian species

For several other mammalian species including the other carnivore analysed in this study, the domestic cat (*Felis catus),* to date only *IRGC*- and *IRGQ*-related sequences could be recovered (cat genome Ensembl release CAT, v44.1, April 2007). The common shrew (*Sorex araneus),* the little brown bat *(Myotis lucifugus)* and the cat (*Felis catus)* possess one gene each that is closely homologous to murine *Irgq* and *Irgc* and a third gene that is more distantly related to *Irgc (IRGC-like)* (shrew genome Ensembl release sorAra1, version 45, Jun 2007; bat genome Ensembl release myoLuc1, v44.1, April 2007). The pig *(Sus scrofa)* also possesses two *IRGC*-related sequences (*IRGC* and *IRGC-like)* but an *Irgq* orthologue could not be retrieved yet (Sscrofa1, http://pre.ensembl.org). Only a single *IRGC* and *IRGQ* gene each is present in the horse *(Equus caballus)* genome (EquCab2, http://pre.ensembl.org).

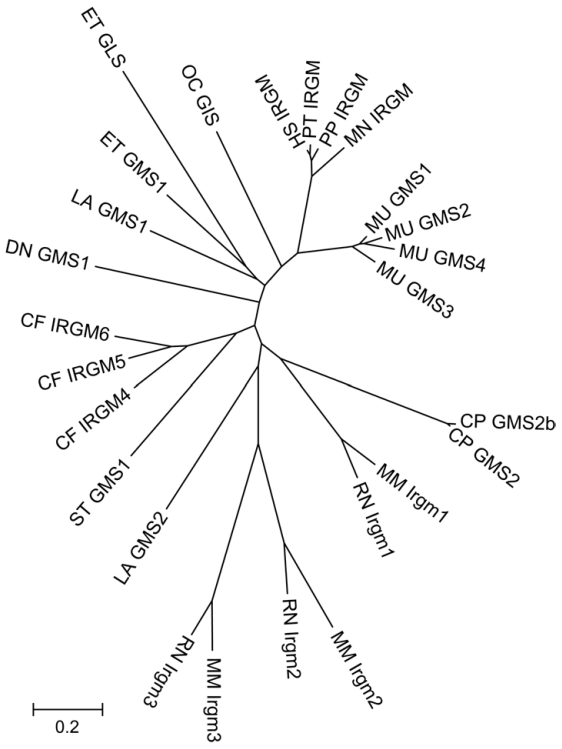


Figure 21A Phylogenetic tree of the mammalian IRGM proteins. Unrooted maximum likelihood consensus bootstrapped tree based on the alignment of the complete IRGM proteins (see B) using the JTT+G amino acid replacement matrix (Jones 1992)

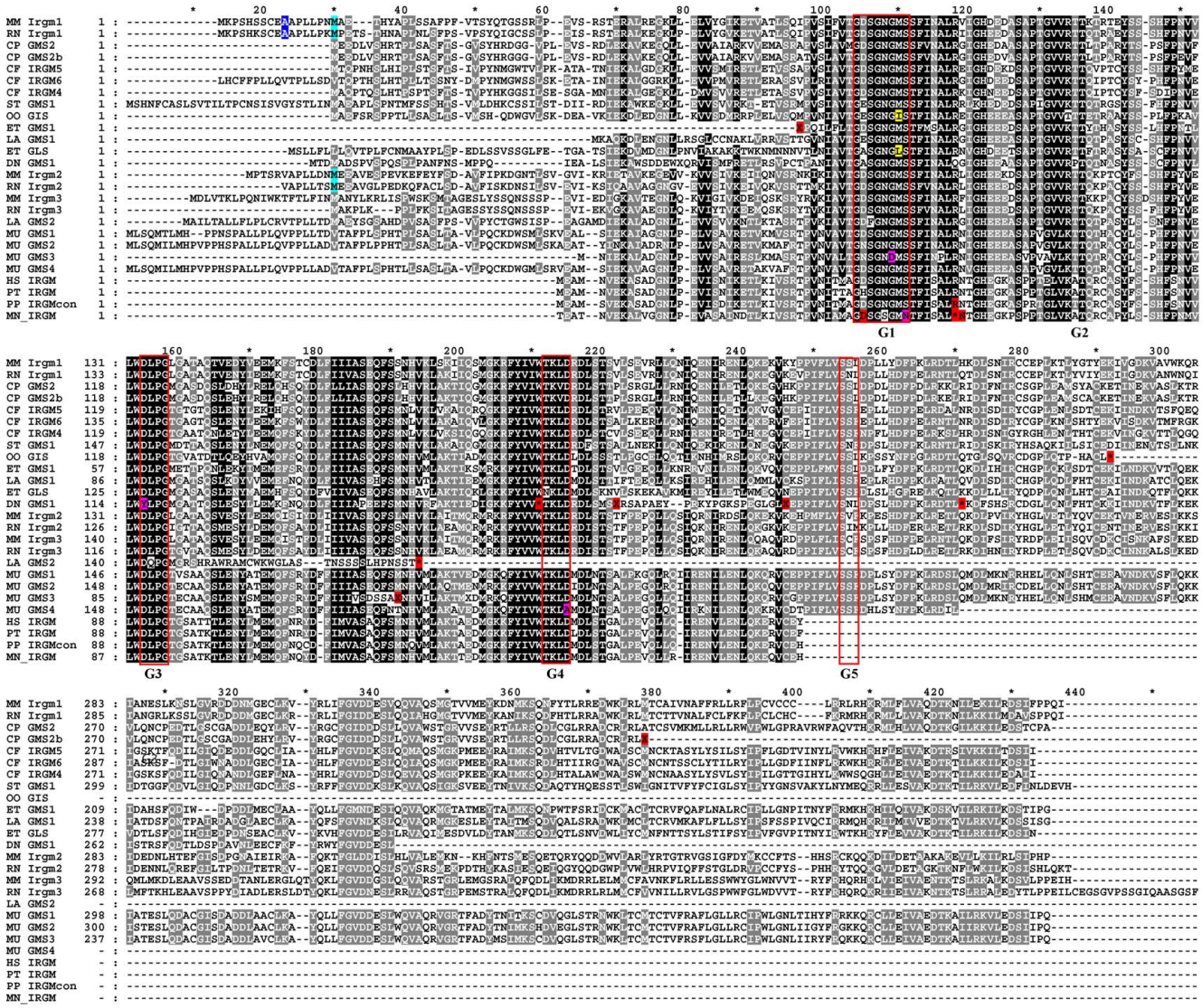


Figure 21B Alignment of the mammalian IRGM proteins. The turquoise-highlighted M in Irgm1 and Irgm2 are initiation codons that are dependent on alternative splicing (highlighted in dark blue). Conservative substitutions of the IRGM typical methionine in the G1 motif to leucine or isoleucine are highlighted in yellow. Substitutions disrupting the conserved GTPase motifs (indicated with red boxes) are highlighted in pink. Red highlighting indicates premature stop codons (*) or frame shifts (X). Red highlighted R in the PP_IRGM consensus sequence marks the position of the premature stop codon present in the majority of the orang-utan sequences (see appendix V.10). Note that a consensus sequence was used for orang-utan IRGM (see main text). Black highlighting: 90% conserved, grey highlighting: 50% conserved. CF: *Canis familiaris*, CP: *Cavia porcellus*, DN: *Dasypus novemcinctus*, ET: *Echinops telfairi*, HS: *Homo sapiens*, LA: *Loxodonta africana,* MM: *Mus musculus,* MN: *Macaca mulatta*, MU: *Microcebus murinus*, OC: *Oryctolagus cuniculus,* PP: *Pongo pygmaeus*, PT: *Pan troglodytes,* RN: *Rattus norvegicus*, ST: *Spermophilus tridecemlineatus.*

However, other species in the respective orders, *Insectivora, Carnivora* and *Artiodactyla* (even-toed ungulates), did posses immune type IRG genes (e.g. tenrec, dog and bull; Table 7). It is therefore unlikely that the absence of homologues of the IFN-inducible mouse IRGs will persist in all cases upon availability of a larger number of sequences in the databases. The case of the carnivores has already been discussed above. In the ungulate *Bos taurus* (cattle) four IRGs are present, namely *IRGQ, IRGD, IRGC* and *IRGC-like* as opposed to the situation in the pig (see above). Two of the three insectivore species analysed contained homologues of the IFN-inducible mouse *IRG* genes. The western European hedgehog (*Erinaceus europaeus*) possesses an *IRGQ* gene

and two members of the IRGB group, one of which clearly belongs to the 'quasi' IRGs (*IRGBQ1*: GNF in G1 motif; Table 8, Figure 18, Figure 19). In the lesser hedgehog tenrec (*Echinops telfairi*) eight IRG sequences were identified, *IRGC, IRGQ,* one *IRGA* gene, one *IRGD* gene, two *IRGB* genes and two GMS genes (Table 7, Figure 9, Figure 10). Interestingly, both the tenrec and the rabbit (see above) possess one GMS related gene containing a conservative substitution for the unusual methionine in the G1 motif, Gx_4GLS and Gx_4GIS respectively, while preserving GMS characteristic features throughout the sequence, thus suggesting that a large hydrophobic residue in this position might not only be permissive but also essential for GMS function (Figure 21, appendix V.13).

Two more species were analysed to cover the deep branches of the Eutheria; the African elephant (*Loxodonta africana)* as a representative of the Proboscidea and the nine-banded armadillo (*Dasypus novemcinctus*) as a representative of the Cingulata. They both possess a variety of IRG genes, *IRGC*, *IRGQ,* two GMS, two *IRGA* and two *IRGD* genes in case of the elephant and three IRGB genes and a single GMS, *IRGD*, *IRGC* and *IRGQ* gene each in case of the armadillo (Table 7, Figure 18, Figure 19). Thus, divergent IRG GTPases are widespread in the Eutheria and at least as old as the rapid eutherian radiation 85-100 million years ago (Bininda-Emonds 2007). On this ground it was analysed whether the reduced IRG status in humans is representative for other primates and where in the evolution of the primate lineage the divergent, immune-type IRG genes have been lost.

III.1.7. IRG homologues in non-human primates

Not only the human but apparently also other higher primates have lost full length, IFN-inducible *IRG* genes. The situation in the chimpanzee is identical to that in humans. The chimp *IRGM* gene is 98% identical to human *IRGM* and is situated on a stretch on chromosome 5 (152,8 Mb) collinear with the syntenic part of human chromosome 5 flanked by the same marker genes (*DCTN4* and *TNIP1*) (chimp genome Ensembl release PanTro 2.1 v44.21c, April 2007) (Figure 21 B). Also the stretch on chromosome 19 harbouring chimp *IRGC* (49.3 Mb) and *IRGQ* is syntenic to the respective region on human chromosome 19. The genome of the orang-utan (*Pongo pygmaeus)* contains 9 sequences related to human *IRGM*, most of which are disrupted by premature stop codons and/or frame shifts (appendix V.10) (BLAST of the incomplete *Pongo* genome via http://www.ncbi.nlm.nih.gov/sutils/genom_tree.cgi). Alignment of the protein sequences revealed that they are all near identical (appendix V.10). Taking into account that the orang-utan genome database is still in a premature state, the sequences might all represent two *IRGM* genes, one of which possesses a stop codon immediately after the G1 motif. For the phylogenetic analysis, a consensus sequence generated from the 9 sequences was used (for sequence see appendix V.13). Further analysis awaits database versions with a better sequence coverage of the genome.

The Rhesus monkey *(Macaca mulatta)* as a representative of the old world monkeys possesses three IRG genes, *IRGC, IRGQ* and a single *IRGM* gene 93% identical to human *IRGM.* The Rhesus monkey *IRGM* gene is situated on chromosome 6 (Mb 147.3) between the same syntenic markers as the human orthologue with comparable

spacing (*DCTN4* and *TNIP1*) (macaque genome Ensembl release MMUL 1.0 v43.10c, February 2007) (Figure 18). Interestingly, this gene is clearly damaged in several ways, as it contains a frame shift before and a stop codon after the first GTP-binding motif (Figure 21 B, appendix V.13). Furthermore, the amino acid sequence of the G1 motif is Gx_4GMN, a mutation that renders other IRG proteins non-functional ((Taylor 1997) and see below, chapter III.2). Despite the absence of confirming transcripts from the databases and the preliminary status of the macaque genome project, a tripartite disruption is unlikely to be due to sequencing errors alone. Furthermore, an independent genomic clone (BV210965) confirms the damaged IRGM sequence.

Recent genomic analysis of the common marmoset *(Callithrix jacchus)* as a representative of the new world monkeys yielded a similar picture. Despite the unfinished status of the genome project, three IRG genes were recovered representing *IRGC, IRGQ* and *IRGM* (Table 7). As in the Rhesus monkey, the first GTP-binding motif of the marmoset IRGM protein is damaged by mutation (Gx_4SIS in this case). The mutation of the universally conserved glycine residue in the G1 motif was reported to interfere with nucleotide binding in H-Ras (Chen 1994; Powers 1989). Thus, all higher primates analysed exclusively possess IRGM GTPases that are significantly shorter than a typical G-domain and terminate before the G5 motif (Figure 14). In case of the orang-utan, the rhesus monkey and the marmoset the *IRGM* genes are even further damaged by truncation or mutation.

protein	clade	motif	protein	clade	motif
MM_Irga1	A	MGQLF**S**L	ST_GKS1	B	MGASS**S**T
MM_Irga2	A	MGQLF**S**S	CP_GKS3	B	MGQAS**S**S
MM_Irga3	A	MGQLF**S**H	CP_GKS6	B	MGQAS**S**S
MM_Irga4	A	MGQLL**S**D	OC_GKS1	A	MGASF**S**A
MM_Irga5p	A	MGQLF**S**G	EE_GKS1	B	MGQSF**S**M
MM_Irga6	A	MGQLF**S**S	EE_GKS3	B	MGQSC**S**T
MM_Irga8	A	MGQLF**S**N	ET_GKS1	A	MGQLF**S**S
MM_Irgb2	B	MGQTS**S**S	ET_GKS2	B	MGQSS**S**K
MM_Irgb5	B	MGQTS**S**S	ET_GKS3	B	MGQSS**S**N
MM_Irgb9	B	MGQTS**S**S	OG_GKS1	A	MGQHS**S**A
MM_Irgb10	B	MGQSS**S**K	MU_GKS2	A	MGQLF**S**A
RN_Irga5	A	MGQLF**S**G	CF_IRGB12	B	MGQSS**S**T
RN_Irga11	A	MGQLF**S**L	CF_GKS1	B	MGQSS**S**T
RN_Irga12	A	MGQWF**S**S	CF_GKS2p	B	MGQSS**S**T
RN_Irga13	A	MGQWF**S**S	DN_GKS3	B	MGQLCFS
RN_Irga14p	A	MGQLF**S**D	AC_GKS4	E	MGIALTK
RN_irga15	A	MGQLF**S**D	AC_GKS7	E	MGGTN**S**H

Table 9 Myristoylation motifs of IRG GTPases. IRG proteins with intact or disrupted potential myristoylation motifs. The consensus sequence is MG{EDRHPFYW}X_2[**S**TAGCNDEF]{P} with curly brackets indicating non-permissive amino acids and square brackets indicating essential amino acids (X: any amino acid) (Maurer-Stroh 2002). Serine (bold) is favoured at position 6. AC: *Anolis carolinensis*, CF: *Canis familiaris*, CP: *Cavia porcellus*, DN: *Dasypus novemcinctus,* EE: *Erinaceus europaeus*, ET: *Echinops telfairi*, LA: *Loxodonta africana,* MM: *Mus musculus,* MU: *Microcebus murinus,* OC: *Oryctolagus cuniculus*, OG: *Otolemur garnettii,* RN: *Rattus norvegicus*, ST: *Spermophilus tridecemlineatus*

In contrast, in the lower primate *Otolemur garnettii* (bushbaby) an *IRGA* gene (OG_*GKS1*) coding for a full-length protein with a functional myristoylation sequence at the N-terminus was found in addition to an *IRGC* and *IRGQ* gene (Table 9) despite the low coverage of the genome assembly. The IRGA protein has an unusual substitution (DxxA instead of DxxG) in the G3 motif but possesses consensus sequences in the other GTPase motifs (Table 8). The analogous mutation prevents bound GTP from changing the protein conformation in EF-Tu, $G_{S\alpha}$ and H-Ras while affecting the nucleotide affinities only moderately (Bourne 1991; Ford 2005; Lee 1992). Thus, it remains to be determined whether *Otolemur* IRGA can function as a GTPase. Further analysis should also aim to identify the putative promoter of this gene and clarify whether it contains GAS and/or ISRE sites and therefore fulfils the prerequisites for functioning in IFN-mediated cell autonomous resistance.

Only shortly before completion of this study, genomic data from another lower primate (Strepsirrhini) has become available. *Microcebus murinus* (grey mouse lemur) possesses at least 9 IRG genes, including a single *IRGC, IRGQ* and *IRGD* gene each, two *IRGA*, and four full-length GMS genes. Furthermore, four short fragments of IRG genes are present, two of which clearly belong to the GMS subfamily. At least one of the *IRGA* genes encodes a protein with a functional myristoylation sequence (Table 9). Thus, resistance IRG GTPases must have been lost sometime between the divide of the Haplorrhini (tarsier and anthropoids) and the Strepsirrhini (loris and lemurs) perhaps 60 million years ago and the division of the Hominoidae and the Old World monkeys (Catarrhini) from the New World monkeys (Platyrrhini) (30-40 million years ago; for phylogenetic tree of the primates see Figure 56).

To understand why the higher primates could afford to lose a resistance system highly active and absolutely essential in rodents might shed light onto important differences in innate immune mechanisms of mice and men.

III.1.8. IRG homologues outside the Eutheria

To gain further insight into the dynamics of the evolution of the IRG family and the phylogenetic age of different subgroups, all available chordate genomes were scanned for IRG genes by homology searches. The results are summarised in Table 7. *Irgc* orthologues can be found throughout the mammals right down to the Marsupialia (*Monodelphis domestica,* opossum) and Monotremata (*Ornithorhynchus anatinus,* platypus) and are frequently accompanied by a second, closely related gene (*IRGC-like*) spatially linked to *IRGC* (see Table 7) (opossum genome Ensembl release MonDom 4.0 v43.3c, Feb 2007; platypus genome Ensembl release Oana-5.0 v43.1a, Feb 2007). In the opossum one fragment representing *IRGC,* one *IRGC-like* gene and two *IRGD*-related sequences were detected. The two *Irgc*-related sequence are situated in a region on chromosome 4 syntenic to human chromosome 19 in the vicinity of the genomic marker *Plaur* that is found adjacent to *Irgc* in other mammals. The N-terminus of the opossum *IRGC* sequence is missing due to a gap in the current genome assembly. Interestingly, the opossum *IRGC-like* gene contains a single intron between the G4 and the G5 motif roughly at the same position as the fish *IRGF* genes (see below; highlighted in Figure 23). In the platypus *IRGC,* three genes distantly related to *IRGC-like* and one *IRGB* gene

could be identified (Figure 20). Notably the *IRGC-like* genes from platypus and other mammals are very divergent compared to the classical IRGCs. It remains to be determined whether *IRGC-like* genes represent another clade of resistance IRGs rather than being functionally related to *Irgc*. Thus, divergent IRG genes are also present in marsupials and monotremes but to date no GMS genes could be detected in either group.

After confirming the presence of IRG GTPases throughout the mammals, species from other classes were analysed concerning their IRG gene content. IRG genes were detected in most vertebrate genomes including reptiles, amphibians and fish, with chicken (*Gallus gallus)* and zebra finch (*Taeniopygia guttata*) being the clear exceptions (chicken genome Ensembl release WASHUC2 v46.2d, Aug 2007; zebra finch genome, http://www.ncbi.nlm.nih.gov/projects/genome/guide/finch). Unfortunately, it still cannot be excluded that the absence of IRG genes from the birds rather reflects a poor coverage of the genomes than a true deficiency.

12 IRG sequences were identified in the freshwater pufferfish *Tetraodon nigroviridis*, 9 IRG sequences in the saltwater pufferfish *Takifugu rubripes*, and 14 partially clustered IRG genes in the zebrafish *Danio rerio* (Table 7, zebrafish genome Ensembl release Zv7 v 46.7, Aug 2007; *Tetraodon* genome Ensembl release TETRAODON 7 v46.1i, Aug 2007 and *Takifugu* genome Ensembl release FUGU 4.0 v46.4g, Aug 2007). The fish IRG genes fall into separate clades from the mammalian genes, termed *IRGE, IRGF* and *IRGG* (Figure 22). A specific *IRGC* homologue is not immediately apparent. *IRGM* genes are absent from fish. All the pufferfish sequences are closely related and belong to the *IRGF* group. As the database is still fragmented and rich in gaps and misreads, it is difficult to come to a final conclusion concerning the number of IRG genes in these species but recurring sequence features support the presence of at least three genes in *Tetraodon* and two in the *Fugu* (appendix V.11-V.13). The pufferfish and the four zebrafish *IRGF* genes have one intron identically positioned at the end of helix 4 of the G-domain (Figure 23; appendix V.13). The six distinct *IRGE* subfamily genes and the single *IRGG* gene of the *Danio* are intronless in the open reading frame, like the majority of the mammalian IRG genes (Figure 22-23). Neither *IRGE* nor *IRGG* genes could be detected in the pufferfish genomes. The zebrafish genome additionally contains three IRG homologues with modified GTP-binding motifs (*irgxq1-irgxq3*; Figure 22-23, Table 8). Their homology to other IRG genes is stronger than that of *IRGQ* but, as with *IRGQ,* their function as GTPases is doubtful. The *irgxq1* gene is clustered on chromosome 16 (26.23-26.67 Mb) with four of the apparently normal *IRGE* genes (*irge1-4)* and lies immediately downstream of a truncated IRG sequence, *irgg*, with which *irgxq1* is transcribed in tandem (accession number: BQ481122). Thus, the hypothetical protein product would be a carboxyl-terminally truncated p47 GTPase, linked to a similarly truncated IRG homologue probably devoid of GTPase function. Though no function for tandem IRG genes has been demonstrated yet, it remains intriguing that they appear in such distant species as rat/mouse, zebrafish and *Xenopus* (see below). *Irgxq3* is also situated on chromosome 16, 20 megabases downstream of the other IRG genes. The two remaining *IRGE* genes, *irge5* and *irge6*, are both located on an unpositioned scaffold (Zv7_NA674) while all the *IRGF* genes cluster on chromosome 19. *Irgxq2* is an isolated gene on chromosome 15.

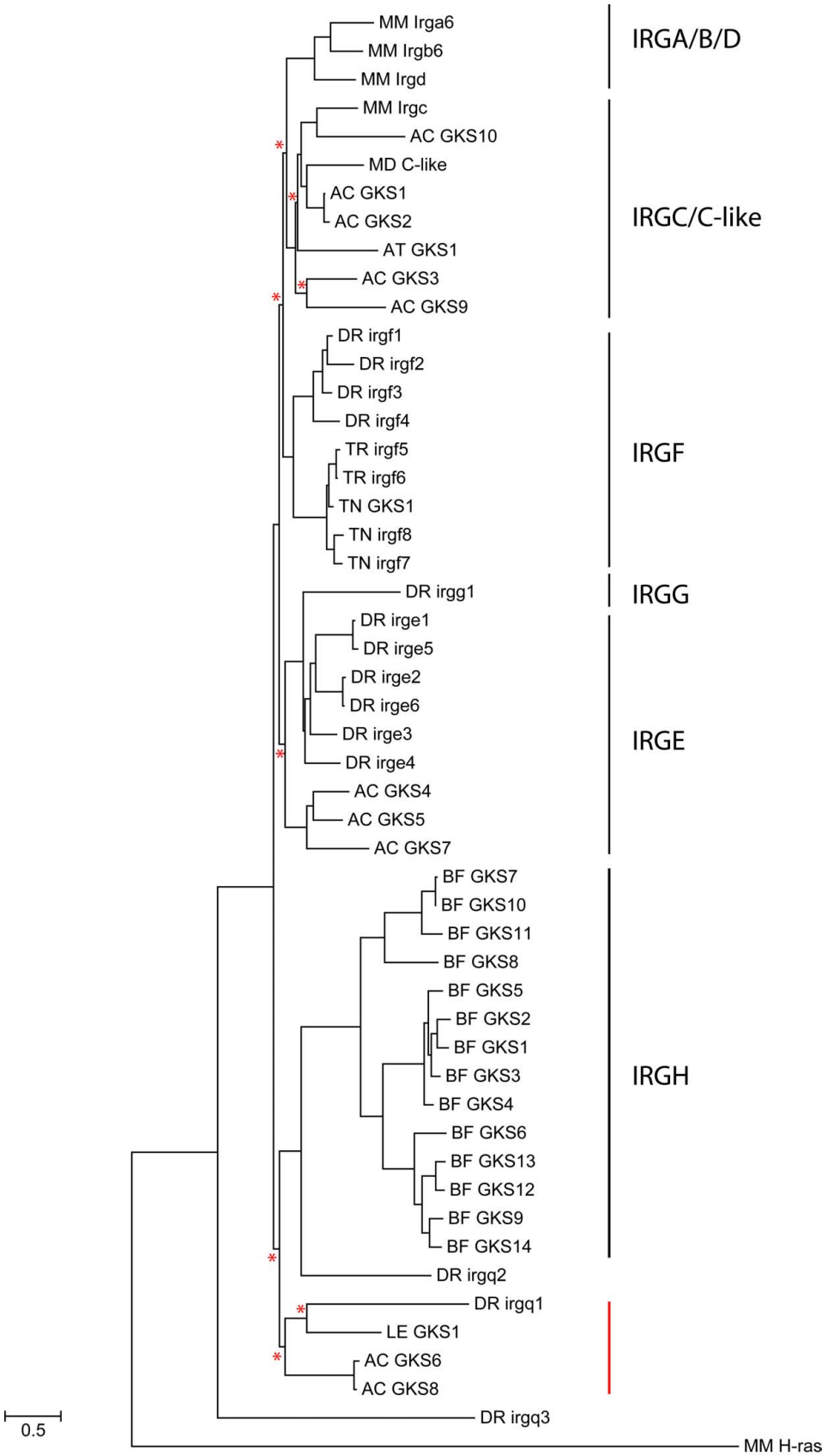


Figure 22 Phylogenetic relationship of fish, reptile, amphibian and Cephalochordate IRG GTPases. Maximum likelihood consensus bootstrapped tree using the WAG+I+G amino acid replacement matrix (Whelan 2001) based on the alignment of the G-domains of mouse, fish, amphibian, reptile and Cephalochordate IRG proteins rooted on H-Ras-1 (see Figure 23). Red asterisks indicate bootstrap values below 20. Red line indicates ambiguous part of the tree with low bootstrap values. AC: *Anolis carolinensis*, AM: *Acropora millepora*, AT: *Ambystoma tigrinum tigrinum,* BF: *Branchiostoma floridae*, DR: *Danio rerio*, MM: Mus musculus

Three *IRG* sequences of cartilaginous fish could be retrieved from the databases: two fragments from the shark *Squalus acanthias*, which are too short for phylogenetic analysis (EE884816, EE049107), and one full length sequence from the little skate (*Leucoraja erinacea),* which forms a new deep branch in the phylogenetic tree together with *Danio irgxq1* and two reptile sequences (see below, Figure 22). However, this part of the tree should be considered with care as the bootstrap values are low and the positions of the branches are instable. The skate sequence was generated as a consensus sequence from five near identical ESTs (DR714243, DR714385, CV221991, DT726565, DR783592).

There are few amphibian IRG sequences available in the databases. One gene distantly related to *Irgc* was found in the eastern tiger salamander (*Ambystoma tigrinum tigrinum*, http://www.ambystoma.org, Tig_NM_019612_Contig_1, EST CN054487) (Figure 22) and two *IRG* sequences transcribed in tandem in the South African clawed frog (*Xenopus tropicalis)* (transcript: DT429555). The 3' half of the tandem, *irgxq4,* is clearly a 'quasi' IRG gene, while the other half is only slightly modified in the G3 motif (Table 8). The two *Xenopus* genes are situated on the same genomic scaffold (scaffold:JGI4.1:scaffold_449) with the *irgxq4 (XT_GKS1)* gene only 4 kb downstream of *XT_GKS2* (Xenopus genome Ensemble release JGI 4.1 v46.41g, Aug 2007). Neither of these two genes falls into one of the known *IRG* clades. The Xenopus sequences were excluded from the phylogenetic analysis displayed in figure 13 as they are rather divergent and destabilised parts of the tree when included in the analysis.

There are at least 10 IRG genes in the lizard *Anolis carolinensis,* five of which are distantly related to *Irgc* and two form a new clade together with zebrafish *irgxq1* and the sequence from *Leucoraja* (Figure 22) (BLAST of *Anolis* genome via http://www.ncbi.nlm.nih.gov/sutils/genom_tree.cgi). As described above, this part of the phylogenetic tree (marked by a red line in Figure 22) should be considered provisional. The remaining three *Anolis* genes belong to the *IRGE* group. The bootstrap values of this particular branch of the phylogenetic tree (marked by red asterisks in Figure 22) are low because the Anolis sequences are rather divergent and also related to the *IRGF* genes. However, the absence of an intron in these Anolis genes supports their affiliation to the likewise intronless *IRGE* genes. In general, the phylogenetic analysis of IRG genes is complicated by the combination of high intra- and inter-species sequence divergence in this gene family, resulting in low bootstrap values and instable branches in less well defined parts of the phylogenetic trees, though other parts of the trees remain reliable. Thus, especially in cases when only few supportive sequences are available, the precise positions of these branches in the phylogenetic tree should be considered provisional until more IRG sequences become available.

In summary, *IRG* genes were identified in many vertebrates, including mammals, reptiles, amphibians and fish. *IRGC*-related sequences are found in all mammalian groups, in reptiles and amphibians but the highly conserved classical *IRGC* gene seems to be specific for mammals. Most genomes analysed encode 'quasi' IRG proteins that are homologous to the IRG GTPases but almost certainly devoid of GTPases function as they are radically modified in the GTP-binding site. *IRGQ* is represented throughout the Eutheria (Figure 20, appendix V.9), however, to date no homologue could be identified

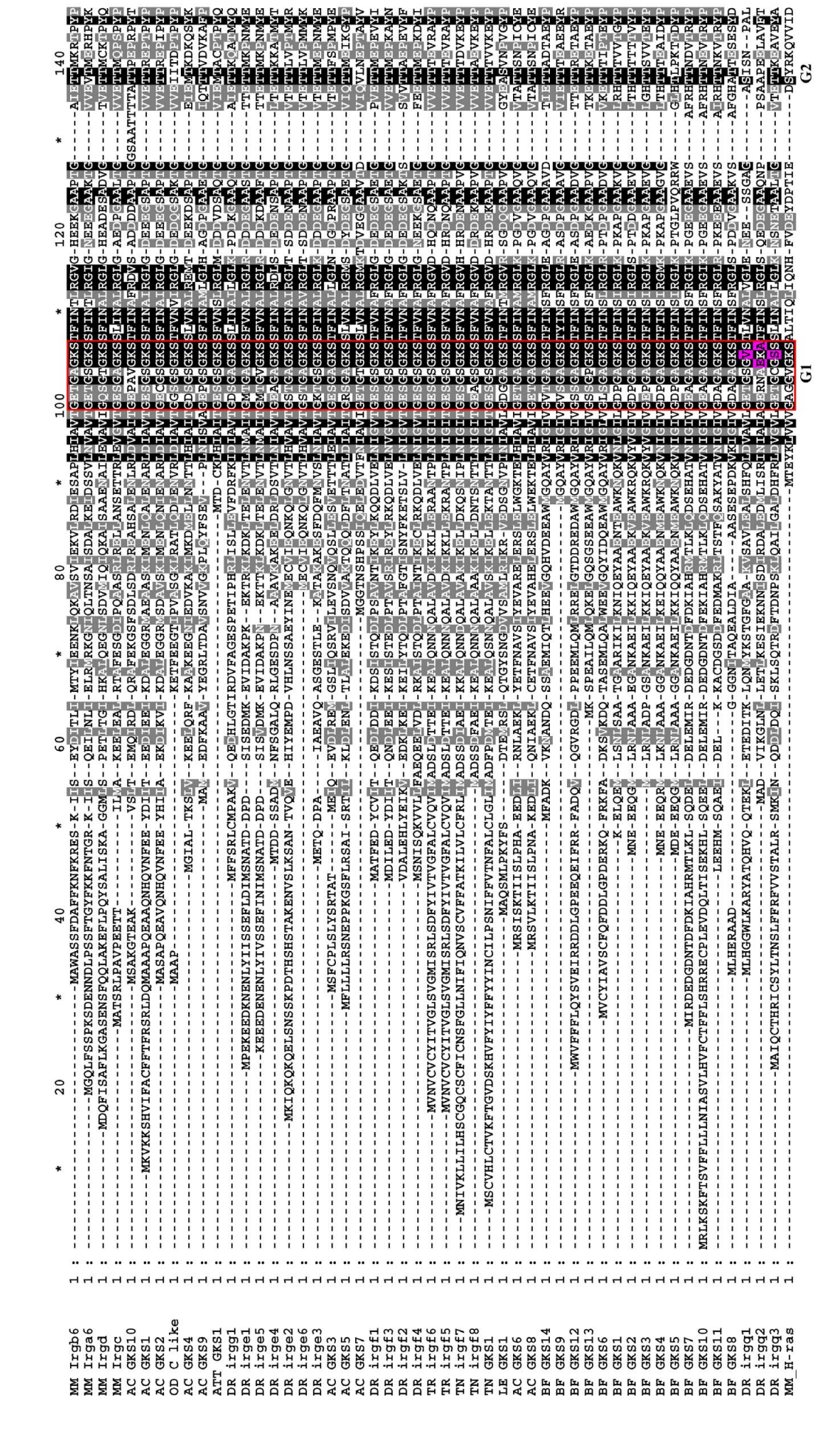
from Metatheria, Prototheria and non-mammalian vertebrates. *IRGA*-related genes are found throughout the Eutheria while *IRGB* and *IRGD* genes can be also found in the Proto- and Metatheria, respectively. *IRGF* and *IRGG* genes seem to be specific for fish while *IRGE* genes are also present in reptiles (Figure 22). A new deep branch on the phylogenetic tree is formed by two *Anolis* genes (AC_GKS8, AC_GKS6), zebrafish *irgxq1* and one sequence from the little skate (*Leucoraja erinacea*). Myristoylation motifs can be found throughout the vertebrates in a subset of the IRGA, -B and -E proteins suggesting that myristoylation is an ancient mechanism of membrane attachment in IRG proteins (Table 9).

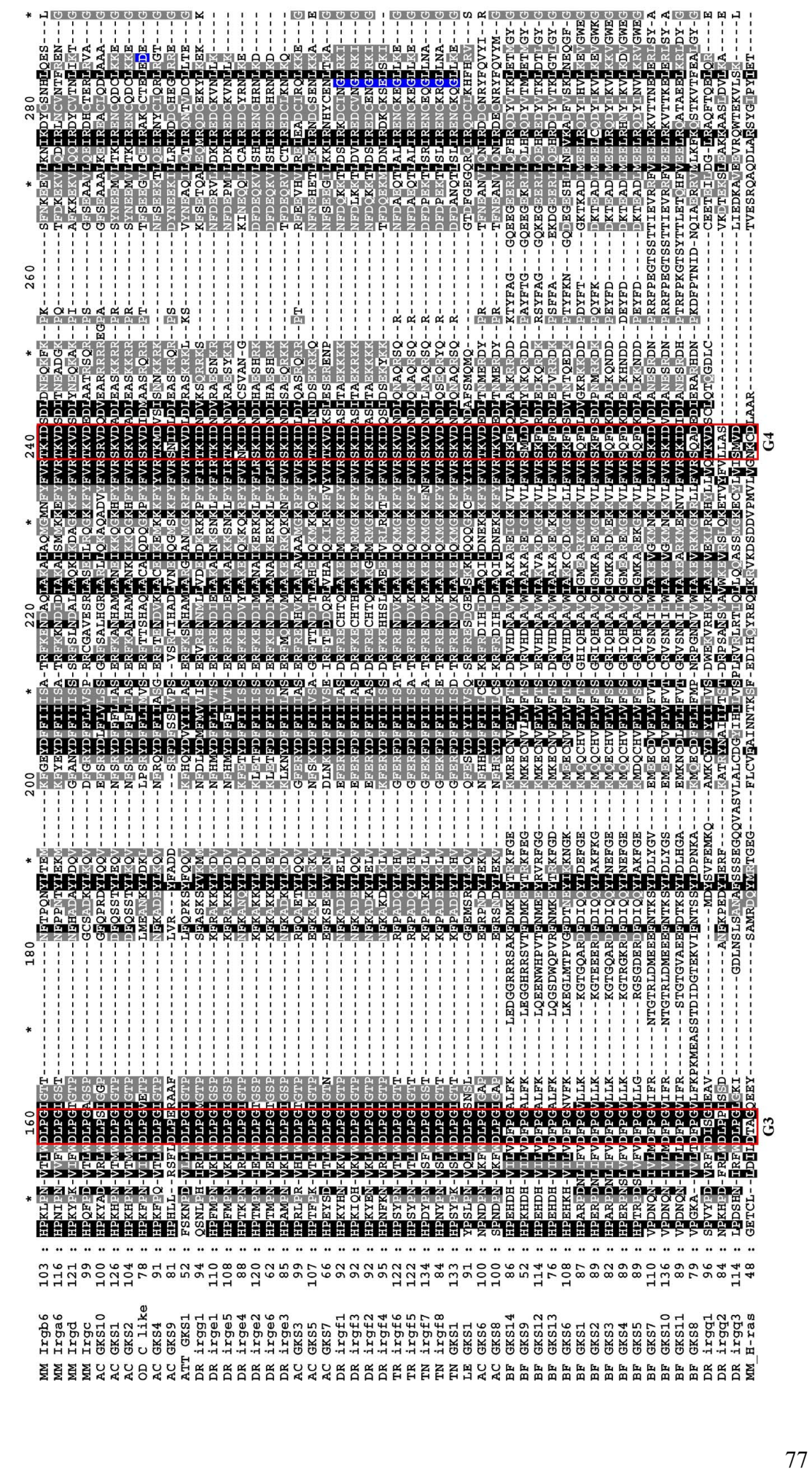
The only unambiguous IRG homologues outside the vertebrates have been found in the Cephalochordate *Branchiostoma floridae.* More than 14 genes were identified in the yet incomplete genome (http://genome.jgi-psf.org/Brafl1/Brafl1.home.html) forming an independent clade within the GKS subfamily most closely related to zebrafish *irgxq2* (IRGH, Figure 22). Two possibly related sequences were recovered from the *Caenorhabditis elegans* genome (C46E1.3, W09C5.2), and several groups of putative GTPases of unknown function with sequence features reminiscent of IRG GTPases exist in bacteria (e.g. BAA10832, BAA18140, BAA18642, BAC08557, BAC08842). These sequences are, however, too distantly related for a clear phylogenetic relationship to IRG proteins to be established from sequence similarity alone. To date, no IRG family members could be identified in the sea urchin and tunicate genomes. A cDNA derived from the coral *Acropora millepora* (DY583583) is closely related to zebrafish *irgg1* and therefore clearly represents a cross contamination with a fish sequence.

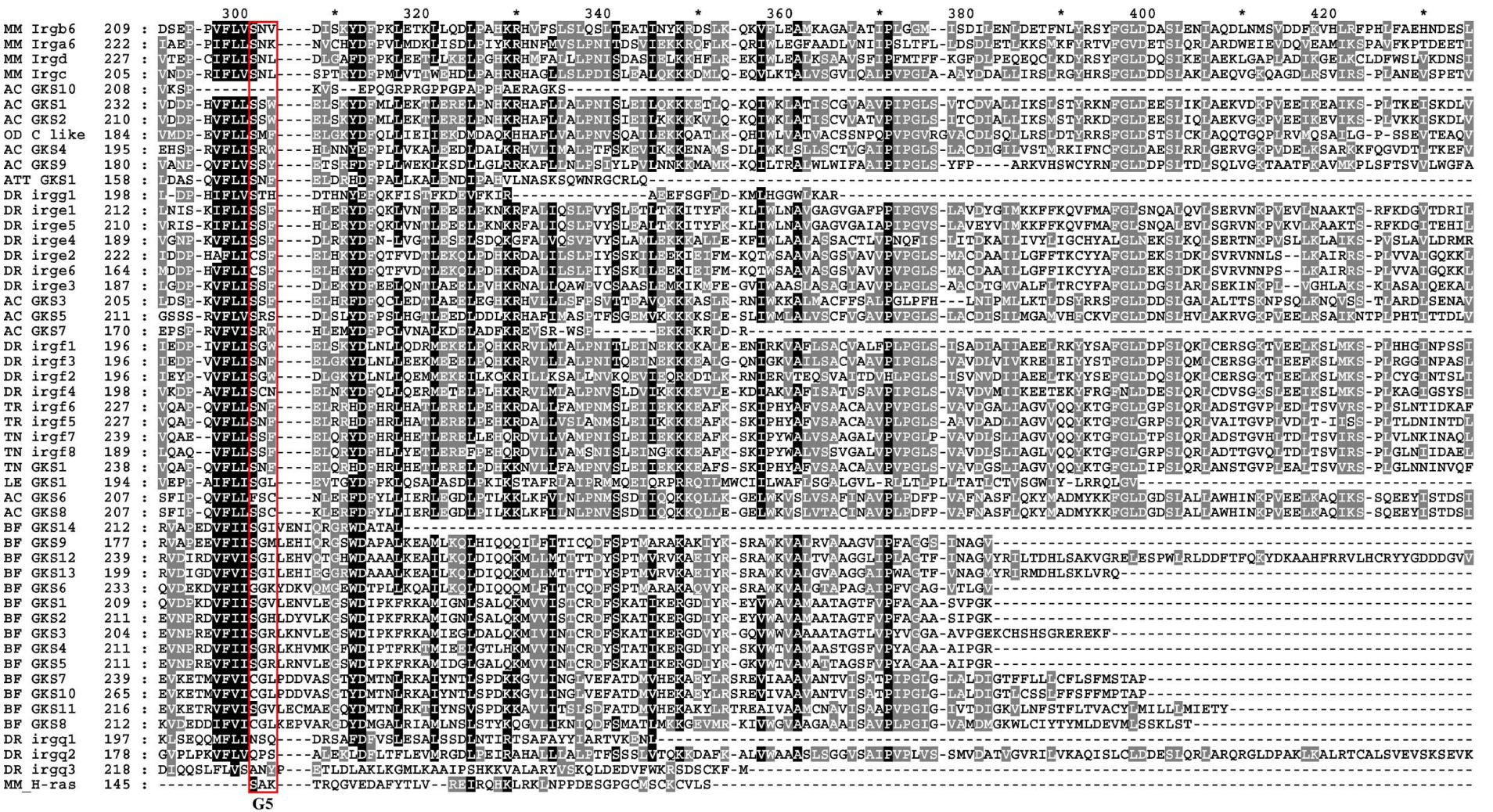
Thus, while GKS IRG proteins are at least as old as the euchordates, IRGM GTPases seem to be specific for Eutheria as they can be found in rodents, lagomorphs, insectivores, primates and carnivores but seem to be absent from marsupials, monotremes, reptiles, amphibians, fish and Cephalochordates. Hence, IRGM genes must have been present in an early eutherian progenitor and therefore must be at least 75 million years old (Bininda-Emonds 2007).

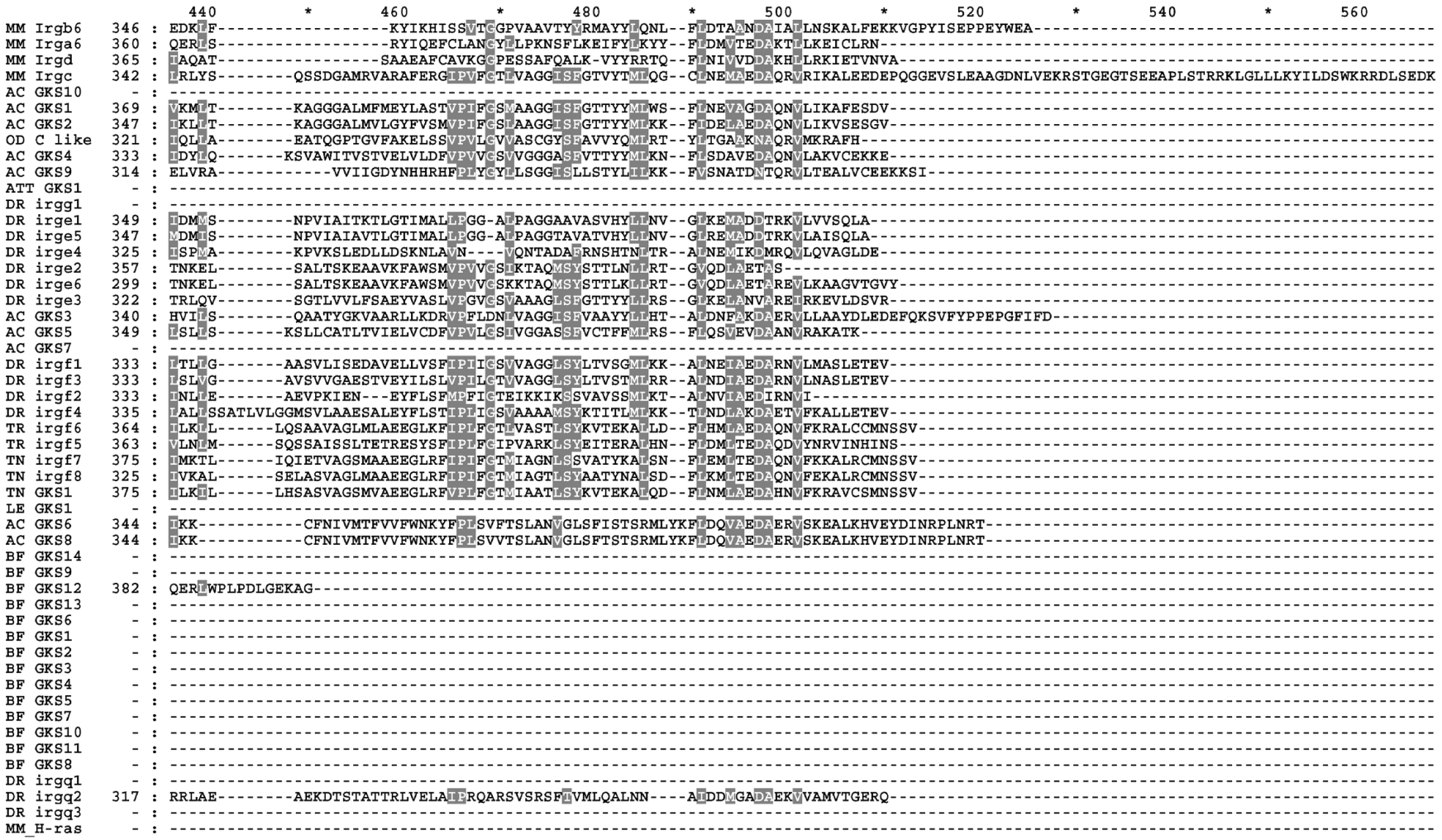
Taken together the IRG resistance GTPases are an ancient family that recently underwent extensive expansion and diversification (as well as contraction) in the euchordates – a feature characteristic for multigene families associated with pathogen resistance due to host-pathogen coevolution (Angata 2004; Borghans 2004; Delarbre 1992; Hood 1975; Klein 1986; Leister 2004; Mashimo 2003; Noel 1999; Trowsdale 2001). The absence of full length IRGM GTPases and IFN-inducible IRGs in general in higher primates indicates that this ancient resistance system, despite its importance for the mouse, became – at least largely - dispensable in this linage and was lost during divergent evolution of the primates. Consequently, mice and humans must deploy their immune resources against vacuolar pathogens in radically different ways.

Figure 23 Alignment of fish, reptile, amphibian and Cephalochordate IRG proteins (see also next page). Alignment of fish, reptile, amphibian and Cephalochordate IRG proteins with representatives of the murine IRGs and the opossum IRGC-like protein. The positions of the introns in fish *IRGF* genes and opossum *IRGC-like* are marked in dark blue. Modifications in the G1 motif of the *Danio IRGXQ* genes are highlighted in pink. GTPase motifs are marked by red boxes. AC: *Anolis carolinensis*, AT: *Ambystoma tigrinum tigrinum,* BF: *Branchiostoma floridae*, DR: *Danio rerio*, MM: *Mus musculus*, MD: *Monodelphis domestica.*


MM Irgb6
MM Irga6
MM Irgd
MM Irgc
AC GKS10
AC GKS1
AC GKS2
OD C like
AC GKS4
AC GKS9
ATT GKS1
DR irgg1
DR irge1
DR irge5
DR irge4
DR irge2
DR irge6
DR irge3
AC GKS3
AC GKS5
AC GKS7
DR irgf1
DR irgf3
DR irgf2
DR irgf4
TR irgf6
TR irgf5
TN irgf7
TN irgf8
TN GKS1
LE GKS1
AC GKS6
AC GKS8
BF GKS14
BF GKS9
BF GKS12
BF GKS13
BF GKS6
BF GKS1
BF GKS2
BF GKS3
BF GKS4
BF GKS5
BF GKS7
BF GKS10
BF GKS11
BF GKS8
DR irgq1
DR irgq2
DR irgq3
MM H-ras
G1
G2


G3
G4


300
320
340
360
380
400
420
MM Irgb6 209 :
MM Irga6 222 :
MM Irgd 227 :
MM Irgc 205 :
AC GKS10 208 :
AC GKS1 232 :
AC GKS2 210 :
OD C like 184 :
AC GKS4 195 :
AC GKS9 180 :
ATT GKS1 158 :
DR irgg1 198 :
DR irge1 212 :
DR irge5 210 :
DR irge4 189 :
DR irge2 222 :
DR irge6 164 :
DR irge3 187 :
AC GKS3 205 :
AC GKS5 211 :
AC GKS7 170 :
DR irgf1 196 :
DR irgf3 196 :
DR irgf2 196 :
DR irgf4 198 :
TR irgf6 227 :
TR irgf5 227 :
TN irgf7 239 :
TN irgf8 189 :
TN GKS1 238 :
LE GKS1 194 :
AC GKS6 207 :
AC GKS8 207 :
BF GKS14 212 :
BF GKS9 177 :
BF GKS12 239 :
BF GKS13 199 :
BF GKS6 233 :
BF GKS1 209 :
BF GKS2 211 :
BF GKS3 204 :
BF GKS4 211 :
BF GKS5 211 :
BF GKS7 239 :
BF GKS10 265 :
BF GKS11 216 :
BF GKS8 212 :
DR irgq1 197 :
DR irgq2 178 :
DR irgq3 218 :
MM_H-ras 145 :
G5


440 * 460 * 480 * 500 * 520 * 540 * 560
MM Irgb6 346 : EDKLF-----------------KYIKHISSVTGGPVAAVTYYRMAYYLQNL--FLDTANDAIALNSKALFEKKVGPYISEPPEYWEA----------
MM Irga6 360 : QERLS-----------------RYIQEFCLANGYLLPKNSFLKEIFYLKYY--FLDMVTEDAKTLLKEICLRN----------
MM Irgd 365 : LAQAT----------------SAAEAFCAVKGGPESSAFQALK-VYYRRTQ--FLNIVVDDAKHLLRKIETVNVA----------
MM Irgc 342 : LRLYS--------QSSDGAMRVARAFERGIPVFGTLVAGGISEGTVYTMLQG--CLNEMAEDAQRVRIKALEEDEPQGGEVSLEAAGDNLVEKRSTGEGTSEEAPLSTRRKLGLLLKYILDSWKRRDLSEDK
AC GKS10 - : ----------
AC GKS1 369 : VKMLT--------KAGGGALMFMEYLASTVPIFGSVAAGGISEGTTYYMLWS--FLNEVAGDAQNVLIKAFESDV----------
AC GKS2 347 : IKLLT--------KAGGGALMVLGYFVSMVPIFGSLAAGGISEGTTYYMLKK--FLDELAEDAQNVLIKVSESGV----------
OD C like 321 : LQLLA--------EATQGPTGVFAKELSSVPVLGVVASCGYSEAVVYQMLRT--YLTGALKNAQRVMKRAFH----------
AC GKS4 333 : LDYLQ-------KSVAWITVSTVELVLDFVPVVGSVVGGGASEVTTYYMLKN--FLSDAVEDAQNVLAKVCEKKE----------
AC GKS9 314 : ELVRA------------VVIIGDYNHHRHFPLYGYLLSGGISLLSTYLILKK--FVSNATDNTQRVLTEALVCEEKKSI----------
ATT GKS1 - : ----------
DR irgg1 - : ----------
DR irge1 349 : LDMVS--------NPVIAITKTLGTIMALLEGG-ALPAGGAAVASVHYLLNV--GLKEMADDTRKVLVVSQLA----------
DR irge5 347 : VDMIS--------NPVIAIAVTLGTIMALLEGG-ALPAGGTAVATVHYLLNV--GLREMADDTRKVLAISQLA----------
DR irge4 325 : ISPMA--------KPVKSLEDLLDSKNLAVN----VQNTADAERNSHTNLTR--ALNEMIKDMRQVLQVAGLDE----------
DR irge2 357 : TNKEL--------SALTSKEAAVKFAWSMVPVVGSLKTAQMSYSTTLNLLRT--GVQDLAETAS----------
DR irge6 299 : TNKEL--------SALTSKEAAVKFAWSMVPVVGSKKTAQMSYSTTLKLLRT--GVQDLAETAREVLKAAGVTGVY----------
DR irge3 322 : TRLQV--------SGTLVVLFSAEYVASLVEGVGSVAAAGLSEGTTYYLLRS--GLKELANVAREIRKEVLDSVR----------
AC GKS3 340 : HVILS--------QAATYGKVAARLLKDRVEFLDNLVAGGISEVAAYYLLHT--ALDNFAKDAERVLLAAYDLEDEFQKSVFYPPEPGFIFD----------
AC GKS5 349 : ISLLS-------KSLLCATLTVIELVCDFVPVLGSLVGGASSEVCTFFMLRS--FLQSVEVDAANVRAKATK----------
AC GKS7 - : ----------
DR irgf1 333 : ITLLG-------AASVLISEDAVELLVSFIPIIGSVVAGGLSYLTVSGMLKK--ALNEIAEDARNVLMASLETEV----------
DR irgf3 333 : ISLVG-------AVSVVGAESTVEYILSLVPILGTVVAGGLSYLTVSTMLRR--ALNDIAEDARNVLNASLETEV----------
DR irgf2 333 : INLLE-------AEVPKIEN---EYFLSFMEFIGTEIKKIKSSVAVSSMLKT--ALNVIAEDIRNVI----------
DR irgf4 335 : LALLSSATLVLGGMSVLAAESALEYFLSTIPLIGSVAAAAMSYKTITLMLKK--TLNDLAKDAETVFKALLETEV----------
TR irgf6 364 : ILKVL------LQSAAVAGLMLAEEGLKFIPLFGTLVASTLSYKVTEKALLD--FLHMLAEDAQNVFKRALCCMNSSV----------
TR irgf5 363 : VLNLM------SQSSAISSLTETRESYSFIPLFGIPVARKLSYEITERALHN--FLDMLTEDAQDVYNRVINHINS----------
TN irgf7 375 : IMKTL------IQIETVAGSMAAEEGLRFIPIFGTMIAGNLSSVATYKALSN--FLEMLTEDAQNVFKKALRCMNSSV----------
TN irgf8 325 : IVKAL------SELASVAGLMAAEEGLRFIPIFGTMIAGTLSYAATYNALSD--FLKMLTEDAQNVFEKALRCMNSSV----------
TN GKS1 375 : ILKIL------LHSASVAGSMVAEEGLRFVPLFGTMIAATLSYKVTEKALQD--FLNMLAEDAHNVFKRAVCSMNSSV----------
LE GKS1 - : ----------
AC GKS6 344 : IKK----------CFNIVMTFVVFWNKYFPLSVFTSLANVGLSFISTSRMLYKFLDQVAEDAERVSKEALKHVEYDINRPLNRT----------
AC GKS8 344 : IKK----------CFNIVMTFVVFWNKYFPLSVVTSLANVGLSFTSTSRMLYKFLDQVAEDAERVSKEALKHVEYDINRPLNRT----------
BF GKS14 - : ----------
BF GKS9 - : ----------
BF GKS12 382 : QERLWPLPDLGEKAG----------
BF GKS13 - : ----------
BF GKS6 - : ----------
BF GKS1 - : ----------
BF GKS2 - : ----------
BF GKS3 - : ----------
BF GKS4 - : ----------
BF GKS5 - : ----------
BF GKS7 - : ----------
BF GKS10 - : ----------
BF GKS11 - : ----------
BF GKS8 - : ----------
DR irgq1 - : ----------
DR irgq2 317 : RRLAE--------AEKDTSTATTRLVELAIERQARSVSRSFMVMLQALNN----AIDDMGADAEKVVAMVTGERQ----------
DR irgq3 - : ----------
MM_H-ras - : ----------

III.2. Regulatory interactions between IRG GTPases controlling activation and function

The exceptional importance of the IFNγ-inducible p47 GTPases in cell autonomous resistance has been demonstrated both *in vivo* (Collazo 2001; MacMicking 2003; Taylor 2000) and *in vitro* (Bernstein-Hanley 2006; Feng 2004; Martens 2005; Nelson 2005), but next to nothing is known about the molecular and cell biological mechanisms mediating these effects. To fill this gap and to disentangle the effect mediated by individual IRG GTPases from the residual IFNγ response, stable cell lines inducibly expressing single IRGs were established. Inducible expression was chosen because an earlier study reported difficulties in obtaining stable clones constitutively expressing Irgb6 and observed a reduced growth rate for cells expressing high levels of this protein (Carlow 1998). This may indicate that the constitutive expression of single IRGs in absence of IFN-induction can be toxic for cells.

III.2.1. Generation and characterisation of stable cell lines inducibly expressing single IRGs

The GeneSwitch (gs) system, based on the NIH3T3 derived gs3T3 cell line stably expressing the GeneSwitch inducer protein (Invitrogen; (Wang 1994)), was used for hormone-inducible expression of the IRG proteins. Upon binding of the synthetic steroid, Mifepristone, the inducer drives the expression of the protein of interest from the pGene/V5-His vector (see material and methods (II.5.3) for details). The complete open reading frames of Irga6, Irgb6, Irgc, Irgd, Irgm1, Irgm2 and Irgm3 were cloned into this vector and the resulting native constructs were transfected individually into gs3T3 cells. Cells with stable integrated IRG expression constructs were selected with Zeocin and cloned twice by limiting dilution. The resulting stable cell lines were characterised for IRG protein expression following Mifepristone induction, using SDS-PAGE and immunoblotting with specific serological reagents (Figure 24). The IRG expression level was not adjustable by using different hormone concentrations (Figure 24; shown for Irga6 (A), Irgm3 (F) and Irgm1 (G)). No induction was detected with a concentration of 10^{-11} M Mifepristone or less, and maximal induction with 10^{-10} M and more. Subsequently, 10^{-9} M Mifepristone was used for induction unless specified elsewise. Clones expressing IRG proteins at comparably high levels following Mifepristone- and IFNγ-induction were chosen for further analysis (Figure 24, see also Figure 34). In the case of Irgb6, all clones analysed expressed considerably less protein following Mifepristone induction than following IFNγ induction (Figure 24 B).

All IRG proteins migrated in SDS-PAGE at the position expected from their calculated molecular weight (MW), which is about 47 kDa for Irga6, Irgb6, Irgd and the IRGM proteins but 50 kDa for Irgc. The exception was Irgm1 that had an apparent molecular weight in SDS-PAGE of around 37 kDa compare with the calculated MW of ~ 47 kDa (Figure 24 G). Furthermore, Irgm3 protein migrated as a double band following IFNγ induction but only the upper band at 47 kDa was detected following Mifepristone induction (Figure 24 F). The second Irgm3 band migrating slightly faster than the predominant band was also seen by Taylor *et al* (Taylor 1996). The difference between the two bands could either be due to the use of an alternative methionine 20 amino acids

downstream of the primarily used one (as for Irgm1 and Irgm2; see above III.1.2) or result from an incomplete covalent protein modification. Yet, no ESTs supporting the use of the alternative methionine could be found in the databases.

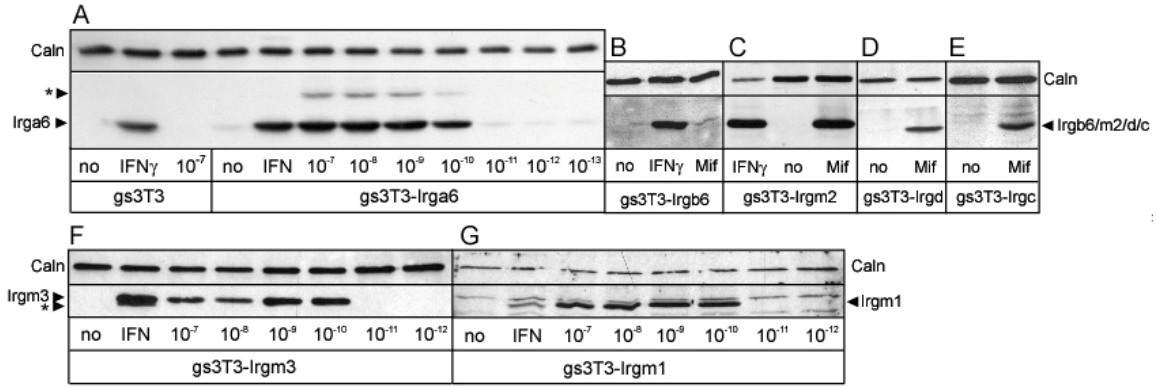


Figure 24 IRG protein expression in gs3T3-IRG cell lines. Gs3T3-IRG and -wt cells were stimulated with Mifepristone (10^{-7} to 10^{-13} M) or IFNγ (200 U/ml) for 24 hrs, postnuclear supernatants of PBS/1% Triton X-100 cell lysates were separated by SDS-PAGE and Irga6 (A), Irgb6 (B), Irgm2 (C), Irgd (D), Irgc (E), Irgm3 (F) and Irgm1 (G) were detected in immunoblot with the specific antibodies 165, A20, H53, 2078, 39/3°, αIGTP clone 7, and L115/B0 respectively. Detection of Calnexin with the SPA-865 serum served as a loading control. (A), (F) and (G) show Mifepristone dose responses for gs3T3-Irga6, -Irgm3 and –Irgm1 cells respectively. In (B-E) cells were induced with 10^{-9} M Mifepristone. Black arrowheads indicate the respective IRG protein band(s), the asterisks in (A) indicates the additional Irga6 band at 55 kDa prominent in Mifepristone-induced gs3T3-Irga6 cells, the asterisks in (F) the second Irgm3 band seen only in IFNγ-induced cells. The asterisk in (G) indicates an unrelated protein band present in gs3T3 cells, the Irgm1 serum (L115/B0) cross-reacted on.

In the case of Irga6, Mifepristone induction resulted in the expression of two forms, one migrating at the same height as the endogenous Irga6 and one form that had an apparent molecular weight of approximately 55 kDa (Figure 24 A asterisks and Figure 25 D red arrowhead). Likewise, transient transfection of an Irga6 expression construct into murine fibroblasts in absence of IFNγ induction resulted in formation of two predominant Irga6 bands in SDS-PAGE, in this case with an apparent molecular weight of about 47 and 48 kDa respectively (Figure 25 C black and green arrowhead respectively; see also (Papic 2007)). The lower band correlates with endogenous IFN-induced Irga6 that migrated predominantly as a single band of 47 kDa (Figure 25 A+B). The 55 kDa Irga6 band seen in Mifepristone-induced cells could also be detected in cells transiently transfected with Irga6 and sometimes even in IFNγ-induced cells, though the signal was much weaker in the last case (Figure 25 A+C+D red arrowhead).

It was previously shown by Triton X-114 phase partitioning that endogenous Irga6 is completely myristoylated at its N-terminal myristoylation site *in vivo* (Martens 2004b). In this assay, hydrophobic proteins partition into the detergent phase upon temperature shift induced phase separation (Bordier 1981). Irga6 from IFNγ-induced L929 cells partitioned roughly equally between the aqueous and the detergent phases in consecutive Triton X-114 extractions, indicating complete lipid modification of Irga6 but only 50% efficiency of partitioning for this protein (Martens 2004b). This result was confirmed with IFNγ-induced gs3T3 and gs3T3-Irga6 cells (Figure 25 A+B). The 47 kDa Irga6 band seen in transiently transfected cells behaved like endogenous Irga6 partitioning equally between the two phases in subsequent rounds of extraction, while the 48 kDa band was completely aqueous (Figure 25 C; see also (Papic 2007)). The latter

band was shown to represent non-myristoylated Irga6 (Papic 2007) behaving like the myristoylation mutant Irga6(G2A) (Martens 2004b). It is noteworthy that the 48 kDa form of Irga6 was also found in tiny amounts in IFN-induced cells (Figure 25 A, A4, green arrow). This indicates that lipid modification of Irga6 by the N-terminal myristoyl transferase is carefully balanced and overexpression of Irga6 by transient transfection of an expression construct with a strong viral promoter probably overburdened the capacity of the myristoylation machinery. In Mifepristone-induced gs3T3-Irga6 cells, both the 47 and the 55 kDa form of Irga6 behaved like the endogenous protein partitioning roughly equally between the aqueous and the detergent phase (Figure 25 D black and red arrowhead). Thus, the 55 kDa form of Irga6 is myristoylated and, whatever modification leads to the shift in apparent molecular weight, it does not influence the hydrophobicity of the protein detectably.

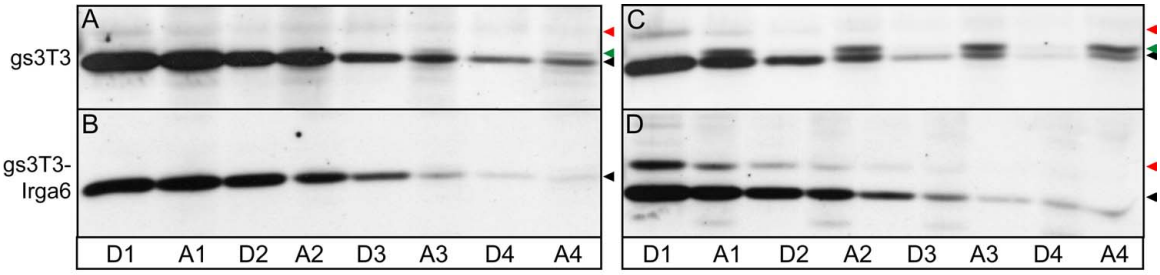


Figure 25 IFNγ- and Mifepristone-induced Irga6 is completely myristoylated *in vivo*. Gs3T3 cells (A, C) were either induced with IFNγ (200 U/ml) for 24 hrs or transiently transfected with the pGW1H-Irga6 (2μg) expression vector. Gs3T3-Irga6 cells (B, D) were either induced with IFNγ (200 U/ml) or Mifepristone (10^{-9} M) for 24 hrs. Cell lysates were generated with PBS/1% Triton X-114/3 mM $MgCl_2$ and the aqueous and detergent phases were separated by shifting the sample to temperatures above the cloud point (22°C) aided by centrifugation. Four subsequent partitionings of aqueous (A1-4) and detergent phase (D1-4) were performed by re-adding detergent to the aqueous phase (A1 partitions into A2 and D2 etc.).

As the Irga6 band at 55 kDa was resistant to reducing and denaturing agents (β-mercaptoethanol, SDS, boiling), a covalent post-translational modification of the protein was likely to be the explanation for the shift of around 8 kDa. This size difference is consistent with mono-ubiquitination, as one ubiquitin molecule has a molecular weight of 8.5 kDa (Schlesinger 1975). To test this possibility, Irga6 was immunoprecipitated from lysates of gs3T3-Irga6 cells induced with IFNγ, Mifepristone or both with the specific polyclonal serum 165 coupled to protein A Sepharose. Gs3T3 cells and untreated gs3T3-Irga6 cells served as controls. The immunoprecipitates were analysed in SDS-PAGE followed by immunoblot with the monoclonal antibodies 10D7 (Figure 26 A) and FK2 (Figure 26 B) specific for Irga6 and ubiquitin respectively. The FK2 antibody readily detected ubiquitinated proteins in the lysate controls (Figure 26 B, lanes 1-4 and 9-12). Furthermore, large amounts of ubiquitin-positive higher molecular weight species were detected following Mifepristone induction indicating strong polyubiquitination of Irga6 (Figure 26 B, lane 7). Upon longer exposure, the higher molecular weight bands were also visible in the Irga6 immunoblot (Figure 26 C, lane 7 and 8). The effect was only partially reversed by co-induction with IFNγ (Figure 26 B, lane 8) whereas no ubiquitination could be detected following IFNγ induction (Figure 26 B, lane 6). The immunoprecipitated sample from Mifepristone-induced gs3T3 wild type lacked the prominent ladder of bands ranging from 55 to 200 kDa seen in Mifepristone-induced gs3T3-Irga6 cells (Figure 26 B, compare lane 15 to lane 7), ruling out unspecific co-

precipitation of unrelated ubiquitinated proteins. Thus, ectopically expressed Irga6 is polyubiquitinated *in vivo*. This effect was not (or only partially) rescued by parallel induction with IFNγ, probably due to the elevated Irga6 expression levels following double induction. However, no obvious difference in protein stability was seen between endogenous Irga6 protein in MEFs and Mifepristone-induced Irga6 protein in gs3T3-Irga6 cells (half life in both cases about 16 hrs; T. Steinfeld, unpublished results). A prominent but diffuse band at 55 kDa was detected with the anti-ubiquitin antibody in immunoprecipitated samples from Mifepristone-induced gs3T3-Irga6 cells (Figure 26 B, lane 7). However, the fact that this band is equally intense even in Irga6 negative samples argues strongly for a cross-reaction of the ubiquitin antibody with the heavy chain (MW ~ 50 kDa) of the serum used for immunoprecipitation (Figure 26 B lane 5, 13, 15). Due to this technical problem, it could not be determined in this assay whether mono-ubiquitination was responsible for the observed size shift of Irga6 protein in SDS-PAGE. The immunoprecipitation was repeated and the samples were subjected to SDS-PAGE under non-reducing conditions (without β-mercaptoethanol) to prevent the separation of the disulfide bridge linked heavy and light chains of the used serum. Thereby the detected antibody band should shift to a higher molecular weight (~150 kDa). Unfortunately, Irga6 did not migrate properly under these conditions (data not shown).

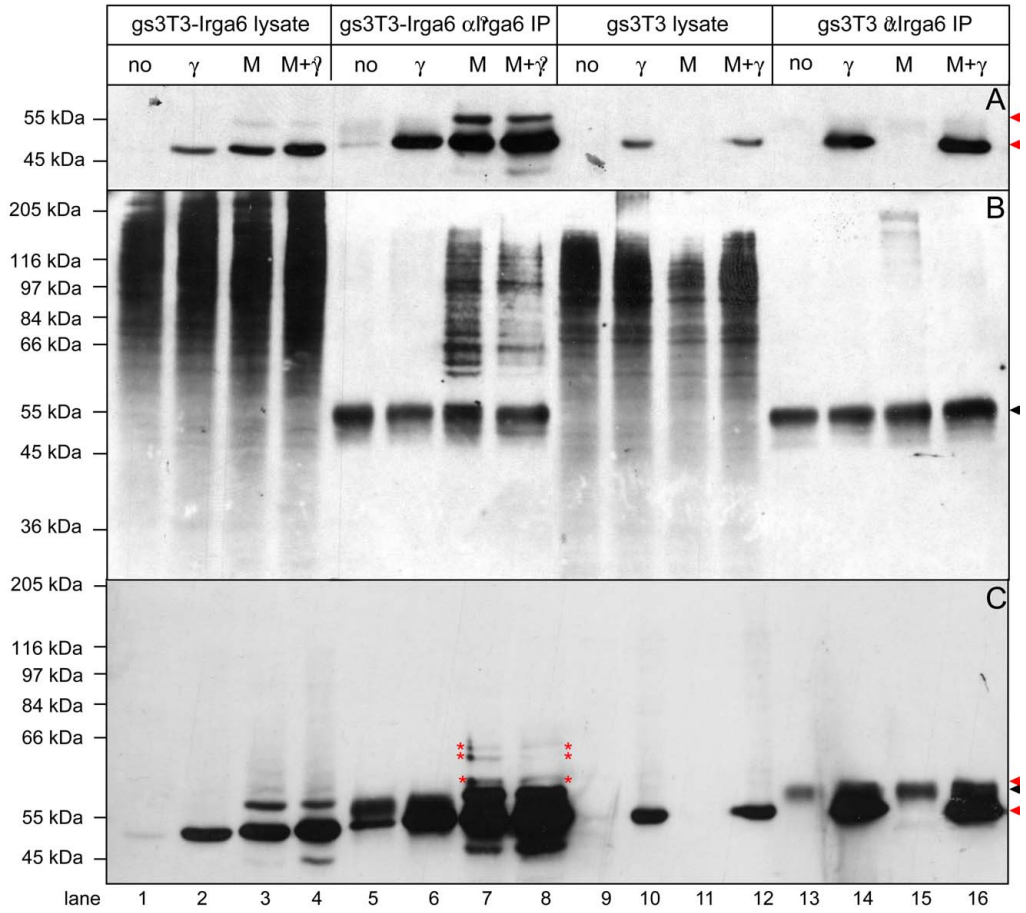


Figure 26 Irga6 is polyubiquitinated in Mifepristone- but not in IFNγ-induced gs3T3-Irga6 cells. Gs3T3 wt (lane 1-8) and -Irga6 cells (lane 9-16) were stimulated for 24 hrs with either IFNγ (200 U/ml), Mifepristone (10^{-9} M) or both or were left untreated before lysis in PBS/1% Triton X-100. Irga6 protein was immunoprecipitated from the postnuclear supernatants with 165 antiserum coupled to protein A Sepharose beads. 2% of the original lysates (lanes 1-4 and 9-12) and 50% of immunoprecipitates (lane 5-8 and 13-16) were analysed in two parallel immunoblots with the monoclonal antibodies directed against (A, C) Irga6 (10D7) and (B) ubiquitin (FK2). (C) Represents a longer exposure of the immunoblot displayed in

(A) in order to make the ubiquitinated higher molecular weight bands visible (γ: induced with IFNγ; M: induced with Mifepristone, M+γ: induced with both IFNγ and Mifepristone, no: untreated). Black arrowheads indicate the heavy chain of the serum used for immunoprecipitation. Irga6 specific band are indicated in red: red arrowheads indicate the 47 and 55 kDa forms of Irga6 and red asterisks indicate ubiquitinated Irga6.

In order to determine the type of protein modification found on the 55 kDa form of Irga6, the protein was immunoprecipitated from Mifepristone-induced cells and separated by SDS-PAGE. The band at 55 kDa was cut from the Coomassie stained gel and was sent for matrix assisted laser desorption/ionisation mass spectrometry (MALDI-MS) analysis (U. Roth, Bioanalytical Laboratory of the CMMC, Cologne). Even though Irga6 derived peptides could be identified, the contamination with the heavy chain of the antibody used for immunoprecipitation resulted in a largely reduced coverage making the identification of a protein modification impossible (data not shown). Also separation of Irga6 by two-dimensional gel electrophoresis failed to yield any further insights as the protein did not focus in the separation by isoelectric point (data not shown).

III.2.2. Influence of IRG expression cell proliferation and survival

Prompted by the previous report on the negative influence of stable ectopic expression of Irgb6 on cell survival and proliferation (Carlow 1998), the growth behaviour of the inducible gs3T3-IRG cell lines was monitored over a period of 10 days using a colorimetric cell proliferation assay (CellTiter96, Promega). Several clones were analysed for each IRG protein in at least two independent experiments. Results of single representative experiments are shown. Gs3T3 wt cells induced with Mifepristone proliferated slightly but repeatedly slower than both untreated and IFNγ-induced cells (Figure 27 B). This effect could not be documented with NIH3T3 cells not containing the inducer plasmid, suggesting that the expression of the GeneSwitch protein might be responsible for the reduction in growth rate (Figure 27 A). On the other hand, NIH3T3 cell growth rate was slightly reduced following IFNγ-induction indicating some general differences between the NIH3T3 and gs3T3 cells used. When induced with Mifepristone, gs3T3-Irga6, -Irgm1, -Irgm2, -Irgm3 and -Irgc cells reproducibly displayed a minor growth retardation compared to non-induced cells in the same range as the effect seen in the wild type gs3T3 cells (Figure 27 C-G). No additional growth inhibitory effect mediated by the expressed proteins could be detected. In contrast to the other cell lines tested, gs3T3 cells expressing Irgd or Irgb6 following Mifepristone induction showed a dramatic growth arrest over the whole course of the experiment (Figure 27 H, I). Microscopically monitoring of the cells during the assay confirmed a lack of proliferation and revealed increasing cell death over time. All gs3T3-Irga6 and -Irgd clones analysed were growth inhibited following Mifepristone induction (3 clones for Irgd and 5 clones for Irgb6; data not shown). As induction of endogenous Irgb6 and Irgd in the context of the IFNγ response did not show a negative effect on cell proliferation (Figure 27 B, K, I), it was hypothesised that in Mifepristone-induced cells an IFNγ-induced regulatory factor is missing that normally prevents growth inhibition by these proteins. To test this possibility, double inductions with Mifepristone and IFNγ were performed, inducing either simultaneously or preinducing 24 hrs with one inducer before adding the other one. Unexpectedly, neither of these treatments resulted in a rescue of the complete growth

inhibition (Figure 27 K, I). In the case of Irgd that could be explained by overexpression of the protein as a result of double induction. As the Irgb6 expression induced by Mifepristone was hardly detectable in Western blot and drastically lower than following induction with IFNγ, such a stoichiometric effect is highly unlikely for Irgb6.

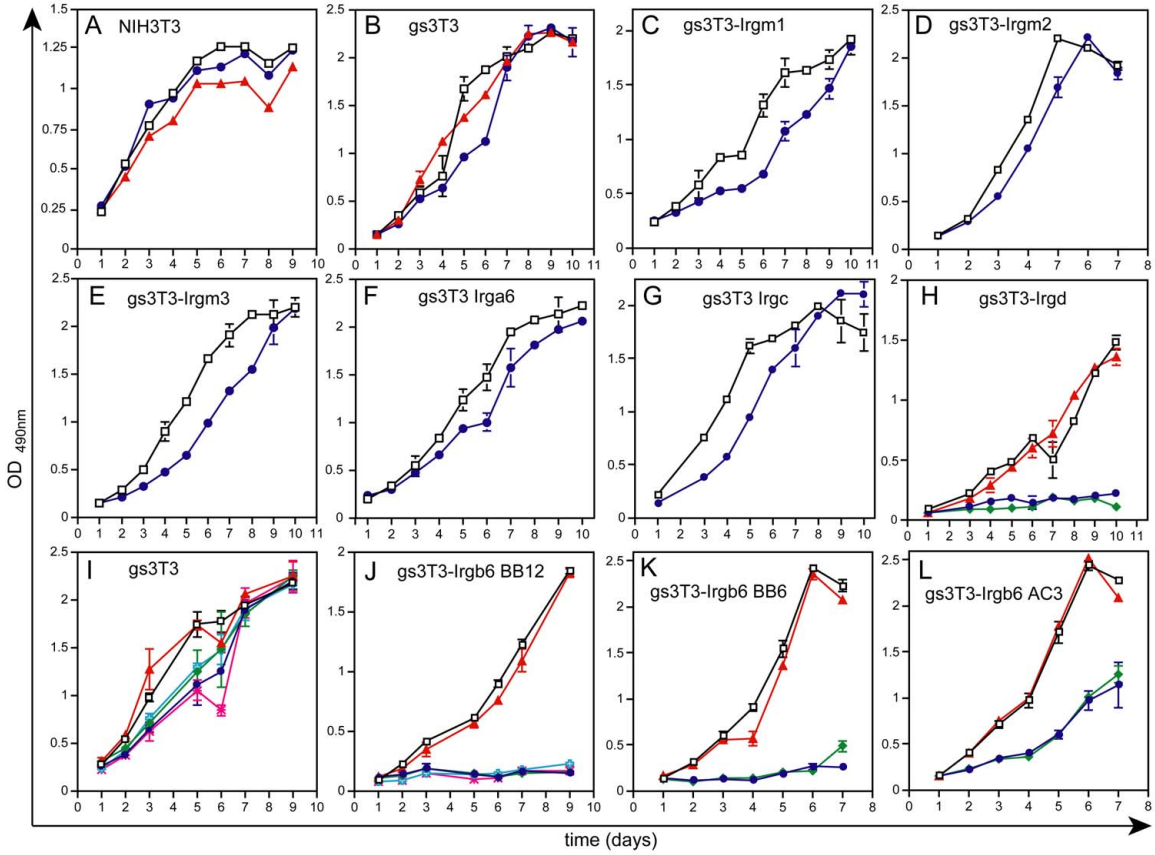


Figure 27 Influence of IRG expression on cell proliferation. Equal numbers of cells from the indicated cell lines (1000/well) plated onto 96well plates in quadruplicates were induced with either IFNγ (200 U/ml), Mifepristone (10^{-9} M) or a combination of both for the indicated time. Cell proliferation was measured in a colorimetric assay detecting the formazan product of a tetrazolium salt bioreduced by living cells at 490 nm (empty squares: non-treated; red triangles: IFNγ-induced; blue circles: Mifepristone-induced; green diamonds: Mifepristone + IFNγ; blue crosses: 24 hrs preinduction with IFNγ followed by double induction; pink asterisks: 24 hrs preinduction with Mifepristone followed by double induction). The small differences in kinetics between the cell lines and individual experiments can be largely attributed to variations in cell number at the beginning of the experiment due to counting errors, as cells grew slower when plated at a lower density. Kinks in the curves were due to disturbance of the cells during feeding. One representative of two performed experiments is displayed.

To determine whether the observed growth inhibition was due to a cell cycle arrest, the DNA content of untreated, IFNγ- and Mifepristone-induced gs3T3 wt and -Irgb6 cells was determined (Figure 28). Independent of treatment, wild type cells displayed a DNA profile typical for proliferating cells (Figure 28 A), and so did non-induced gs3T3-Irgb6 cells (Figure 28 B upper panel). By contrast, gs3T3-Irgb6 cells were blocked in cell cycle progression before reaching the G2/M phase as early as 24 hrs after Mifepristone induction (Figure 28 B lower panel).

In order to exclude that the site of chromosomal integration of the Irgb6 expression construct was responsible for the observed growth inhibition and cell cycle arrest, the stable transfection was repeated and the growth behaviour of newly generated clones was determined. While the clone BB6 was also completely growth inhibited

following both Mifepristone and double induction (Figure 27 K), another clone, AC3, was only partially inhibited (Figure 27 L). Thus, further experiments need to be performed to confirm the specificity of the observed growth inhibitory effects, to analyse the suggested effect of Irga6 (and probably also Irgd) on the cell cycle and to establish whether apoptosis or necrosis is induced in cells expressing these proteins. Consequently, only short-term experiments were performed when expressing Irgb6 or Irgd in absence of IFNγ induction.

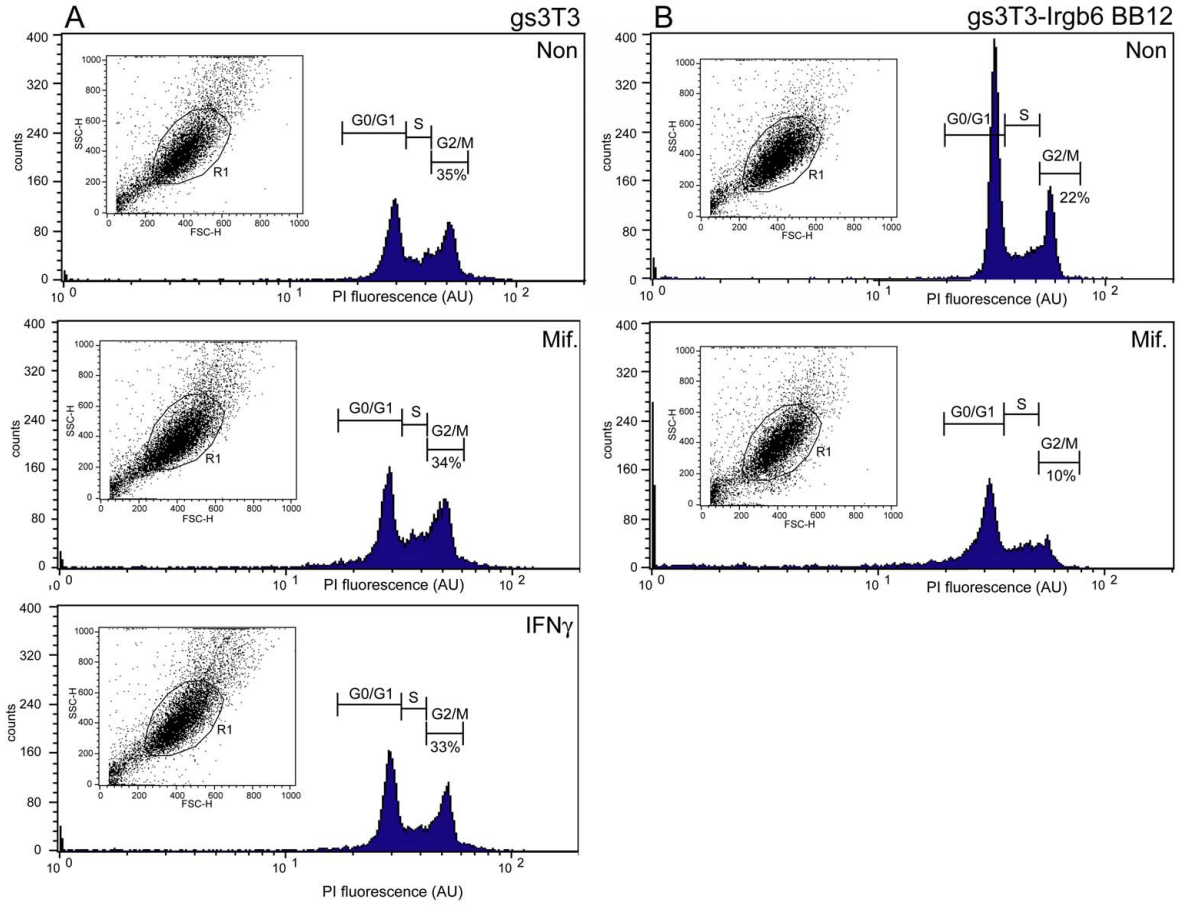


Figure 28 Influence of Irgb6 expression on cell cycle progression. Cell cycle progression of gs3T3 wt (A) and -Irgb6 cells (B) was monitored by analysing the cellular DNA content using propidium iodide (PI) staining of ethanol fixed cells and flow cytometry. Non-induced cells were compared with 24 hrs Mifepristone (10^{-9} M) and IFNγ (200 U/ml) induced ones. Single cells were gated (R1 in small inlay; cell size (forward scatter, FSC) is plotted against cell granularity (side scatter, SSC) and the DNA content measured by intensity of PI fluorescence was plotted against cell counts. The duplication of the cellular DNA content during the S phase can be see as an increase in PI fluorescence. The different phases of the cell cycle are indicated above the peaks. Percentage of gated cells in the G2/M phase is given.

III.2.3. Influence of IFNγ on the subcellular localisation of Irgm1-3, Irgc, Irgd and Irga6

The subcellular localisation of single IRG proteins induced by Mifepristone and IFNγ, respectively, was examined in gs3T3 mouse fibroblasts by immunofluorescence using specific immunoreagents. Of the three IRGM proteins, Irgm3 has been reported to localise to the endoplasmic reticulum and to undefined circular cytoplasmic structures (Taylor 1997). Irgm1 is found at a high intensity on Golgi membranes (Martens 2004b), and on the endolysosomal system (Zhao unpublished results), while Irgm2 is localised

rather precisely to Golgi membranes (Martens 2006). All three of these distinctive localisations were accurately reproduced in IFNγ-induced gs3T3 cells as seen by partial colocalisation with the organellar markers calnexin (ER), giantin and GM130 (both Golgi) (A-C of Figure 29-31. Furthermore, individual IRGM proteins expressed in absence of IFN by Mifepristone induction in the respective gs3T3-IRG cells (D-F of Figure 29-31) or by transient transfection of gs3T3 cells with single cDNA expression constructs (G-I of Figure 29-31; see also (Taylor 1997) (Irgm3) and (Martens 2004b) (Irgm1)) displayed the same localisation. Taylor and colleagues showed that the G1 mutant of Irgm3, S98N, did not bind GTP *in vivo* and localised normally to the ER when transfected as a GFP fusion into HeLa cells not stimulated with interferon (Taylor 1997). A similar result has been reported in transfected mouse L cells for the homologous mutant, Irgm1(S90N), in an otherwise native construct (Martens 2004b). These observations were both confirmed in gs3T3 fibroblasts transfected with cDNA expression constructs of the mutant GTPases singly (J-L of Figure 29, Figure 31) and were extended to include Irgm2(S78N), which is also correctly localised in uninduced gs3T3 fibroblasts (Figure 30 J-L). The localisation of the three GMS GTPases is thus independent of IFNγ stimulation, the presence of other IRG proteins and GTP binding.

Mifepristone-induced Irgc (CINEMA) in gs3T3-Irgc cells localised in a distributed reticular pattern both in absence (Figure 32 A) and presence of IFNγ (Figure 32 B-C). This is consistent with the primarily cytosolic localisation indicated by the predominant presence of Irgc in the supernatant of hypotonic lysates of seminiferous tubules from murine testes (Rhode 2007). Direct comparison of the subcellular localisation of hormone-induced Irgc in fibroblasts with endogenous Irgc was not possible, as Irgc is exclusively expressed in haploid spermatids *in vivo* (Rhode 2007). There are no immortalised cell lines that adequately represent this cell type and cell lines of other origin did not express detectable levels of the protein, either constitutively or after induction with interferons (Bekpen 2005b; Rhode 2007). However, endogenous Irgc was detected the whole cytoplasm of haploid spermatids with the specific rabbit serum 39/°3 in testis sections (Rhode 2007), thus arguing that Irgc localisation following ectopic expression in gs3T3 cells did not differ radically from the normal subcellular localisation.

The subcellular localisation of endogenous Irgd induced by IFNγ and Irgd ectopically expressed by Mifepristone induction could not be determined with absolute certainty as the Irgd specific serum 2078 worked poorly in immunofluorescence of uninfected cells (Figure 32 D-F). However, a weak, distributed, reticular signal could be detected in IFNγ as well as in Mifepristone-induced cells (Figure 32 E-F) but not in uninduced cells (Figure 32 D). To ascertain Irgd localisation, gs3T3 cells were transiently transfected with an expression construct of Irgd C-terminally tagged with ctag1 in presence and in absence of IFNγ. Irgd detected with the ctag1 specific 2600 antiserum localised in a distributed reticular pattern independent of IFN-induction and partially colocalised with Irga6 in IFN-induced cells (Figure 32 G-I). Additionally, Irgd was found inside the nucleus, especially in cells expressing high levels of the protein, whereas Irga6 was excluded from the nucleus (Figure 32 G-I). These results suggest that Irgd is largely cytosolic *in vivo,* as was indicated by its predominant presence in the supernatant of hypotonic lysates from IFNγ-induced fibroblasts (~ 90%; (Martens

2004b)), allowing detectable diffusion of the protein into the nucleus through the nuclear pores (permeable for proteins smaller than ~60 kDa; (Paine 1975)). In absence of any other conspicuous signal, it is likely that the residual membrane associated 10% of the endogenous Irgd protein localise to the ER.

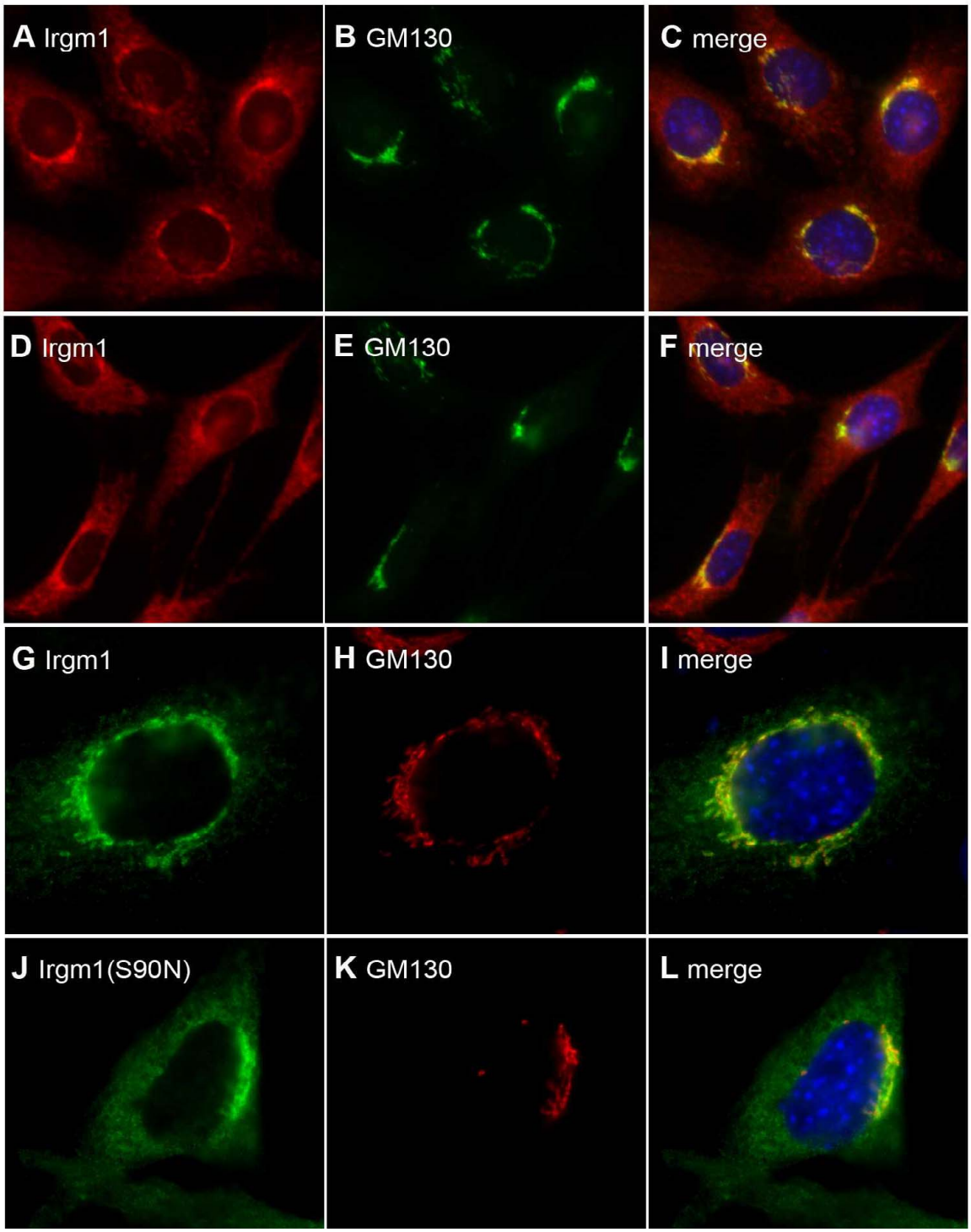


Figure 29 Subcellular localisation of Irgm1 is independent of IFNγ and nucleotide binding. (A-F) Gs3T3-Irgm1 cells were induced with IFNγ (200 U/ml) (A-C) or Mifepristone (10^{-9} M) (D-F) for 24 hrs. (G-L) Gs3T3 cells were transiently transfected with an (pGW1H, 1µg) expression construct for either Irgm1 wt (G-I) or for the G1 motif mutant Irgm1(S90N) (J-L) in absence of IFN. Irgm1 was detected in IF with the Irgm1 specific antiserum A19 (red in A, D; green in G, J) and the Golgi protein GM130 with G65120 monoclonal antibody (green in B, E; red in H, K). Overlay with the nuclear counterstain (DAPI, blue) is shown in C, F, I and L. The magnification is 400x in A-F and 630x in G-L.

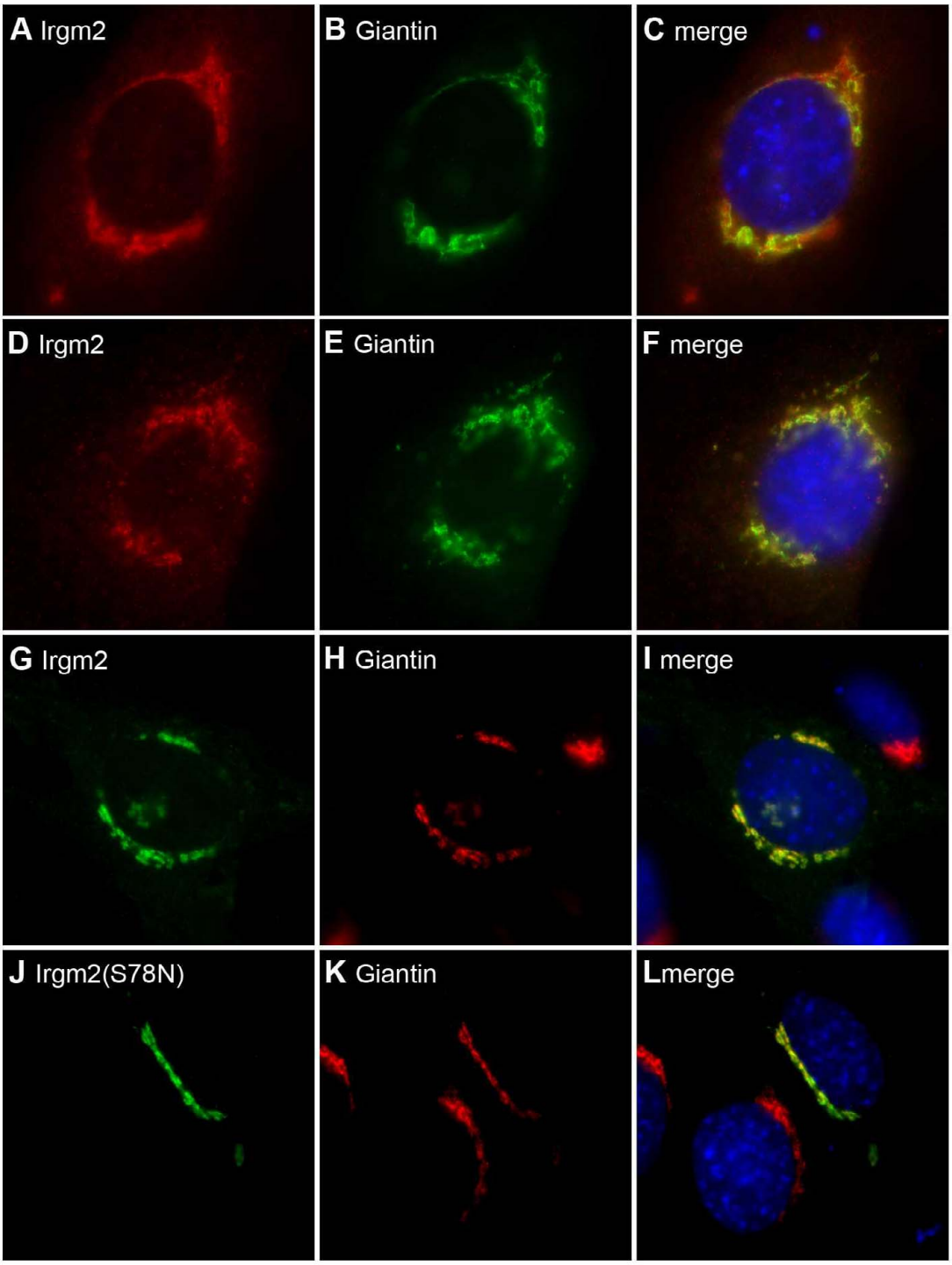


Figure 30 Subcellular localisation of Irgm2 is independent of IFNγ and nucleotide binding. (A-F) Gs3T3-Irgm2 cells were induced with IFNγ (200 U/ml) (A-C) or Mifepristone (10^{-9} M) (D-F) for 24 hrs. (G-L) Gs3T3 cells were transiently transfected with an (pGW1H, 1µg) expression construct for either Irgm2 wt (G-I) or for the G1 motif mutant Irgm2(S78N) (J-L) in absence of IFN. Irgm2 was detected with in IF with the Irgm2 specific antiserum H53 (red in A, D; green in G, J) and the Golgi protein Giantin with αGiantin monoclonal antibody ((green in B, E; red in H, K). Overlay with the nuclear counterstain (DAPI, blue) is shown in C, F, I and L. The magnification is 630x.

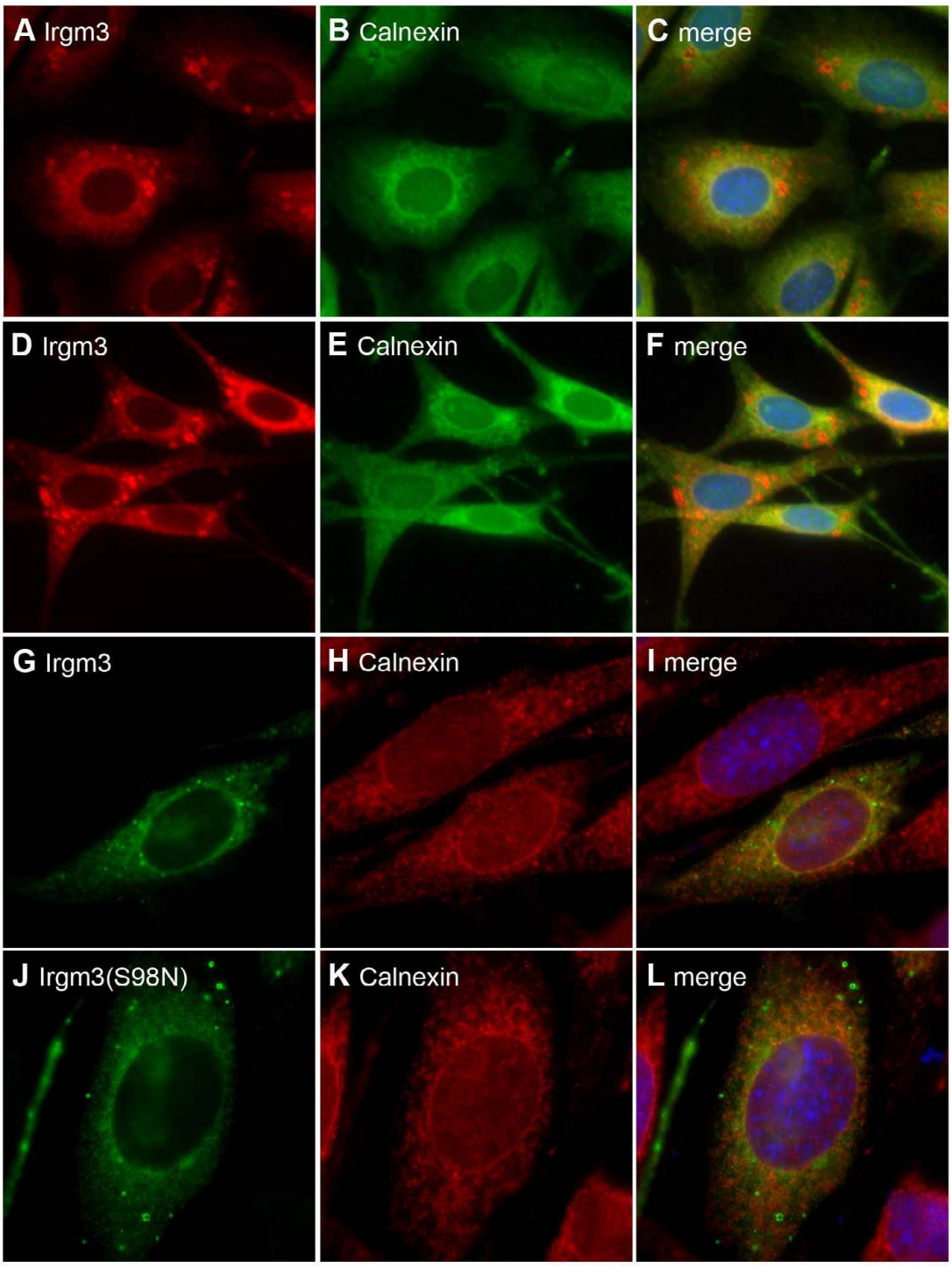


Figure 31 Subcellular localisation of Irgm3 is independent of IFNγ and nucleotide binding. (A-F) Gs3T3-Irgm3 cells were induced with IFNγ (200 U/ml) (A-C) or Mifepristone (10^{-9} M) (D-F) for 24 hrs. (G-L) Gs3T3 cells were transiently transfected with an (pGW1H, 1µg) expression construct for either Irgm3 wt (G-I) or for the G1 motif mutant Irgm3(S98N) (J-L) in absence of IFN. Irgm3 was detected with in IF with the Irgm3 monoclonal antibody αIGTP clone 7 (red in A and D; green in G and I) and the ER protein Calnexin with SPA-865 antiserum (green in B, E; red in H, K). Nuclei/cellular DNA was stained with DAPI (blue). The magnification is 400x in A-F and 630x in G-L.

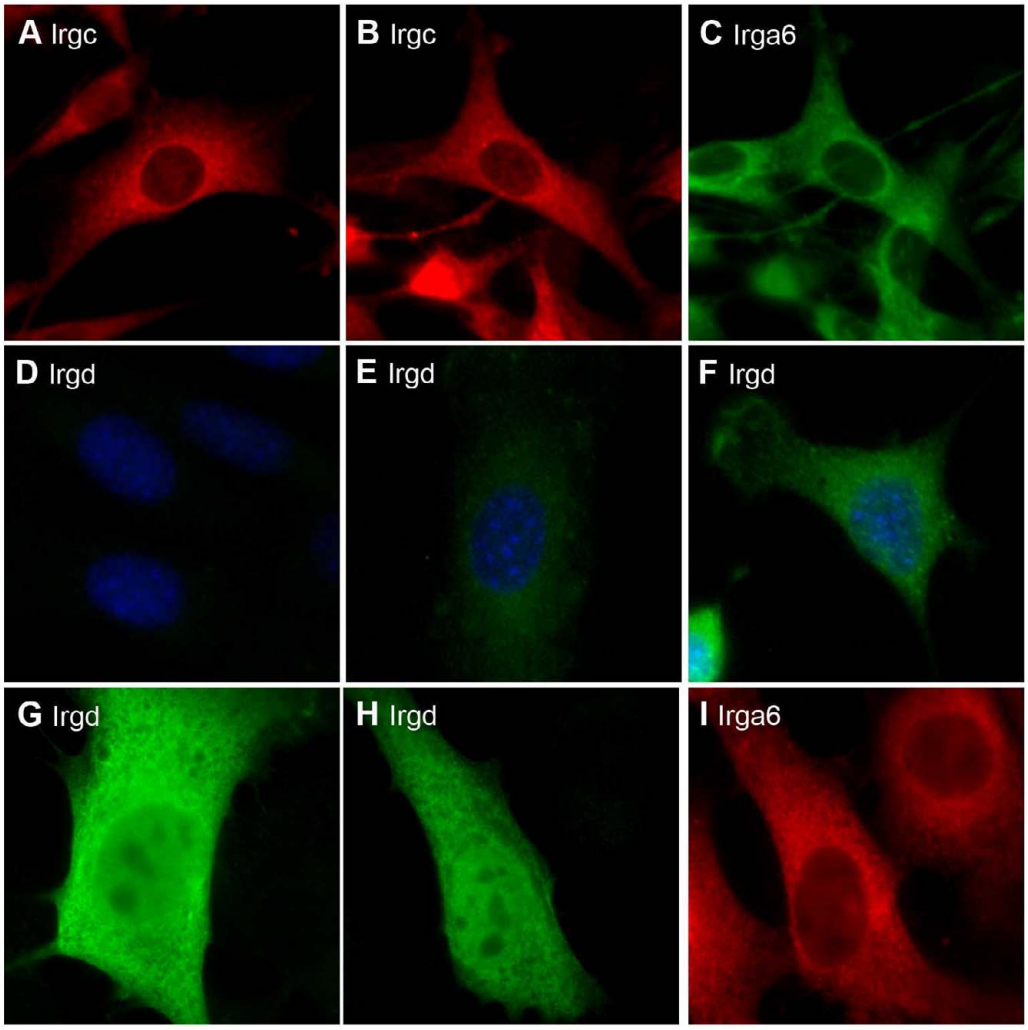


Figure 32 Subcellular localisation of Irgc and Irgd are independent of IFNγ. Gs3T3-Irgc cells were induced with (A) (10^{-9} M) Mifepristone alone (A) or together with (200 U/ml) IFNγ (B-C) for 24 hrs and stained with Irgc specific antiserum 39/3° (red in A and B). IFN-induced cells were counterstained for Irga6 with 10E7 monoclonal antibody (green in C). (D-F) Gs3T3-Irgd cells were either stimulated with IFNγ (E) or Mifepristone (F) as described above or left untreated (D), and stained with the Irgd specific antiserum 2078 (green) and DAPI for nuclear counter staining (blue). (G-I) Gs3T3 cells were transiently transfected with an (pGW1H, 1µg) expression construct for Irgc-ctag1 in the absence (G) or presence of IFNγ (H-I). Irgd was detected with the ctag1-specific serum 2600 (green in G-I). 10E7 monoclonal antibody was used for Irga6 counterstaining (red in I). The magnification is 400x in A-C and 630x in D-F.

Irga6 expressed in absence of IFNγ by transient transfection of a cDNA expression construct (Martens 2004b) or by Mifepristone-induction of gs3T3-Irga6 cells departed significantly from the typical dispersed ER distribution of endogenous, IFN-induced Irga6 (Figure 33 A) forming small cytoplasmic aggregates (Figure 33 B). Thus, this mislocalisation behaviour was not an artefact of the transient transfection protocol and it was also not due to over-expression, since gs3T3 cells expressed Irga6 at comparable levels following interferon and Mifepristone induction (Figure 34). When gs3T3-Irga6 cells were treated simultaneously with IFNγ and Mifepristone the wild-type localisation of Irga6 was largely restored (Figure 33 C). Hence, subcellular localisation of Irga6 was dependent on other IFN-inducible factors. The presence of residual aggregates in some cells suggest that this regulatory effect is close to stoichiometric and slightly overburdened in cells expressing twice the normal amount of Irga6 (see also Figure 33 M-R).

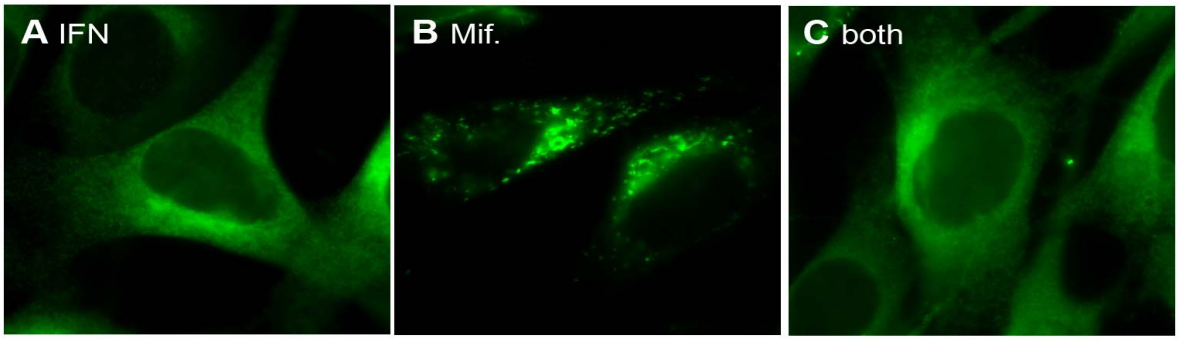


Figure 33 Subcellular localisation of Irga6 is dependent on IFNγ. Gs3T3-Irga6 cells were induced with (A) 200 U/ml IFNγ, (B) 10^{-9} M Mifepristone or (C) both for 24 hrs and Irga6 was detected in immunofluorescence with the 10E7 monoclonal antibody (green).

Since Irga6 was shown to form GTP-dependent homo-oligomers *in vitro,* it was conceivable that the Irga6 aggregates seen *in vivo* in absence of IFNγ represent GTP-bound oligomers. Thus, the monoclonal Irga6 antibody 10D7 that was shown to exclusively recognise the active, GTP-bound conformation of Irga6 in immunofluorescence (Papic 2007) was used to determine the nucleotide status of Irga6 protein *in vivo.* The polyclonal rabbit serum 165 (Figure 35 A, D, G, J, M, P) and the monoclonal antibody 10E7 directed against Irga6 both recognised Irga6 in immunofluorescence independent of the expression conditions (Figure 35 B, H, N). The monoclonal antibody 10D7, by contrast, failed to detected endogenous Irga6 in IFN-induced cells (Figure 35 E) but efficiently bound to the Irga6 aggregates formed following ectopic expression by Mifepristone induction of gs3T3-Irga6 cells (Figure 35 K). Upon co-induction with IFNγ and restoration of the wild type localisation, both the ectopic expressed and the endogenous Irga6 protein became undetectable for 10D7 (Figure 35 Q). Only residual aggregates were stained by 10D7. Thus, Irga6 is in the inactive, GDP-bound conformation in IFNγ-induced cells and in the active, GTP-bound conformation when expressed in absence of IFNγ. To further ascertain the nucleotide-dependence of Irga6 mislocalisation *in vivo,* the biochemical properties of two mutants of the first guanine nucleotide-binding motif, Irga6(K82A) and Irga6(S83N), were characterised *in vitro* and analysed respecting their subcellular localisation.

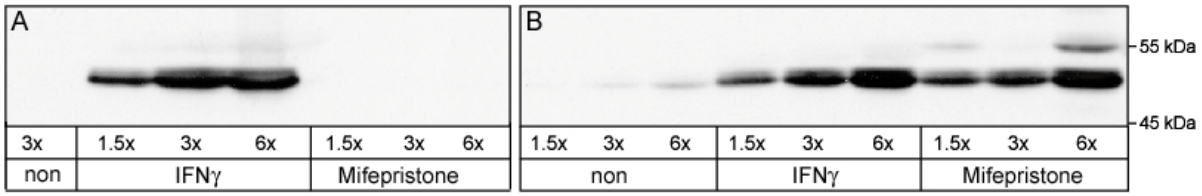


Figure 34 Comparable Irga6 expression levels following IFNγ-and Mifepristone-induction. Gs3T3 wt (A) and gs3T3-Irga6 cells (B) were stimulated with 200 U/ml IFNγ or 10^{-9} M Mifepristone for 24 hrs or left untreated. Postnuclear supernatants of PBS/1% Triton X-100 lysates from 1.5, 3 and $6*10^4$ cells (1.5x, 3x, 6x) were separated by SDS-PAGE. Irga6 was detected in immunoblot with the 165 serum. Note the presence of the 48 kDa band of Irga6 protein in both IFN- and Mifepristone-induced cells.

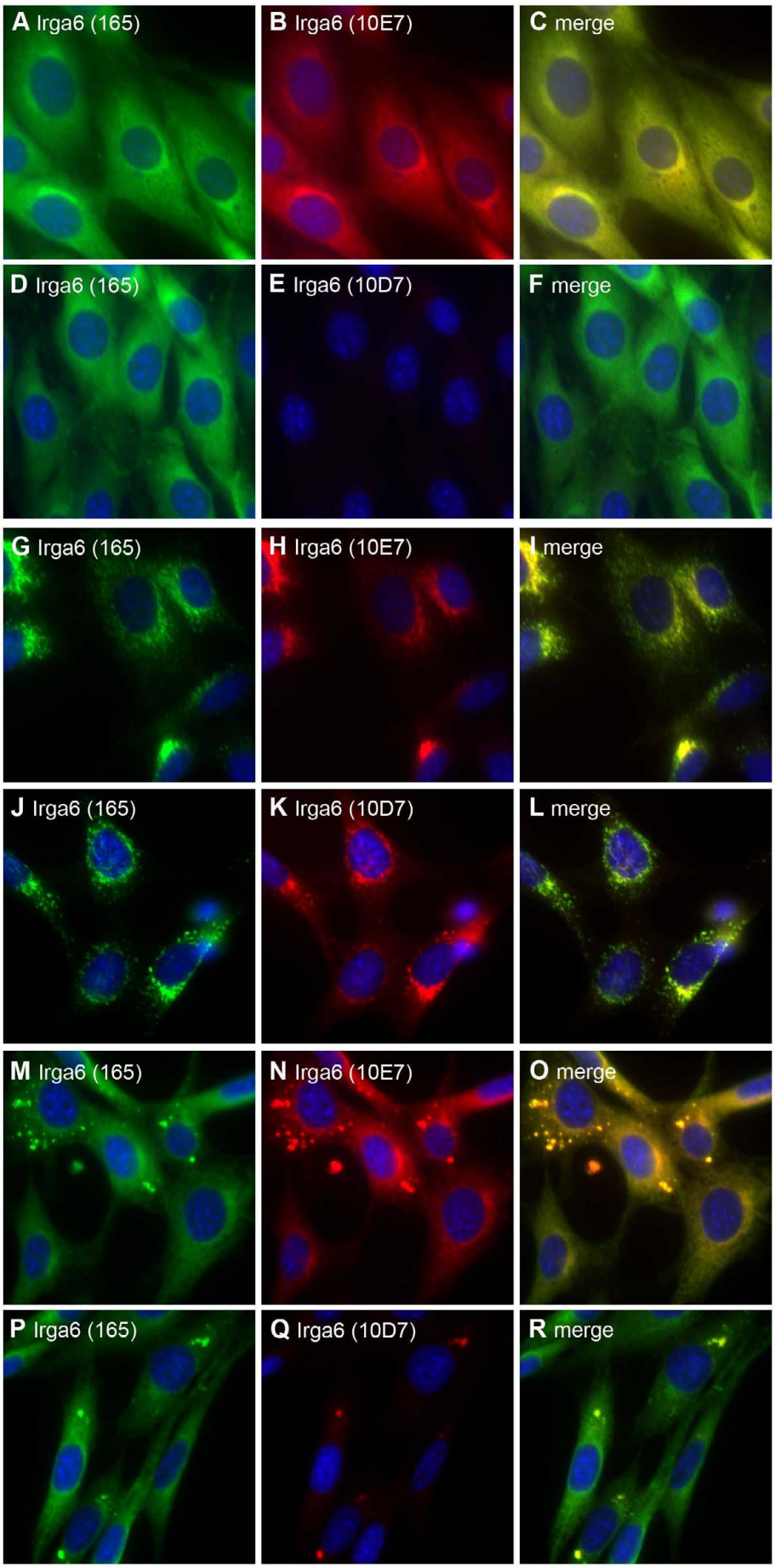
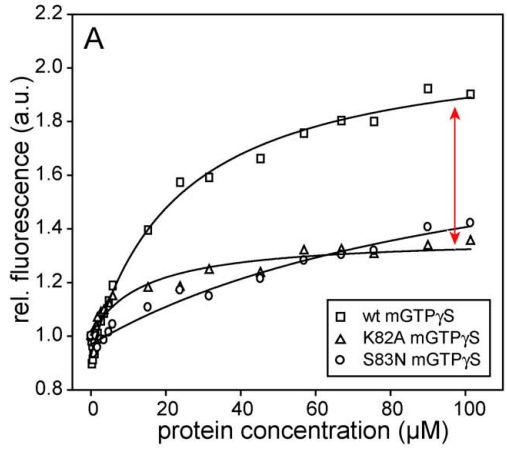

A Irga6 (165)
B Irga6 (10E7)
C merge
D Irga6 (165)
E Irga6 (10D7)
F merge
G Irga6 (165)
H Irga6 (10E7)
I merge
J Irga6 (165)
K Irga6 (10D7)
L merge
M Irga6 (165)
N Irga6 (10E7)
O merge
P Irga6 (165)
Q Irga6 (10D7)
R merge

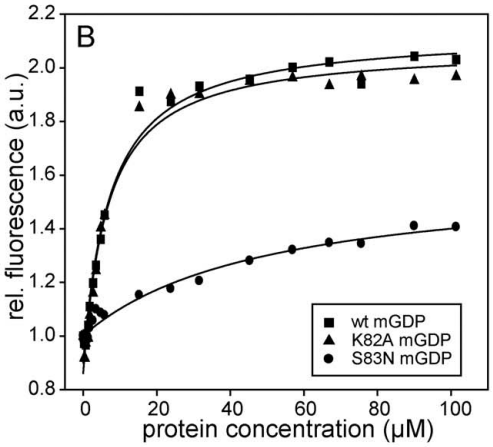
Figure 35 Irga6 conformation is dependent on the presence IFNγ (see previous page). Gs3T3-Irga6 cells were induced with 200 U/ml IFNγ (A-F), 10^{-9} M Mifepristone (G-L) or both (M-R) for 24 hrs and Irga6 was detected in immunofluorescence with the polyclonal rabbit serum 165 (A, D, G, J, M, P; green) and in parallel with either 10E7 (B, H, N; red) or 10D7 monoclonal antibodies (E, K, Q; red). Overlay of both channels with DAPI used for nuclear staining (blue) is shown in C, F, I, L, O and R. 10D7 is specific for the GTP-bound active conformation of Irga6 in immunofluorescence while 165 and 10E7 are not conformation sensitive.

III.2.4. Biochemical properties of Irga6 G1 mutants

For the biochemical analysis, Irga6(K82A) and Irga6(S83N) were expressed as N-terminal GST-fusions in bacteria and purified via a glutathione Sepharose affinity column followed by size exclusion chromatography (for documentation see appendix V.1). K82 is homologous to K16 of p21 Ras and mutations of this P-loop residue have been shown in several GTPases to be deficient in GTP binding (Pitossi 1993; Praefcke 2004a; Sigal 1986). Irga6(K82A) was previously reported to have wild type affinity for mant-(m)GDP but its interaction with mGTPγS has not been detected (Uthaiah 2002). Unexpectedly, Irga6(K82A) had essentially wild type affinity for mGTPγS as measured in equilibrium titration (Figure 36 A, C), but with an unusually weak signal on mant-nucleotide binding (red arrow in Figure 36 A). The increase in the intensity of the fluorescence signal from the mant group normally observed upon binding of labelled nucleotides to GTPases is due to reduced quenching by exclusion of solvent from the binding grove (Rojas 2003). Thus, a decreased change in fluorescence presumably reflects some restructuring of the nucleotide-binding site leading to a more open conformation. As previously reported, the affinity of Irga6(K82A) for mGDP was also wild-type (Figure 36 B, C; (Uthaiah 2002)). Unlike wild-type protein, however, Irga6(K82A) had no detectable GTPase activity over a large protein concentration range (Figure 36 E-F) while wt Irga6 showed cooperative hydrolysis of αP^{32}-GTP as documented before (Figure 36 F; (Uthaiah 2003)). These surprising results implicate Irga6 lysine 82 not in nucleotide binding as in H-Ras (Sigal 1986) but rather in nucleotide hydrolysis alone. Wild-type Irga6 rapidly forms oligomers *in vitro* on addition of GTP that resolve as GTP is hydrolysed. Oligomerisation can be monitored by conventional (LS) and dynamic light scattering (DLS) (Figure 36 G-H) (Uthaiah 2003). Irga6(K82A), in contrast, slowly and continuously formed GTP-dependent oligomers that, consistent with its failure to hydrolyse GTP, did not resolve over the time-course of the experiment. Thus Irga6(K82A) is biochemically constitutively active, locked into the GTP-bound state.

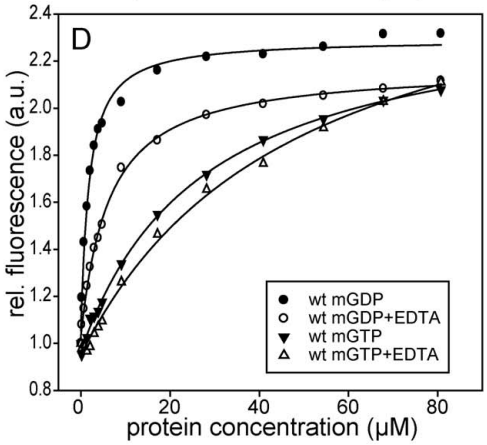
Figure 36 Biochemical properties of Irga6 wt, Irga6(K82A) and Irga6(S83N) (see next page). (A-D) Nucleotide affinities of wt Irga6, Irga6(K82A) and Irga6(S83N) measured by equilibrium titration with 0.5 µM mant nucleotides and protein concentrations of 0-100 µM. (A) Equilibrium titration with mGTPγS. (B) Equilibrium titration with mGDP. (C) Average dissociation constants (K_d) from 2 independent experiments. (D) Effect of magnesium ion depletion with 10 mM EDTA on the nucleotide affinities of wt Irga6 protein. (E) Kinetics of α^{32}P-GTP hydrolysis by Irga6 and Irga6 mutants (80 µM) displayed as ratio of the GTP concentration to the starting concentration (10 mM) against time. Similar results were obtained with 1 mM GTP and 50 µM protein (data not shown). (F) Specific GTP hydrolysis activity of Irga6wt and Irga6(K82A) measured over 30 min in the presence of 10 mM GTP and 20-160 µM protein (for wt see also (Uthaiah 2003) (G-H) Nucleotide dependent oligomerisation of 80 µM Irga6 proteins in the presence of 10 mM nucleotide measured by conventional light scattering at 350 nm (F) and by dynamic light scattering at 650 nm (G). (a.u.: arbitrary units).

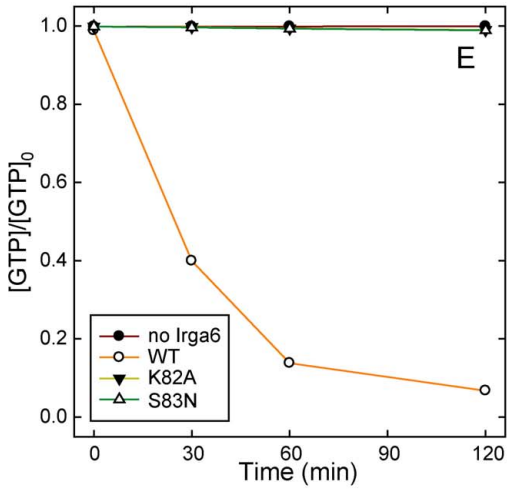


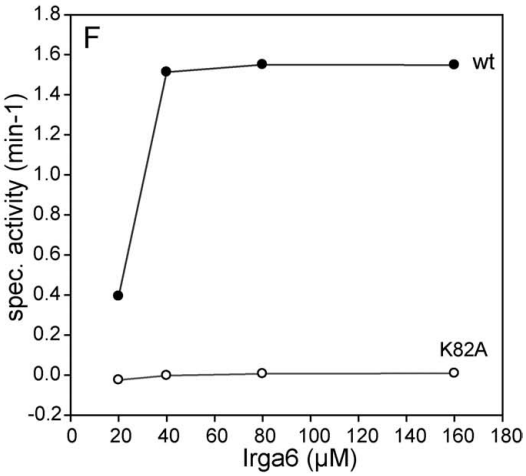


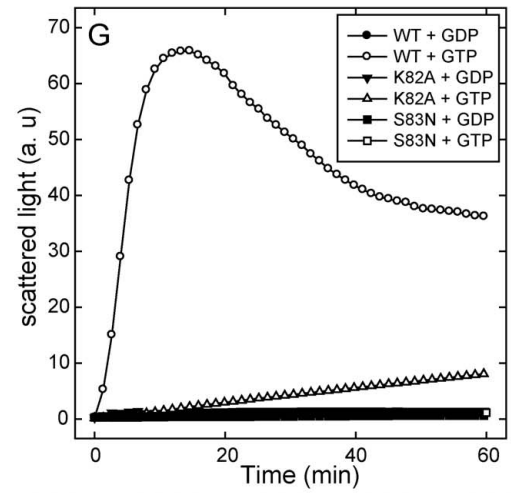
C

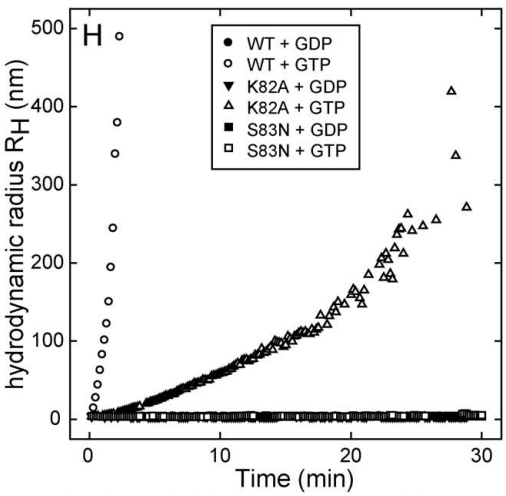
	mean K_d (µM)	
Irga6/Nucleotide	mGDP	mGTPγS
Wild-type	4.3 +/-2.7	24.8 +/- 2.1
K82A	4.0 +/- 2.2	11.7 +/- 2.1
S83N	90 +/- 40	200 +/- 90
Wild-type+EDTA	5.9 +/- 0.3	47.6 +/- 6.3











The mutation S17N of p21 Ras retains normal affinity for GDP but is unable to bind GTP, and is therefore dominant negative, locked in the inactive state (Feig 1988). The homologous mutation of Irga6, S83N, generated in this study, however, had greatly reduced binding affinities for both nucleotides (Figure 36 A-C), did not hydrolyse αP^{32}-GTP (Figure 36 E), and did not form GTP-dependent oligomers (Figure 36 G-H). The hydrodynamic radius of Irga6(S83N) measured in DLS was consistent with a monomer. Irga6(S83N) therefore provided a negative control protein for the documentation of nucleotide-dependent processes mediated by Irga6.

Binding of both GDP and GTP is strongly dependent on the presence of Mg^{2+} for H-Ras (Feuerstein 1987; Hall 1986; John 1988; Tucker 1986) and other small GTPases such as RalA (Frech 1990) and nucleotide dependent processes can therefore be easily inhibited *in vitro*. Thus, the influence of Mg^{2+} depletion by addition of EDTA on nucleotide binding of Irga6 was determined *in vitro.* Surprisingly the binding affinities for both mGDP and mGTPγS were only marginally lower in presence than in the absence of EDTA (Figure 36 C-D), hence divalent cations are dispensable for nucleotide binding by Irga6. A thorough literature search revealed that numerous other GTPases like SRP (Shan 2005), ARL3 (Hillig 2000) and the Rho family members Cdc42, Rac1 and RhoA (Zhang 2000) have also been reported to bind nucleotide in absence of magnesium ions indicating that *a priori* assumptions for the behaviour of GTPases on the basis of H-Ras data should be considered with caution.

III.2.5. Influence of nucleotide binding on the subcellular localisation of Irga6

Having characterised the Irga6 G1 mutants K82A and S83N as biochemically dominant negative and inactive respectively, these mutants were employed to confirm that aggregated wild-type Irga6 in cells not treated with IFNγ was trapped in the GTP-bound state. Like wt protein, Irga6(K82A)-ctag1 expressed in uninduced murine fibroblasts by transient transfection formed cytoplasmic aggregates (Figure 37 A) but unlike endogenous and Mifepristone-induced wild type protein (Figure 33 A, B) as well as transiently transfected Irga6-ctag1 (data not shown), this mutant made similar aggregates in IFNγ-induced cells (Figure 37 B). A similar anomalous localisation of transfected Irga6(K82A) was reported earlier in interferon-induced astrocytes (Martens 2005). Furthermore, the endogenous Irga6 apparently co-localised in aggregates with the transfected Irga6(K82A) protein (Figure 37 C-D). These results suggested that cytoplasmic aggregation is a property of GTP-bound Irga6, constitutively in the case of the constitutively active protein, and, in the absence of IFNγ, also for the wild type. Furthermore it appeared that the constitutively active mutant could capture wild-type Irga6, presumably also in the GTP-bound form, in mixed aggregates. The significance of nucleotide binding for the formation of Irga6 aggregates was indicated by the failure of transfected, ctag1-tagged Irga6(S83N) to form aggregates either in IFN-induced or uninduced cells (Figure 37 E-H). Irga6(S83N)-ctag1 protein was distributed smoothly on cytoplasmic membranes.

These results suggest that correct positioning of Irga6 to the ER is prohibited by GTP binding. When GTP binding occurs but hydrolysis is inhibited, as in the K82A

mutant, Irga6 delocalises and tends to aggregate. However, since the wild-type protein in interferon-induced cells localises correctly to the ER, but forms aggregates when expressed alone in cells, either following transfection or in the Mifepristone-inducible cells, the binding or hydrolysis of GTP in the natural situation must be regulated by other interferon-inducible factors. This conclusion was supported by the largely restored localisation of Irga6 observed in gs3T3-Irga6 cells induced simultaneously with Mifepristone and IFNγ (Figure 33 C, (Figure 35 M-R). It is, however, interesting that, despite the obvious normalisation of Irga6 in most doubly induced cells, the correct localisation was not completely restored in all cells. This could suggest that the control of Irga6 localisation in the interferon-induced cells is nearly limiting, and that the second "aliquot" of Irga6 expression due to the Mifepristone induction exceeds the capacity of the regulatory system. The behaviour of Irga6(K82A), forming aggregates whether the cell is induced with IFNγ or not, suggests that this mutant protein cannot be controlled by the regulatory system.

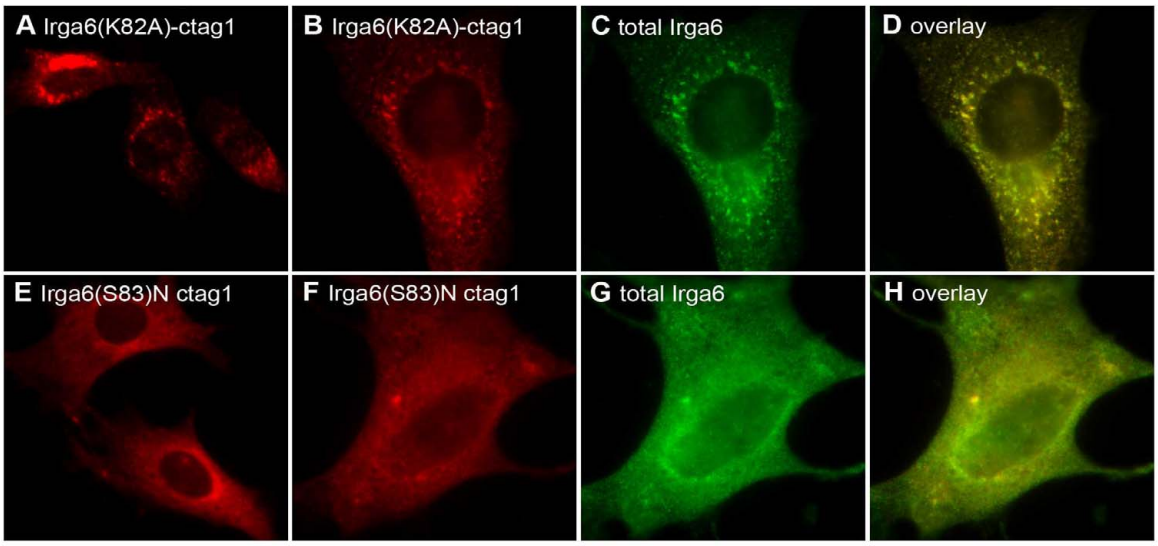


Figure 37 Subcellular localisation of Irga6 is dependent on nucleotide binding. (A-D) Fibroblasts transiently transfected with "dominant negative" Irga6(K82A)-ctag1. (A) Gs3T3 cells not induced with IFNγ show typical cytoplasmic aggregates. (B-D) wild type MEFs induced with IFNγ. Irga6 is aggregated in IFNγ-induced cells transfected with Irga6(K82A). (E-H) Fibroblasts transiently transfected with inactive Irga6(S83N)-ctag1. (E) Gs3T3 cells not induced with IFNγ show essentially wild type, dispersed distribution. (F-H) wild type MEFs induced with IFNγ. Irga6 is not aggregated in cells transfected with Irga6(S83N). Transfected proteins were detected with anti-ctag1 antiserum 2600 (red in A, B, E, F) and total Irga6 with 10E7 monoclonal antibody (green in C, D).

III.2.6. Influence of IFNγ and nucleotide on the subcellular localisation of Irgb6

Subcellular localisation of Irgb6 in IFNγ and Mifepristone-induced gs3T3-Irgb6 cells could not be determined due to the low expression level following Mifepristone induction and the fact that the goat anti-Irgb6 antiserum, A20, detected interferon-induced cellular Irgb6 inefficiently in immunofluorescence (Figure 38 B-D). So an Irgb6 construct with a C-terminal FLAG-tag was used for subsequent experimentation. This tag did not interfere with the properties of Irgb6 in any of the performed analyses (Figure 38 see below, Figure 51).

Wild-type Irgb6, and Irgb6 with a C-terminal FLAG tag, transiently transfected into non-induced fibroblasts were both even more strikingly mislocalised than Irga6, forming long fibre-like cords that surrounded the nucleus and appeared to entangle other cellular organelles (Figure 38 A). Irgb6-FLAG transfected into IFNγ-induced gs3T3 fibroblasts showed a fine reticular pattern with a few small residual structures (Figure 38 B-D). The striking fibre-like cords associated with transfection of Irgb6 into unstimulated cells were absent. Unlike the endogenous Irgb6 and the redistributed transfected protein, the residual aggregates as well as the severe aggregates observed in absence of IFN could be visualised by use of the A20 antiserum (Figure 38 B, C). Thus, like for Irga6 subcellular localisation of Irgb6 was dependent on other IFN-inducible factors. The reticular distribution seen in IFNγ-induced cells transiently transfected with Irgb6-FLAG presumably reflects the subcellular localisation of endogenous Irgb6, which is in accordance with the majority of the protein being soluble in hypotonic lysates (Martens 2004b). In absence of any other conspicuous signal in immunofluorescence, it is likely that the membrane-associated part of Irgb6 protein (20-30%) localise to the ER.

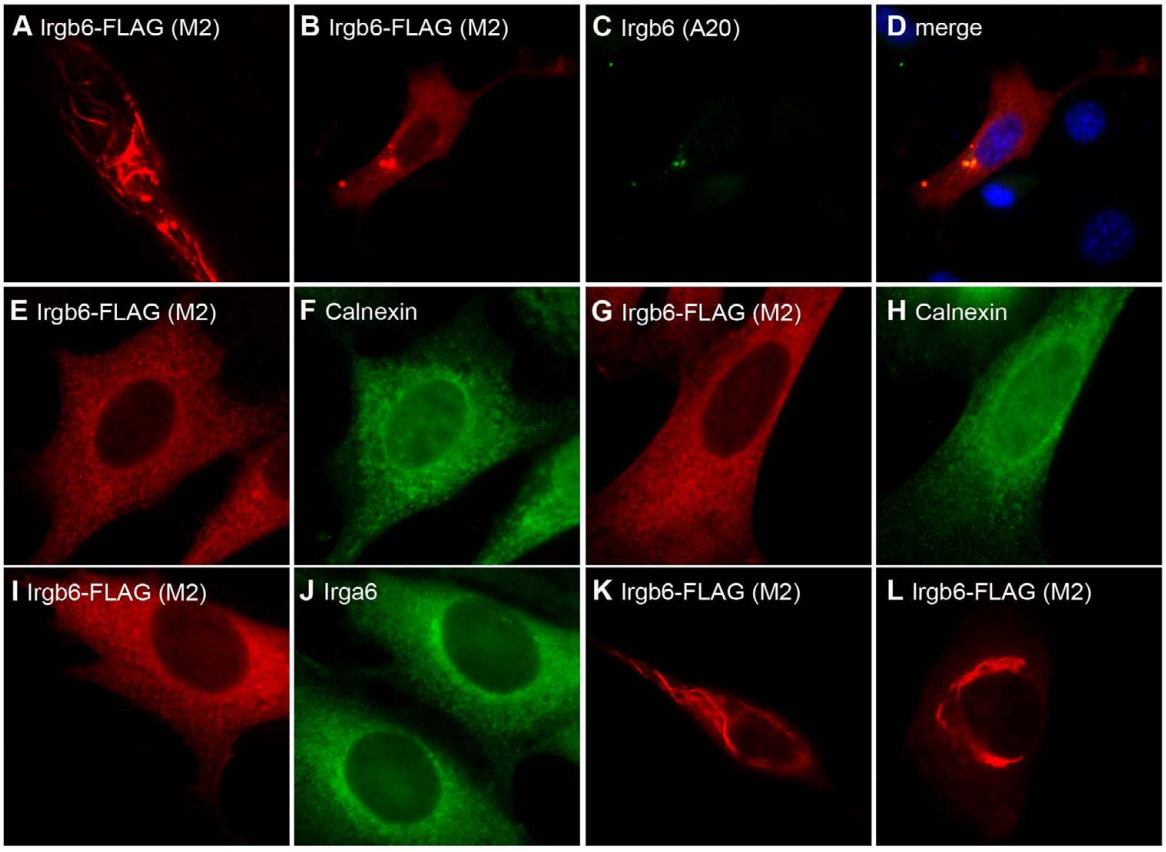


Figure 38 Subcellular localisation of Irgb6 is dependent of IFNγ and nucleotide binding. Gs3T3 cells transiently transfected with FLAG-tagged Irgb6 wt and mutant expression constructs (pGW1H, 1 µg). Transfected proteins were detected with the anti-FLAG monoclonal antibody M2 (red, A-B, D, E, G, I, K-L). (A-D) Irga6 wt-FLAG transfected into (A) uninduced and (B-D) IFNγ-induced cells (200U/ml). (C) Irgb6 specific A20 antiserum (green) does not detect distributed Irgb6 protein efficiently. (D) Overlay of B and D with nuclear counterstain (DAPI). (E-J) Irga6(S70N)-FLAG transfected into (E-F) uninduced and (G-J) IFN-induced cells. (F, H) Calnexin staining with SPA-865 serum. (J) Irga6 detected with 165 serum. (K-L) Irgb6(K69A)-FLAG transfected into (K) uninduced and (L) IFN-induced cells.

As with Irga6, the formation of aggregates by Irgb6 in absence of IFNγ was dependent on the integrity of the GTP-binding site and hence nucleotide binding. The G1

motif mutant Irgb6(S70N) like the corresponding nucleotide binding-deficient Irga6 mutant, S83N, displayed wild-type localisation, apparently overlapping with Calnexin, both in IFNγ-induced (Figure 38 G-H) and uninduced cells (Figure 38 E-F) and with Irga6 in induced cells (Figure 38 I-J). The mutant Irgb6(K69A), like the corresponding nucleotide-binding but hydrolysis-deficient Irga6 mutant, K82A, formed typical aggregates in the absence and presence of IFNγ (Figure 38 K-L).

III.2.7. Behaviour of IRG proteins in size exclusion chromatography

Since the IRG proteins displayed such a distinct behaviour in cells their behaviour in analytical size exclusion chromatography was determined. Furthermore, it was analysed whether the formation of Irga6 and Irgb6 aggregates in absence of IFNγ was reflected in a shift of the proteins to higher molecular weight fractions *ex vivo*. Lysates of IFNγ- or Mifepristone-induced gs3T3 and gs3T3-IRG cells were separated over a Superose6 HR column and fractions were collected and analysed for the IRG proteins in SDS-PAGE followed by Western blot with specific immunoreagents (Figure 39). The analysed members of the IRG family displayed a distinct running behaviour following IFNγ induction in analytical gel filtration. Irgb6 and Irgd signals were strongest at a size that roughly equals a monomer (Figure 39 C+F), while all the IRGM proteins migrated at sizes consistent with a dimer or trimer: 70 kDa for Irgm1 (monomer 37 kDa in WB, Figure 39 G), 100 kDa for Irgm2 (Figure 39 H, Mifepristone-induced) and 150 kDa for Irgm3 (Figure 39 I). Irgm2 was analysed only following Mifepristone-induction, as the polyclonal H53 serum did not recognise endogenous, IFN-induced Irgm2 in NIH3T3/gs3T3 cells, probably due to a polymorphism present in different mouse strains (H53 was generated against a C-terminal peptide from C57BL/6 Irgm2, see material and methods II.1.8). Heteromeric interactions cannot formally be excluded for the IRGM proteins, though at least for Irgm2 IFN-induced proteins are not involved. Irga6 trailed over a large range from the size of a monomer to the size of a tetramer (50-200 kDa) (Figure 39 A) indicating that Irga6 is present in different states in the cell.

Even though Irga6 and Irgb6 formed large, nucleotide-dependent intracellular aggregates upon expression in absence of IFNγ, Mifepristone-induced proteins migrated identical to IFNγ-induced endogenous proteins in size exclusion chromatography (Figure 39 B, D, E). The covalently modified 55 kDa form of Irga6 found in Mifepristone-induced cells also behaved like IFNγ-induced 44 kDa form of Irga6. The two analysed gs3T3-Irgb6 clones (AC3 and BB6) that behaved somewhat differently in the cell proliferation assays (partially and completely growth inhibited respectively) displayed identical elution profiles (Figure 39 D, E). This suggests that *in vivo* aggregation is largely reversible upon cell lysis and that no other IFN-inducible factors participate in forming the higher molecular weight fraction of Irga6. To exclude that the aggregated pool of Irga6 protein is actually lost by pelleting the nuclei after cell lysis with non-ionic detergents, this pellet was solubilised in boiling protein loading buffer and separated in SDS-PAGE. As seen in Figure 40, all the Mifepristone-induced Irga6 protein is found in the detergent soluble fraction and no protein was found in the pellet fraction.

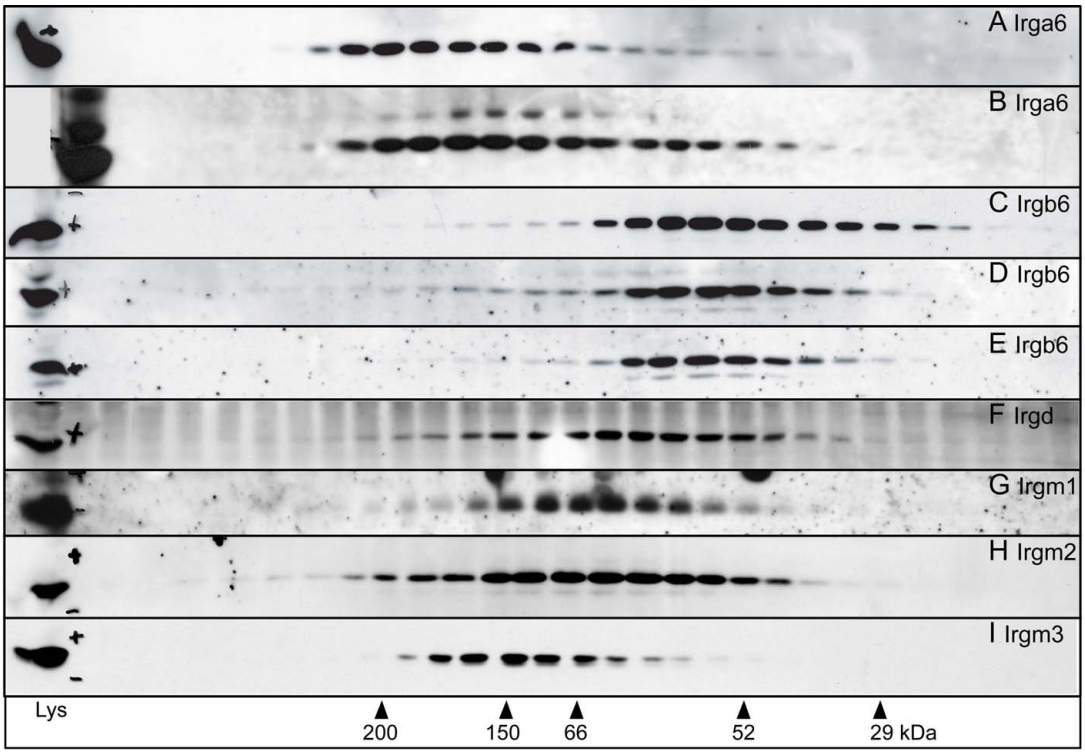


Figure 39 Behaviour of IRG proteins in analytical size exclusion chromatography. Gs3T3 cells induced with 200U/ml IFNγ or 10^{-9} M Mifepristone respectively were lysed in PBS/0.1% Thesit and the postnuclear supernatants were subjected to analytical gel filtration on a Superose 6 HR column in the same buffer. The eluted fractions were analysed by SDS-PAGE followed by Immunoblot with specific antisera ((A-B) Irga6 with 10D7, (C-E) Irgb6 with A20, (F) Irgd with 2078, (G) Irgm1 with A19, (H) Irgm2 with H53, (I) Irgm3 with αIGTP clone7). (A, C, F, G, I) IFNγ-induced gs3T3-Irga6 cells (B) Mifepristone-induced gs3T3-Irga6 cells, (D) Mifepristone-induced gs3T3-Irgb6 clone AC3 cells, (E) Mifepristone-induced gs3T3-Irgb6 clone BB6 cells. (H) Mifepristone-induced gs3T3-Irgm2 cells. The molecular mass correlating with a certain elution volume derived from calibration of the column with marker proteins (see material and methods) is indicated at the bottom of the figure.

In contrast to cellular Irga6, non-myristoylated recombinant bacterial Irga6 (Uthaiah 2002) and recombinant Irga6(G2A) from insect cells (Papic 2007) fractionated as a monomer. Myristoylated recombinant Irga6 from insect cells, however, ran at 200-250 kDa in size exclusion chromatography (Papic 2007). Thus, the high molecular weight fraction of Irga6 is dependent on the myristoyl group itself and seems to be independent of other cellular proteins. This effect is probably either due to the intrinsic properties of the lipid moiety or to myristoyl-induced conformational changes that might enhance homomeric interaction.

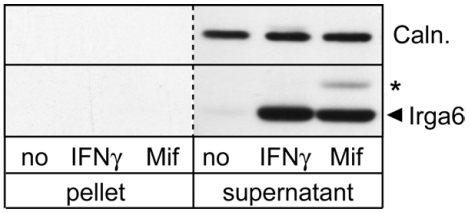


Figure 40 Irga6 protein from Mifepristone and IFNγ-induced cells is completely detergent soluble. Untreated, IFNγ- and Mifepristone-induced gs3T3-Irga6 cells were lysed in PBS/1% Triton X-100. Unsolubilised material pelleted by centrifugation and boiled in protein loading buffer (50 mM Tris/HCl pH 6.8/0.7 % β-ME/1% SDS/5 % glycerol/0.0025 % (w/v) bromephenol blue) as well as the postnuclear supernatants were subjected to SDS-PAGE followed by immunoblot with the anti-Irga6 serum 165 and calnexin antiserum SPA-865. Both IFN- and Mifepristone-induced Irga6 was completely detergent soluble. Similar results were obtained with 0.1% Thesit and fibroblasts transiently transfected with expression constructs for Irga6 (Papic unpublished data).

III.2.8. Regulation of Irga6 and Irgb6 positioning by the GMS proteins

In view of the complex representation of the IRG proteins in the mouse and their extensive functional non-redundancy, it was conceivable that IRG proteins themselves are the IFNγ-dependent regulators of Irga6 and Irgb6 and interact with each other to maintain some kind of functional equilibrium in the interferon-induced cell, reflected in the correct ER localisation of these two proteins. To examine this hypothesis directly, the 5 other IRG GTPases Irgm1, Irgm2, Irgm3, Irgd and Irgb6 were transfected transiently into the Mifepristone-inducible gs3T3-Irga6 fibroblasts that normally express mislocalised Irga6. After Mifepristone induction, 84% of the cells expressing the other IRG proteins showed wild type Irga6 localisation, while 87 % of the control cells transfected with EGFP alone showed typical Irga6 aggregates (Figure 41 A). The effect of the transfected IRG GTPases on Irga6 localisation correlated positively with the DNA amount used for transfection (compare Figure 41 B with D). Not all 5 IRG members were necessary to correct the localisation of Irga6. In single transfections, the GKS subfamily members Irgd and Irgb6 did not promote the wild type localisation of Irga6, while the three GMS proteins Irgm1, Irgm2 and Irgm3 all did so (Figure 41 B, D). Transfection of Irgb6 into Mifepristone-induced gs3T3-Irga6 cells caused a further maldistribution of Irga6 (Figure 41 B). Irgb6 was itself massively mislocalised in these transfectants (see above). Single IRGM GTPases mediated a less pronounced effect than transfection of all 5 p47 GTPases but all were clearly active (Figure 41 B, D). The deficient normalising activity of single IRGM proteins could not be compensated by an increase in the amount of DNA transfected (compare Figure 41 B with D). However, simultaneous transient transfection of 3 GMS GTPases into Mifepristone-induced gs3T3-Irga6 fibroblasts restored wild type Irga6 localisation as efficiently as co-transfection with all 5 additional GTPases (Figure 41 A, C and Figure 42 A-C). Hence, intracellular localisation of Irga6 is indeed regulated by other IRG proteins, and the regulatory function is specifically a property of the three IRGM proteins, working in concert and independently of Irgb6 and Irgd. The ability of the IRGM GTPases to rescue Irga6 localisation was dependent on an intact G1 motif. When the three IRGM proteins carrying the GMS to GMN mutation in the G1 motif of the nucleotide binding site or EGFP were co-transfected into Mifepristone-induced gs3T3-Irga6 cells no normalisation of Irga6 localisation was seen (Figure 41 A, C, Figure 42 D-F). An intact G1 motif was required in all three IRGM proteins for full normalisation of Irga6 localisation but individual GMN mutants had no dominant negative effect on the normalisation activity of co-transfected wild-type IRGM proteins (data not shown).

In transient transfection of the 5 IRG proteins Irgm1-3, Irgd and Irgb6 into Mifepristone-induced gs3T3-Irga6 fibroblasts, it was already observed that not only that Irga6 aggregates were resolved, but also that the transfected Irgb6 was localised in a distributed, smooth, cytoplasmic pattern. No cell with Irgb6 fibres was detected in any of these experiments. Hence, it was analysed whether the regulatory activity of the IRGM proteins could be generalised on Irgb6 as well as Irga6. Aggregate formation in uninduced cells by transfected, FLAG-tagged wild type Irgb6 was eliminated by co-transfection with the 3 GMS proteins (Figure 42 G-I, Figure 43 A) but not by co-transfection of the 3 GMN mutants or EGFP (Figure 42 J-L, Figure 43 A). As shown above for Irga6, GKS proteins (Irgd and Irga6) had no beneficial effect (data not shown)

whereas Irgm2 and Irgm3 both partially restored Irgb6 localisation when co-transfected one at a time with Irgb6, however Irgm1 alone was ineffective for Irgb6 (Figure 41 B). Transfection of Irgm2 and Irgm3 together could not normalise Irgb6 localisation completely, and, as with Irga6, only co-transfection of all 3 IRGM proteins was fully effective (Figure 41 B). The situation with Irgb6 is thus very similar to that with Irga6. Both proteins form distinctive aggregates and mislocalise in the absence of other interferon-induced proteins. This mislocalisation is dependent on the integrity of the GTP binding site. The mislocalisation can be completely prevented if the 3 IRGM proteins are co-expressed, and this normalising activity is in turn dependent on the integrity of the GTP binding sites of the IRGM proteins.

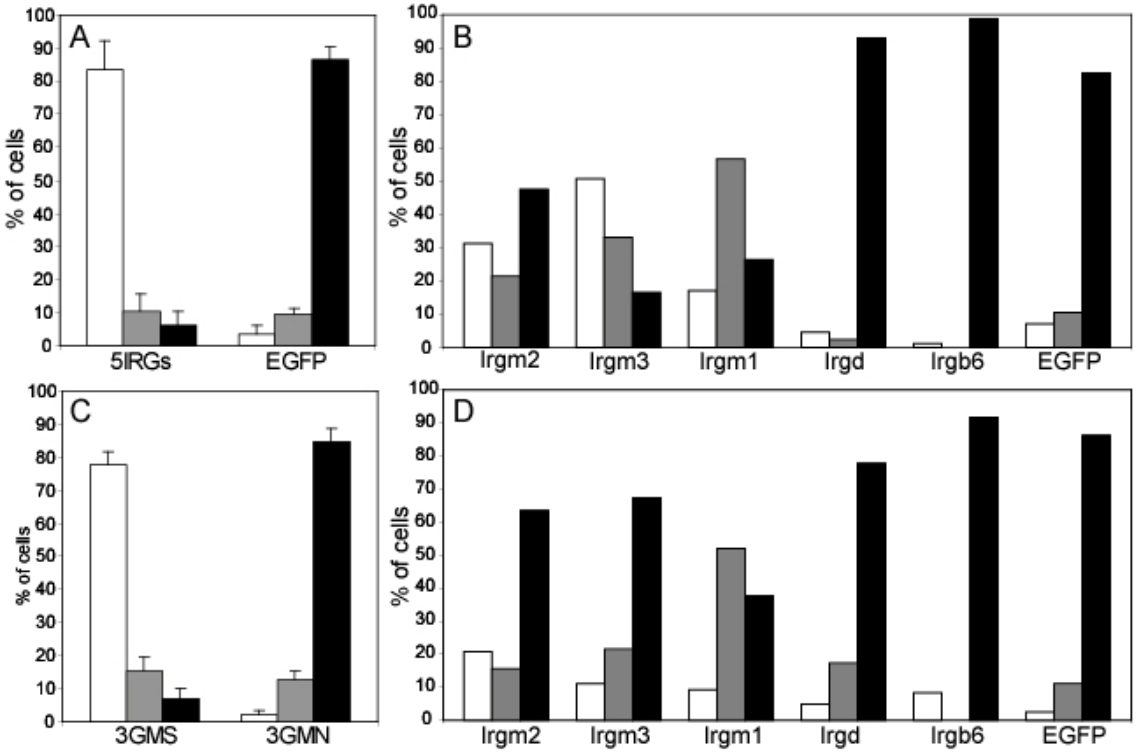


Figure 41 Quantification of the effect of other IRG proteins on Irga6 localisation. Irga6 localisation was assayed in Mifepristone-induced gs3T3-Irga6 cells transiently transfected with the indicated (pGW1H) expression constructs either individually or in pools (see also Figure 42). Effect of presence of GMS proteins on resting localisation of Irga6, recorded as smooth reticular (wild type) staining (white), partly aggregated (grey) and strongly aggregated (black). 150 cells were counted blind per data point. (A) Cells transfected with a pool of (pGW1H) expression constructs of Irgm1, Irgm2, Irgm3, Irgb6 and Irgd (5IRGs; 400 ng each, total 2 µg) were compared with cells transfected with an EGFP expression vector (2 µg). (B, D) Cells were transfected with (pGW1H) expression constructs encoding single IRG proteins: (B) 2 µg each, (D) 400 ng each. Each IRGM protein alone showed a significant dose-dependent beneficial effect on Irga6 localisation, but this was in all cases incomplete. Similar effects were also seen with cells transfected with constructs encoding pairs of GMS proteins (data not shown). Irgd was inactive. Cells transfected with the Irgb6 expression construct showed striking Irgb6 aggregates, which largely colocalised with aggregated Irga6 (see also Figure 38). (C) Cells transfected with a pool of (pGW1H) expression constructs of the three wild type GMS proteins were compared with cells transfected with a similar pool of the three inactive GMN mutants (667 ng each).

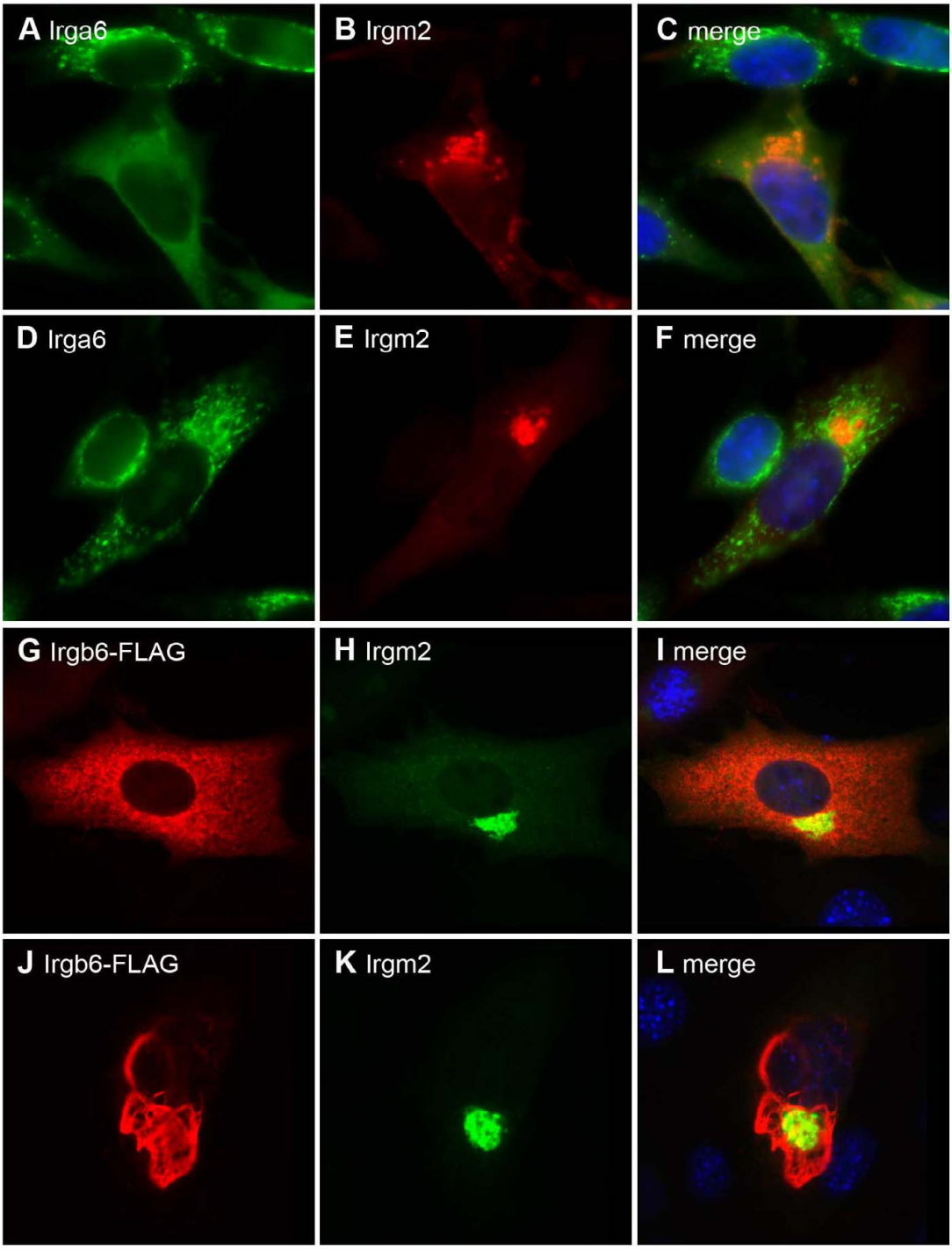


Figure 42 Subcellular localisation of Irga6 and Irgb6 are dependent on the 3 wild-type GMS proteins. (A-F) Gs3T3-Irga6 cells were induced for 24 hrs with 10^{-9} M Mifepristone and transfected transiently with other IRG expression constructs (pGW1H). Irga6 (green) was identified in immunofluorescence by 10E7 monoclonal antibody and Irgm2 (red) with H53 antiserum. In (C, F) nuclei are counterstained with DAPI (blue). (A-C) Cells transiently transfected with 3 expression constructs encoding Irgm1, Irgm2 and Irgm3 (667 ng DNA each). Note the diffuse localisation of Irga6 only in cells transfected with the 3 functional GMS proteins. (D-F) Cells transiently transfected with 3 expression constructs encoding Irgm1(S90N), Irgm2(S78N) and Irgm3(S98N) (667 ng DNA each). (G-L) Uninduced gs3T3 wt cells transiently co-transfected with (pGW1H) expression constructs for FLAG-tagged Irgb6 and either 3 GMS wt (G-I) or GMN mutant (J-L) proteins (0.5 µg each). Irgb6-FLAG was detected with the FLAG-specific monoclonal antibody M2 (red) and Irgm2 (green) was detected with the rabbit serum H53 (green). In (I, L) nuclei are counterstained with DAPI (blue).

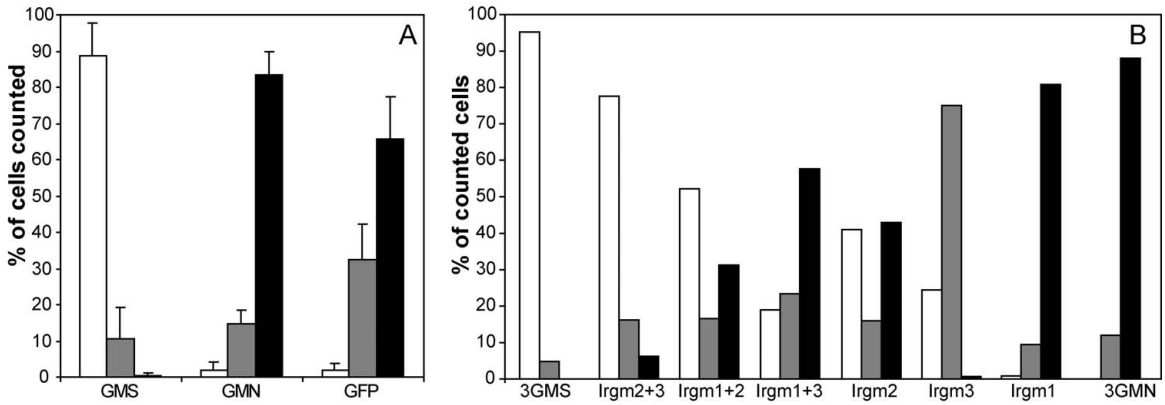


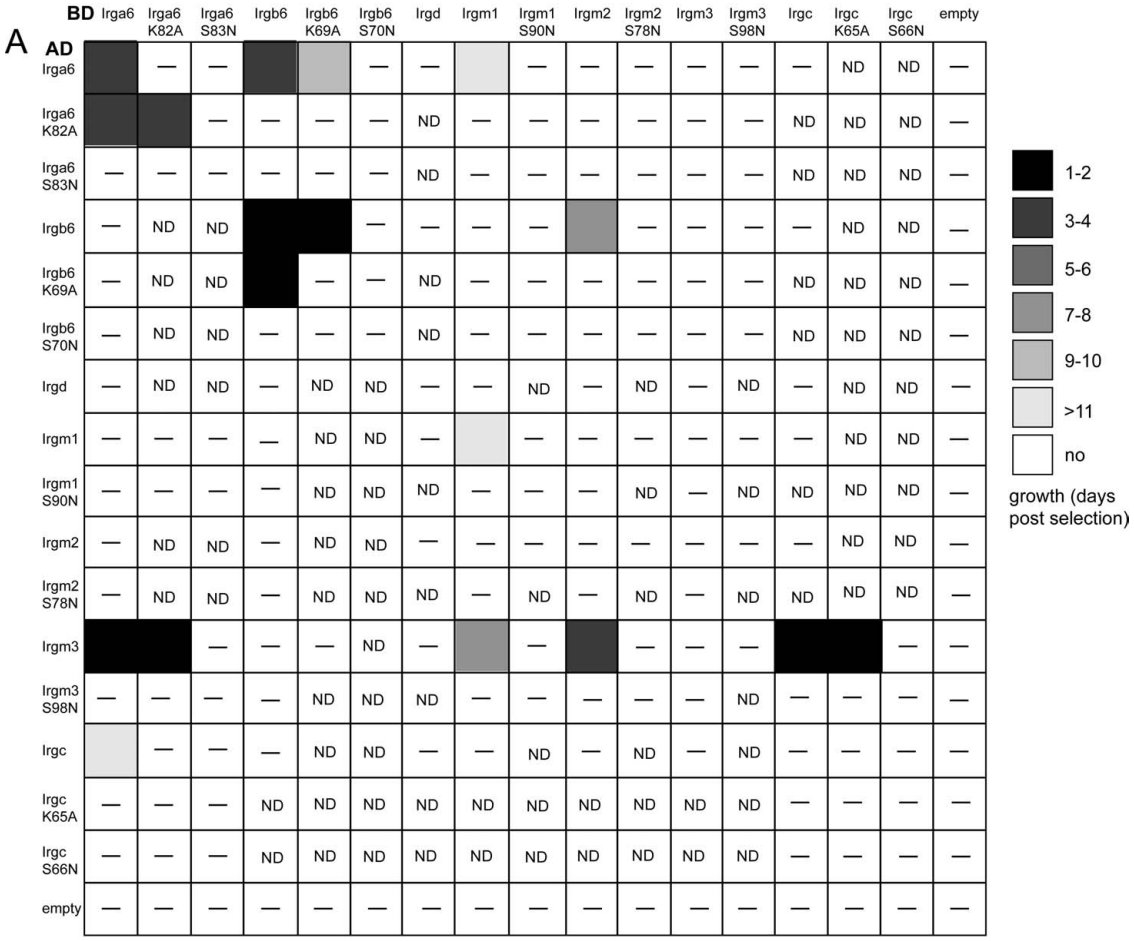
Figure 43 Quantification of the effect of GMS proteins on Irgb6 localisation. Quantification of Irgb6 localisation in gs3T3 wt cells transiently transfected with an Irgb6-FLAG expression construct (0.5 µg) recorded as smooth reticular (wild type) staining (white), partly aggregated (grey) and strongly aggregated (black). 150 cells were counted blind per data point. (A) Cotransfection of the 3 GMS proteins (0.5 µg each), the 3GMN proteins (0.5 µg each) or EGFP (1.5 µg) respectively (see also Figure 42 G-L) (B) Cotransfection of single GMS protein expression constructs compared to cotransfection with different combinations of these constructs (0.5 µg per construct, adjusted to 1.5 µg of DNA with empty vector).

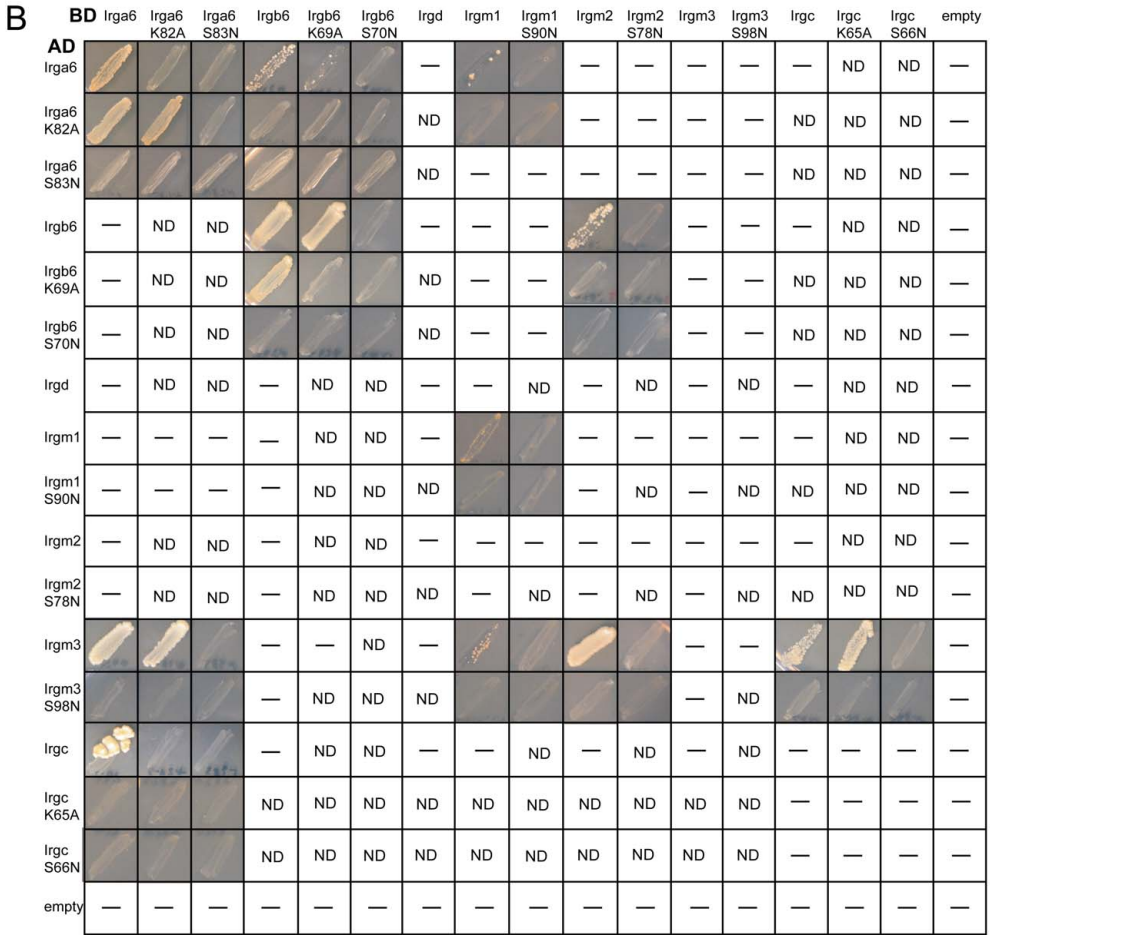
III.2.9. Nucleotide-dependent direct interactions of IRG proteins in Y2H

To determine whether nucleotide-dependent regulation of Irga6 by the IRGM proteins involved direct interaction between the proteins, a yeast two-hybrid (Y2H) assay was performed. The complete coding regions of the IRG proteins studied were expressed as N-terminal fusions with the Gal4 DNA-binding (BD) and Gal4 activation domain (AD) in yeast reporter strains (James 1996). Protein-protein interactions were detected by growth on selective media (see Materials and Methods for details). The IRG proteins showed multiple interactions in this assay system with each family member displaying a distinct pattern of behaviour (Figure 44).

The strong homotypic interaction of Irga6 observed in the Y2H system reproduced the *in vitro* oligomerisation of Irga6 in the presence of GTP (Figure 36, (Uthaiah 2003)) strengthening the potential *in vivo* relevance of this interaction. Very strong homotypic interaction was also observed for Irgb6. Furthermore, Irgb6 interacted strongly with Irga6 and more weakly with Irgm2. As anticipated, Irga6 displayed a very strong interaction with Irgm3 and a rather weak one with Irgm1. Irgm3 interacted with Irgm2 and Irgm1 and Irgm1 displayed a weak homotypic interaction. Furthermore, Irgc interacted with Irgm3 and Irga6, though the *in vivo* relevance of these interactions for Irgc function in testis remains elusive. As frequently seen in Y2H, some of the observed positive interactions were unidirectional, probably as a result of the different structural impacts of the AD and BD on the fused protein (Estojak 1995; Van Criekinge 1999).

Figure 44 Nucleotide-dependent interactions of IRG proteins in yeast two hybrid (see next page). Interaction behaviour of full-length IRG proteins and mutants observed in a Gal4 based Y2H assay. Interactions were measured by growth on selective medium (synthetic defined medium lacking leucine, tryptophan, histidine and adenine). (A) Graphic representation of the strengths of the interactions indicated in terms of intensity of shades of grey following the key given next to the table. (B) Photographic documentation of the observed interactions and the respective mutants controls. ND: not determined. -: no interaction.


A
BD
AD
Irga6
Irga6 K82A
Irga6 S83N
Irgb6
Irgb6 K69A
Irgb6 S70N
Irgd
Irgm1
Irgm1 S90N
Irgm2
Irgm2 S78N
Irgm3
Irgm3 S98N
Irgc
Irgc K65A
Irgc S66N
empty
1-2
3-4
5-6
7-8
9-10
>11
no growth (days post selection)


B
BD
AD
Irga6
Irga6 K82A
Irga6 S83N
Irgb6
Irgb6 K69A
Irgb6 S70N
Irgd
Irgm1
Irgm1 S90N
Irgm2
Irgm2 S78N
Irgm3
Irgm3 S98N
Irgc
Irgc K65A
Irgc S66N
empty

To confirm the specificity of the interactions seen and their relevance to the *in vivo* findings, the Y2H assay was repeated with nucleotide binding site mutants of the individual proteins. All Y2H interactions of GMS proteins with other family members were completely abolished by the GMS to GMN mutants, Irgm3(S98N), Irgm2(S78N) and Irgm1(S90N) (Figure 44). Likewise, Irga6(S83N) was also completely negative for all interactions shown by wild-type Irga6 (Figure 44). These results are consistent with the essentially passive behaviour displayed by Irga6(S83N) and the GMN mutants in fibroblasts (see above III.2.5 and III.2.8). The corresponding mutants of Irgb6(S70N) and Irgc(S66N) similarly failed to show any of the interactions observed with the respective wt proteins. The constitutively active mutant Irga6(K82A) retained many interactions in Y2H (Figure 44) but lost bi-directionality in its interaction with wt Irga6 and interaction with Irgb6 and Irgc were completely abolished. This observation is consistent with the slightly altered conformation of the mutant protein suggested by the anomalous behaviour upon mGTPγS binding (see above III.2.4). The corresponding mutant of Irgb6, K69A, preserved its interactions with wt Irgb6 and Irga6, though the latter was weakened significantly, and the interactions with the GMS proteins Irgm2 and Irgm3 were lost (Figure 44). Similarly, Irgc(K65A) retained its interaction with Irgm3 but failed to bind to Irga6. Thus, the multiple interaction between IRG proteins observed in Y2H were all sensitive to mutation of the G1 motif in either of the two interaction partners and hence presumably nucleotide dependent.

An independent Y2H study based on N-terminal fusions with the LexA DNA-binding domain and B42 activation domain also documented multiple interactions within the IRG family though nucleotide dependence was not analysed (Kaiser 2005). In contrast to the Gal4 system used in this study, which is based on low, constitutive expression, the LexA system employs high, inducible expression of the constructs. As anticipated from the different properties possessed by the two assay systems (Van Criekinge 1999) they highlighted distinct interactions, though the observed interaction profile is partially overlapping. Interestingly, both Irgb10, which was not tested in this study, and Irgd, which did not show any detectable interaction in the Gal4 system, interacted with other IRG family members in the LexA Y2H assay (Kaiser 2005). Since the AD- and BD-Irgd proteins were expressed in yeast cells as shown in Western blot (data not shown) it is likely that the absence of any interaction of these proteins with other IRG family members in the Gal4 based Y2H assay is due to structural constraints imposed by the fused Gal4 domains.

III.2.10. Direct GDP-dependent interaction of cellular Irgm3 with Irga6

The results shown above document that functionally significant nucleotide-dependent interactions between Irga6 and the GMS proteins occur in IFNγ-induced cells (III.2.8), and the yeast 2-hybrid data suggested that at least some of these interactions were likely to be direct (III.2.9). The strongest interaction of Irga6 with a GMS protein in yeast 2-hybrid was with Irgm3 and both proteins localise to the ER membrane in IFN-stimulated cells. Accordingly, co-precipitation between Irga6 and Irgm3 from IFN-stimulated cells was examined employing the Irga6-specific antiserum 165 coupled to protein A Sepharose beads. Clear but weak co-precipitation of Irgm3 with Irga6 was

documented in the absence of exogenous nucleotide (Figure 45). In the presence of exogenous GDP, the signal was noticeably weakened. The signal observed for Irgm3 was markedly sub-stoichiometric compared to the amount of immunoprecipitated Irga6. No signal was detected from lysates of uninduced cells. Addition of exogenous GTPγS, which favours Irga6 homomeric interaction (see above III.2.4 and (Papic 2007), completely abolished interaction of Irga6 with Irgm3. Together with the observation that both Irga6(S83N) and the Irgm3(S98N) were inert both in Y2H and in fibroblasts, this suggests that the Irga6-Irgm3 interaction requires GDP bound to both interaction partners. Thus, in absence of exogenous nucleotide GDP is presumably trapped in the co-precipitated complex preformed inside the cell. Depletion of divalent cations by addition of EDTA did not prevent Irga6-Irgm3 interaction, suggesting that nucleotide binding for Irgm3, like for Irga6 (see above, Figure 36), is largely independent of the presence of Mg^{2+}. A possible explanation for the noticeable inhibition of the co-immunoprecipitation of Irgm3 by exogenous GDP is the complex network of (regulatory) interactions between members of the IRG family observed in murine cells and in Y2H, as the lysate from IFNγ-induced cells also contains the other IFN-inducible IRG proteins. This might also apply to the strong inhibitory effect mediated by GTPγS. Yet, it cannot formally be excluded that one of the interaction partners, presumably Irgm3, needs to be in the GTP-bound state for Irga6-Irgm3 interaction.

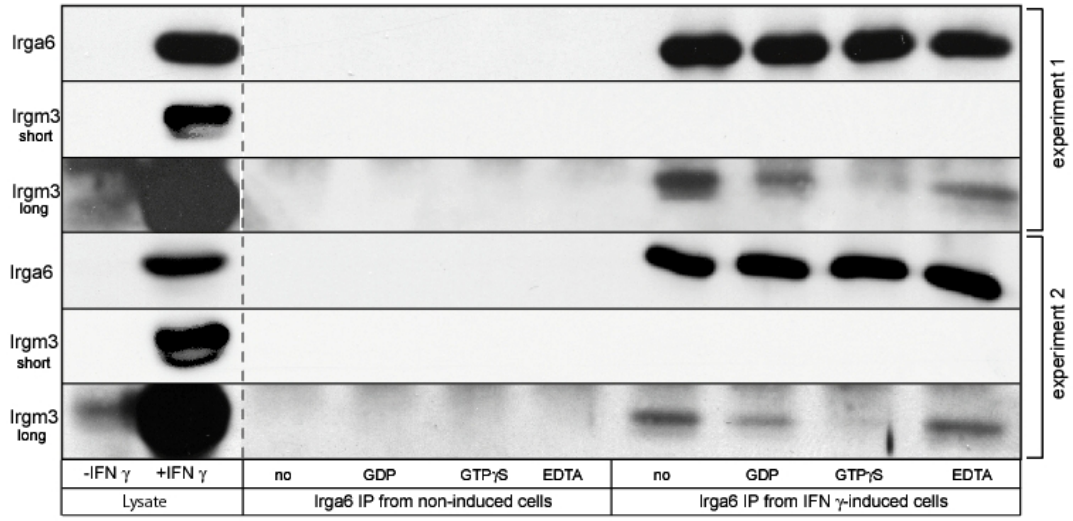


Figure 45 Nucleotide dependent interaction of Irga6 with Irgm3 from mammalian cells in co-IP. Co-immunoprecipitation of Irgm3 with Irga6. Irga6 was immunoprecipitated from IFNγ-induced gs3T3 cells lysed in the presence or absence of nucleotide (0.5 mM) using 165 serum. Irga6 and co-immunoprecipitated Irgm3 were detected in immunoblot with 10D7 and anti-IGTP (Irgm3) antibody respectively. Non-induced cells served as negative control. To visualise Irgm3 in the lysate and in the IP samples, different exposure times were used (middle row: 10 sec; bottom row: 5 min). Two independent experimental are displayed. Loading: 5% of lysates, 25% of IPs for Irga6 blot, 75% for Irgm3 blot. (The shadow above the Irgm3 band in the long exposure is derived from the heavy chain of the 165 antibody used for IP).

To confirm the results from the co-immunoprecipitation and to further clarify the nucleotide status of the observed interaction, purified bacterially expressed GST-Irga6 bound to glutathione-Sepharose beads was used to pull-down Irgm3 from lysates of interferon-treated gs3T3 fibroblasts. In the absence of nucleotide, a very weak Irgm3 signal was detected indicating a basal affinity of the two proteins for each other in absence of nucleotide. However efficient pull-down of Irgm3 followed the addition of

GDP to the binding reaction while GTPγS inhibited the interaction completely suppressing even the basal interaction seen in absence of nucleotide (Figure 46 A). Similar to the co-immunoprecipitation only a small proportion of the cellular Irgm3 protein was trapped by a vast excess of recombinant Irga6 protein. When the Irga6 GTP-binding deficient mutant, Irga6(S83N), was employed in the same assay, no increment in Irgm3 pull-down above background was seen on addition of GDP. Thus Irga6 probably interacts with Irgm3 in the IFNγ-treated cell in the GDP-bound state. As a mixture of GDP and GTPγS reflecting the intracellular concentrations of the nucleotides (100 µM GDP + 300 µM GTP; (Kleineke 1979)) completely inhibited Irgm3 precipitation, it seems unlikely that Irgm3 needs to be in the GTP-bound state in this interaction. GTP labelled with a mant fluorescent group has been shown to interfere with *in vitro* oligomerisation of recombinant Irga6 by sterically hindering the formation of the essential G-domain-G-domain interaction interface involving the bound nucleotides (Pawlowski, unpublished data). The formation of this interaction interface was crucial for substrate-aided hydrolysis of GTP by Irga6. Interestingly, addition of mantGDP to the pull down reaction similarly reduced Irga6-Irgm3 interaction to basal levels, implicating the same G-domain-G-domain interface in this interaction.

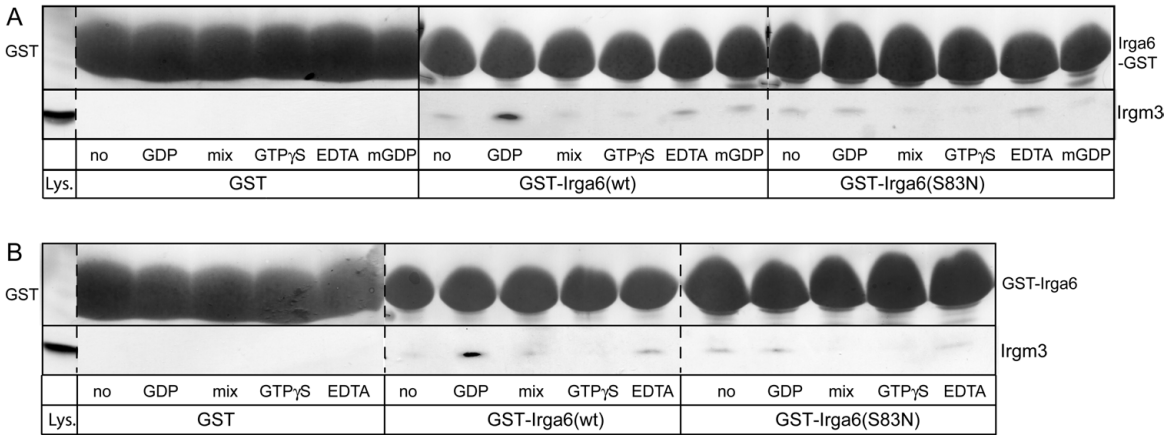


Figure 46 Nucleotide dependent interaction of Irga6 with Irgm3 from mammalian cells in pull-down. Pull-down of cellular Irgm3 with recombinant GST-Irga6 and -Irga6(S83N) protein from IFNγ-induced gs3T3 cells (A) and from Mifepristone-induced gs3T3-Irgm3 cells (B) lysed in the presence or absence of guanine nucleotides (0.5 mM GDP, GTPγS or mGDP; mix: 100 µM GDP + 300 µM GTPγS; EDTA: 10 mM). GST alone served as negative control. Recombinant fusion protein was visualised by Ponceau S staining of Western blot membranes (top row). Irgm3 was detected by immunostaining with anti-IGTP (Irgm3) antibody (bottom row). Lys: lysate control (4% of input). Dotted lines indicate positions where irrelevant lanes were removed from the image of the original gel.

That the GDP-dependent interaction between Irga6 and Irgm3 does not depend on other IFNγ-induced intermediaries was shown by a similar pull down by Irga6 from Mifepristone-induced gs3T3-Irgm3 cell lysates (Figure 46 B). Again, GDP induced a strong interaction with wild type Irga6 but not with Irga6(S83N). The identity of the nucleotide required by Irgm3 for this interaction with Irga6 is not clear, but seems likely also to be GDP in view of the high concentration (0,5 mM) of the nucleotide maintained throughout the experiment. Additionally, the failure of Irgm3(S98N) to influence Irga6 localisation in mammalian cells and to interact with Irga6 in yeast two-hybrid strongly argues against a functional interaction in absence of nucleotide. However, Taylor *et al*

reported some years ago that Irgm3 immunoprecipitated from IFNγ-induced cells predominantly co-precipitated GTP (Taylor 1997).

Having documented the direct, regulatory interactions between GKS and GMS members of the IRG family governing the nucleotide status and localisation of GKS proteins in uninfected cells, the requirement of these interactions for IRG function in resistance to *Toxoplasma gondii* were explored.

III.2.11. Regulation of Irga6 in *Toxoplasma*-infected cells

Data are accumulating that IRG GTPases cause the vesiculation and destruction of *Toxoplasma gondii* parasitophorous vacuole (PV) membranes (Ling 2006; Martens 2005). At least five members of the family (Irga6, Irgb6, Irgd, Irgm2 and Irgm3) accumulate at the PV membrane in interferon-induced astrocytes (Martens 2005), a result that was confirmed for IFNγ-induced gs3T3 and mouse embryonic fibroblasts in this study (see below). However, nothing is known about the transition from the "resting" localisation of IRG proteins in the uninfected, interferon-induced cell to the (presumably) active forms that accumulate at the PV membrane shortly after infection. To shed more light onto this process, it was investigated whether nucleotide-dependent regulatory interactions between IRG proteins, like those required for the resting localisation of Irga6 and Irgb6, are required for transfer of IRG proteins to the PV membrane.

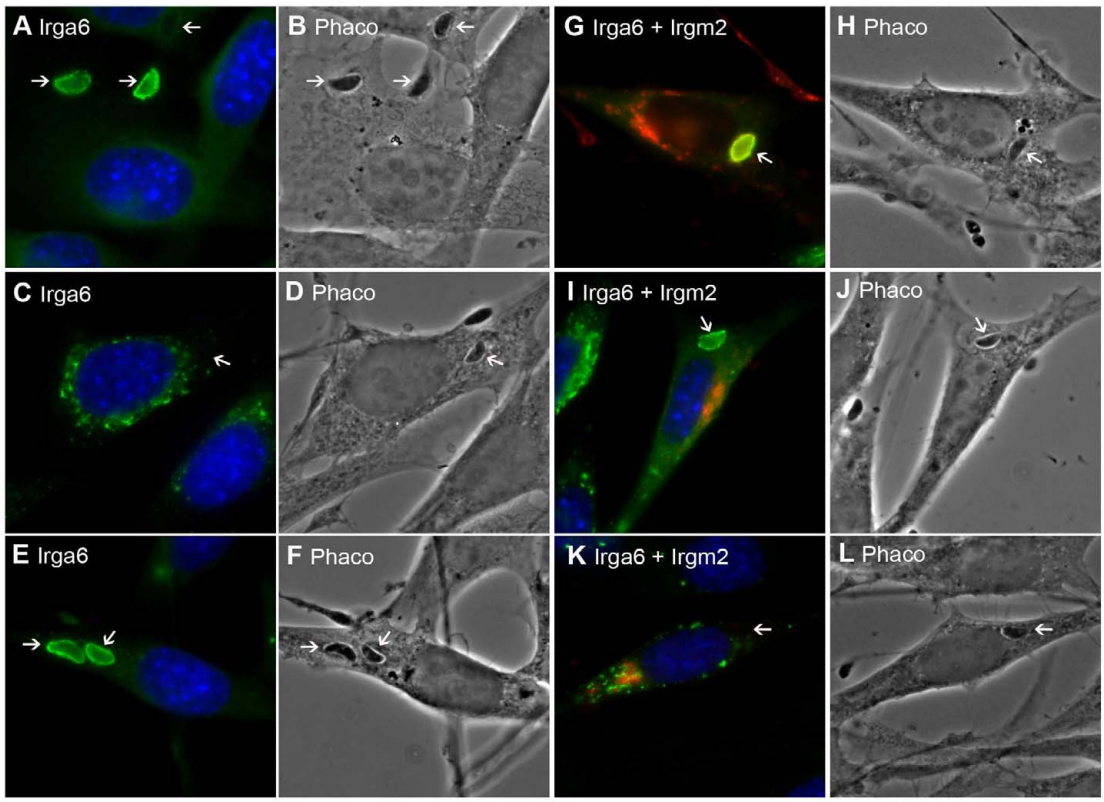


Figure 47 Irga6 localisation to the *T. gondii* PV is IFNγ and GMS protein dependent. Gs3T3-Irga6 cells were induced for 24h with 200/ml IFNγ or 10^{-9} M Mifepristone or both or induced with Mifepristone and transfected transiently with other IRGs. Subsequently, cells were infected with *T. gondii* strain ME49 for 2 hrs (MOI 6). Irga6 was identified by immunofluorescence with 10E7 monoclonal antibody (green) and Irgm2 with H53 antiserum. Nuclei were stained with DAPI. Intracellular *Toxoplasma* identified in phase contrast are marked with white arrows (B, D, F, H, J, M). (A-B) IFNγ-induced cells. (C-D)

Mifepristone-induced cells. (E-F) Cell induced with both Mifepristone and IFNγ. (G-H) Cells induced with Mifepristone and transiently transfected with (pGW1H) expression constructs for 5 IRGs (Irgm1-3, Irgb6, Irgd, 400 ng each). Note that the blue channel was used to detect Irgb6 with the A20 antiserum and the displayed PV was also Irgb6 positive (data not shown). (I-J) Cells induced with Mifepristone and transiently transfected with expression constructs for 3 GMS proteins (667 ng each). (K-L) Cells induced with Mifepristone and transiently transfected with expression constructs for the 3 GMN proteins (667 ng each).

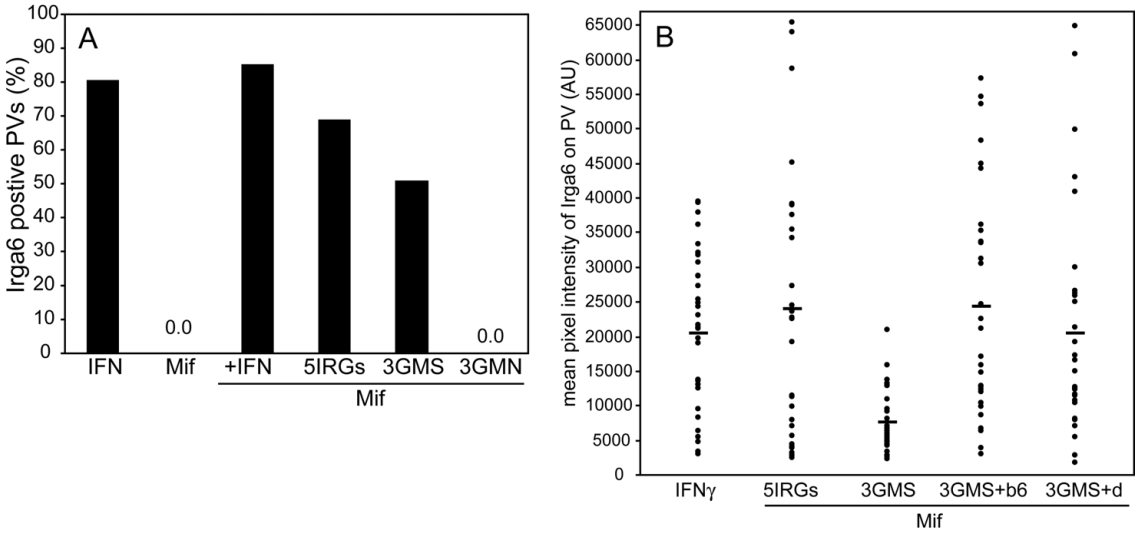


Figure 48 Influence of the other IRG proteins on the Irga6 signal at the *T. gondii* PV. (A) Number of Irga6 positive *T. gondii* PVs identified with the 10E7 antibody in gs3T3-Irga6 cells following 24 hrs induction with IFNγ (200 U/ml), Mifepristone (10^{-9} M) or both, or Mifepristone induction and transfection with either a pool of expression constructs encoding Irgm1, Irgm2, Irgm3, Irgb6 and Irgd (5IRGs), the three GMS proteins or the 3 inactive GMN mutants (in all cases total of 2 µg DNA; see also Figure 47). Data are recorded as the percentage of Irga6-positive PVs per 100 intracellular *T. gondii*. Around 100 vacuoles per data point were counted blind. (B) Quantification of the mean pixel intensity of the Irga6 signal observed in gs3T3-Irga6 cells infected with *Toxoplasma gondii* strain ME49 for 2 hrs at a MOI of 6. Cells were induced with IFNγ or with Mifepristone and simultaneously co-transfected with a pool of the indicated IRG expression constructs (400 ng each, total of 2 µg DNA; empty vector was used to adjust DNA amounts): Irga6, Irgb6, Irgm1, Irgm2 and Irgm3 (5IRGs), Irgm1, Irgm2 and Irgm3 (3GMS), the 3 GMS plus Irgb6 (3GMS+b6) or the 3 GMS plus Irgd (3GMS+d). Irga6 was detected in immunofluorescence with 10E7 antibody. Pixel intensities of the Irga6 signal at the PV at a constant exposure time were determined at 6 points across the vacuole with Image J software (http://rsb.info.nih.gov/ij/). The average pixel intensity is indicated by black lines (35 vacuoles/treatment). Values of 5000 pixels and below cannot be distinguished from the background of the ER signal and can be considered as negative. The 3 GMS proteins are not sufficient to enable wild type intensity of Irga6 accumulation at the PVM. Transfected cells in (A) and (B) were identified by staining for Irgm2 with the H53 serum.

Indeed, in contrast to interferon-induced gs3T3 cells (Figure 47 A-B), Irga6 expressed alone in cells, either by Mifepristone induction in gs3T3-Irga6 cells (Figure 47 C-D) or Irga6-ctag1 expressed in fibroblasts transiently by transfection (data not shown) failed to accumulate at *Toxoplasma gondii* parasitophorous vacuoles. If, however, gs3T3-Irga6 cells were simultaneously induced with IFNγ and Mifepristone, the ability of the Irga6 to accumulate at the PV was restored (Figure 47 E-F). The effect of IFNγ on Irga6 localisation at the PVM was mirrored by transient co-transfection of the 5 other IRG GTPases into Mifepristone-treated gs3T3 fibroblasts (Figure 47 G-H) as well as by transient co-transfection of the 6 IRG GTPases into untreated gs3T3 fibroblasts (data not shown). Under these conditions also all the other IRG proteins except for Irgm1 localised to the PVM (Figure 47 G-H and see below, III.2.12). Co-transfection of the 3 IRGM GTPases into Mifepristone-induced gs3T3-Irga6 fibroblasts was sufficient to allow

accumulation of Irga6 at the PVM (Figure 47 I-J). However fewer Irga6 positive PVs were found than in IFNγ-induced cells (Figure 48 A), and the average intensity of the Irga6 accumulation at the vacuoles was lower than in interferon-induced cells (Figure 48 B). This suggested that other IFN-inducible components might affect the efficiency of vacuolar accumulation of Irga6. Indeed, co-transfection of one or more further GKS IRG proteins along with the 3 GMS proteins (Irgb6 or Irgd or both) increased the average intensity of Irga6 on the vacuoles to a level comparable with that in IFNγ-induced cells (Figure 48 B). In a similar co-transfection, the 3 GMS to GMN G1 motif mutants of Irgm1, Irgm2 and Irgm3 were unable to mediate Irga6 accumulation at the *Toxoplasma* vacuoles (Figure 47 K-L, Figure 48 A), demonstrating nucleotide dependence of this process on the side of the IRGM proteins.

The monoclonal antibody 10D7 that is specific for the active, GTP-bound conformation of Irga6 was used to analyse the nucleotide status of Irga6 at the *T. gondii* parasitophorous vacuolar membrane (PVM). As shown above (Figure 35), 10D7 failed to detect normally distributed Irga6 at the ER of IFNγ-induced gs3T3 cells (Figure 49 C-D). However, 10D7 efficiently recognised the Irga6 protein accumulated at the PVM (Figure 49 C-D). 10E7, in contrast, recognised both pools of Irga6 efficiently (Figure 49 A-B).

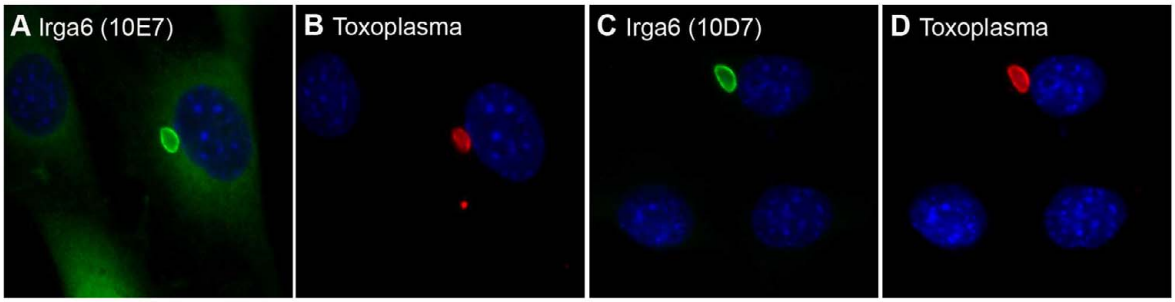


Figure 49 Irga6 is in the active, GTP-bound conformation at the *T. gondii* PV. (A-D) IFNγ-induced (200U/ml) gs3T3 cells were infected with *T. gondii* strain ME49 (MOI 8) for 2 hrs and Irga6 was detected with the monoclonal antibodies (A-B) 10E7 or (C-D) 10D7 (green). Intracellular parasites were identified by containing with anti-*Toxoplasma* serum (B, D, red). Note that the 10D7 antibody that is specific for Irga6 in the active, GTP-bound conformation exclusively recognised the protein at the PV.

To further explore the nucleotide dependence of Irga6 localisation to the PVM embryonic fibroblasts from wt and Irga6 deficient mice (Parvanova 2005) were infected with *Toxoplasma* and transiently transfected with expression constructs of ctag1-tagged wt and mutant Irga6 proteins. The Irga6-deficient cells were used to dissect whether possible association of the mutants with the PVM were due to inherent properties of the mutants or due to interaction with the endogenous wild type protein. C-terminally epitope-tagged Irga6-ctag1 transiently transfected into IFNγ-induced Irga6-deficient fibroblasts (Parvanova 2005) showed essentially normal resting localisation and relocalisation to the *T. gondii* PVM (Figure 50 A-C) showing that this tag does not interfere with normal Irga6 protein function. In contrast, both the biochemically dominant active K82A and the inactive S83N mutant of Irga6 failed to relocalise to the PVM in IFNγ-induced Irga6$^{-/-}$ MEFs (Figure 50 D-F, G-I). Irga6(S83N) also did not accumulate at the PVM in IFNγ-induced wt MEFs nor did it prevent wild-type Irga6 from doing so (Figure 50 J-L). However, the constitutively active mutant Irga6(K82A), itself also impaired in vacuolar targeting in IFN-induced wt MEFs, almost completely

prevented the endogenous Irga6 protein from reaching the *T. gondii* vacuole (Figure 50 M-O), thus acting functionally as a dominant negative as previously reported for *T. gondii*-infected, IFNγ-induced astrocytes (Martens 2005).

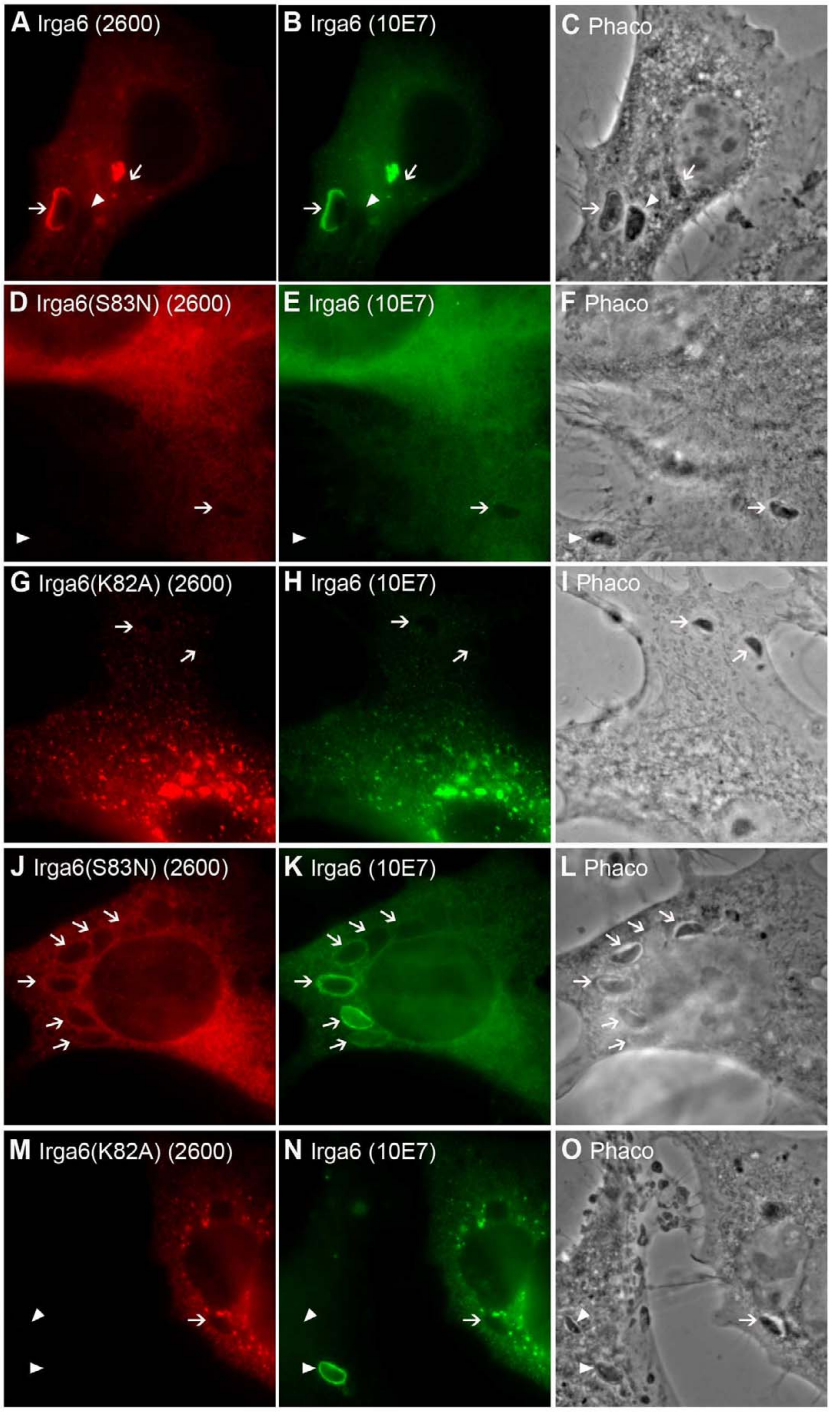


Figure 50 Irga6 localisation to the *T. gondii* PVM is nucleotide-dependent. Localisation of Irga6 in mouse embryonic fibroblasts transiently transfected with ctag1-tagged Irga6 expression constructs (pGW1H, 1 µg each) and infected for 2 hrs with *T. gondii* strain ME49 (MOI 6). White arrows indicate intracellular *T. gondii* identified in phase contrast (Phaco, C, F, I, L, O). Irga6 was detected by immunofluorescence with the monoclonal antibody 10E7 (green). Transfected Irga6-ctag1 protein was also detected by anti-ctag1 antiserum 2600 (red). (A-I) Embryonic fibroblasts (MEFs) from Irga6-deficient mice (A-C) transiently transfected with Irga6-ctag1. One of the three intracellular parasites did not acquire Irga6 (white arrowhead), the two other PVs are Irga6 positive and in the process of destruction. Differences in coating of individual PVs with IRG proteins seem to be related to the time of infection (Khaminets

unpublished data). Irga6$^{-/-}$ MEFs transiently transfected with Irga6(S83N)-ctag1 (D-F) or Irga6(K82A)-ctag1 (G-I). Wt MEFs transiently transfected with Irga6(S83N)-ctag1 (J-L). . Wild type Irga6 (green in G) relocated normally to the *T. gondii* PVM (white arrows) while the Irga6(S83N) (red in H) fails to transfer to the PVM. Irga6 was not aggregated in cells transfected with Irga6(S83N) (see also Figure 37 E-H). Wt MEFs transiently transfected with Irga6(K82A)-ctag1 (M-O). The transfected protein (red in E) did not transfer to the *T. gondii* PVM and also inhibited the transfer of the endogenous protein (green in D (includes Irga6-ctag1)). The non-transfected cell on the left of the image shows a normal PVM stain with endogenous Irga6 (white arrowhead) despite the low expression of Irga6. Irga6 is aggregated in IFNγ-induced cells transfected with Irga6(K82A) (see also Figure 37 A-D).

Together these results suggest that cytoplasmic aggregation and failure to reach the *T. gondii* PVM are properties of GTP-bound Irga6, constitutively in the case of the constitutively active mutant, and, in the absence of IFNγ, also for the wild-type protein. Furthermore it appeared that the constitutively active mutant could capture wild type Irga6, presumably also in the GTP-bound form, in mixed aggregates. On these grounds it was hypothesised that wild type Irga6 is normally maintained in the cytoplasm in the GDP-bound state in IFNγ-induced cells, but accumulates at the *T. gondii* PVM in the active GTP-bound state. In the absence of IFNγ, wild-type Irga6 activates prematurely in the cytoplasm and forms "sterile" aggregates that cannot localise to the PVM. Constitutively active Irga6(K82A) activates constitutively in the cytoplasm and forms ectopic GTP-bound aggregates, which can also capture wild-type Irga6, thus acting as a functional dominant negative. Above all, these experiments strongly suggested that the nucleotide-binding status of Irga6, and hence its adaptive behaviour in *T. gondii* resistance, is normally regulated by the unusual GMS proteins.

III.2.12. Regulation of Irgb6, Irgd, Irgm1-3 and Irgc in *Toxoplasma*-infected cells

In view of the regulation of Irgb6 resting localisation by the GMS proteins, it was unexpected that Irgb6 relocated apparently normally - though somewhat less efficiently - to the *Toxoplasma gondii* parasitophorous vacuole the in absence of other IRG proteins after transfection into unstimulated gs3T3 fibroblasts (Figure 51 A-B) and in Mifepristone-induced gs3T3-Irgb6 cell lines (data not shown). This relocalisation occurred even when the majority of the Irgb6 protein was strikingly mislocalised elsewhere in the cytoplasm (Figure 51 A-B). Irgb6 was the only IRG family member studied whose association with the PV membrane was found to be independent of IFN-signalling or any other IRG GTPase (see below). The G1 motif mutant Irgb6(S70N), like the corresponding S83N mutant of Irga6, failed to associate with *T. gondii* parasitophorous vacuoles in transiently transfected gs3T3 cells (Figure 51 C-F). In contrast, Irgb6(K69A), corresponding to Irga6(K82A), was capable of targeting the *T. gondii* PV efficiently in IFNγ-treated but not in untreated cells, despite mislocalisation elsewhere in the cell (Figure 51 G-J). The ability of transfected wild-type Irgb6 to accumulate on the *T. gondii* PVM in uninduced cells may suggest that in the absence of GMS proteins an equilibrium, perhaps between GTP-bound, aggregated and GDP-bound free forms, is slightly more favourable to GDP in the case of Irgb6. In IFNγ-induced cells, the by analogy to Irga6(K82A) constitutively active Irgb6(K69A) can perhaps also be partially repaired by co-expressed GMS proteins, which push the equilibrium sufficiently towards the GDP-bound state.

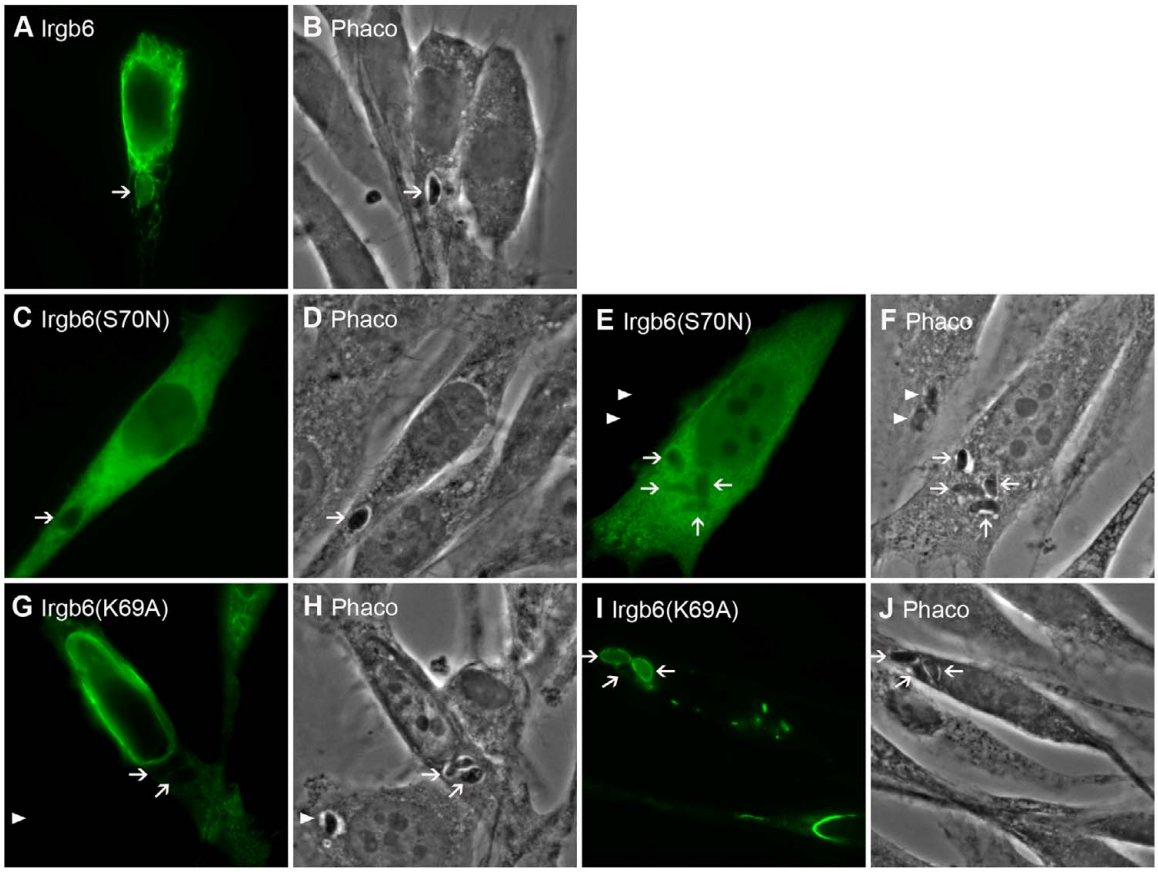


Figure 51 Localisation of Irgb6 to the *T. gondii* PVM is nucleotide-dependent. Gs3T3 cells were transiently transfected with (pGW1H) expression constructs of FLAG-tagged wt and mutant Irgb6 (0.5 µg/construct) in absence (A-B, C-D, G-H) and presence of 200U/ml IFNγ (G-F, I-J) and infected with *T. gondii* strain ME49 at an MOI of 6 for 2 hrs. Transfected protein was detected with the FLAG-specific monoclonal antibody M2 (green). White arrows indicate intracellular *T. gondii* identified in phase contrast (Phaco, B, D, F, H, J). Wild type Irgb6 in uninduced cells (A-B). Irgb6(K69A) in uninduced (C-D) or IFN-induced (G-F) cells. Irgb6(S70N) in uninduced (G-H) or IFN-induced cells (I-J).

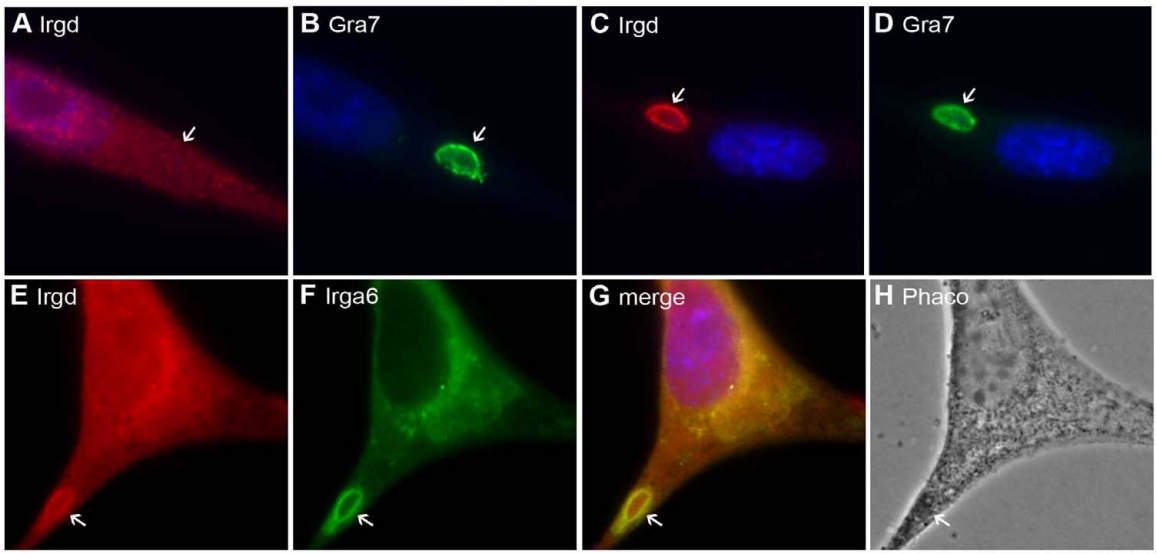


Figure 52 Irgd accumulation at the *T. gondii* PV is dependent of other IRG proteins. Localisation of Irgd detected with 2078 antiserum (red in A, C, E) in gs3T3-Irgd cells infected with *T. gondii* strain ME49 at a MOI of 6 for 2 hrs. Note that the serum weakly recognises the distributed Irgd at normal expression levels, but efficiently stains Irgd on the PVM. Intracellular parasites are indicated by white arrows. (A-B) Cells induced with 10^{-9} M Mifepristone. (C-D) cells induced with 200 U/ml IFNγ. Parasites are stained with Gra7 antiserum (green in B and D). (E-F) Mifepristone-induced (10^{-9} M) cells transiently transfected with (pGW1H) expression constructs for Irga6, Irgb6, Irgm1, Irgm2 and Irgm3 (500 ng each). Transfected

cells were identified by counterstained for Irga6 with the 10E7 monoclonal antibody (green in B). Note that the Irgd positive vacuole also contains Irga6. The Irgd positive vacuole also contains Irga6. Overlay together with nuclear counterstain (DAPI, blue) is shown in (C). Intracellular parasites were identified in phase contrast (H).

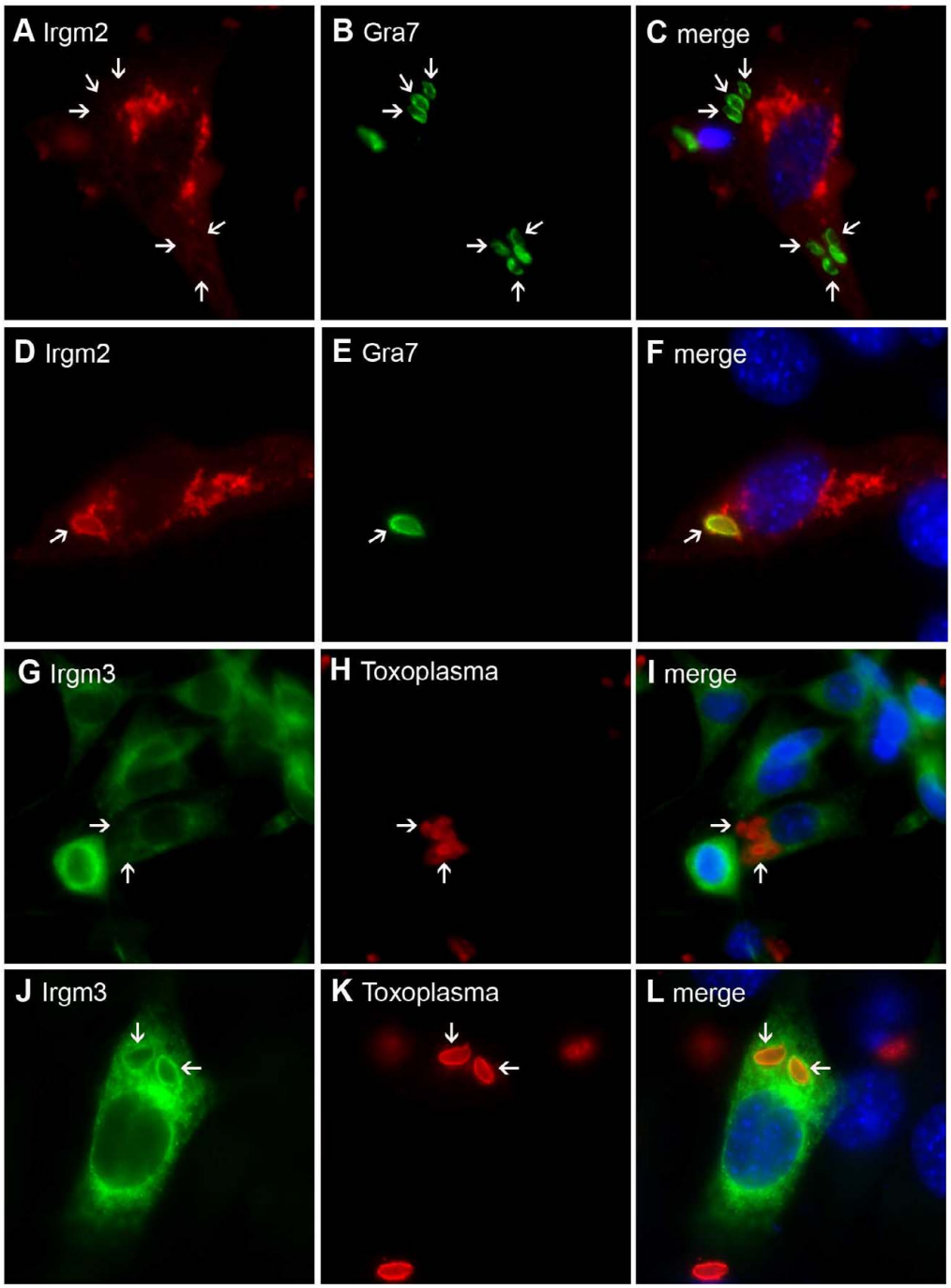


Figure 53 PVM localisation of Irgm2 and Irgm3 is dependent on other IRG proteins. Cells infected with *T. gondii* strain ME49 at a MOI of 6 for 2 hrs. Parasites were detected with anti-Gra7 monoclonal antibody (B, E, green) or *T. gondii* specific rabbit serum (H, K, red), Irgm2 with H53 serum (A, D, red) and Irgm3 with anti-IGTP clone7 mAb (G, J, green). Nuclei are counterstained with DAPI (blue). White arrows indicate intracellular *T. gondii*. (A-C) gs3T3-Irgm2 cells induced with Mifepristone. (D-F) Relocalisation of Irgm2 (red) to the PVM in uninduced gs3T3 cells transiently transfected with a pool of 6 IRG expression plasmids (Irgm1, Irgm2, Irgm3, Irga6, Irgb6, Irgd, total 2 μg of DNA). (G-I) gs3T3-Irgm3 cells induced with Mifepristone. (J-L) Relocalisation of Irgm3 (green) to the PVM in uninduced gs3T3 cells transiently transfected with a pool of 6 IRG expression plasmids (Irgm1, Irgm2, Irgm3, Irga6, Irgb6, Irgd, total 2 μg of DNA)

Unlike Irgm1, both Irgm2 and Irgm3 have been shown to accumulate on the *T. gondii* PVM in IFNγ-induced, infected cells (Martens 2005). Even though the subcellular localisation of these two proteins in uninfected cells was independent of IFN (see above, Figure 30-31), the accumulation at the PVM did not occur in Mifepristone-induced stable Irgm2 and Irgm3 gs3T3 cell lines cells infected with *T. gondii* in the absence of IFNγ (Figure 53 A-C and G-I). Normal localisation of Irgm2 and Irgm3 to the *T. gondii* PVM was, however, seen in cells transfected with Irga6, Irgb6 and Irgd in addition to the three GMS proteins (Figure 53 D-F and J-L). Furthermore, the ability of Irgm2 and Irgm3 to accumulate on the PVM was dependent on the integrity of the G1 motif (data not shown). Thus, correct localisation of Irgm2 and Irgm3 to the PV was dependent on the presence of other IRG proteins and nucleotide binding.

Surprisingly, Irgc expressed by Mifepristone induction in gs3T3-Irgc fibroblasts accumulated at the Toxoplasma PVM in an IFNγ dependent manner (Figure 54 A-D). Transient cotransfection of the IRG proteins Irgm1, Irgm2, Irgm3, Irga6, Irgb6 and Irgd also promoted Irgc association with the PVM (Figure 54 E-F) and so did the 3 GMS proteins alone, though markedly less efficiently (Figure 54 G-H). These features are reminiscent of the behaviour of Irga6 in *Toxoplasma* infected cells. Whether and how the localisation of Irgc to the PVM relates to the *in vivo* function of the protein is unclear.

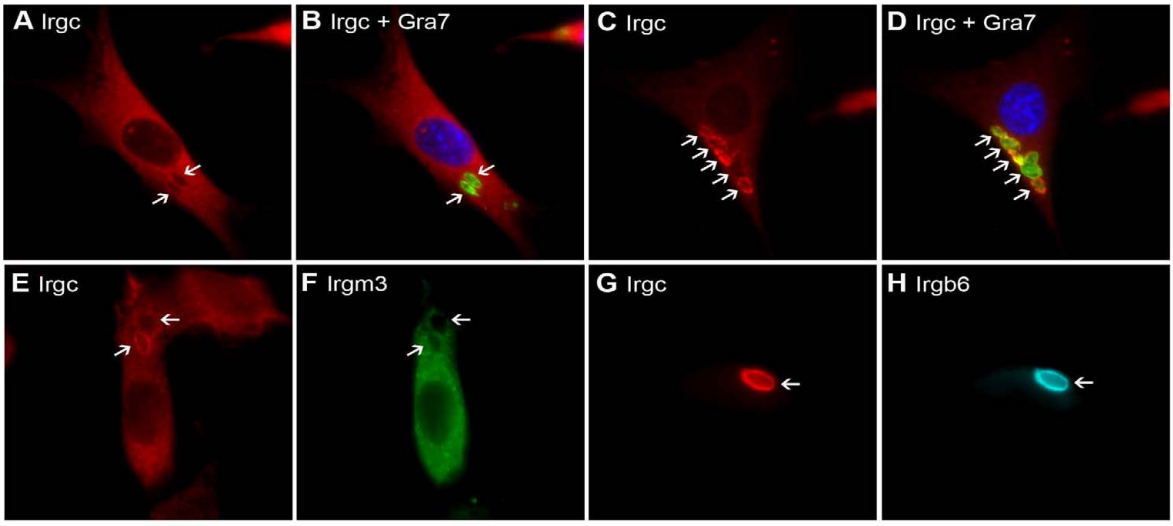


Figure 54 Irgc accumulates at the PVM in the presence of other IRG proteins. Irgc localisation was detected by the 39°3 serum (red in A-D, E, G) in cells infected with *T. gondii* strain ME49 for 2 hrs at a MOI of 6. Intracellular parasites identified in phase contrast (not shown) are marked with white arrows. (A-B) gs3T3-Irgc cells induced with10^{-9} M Mifepristone alone (A-B) or together with IFNγ (200 U/ml) (C-D). Parasites were counterstained with Gra7 antibody (green in B and C). (E-F) gs3T3-Irgc cells induced with Mifepristone (10^{-9} M) and transiently transfected with (pGW1H) expression constructs for the 3 GMS proteins (667 ng each). Transfected cells were identified by co-staining with anti-Irgm3 antibody (green in F). (G-H) gs3T3-Irgd cells induced with Mifepristone (10^{-9} M) and transiently transfected with (pGW1H) expression constructs for Irga6, Irgb6, Irgd, Irgm1, Irgm2 and Irgm3 (333 ng each). Cells were counterstained for Irgb6 with the A20 antiserum (blue in H).

III.2.13. Virulent *T. gondii* counteract IRG protein accumulation at the PVM

As IRG proteins are essential resistance factors in mouse active against *Toxoplasma gondii,* it was asked whether virulent strains of the parasite evolved mechanisms to counteract IRG proteins. Indeed, accumulation of Irgb6 on the PVM of virulent YFP-

labelled RH strain parasite was strongly impaired (Figure 55). Only 8 % of the PVs from containing RH *T. gondii* accumulated detectable amounts of Irgb6 protein as opposed to 60-70 % in case of the avirulent ME49 strain (Figure 55 E). The accumulation of other IRG proteins at the PVM was also diminished in the case of the virulent strain, but the effect was much less pronounced (Figure 55 E, shown for Irga6 and Irgd). There was only a reduction of ~15% in the number of detectably Irga6 positive PVs, yet the amount of Irga6 protein found on the vacuoles was largely reduced (Figure 55 F, white arrows in A-D). Interestingly, only those vacuoles containing large amounts of Irgb6 also acquired Irga6 efficiently (Figure 55 A-D, white arrow head) while the Irgb6 negative PVs were Irga6 dim (Figure 55 A-D, white arrows). Thus, virulent *Toxoplasma* possess a mechanism to counteract IRG protein accumulation on the parasitophorous vacuole that presumably targets Irgb6. As data are accumulating that Irgb6 is a pioneer in the colonisation of the PV and necessary for efficient vacuolar recruitment of other IRG proteins (Figure 55, Figure 51 and Könen-Waismann, Khaminets unpublished data), the effect on the other IRG proteins is presumably secondary. The mechanism by which virulent *T. gondii* inhibits Irgb6 remains to be determined.

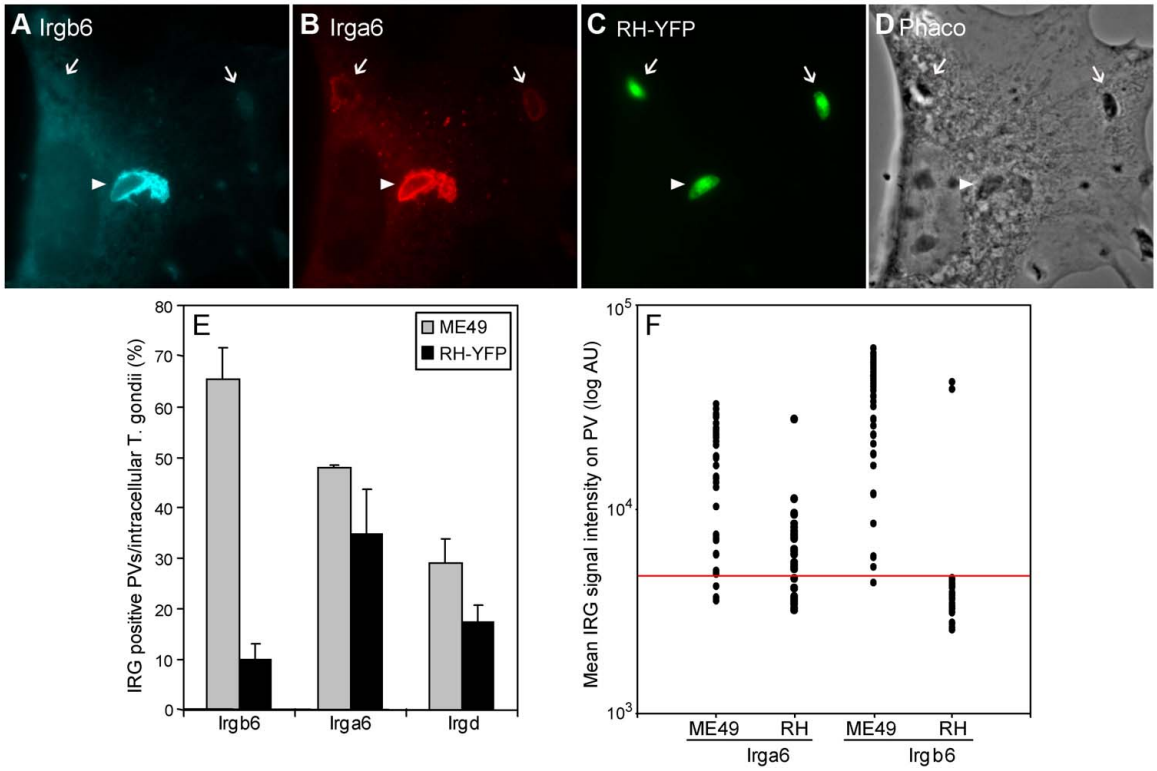


Figure 55 Accumulation of IRG proteins on the PVM is hindered by virulent *T. gondii*. IFNγ-induced (200 U/ml) gs3T3 cells were infected with the *T. gondii* strains ME49 or RH-YFP for 2 hrs at a MOI of 6. Irgb6 (A, blue) and Irga6 (B, red) were detected in immunofluorescence with the A20 serum and the 10E7 monoclonal antibody respectively. Intracellular, yellow fluorescent parasites (green in C) identified in phase contrast (D) are indicated by white arrows heads (strongly IRG positive) and arrow (weakly IRG positive). (E-F) Quantification of data displayed in (A-D). (E) Number of IRG protein positive PVs per 100 intracellular parasites (grey bars: ME49 strain; black bars: RH-YFP strain). 50 intracellular parasites were counted blind per data set. (F) Mean pixel intensities of the Irga6 and Irgb6 signal at the PVM quantified as described in Figure 48. The red line marks the approximate visible threshold. 35 PVs were per data set were quantified blind.

IV. Discussion

The immunity-related GTPases are a large family of essential, cell autonomous resistance factors in the mouse against intracellular pathogens with diverse life styles. All family members identified thus far contribute in a unique way to this resistance, mediating the immune control of distinct pathogens in different phases of the infection by – at least partially - different mechanisms. Thus, newly identified family members are promising candidates for an important innate immune function. Furthermore, analysis of their role in resistance will help to elucidate common mechanisms and to understand the IRG resistance system as such. This study identified the complete set of IRG genes present in the mouse and also in other animal species. Phylogenetic analyses of the large set of newly identified genes (more than 180 in over 30 species) allowed insights into the dynamics of evolution in the IRG family, which underwent rapid expansion, diversification and contraction in euchordates, presumably driven by host-pathogen interactions.

Furthermore, the cellular properties, regulation and function of seven mouse IRG proteins were analysed to shed light onto the molecular mechanism of IRG resistance function and the reasons for the high degree of genetic diversity and plasticity in the evolution of the IRG family. Six of the analysed IRG proteins localise to the parasitophorous vacuole upon infection with avirulent *Toxoplasma gondii*. This behaviour is dependent on nucleotide binding and on a complex network of direct nucleotide-dependent regulatory interactions between IRG family members. Irga6 and Irgb6 self-interaction leads to activation and formation of oligomers on the PVM. In contrast, Irga6/Irgb6-GMS protein interaction is inhibitory, keeping the GKS proteins in the GDP-bound state and preventing premature activation prior to infection. The necessity of complex patterns of regulatory interactions within the IRG family for normal function of the proteins delivers an explanation for the pronounced functional non-redundancy of the IRGs.

IV.1. Signs for divergent evolution of the IRG genes in rodents

The IRG family constitutes an essential, cell autonomous resistance system against intracellular bacterial and protozoal pathogens in the mouse involving all of the six previously published IRG proteins (Irga6, Irgb6, Irgd, Irgm1, Irgm2, Irgm3) (Martens 2006; Miyairi 2007). In the course of this study, 15 additional murine IRG genes were identified and mapped in the mouse genome to four clusters on three chromosomes (Figure 9-10). One of the newly identified genes, Irgb10, was subsequently shown to be involved in resistance against *Chlamydia* species (Bernstein-Hanley 2006; Miyairi 2007). This strongly suggests that also other newly identified family members contribute to pathogen resistance. The IRG genes in the individual clusters display various degrees of sequence conservation arguing for different evolutionary ages of the underlying duplication events (Figure 11-12). The eight Irga6-related genes (IRGA clade) clustered on chromosome 18 and seven of the Irgb6-like genes (IRGB clade) clustered on chromosome 11 show the least divergence amongst each other and therefore presumably originate from a more recent gene duplication. The entire amino acid sequence of IRG

proteins is typically encoded on a single long exon. Four of the IRGB genes, however, consist of 2 IRG coding units transcribed in tandem and thus give rise to proteins twice the size of a usual IRG (appendix V.5). The fact that IRG proteins form functionally relevant dimers and oligomers *in vitro* (Figure 36, (Uthaiah 2003)) and *in vivo* (Figure 44-45, (Papic 2007)), suggests that these tandems may actually have their interaction partners covalently linked.

Irgc, Irgd, Irgb10, Irgm1, Irgm2 and *Irgm3* are all more divergent than the other IRG members proposing that these genes are phylogenetically older (Figure 12). Comparison with the IRG genes from rat confirmed this interpretation. The 15 rat IRG coding units are situated on three chromosomes in four clusters, which are syntenic to the respective mouse chromosomal clusters. Rat *Irgc, Irgd, Irgb10, Irgm1, Irgm2* and *Irgm3* are clearly orthologous to the respective mouse genes. For the seven closely related IRGA genes, however, no clear mouse orthologues could be identified with the exception of *Irga5* (Figure 14). The set of IRGB genes in the proximal cluster on rat chromosome 10 is reduced relative to the mouse (Figure 16). Only two coding units, probably representing a single tandem gene, are present (Irgb14-Irgb13). Therefore, the numerous mouse IRGB genes must either represent a very recent gene duplication of an ancestral IRGB tandem gene or alternatively the rat must have lost the rest of the IRGB genes. The presence of the mouse IRGB coding units in pairs of two, which cluster phylogenetically with the rat tandem pair, and the high degree of sequence conservation in the intergenic spaces between the mouse IRGBs favour the first hypothesis. Interestingly, one of the first mouse IRG genes identified, Irgb6 (TGTP) (Carlow 1995), is absent from the rat and is the only IRGB sequence in the cluster, which is not part of a tandem.

The *Irgc* genes exclusively expressed in haploid spermatids in the testis (Rhode 2007) as well as the distantly related IRG gene with the radically modified GTPase domain sequence, *Irgq*, are both highly conserved between rat and mouse, arguing for a housekeeping rather than an immune function. This view is supported by the high degree of conservation of *Irgc* and *Irgq* throughout the mammals and the *Eutheria*, respectively (Figure 20). Despite their phylogenetic relationship to the classical IRG proteins, IRGQ proteins probably lack nucleotide binding and hydrolysis function owing to the absence of conserved GTP-binding motifs and the low degree of sequence conservation in the G-domain (Table 6, Figure 20).

IV.2. Divergent IRG genes are found throughout the mammals

IRG genes were detected in all mammalian orders, indicating that this family of GTPases was widely distributed at the origin of the mammalian radiation. There is, however, plenty of evidence for diversifying evolution in the mammalian IRG proteins.

Interestingly, all IRGB genes identified in mammals other than rat and mouse cluster phylogenetically with the isolated mouse *Irgb10* gene (Figure 14, Figure 18). Together with the chromosomal positioning of *Irgb10* in mouse and rat in a separated cluster with Irgm2 and Irgm3, this suggests that an ancestral *Irgb10*-like gene probably gave rise to the second cluster of IRGB genes in rat and mouse by gene duplication (Figure 10, Figure 16). The association of the *Irgb10* gene cluster with *Irgm2* and *Irgm3* and of the other IRGB gene cluster with *Irgm1* suggests that an ancestral *Irgb10*-like

gene in close proximity to a single IRGM gene gave rise to these two IRG gene clusters by a series of gene duplication followed by diversification.

In agreement with this, all dog IRGM and IRGB genes (except for one pseudogene on chromosome 29) are grouped together in one cluster of only 400 kb on chromosome 10 (Figure 17). This cluster also contains two IRGD-related genes, one of which possesses a drastically damaged G1 motif. Thus, the common mammalian ancestor must have possessed at least a single copy of an IRGM, IRGD and IRGB10-like gene each. The absence of IRGA genes from the dog seems to be a specific loss in this lineage as not only rodents and rabbits but also primates, insectivores and elephants possess IRGA genes. Dog IRGB12 and the three dog IRGM genes have been shown to be IFNγ-inducible *in vitro* (Bekpen 2005a; Bekpen 2005b), classifying them also experimentally as immune-type IRGs.

IRGC, IRGB and IRGD genes were found in all mammalian groups, whereas IRGQ, IRGM and IRGA genes were only detected in the *Eutheria*. All the clades originally defined on the basis of a phylogenetic comparison of mouse and dog IRG sequences alone (Bekpen 2005b), are well defined throughout the mammals. Interestingly, in two of the newly identified members of the IRGM subfamily, the unusual methionine in the first GTP-binding motif was exchanged for a leucine or isoleucine, respectively. This conservative exchange suggests that a large hydrophobic residue at this position might be essential for the function of the proteins. Though the mutation of the conserved P loop lysine interferes with nucleotide binding and hydrolysis in many GTPases (e.g. hGBP1, H-Ras and Mx1/MxA (Pitossi 1993; Praefcke 2004a; Sigal 1986)), the analogous lysine to methionine substitution in the ATPase shikimate kinase has been reported to increase the ATP affinity of the protein (Krell 2001). However, nucleotide hydrolysis was abolished and the organisation of the P loop flanking regions of the protein was disturbed. Biochemical analysis of recombinant IRGM proteins is therefore essential to gain insight into the role of the P loop methionine in the activity of these unusual GTPases.

The IRGM genes of different mammalian species have either been duplicated independently or have diversified so much, that clear orthologues of individual genes can not be identified (Figure 18, Figure 20). Different clades of the IRG gene family have expanded in different mammalian species. Dog, hedgehog, armadillo, guinea pig and tenrec contain several IRGB10-related genes, while the other IRGB genes represent a rodent specific expansion (Figure 18, Figure 14). IRGA genes have been massively expanded in mouse and rat whereas the guinea pig possesses multiple IRGD genes. This indicates that the immune-type IRG genes are under strong diversifying selection, probably due to host-parasite coevolution.

Another sign of diversity in the IRG family is the presence of a single, *IRGC*-related gene (*IRGC-like*) in numerous but not all mammals. The *IRGC-like* was found in shrew, bat, cat, dog, bull, pig as well as in opossum (Table 7). It is phylogenetically related to and spatially closely linked to *IRGC* (Figure 17, Figure 20). The *IRGC-like* genes are highly divergent suggesting an immune function. Apart from *IRGQ*, which contains an N-terminal extension encoded on a separate upstream exon, *IRGC-like* is the only mammalian IRG gene that contains an intron (see also below).

IV.3. Loss of IRG resistance system in the anthropoids

From the essential role of IRG proteins in pathogen resistance of mice and the presence of immune-type divergent IRG proteins (IRGA, IRGB, IRGD and IRGM clades) in many mammalian orders, it was surprising to find that the IRG resistance system has apparently been lost in the human (Table 7, Figure 15). The argument for the absence of the IRG resistance system in humans relies on several findings. The system is reduced from 22 genes in mouse to two full-length genes plus a transcribed G-domain in humans, and the residual genes lack the character of functional resistance genes. Thus, *IRGC* is highly conserved in mammals (Figure 20), is not interferon or infection inducible, and is expressed constitutively in mature testis (Bekpen 2005b; Rhode 2007). *IRGQ* is also highly conserved in mammals plus it contains a radically modified G-domain (Table 8). *IRGM*, although clearly derived from a typical GMS subfamily resistance gene, is transcribed constitutively from an endogenous retroviral LTR, is unresponsive to interferon (Bekpen 2005b), and is structurally damaged by truncation and loss of the fifth GTP-binding motif ().

Similarly, all anthropoids (old world monkeys, new world monkeys and hominoids; Figure 56) contain the reduced set of IRG genes (*IRGC*, *IRGQ* and *IRGM*). Interestingly, the IRGM protein of both macaque and marmoset possess a damaged first GTP-binding motif, Gx_4GMN and Gx_4SIS respectively (Table 8). In analogy to H-Ras, the loss of the universally conserved glycine in the marmoset IRGM is expected interfere with nucleotide binding (Chen 1994; Powers 1989). The serine to asparagine mutation occurring in the macaque has been reported to interfere with nucleotide binding in Irgm3 (Taylor 1997) and rendered all IRG proteins analysed in this study (and other GTPases; (Feig 1988; Nuoffer 1994; Praefcke 2004a; Stenmark 1994)) non-functional *in vivo* (Figure 42, 44, 47, 50 and 51). The macaque *IRGM* gene additionally contains a premature stop codon shortly after the G1 motif (Figure 21). Thus, it seems very unlikely that any of these primate IRG proteins carries out an immune function. Consequently, effective resistance to vacuolar pathogens in the anthropoids must be organised on radically different principles than in the mouse. Recently, however, the human IRGM protein has been implicated in resistance to *Mycobacterium tuberculosis* by induction of autophagy (Singh 2006). However, the observed effect was small and the data are not conclusive. In particular, no explanation is provided for how an IRG protein with an incomplete G-domain that lacks the complete N- or C-terminal helical domains and is constitutively expressed under the control of the long terminal repeat of the endogenous retrovirus ERV9 (Bekpen 2005b), should mediate an immune function. The involvement of IRGM in autophagy-mediated mycobactericidal activity has been analysed in analogy to that observed for the murine Irgm1 protein (Gutierrez 2004). Irgm1, however, is part of the highly interactive IRG resistance system (III.2.11, III.2.12 and see below) and at least its function in resistance to *Toxoplasma gondii* is dependent on an intact nucleotide-binding site (Figure 47-48). Human *IRGM* was recently identified as a susceptibility locus associated with Crohn's disease (Jacobs 2007; WTCCC 2007). Both the suggested involvement of IRGM in *Mycobacterial* resistance and in Crohn's disease requires further experimental analysis.

Since lower primates (lemurs and bushbabies, Strepsirrhini*)* possess diverse IRG genes belonging to the IRGD and IRGA clade as well as full-length GMS subfamily genes, it is clear that the IRG resistance system has been lost from the higher primates. This loss happened after the divide of the Haplorrhini from the Strepsirrhini around 60 million years ago and before the split of the old world monkeys about 30-40 million years ago (Figure 56). It will be of interest to analyse the tarsier genome *(Tarsius syrichta*, currently being sequenced; http://genome.wustl.edu, Genome Sequencing Centre) to further pinpoint the loss of divergent IRG genes in the primate lineage (Figure 56).

The IRG resistance system seems to have been lost independently also from some other species outside the primates. For cat, pig, bat and shrew only *IRGC* and *IRGQ* genes were detected in the database, while other species belonging to the same orders of mammals (e.g. dog as a carnivore, tenrec and hedgehog as insectivores) were shown to possess immune-type IRG genes (Table 7). Unless the absence of divergent IRGs from the listed species reflects a failure of the databases, the IRG resistance system has been repeatedly lost during the evolution of the mammals, suggesting a balance between the benefit of immunity against intracellular pathogens and the evolutionary cost of possessing the IRG system. Strikingly, in the cat as the primary host of *Toxoplasma gondii,* the IRG resistance system seems to be absent, possibly allowing unrestricted replication. A very recent sequence analysis, however, discovered a fragment of a GMS gene closely related to dog *IRGM4* in the cat genome database. Thus, more immune type IRG genes may remain to be discovered in the growing databases.

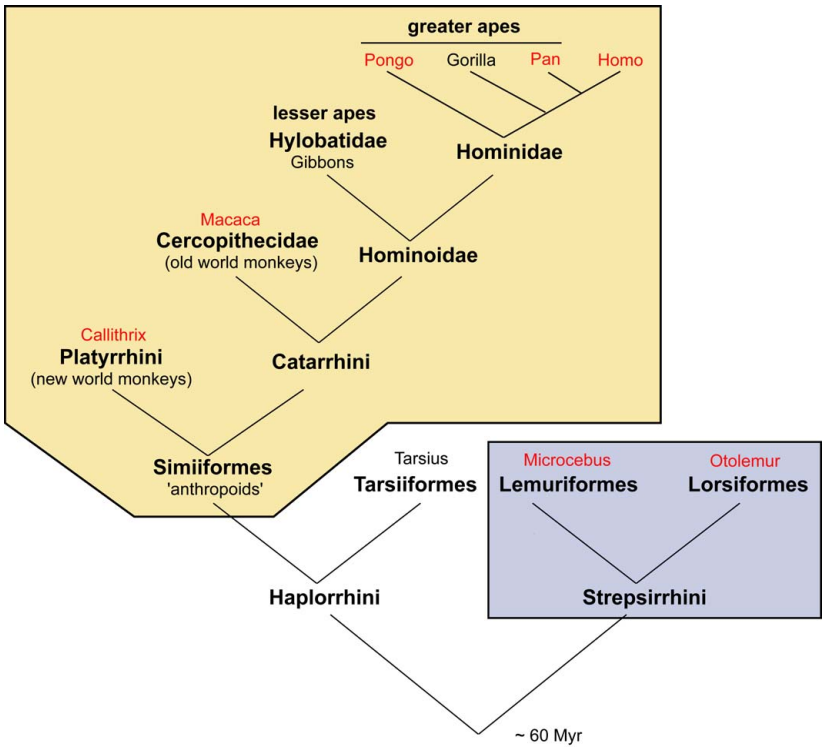


Figure 56 Schematic phylogenetic tree of the primates. Phylogenetic relationships between species are based on (Shoshani 1996). Species analysed for IRG genes are indicated in red (human *(Homo sapiens),* chimpanzee *(Pan troglodytes),* orang-utan *(Pongo pygmaeus),* marmoset *(Callithrix jacchus),* Rhesus monkey *(Macaca mulatta),* grey mouse lemur *(Microcebus murinus),* bushbaby *(Otolemur garnettii).* Note that the *Tarsius syrichta* genome database is not available yet. Primates with a reduced set of IRG proteins are highlighted in yellow, those with immune type IRGs in light blue.

In line with the argument above, the frequent presence of IRG genes with altered nucleotide binding motifs may reflect the evolutionary pressure to lose redundant gene copies. This is reminiscent of the balanced frequency of minus over wild type alleles of Mx1 in mice (Haller 1987; Jin 1998; Staeheli 1988a). Alternatively, IRG proteins that have lost their GTPase activity by mutation might function as regulators of intact IRG proteins by forming regulatory heterodimers influencing the interaction equilibria between other family members (see below). A case of a GTPase-like protein regulating the activity of a functional GTPase has recently been reported. The GTPase-activating protein of the small GTPase, Rap1, is itself probably derived from a GTPase ancestor and retains the G-domain structure but not the sequence to reveal its origin (Daumke 2004).

IV.4. The IRG family is ancient and present in euchordates

IRG genes are not only represented in the mammals but in all vertebrate classes except for the birds. Whether this represents another, yet even more complete loss of the IRG system or is rather due to incompleteness of the available databases remains to be determined. Analysis of IRG genes in zebrafish and the anole lizard revealed an array of complexity comparable to the mouse situation (Figure 22, Table 7) and zebrafish IRG genes have been shown to be IFN-inducible (D. Sieger, C. Stein personal communication). IRG genes were also recovered from cartilaginous fish and amphibians (Table 7). Most of the non-mammalian IRG genes formed new, independent clades in the phylogenetic analyses (Figure 22; IRGF, IRGE, IRGG). Myristoylation seems to be an ancient mechanism of membrane attachment in IRG proteins as myristoylation motifs can be found in subsets of the IRGA, -B and -E proteins throughout the vertebrates (Table 9). IRGF genes were exclusively detected in the bony fish, and the zebrafish genes diverged independently from the pufferfish ones, following the same principle already observed in mammals (see above). The IRGE clade contains zebrafish and anole genes and also here no clear orthologues can be identified. To decide whether the zebrafish *irgg* gene is really characteristic for an independent clade or rather represents a divergent IRGE gene, more fish sequence need to be analysed. One sequence from salamander and several anole sequences cluster phylogenetically with the IRGC-like/IRGC genes, suggesting a common ancestor for these proteins. Similar to the mammalian IRGC-like genes, the anole sequences show a high degree of divergence but by contrast they are encoded on a single exon. IRGF genes, however, contain a single intron situated between the G4 and the G5 motif of the G-domain, roughly at the same position as the intron of the mammalian IRGC-like genes (Figure 23, appendix V.13). Furthermore, the IRGF genes are more closely related to the IRGC-like genes and the other mammalian IRGs than to the IRGE group. Thus, one can hypothesise that an intron-containing common ancestor of the IRGF and IRGC-like genes has given rise to the divergent sets of IRG genes in mammals, reptiles, amphibians and fish.

Interestingly, ‘quasi’ IRG proteins with modified GTP-binding motifs were detected in both fish and amphibians. Thus, loss of GTPase activity from members of the IRG family has taken place repeatedly during evolution also outside the mammals. Strikingly, the *Xenopus irgxq4* gene is transcribed in tandem with an other IRG gene that might have retained GTPase activity. Similarly, in zebrafish the canonical *irgg* gene, is

apparently expressed in tandem with the adjacent downstream gene *irgxq1*, which is a modified quasi-GTPase gene (Table 8). Thus, IRG genes that occur in pairs in a head-to-tail arrangement, are transcribed in tandem and are presumably expressed as dimeric proteins, are not a sole invention of the rodents but have occurred repeatedly and probably independently during vertebrate evolution. The basic unit of IRG protein function may therefore be a dimer. This conclusion is consistent with the G-domain-G-domain homodimer shown to be essential for GTP-dependent oligomerisation and cooperative hydrolysis Irga6 (Pawlowski unpublished data), and the regulatory, GDP dependent Irga6-Irgm3 interaction documented in this study (chapters III.2.10 and III.2.11). Unlike the observed homodimer of Irga6, the products of the tandem genes are heterodimers, implying that the two IRG subunits serve distinct functions in the protein. In the cases of zebrafish and *Xenopus* the parts of the tandems without functional G-domains may be regulatory for the canonical IRG domains. Other 'quasi' IRG proteins may also be regulators of IRG proteins, interacting with the functional IRG proteins presumably via the G-domain-G domain interface used by the activating Irga6 self interaction (Pawlowski unpublished data) and the inhibitory Irga6-Irgm3 interaction (Figure 45). Thus, 'quasi' IRG proteins would coevolve with IRG proteins. Curiously, a rat transcript that codes for a triple IRG protein constituted of Irgb10, Irgm3 and Irgm2 is present in the databases (AY321344; see (Bekpen 2005a)). It is unclear whether that protein is functional *in vivo* or whether the transcript represents an anomalous splicing event occasionally occurring for these genes closely clustered on rat chromosome 10 (within 33 kb; Figure 16).

Multiple divergent IRG genes forming an independent phylogenetic clade (IRGH) were identified in the lancet fish. Thus, the IRG resistance system is at least as old as the cephalochordates (>550 million years; (Khalturin 2004; Kumar 1998)) and therefore pre-dates the development of an adaptive immune system in jawless and jawed vertebrates (Cooper 2006; Flajnik 2001). This argues against the proposed role in lymphomyeloid development (Feng 2004; Santiago 2005) being the primary function of IRGs proteins (see also below).

There are no convincing IRG homologues in tunicates, sea urchins or other invertebrates. Two IRG similar sequences recovered from *Caenorhabditis elegans* miss many of the IRG protein characteristic features including some of the conserved GTP-binding motifs (accession numbers: W09C5, C46E1.3; for alignment see (Bekpen 2005a)). Several groups of 45-50 kDa putative GTPases of unknown function with sequence features reminiscent of IRG GTPases, however, exist in bacteria (e.g. *Cyanobacteria*, see also (Bekpen 2005b; Martens 2006)). These bacterial enzymes show a plausible homology to the chordate IRGs in the G-domain, which is located within the proteins at roughly the same, characteristic position. Secondary structure predictions analysis indicated that the bacterial IRG-like proteins have a secondary structure similar to Irga6 (Howard unpublished data; (Ghosh 2004)). These observations raise the possibility that the IRG genes in the euchordates may originate from horizontal transfer of bacterial IRG-like genes. Biochemical or structural studies are, however, necessary to clarify the validity of the bacterial genes as members of the IRG family.

In summary, GKS IRG proteins are an ancient family at least as old as the euchordates (>550 million years; (Khalturin 2004; Kumar 1998)), while GMS (IRGM) GTPases seem to be specific for the Eutheria. Hence, IRGM genes must have already been present in an early eutherian progenitor some 75 million years ago (Bininda-Emonds 2007). The IRG resistance GTPases underwent rapid expansion and diversification as well as contraction in the euchordates probably due to host-pathogen coevolution. This feature has frequently been documented in multigene families associated with pathogen resistance in both animals and plants (Angata 2004; Borghans 2004; Delarbre 1992; Hood 1975; Klein 1986; Leister 2004; Mashimo 2003; Noel 1999; Trowsdale 2001). In these cases, however, the resistance mechanism itself has been retained as its protein mediators have evolved or even, in case of the natural killer cell receptor case, been replaced by a different molecular species (Barten 2001). The IRG case may be different because here the resistance mechanism itself has apparently been lost during primate evolution. It is a key question, why this ancient resistance system became dispensable and was lost during divergent evolution of the primate lineage, despite its importance for mouse resistance against pathogens infecting both rodents and primates.

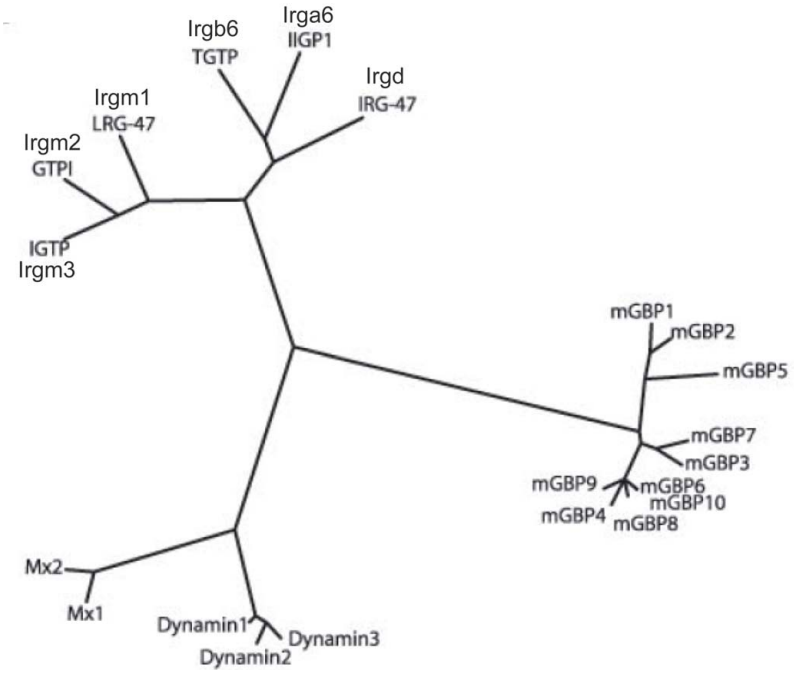


Figure 57 Phylogenetic tree of mouse GBPs, Mx proteins, dynamins, and selected IRGs. Unrooted maximum likelihood phylogenetic tree generated with Treepuzzle (www.tree-puzzle.de) based on a full length protein alignment generated with ClustalW (Thompson 1994). Branch lengths are measured relative to the estimated number of substitutions. Note that the IRG proteins display a markedly higher degree of sequence divergence than the other GTPases (from (Degrandi 2007)).

Since some of the murine GBP proteins have very recently been shown to target the *Toxoplasma gondii* parasitophorous vacuole similarly to the IRG proteins (Degrandi 2007) and the GBP family is conserved in humans, one could argue that the GBPs fulfil the resistance function of IRG proteins in humans. This hypothesis surely deserves experimental testing. It has, however to be noted, that the sequence divergence among the GBP family is much lower than within the IRG family (Boehm 1998; Degrandi 2007; Olszewski 2006) (Figure 57). Thus, the IRG family is either older or more rapidly

evolving than the GBP family or both. In view of this, it seems less likely that the GBPs are involved in direct interaction with pathogen-derived components and that they target a comparably broad pathogen range. Better understanding of the mechanism of action and regulation of the IRG proteins is needed before their complex evolutionary history can be put in context.

IV.5. IRG proteins in cell proliferation and survival

Most of the IRG proteins (Irga6, Irgc, Irgm1, Irgm2 and Irgm3) did not alter cell proliferation and survival when expressed in the absence of IFN (Figure 27). In contrast, both Irgb6 and Irgd expression resulted in a severe growth arrest and cell death. In the case of Irgb6, cells were shown to be blocked in cell cycle progression before the G2/M phase (Figure 28). A simple integration effect as explanation for these effects was ruled out by the repetition of the stable transfection of Irgb6. Thus, this study documents at least three independent events of IRG protein mediated inhibition of cell proliferation (two stable transfections for Irgb6, one for Irgd). Furthermore, similar effects were observed by Carlow and colleagues for Irgb6 (Carlow 1998). They reported problems in generating stable clones constitutively expressing Irgb6 and the few clones obtained showed a reduced growth rate and lost expression over time. It is easily imaginable how the severe fibre-like Irgb6 aggregates that entangle many cellular organelles in absence of IFNγ would negatively influence cell proliferation and survival. Yet, co-induction with IFNγ did resolve the majority of these aggregates (Figure 38) but did not lead to a rescue of the growth inhibition in any way (Figure 27). Saturation of the IFN system by overexpression of Irgb6 following double induction is unlikely to play a role, since Mifepristone-induced Irgb6 expression levels were very low (Figure 24). Irgd expressed in absence of IFNγ showed no signs of mislocalisation (Figure 32) but the expressing cells were completely growth inhibited. As for Irgb6, no beneficial effect of co-induction with IFNγ could be observed. In this case, however, expression of twice the normal amount of Irgd following double induction could play a role.

Taken together, the phenomenon of growth inhibition by Irgb6 and Irgd remains obscure. But given the number of independent events in which this effect has been observed and the absence of any other proposed mechanism of action for these two proteins, it is conceivable that Irgb6 and Irgd function by inducing the death of infected cells, killing the pathogen together with the host cell. Recent data indicate that death of *Toxoplasma gondii* infected, IFN-induced murine fibroblasts is indeed involved in elimination of the parasites (Zhao unpublished data).

IV.6. Covalent protein modification of Irga6

Upon expression of Irga6 in absence of IFNγ in mouse fibroblasts, a proportion of the protein is covalently modified by a yet unidentified moiety, indicated by a shift of about 8 kDa in apparent molecular weight (Figure 24). The observed protein modification did neither alter the hydrophobicity of Irga6 (Figure 25) nor the behaviour in size exclusion chromatography (Figure 39 see below). Both the modified 55 kDa (Figure 25) and the endogenous 47 kDa form of Irga6 were completely myristoylated *in vivo* (Figure 25,

(Martens 2004b)). Transient transfection of Irga6, in contrast, gave rise to considerable amounts of unmyristoylated protein, presumably due to overburdening of the cellular myristoyl transferase by overexpression (Figure 25, see also (Papic 2007)). Thus, efficient myristoylation of Irga6 is independent of IFNγ but dependent on the expression level of the protein. Additional modifications altering the protein hydrophobicity can be excluded to be responsible for the shift in apparent molecular weight of Irga6.

The same modified form of Irga6 was present to a somewhat lesser extent upon co-induction of ectopically Irga6 expressing cells with IFNγ and was observed only in trace amounts for endogenous protein in IFNγ-induced cells (Figure 25). Irga6 is predominantly aggregated in the GTP-bound, active state when expressed ectopically in absence of IFNγ and partially so in the presence of IFNγ, whereas endogenous, IFN-induced Irga6 is predominantly in the GDP-bound, inactive conformation due to negative regulation by the three GMS proteins (Figure 35). Furthermore, Irga6 accumulating at the parasitophorous vacuolar membrane of *Toxoplasma gondii* in infected cells is also GTP-bound (Figure 49). One could therefore hypothesise that the protein modification is linked to Irga6 activation and might function in resistance, for instance by facilitating translocation to the parasitophorous vacuole either directly by influencing Irga6 trafficking or indirectly by shifting the ER-cytoplasm equilibrium of Irga6 more towards the soluble state (Figure 58). Consistent with this idea, the modification did not occur on the nucleotide binding deficient, inactive mutant Irga6(S83N) (Figure 36; (Kaiser 2005)). The protein modification might prevent negative regulation of Irga6 by the GMS proteins upon infection-induced activation by directly or indirectly hindering interaction of these proteins. This hypothesis would also explain why Irga6 and two of its negative regulators, Irgm2 and Irgm3, can co-exist on the PV (see below) and still exert their function in pathogen resistance. An alternative explanation for this phenomenon would be that GKS protein accumulation at the PV precedes the arrival of the GMS proteins. Recent analysis, however, showed that both Irga6 and Irgm2 accumulation can occur within minutes of infection but the relative timing of these events has not been explored (Zhao unpublished data). The IRG resistance system could then be switched back to the inactive state by removal of the protein modification on Irga6 (and maybe also other GKS IRG proteins) allowing access of the GMS proteins that are tightly packed together with Irga6 on 30-40% of the PVs. Alternatively, the reaction could of course be terminated by degradation of the IRG proteins accumulated at the PV together with the parasite by autophagy or related mechanisms (Ling 2006). Thus, the covalent modification of Irga6 documented here could be of central importance in pathogen resistance mediated by IRG proteins. Therefore, it is of great interest to identify the type of modification and to analyse whether this modification also accompanies Irga6 activation in the context of *Toxoplasma gondii* infection. Ubiquitin and ubiquitin-like proteins such as SUMO (small ubiquitin-related modifier), ISG15 (interferon-stimulated gene 15), NEDD8 (neural precursor cell expressed, developmentally down-regulated 8) and FAT10 (HLA-F adjacent transcript 10) (Welchman 2005) are possible candidates as they are all between 6 and 18 kDa in size and thus could induce a shift in apparent molecular weight in SDS-PAGE in the same range as was observed experimentally for modified Irga6.

Monoubiquitination as a candidate for modification of the 55 kDa form of Irga6 was tested experimentally. For technical reasons, this modification could unfortunately neither be confirmed nor be ruled out. Considerable amounts of ubiquitin positive, high molecular weight species were, however, detected in immunoprecipitation of Irga6 from Mifepristone- but not from IFNγ-induced gs3T3-Irga6 cells in absence of proteasome inhibitors (Figure 26). At least a part of the ubiquitinated protein was shown to be Irga6. Thus, Irga6 is ubiquitinated when it is in the GTP bound state in absence of IFNγ, though co-immunoprecipitation of a tightly bound ubiquitinated interaction partner of Irga6 may contribute to the observed ubiquitin signal. IFNγ-induction in parallel with ectopic expression of Irga6 reduced the amount of ubiquitinated protein somewhat (Figure 26). This is consistent with the incomplete restoration of the GDP-bound state and wild type intracellular localisation under these conditions (Figure 35). Interestingly, a subset of the GTP-bound Irga6 aggregates, formed due to ectopic expression, was found to be positive for ubiquitin in immunofluorescence (Kaiser 2005). Despite ubiquitination of ectopically expressed Irga6 in Mifepristone-induced gs3T3-Irga6 cells, the protein half-life was comparable to that observed for endogenous Irga6 in IFNγ-induced mouse embryonic fibroblasts (~16 hrs; Steinfeld unpublished data). Thus, Irga6 might be subject either to monoubiquitination at multiple lysine residues or to lysine 63-linked polyubiquitination, both of which modulate protein function rather than mediate degradation (Pickart 2004). Whether mono- or polyubiquitination plays a role in the regulation of endogenous, IFN-induced Irga6 protein during infection remains to be determined. Interestingly, proteins with deubiquitinating activity have been identified in several virus families including the *Herpesviridae*, in pathogenic bacteria like *Chlamydia trachomatis*, *Salmonella typhimurium* and *Yersinia* as well as in eukaryotic pathogens such as *Plasmodium falciparum* and *Toxoplasma gondii* (Love 2007). In many cases these enzymes have been implicated in host cell manipulation and immune evasion, not least because the machinery for ubiquitin conjugation is missing from non-eukaryotic pathogens.

IV.7. IRG localisation is regulated by IFN and nucleotide

The intracellular behaviour of the seven IRG proteins analysed in this study is summarised in Table 10. The three GMS proteins Irgm1, Irgm2 and Irgm3 localised to the Golgi plus endolysosomes, Golgi and ER, respectively, in IFNγ-induced cells (Figure 29-31, (Martens 2006; Martens 2004b; Taylor 1997), Zhao unpublished data). These localisations were independent of other IFN-inducible factors and nucleotide binding in gs3T3 cells (Figure 29-31), as described previously for Irgm3-GFP in HeLa cells (Taylor 1997) and for native Irgm1 in L929 fibroblasts (Martens 2004b).

Intracellular localisation of Irgc and Irgd was determined in immunofluorescence for the first time in this study. As for Irgd no appropriate immunoreagent was available and Irgc was exclusively expressed in haploid spermatids, the localisation was analysed in mouse fibroblasts ectopically expressing ctag1-tagged Irgd and native Irgc, respectively. Endogenous Irgc and Irgd are predominantly cytosolic and only partially membrane associated according to analysis of hypotonic lysates of mouse fibroblasts and seminiferous tubules, respectively (Martens 2004b; Rhode 2007). Consistent with this observation, both Irgc and Irgd displayed a distributed reticular staining in

immunofluorescence when expressed by transient transfection (Figure 32). In absence of any other conspicuous signal, it is likely that the membrane-associated part of the proteins locates to the ER. This interpretation was supported by partial colocalisation of the Irgd and Irgc with the ER localised Irga6 protein in IFN-induced cells (Figure 32). The presence of Irgd in the nucleus of transfected cells is consistent with the diffusion of the largely soluble protein through the nuclear pores, which are permeable for proteins smaller than 60 kDa (Paine 1975). Both Irgc and Irgd localisation were independent of IFNγ-induction (Figure 32).

Irga6 localised in a distributed reticular pattern in IFNγ-induced cells (Figure 33, Figure 35), and is largely ER associated (Martens 2004b). Localisation of endogenous Irgb6 could not be determined due to the lack of an appropriate immunoreagent but analysis of hypotonic lysates suggested that the protein is predominantly cytosolic (Martens 2004b). Native Irga6 and FLAG-tagged Irgb6 expressed in absence of IFNγ resulted in the formation of large intracellular aggregates (Figure 33, Figure 38). The formation of aggregates was due to the absence of IFN-induced regulatory factors, as costimulation of cells ectopically expressing the IRG proteins with IFNγ largely restored the wild type localisation of Irga6 (Figure 33, Figure 35). Similarly, Irgb6-FLAG localised in a distributed reticular pattern under these conditions, presumably reflecting the wild type localisation of endogenous Irgb6 (Figure 38). The beneficial effect of interferon was clearly dose dependent, as the redistribution of Irga6 was much more complete for Mifepristone-induced gs3T3-Irga6 cells (Figure 33, Figure 35) than following transient transfection, which results in a much higher expression level in most of the cells (Papic 2007). In transient transfection, Irga6 wild type localisation was observed only in cells expressing relatively low amounts of the protein (Papic 2007).

		Localisation in uninfected cells				Localisation in *Toxoplasma* infected cells				
		compartment	IFNγ -dep.	GMS -dep.	Nucl.- dep.	PV asso- ciation	IFNγ- dep.	IRG- dep.	GMS -dep.	Nucl.- dep.
GMS	Irgm2	Golgi	no	no	no	yes	yes	yes	nd	yes
	Irgm3	ER	no	no	no	yes	yes	yes	nd	yes
	Irgm1	Golgi/endolysomes	no	no	no	no	----	----	----	----
GKS	Irga6	ER	yes	yes	yes	yes	yes	yes	yes	yes
	Irgb6	ER (Cytosol)	yes	yes	yes	yes	no	no	no	yes
	Irgd	Cytosol (ER)	no	nd	nd	yes	yes	yes	no	nd
	Irgc	Cytosol (ER)	no	no	nd	yes	yes	yes	yes	nd

Table 10 Summary of the intracellular behaviour of IRG proteins. Table summarising the intracellular behaviour of the seven IRG proteins analysed in this study in uninfected and *Toxoplasma gondii* (strain ME49) infected cells (nd: not determined, dep: dependent, nucl: nucleotide).

Irga6 expressed in absence of IFN did not accumulate in aggresomes, a place in the cell where misfolded proteins accumulate, as the observed aggregates neither fulfil the morphological critera nor the characteristic of detergent insolubility (TritonX-100 (Figure 40), Thesit (Papic 2007) and NP-40 (Kaiser 2005)) (Garcia-Mata 1999; Garcia-Mata 2002; Johnston 1998). The absence of Irga6 (and Irgb6) aggregates upon coinduction with IFNγ despite the expression of twice the amount of Irga6 protein further argued against the involvement of aggresomes (Figure 33, Figure 35, Figure 38). In an

independent study (Kaiser 2005), Irga6 aggregates following ectopic expression were shown also to lack other characteristic of aggresomes, including association with the microtubule organisation centre and restructuring of the intermediate filament cytoskeleton (Garcia-Mata 1999; Garcia-Mata 2002; Johnston 1998). The morphology of Irgb6 aggregates formed upon ectopic expression was also not consistent with localisation in aggresomes (Figure 38). The vimentin containing intermediate filaments were shown to form cage-like structures in HeLa cells expressing high amounts of Irgb6-FLAG following transient transfection but Irgb6 itself was not aggregated inside these cages where aggresomes would be localised (Martens 2004a).

IV.8. Irga6 and Irgb6 aggregates formed in absence of IFNγ represent GTP-bound homomers

Irga6 – and by analogy probably also Irgb6 – self-activates by GTP-dependent oligomerisation *in vitro* (Figure 36, (Uthaiah 2003)). Thus, it was conceivable that the Irga6 and Irgb6 aggregates observed in cells in the absence of IFNγ represented GTP-bound homo-oligomers. Indeed, the formation of the intracellular aggregates was nucleotide dependent, as the nucleotide binding deficient mutants of Irga6(S83N) and Irgb6(S70N) did not aggregate in absence of IFN (Figure 37-38). By contrast, the Irga6(K82A) mutant, which is probably locked in the GTP-bound, active state (as it binds but cannot hydrolyse GTP (Figure 36)), and the analogous Irgb6(K69A) mutant developed aggregates even in IFNγ-induced cells (Figure 37-38). IFNγ failed to resolve the aggregates formed by the two Gx_4GAS mutants either because incipient oligomerisation events cannot be terminated by GTP hydrolysis or because GMS proteins are unable to interact normally with the mutants, as indicated by the altered interaction profile in yeast 2-hybrid assays (Figure 44).

That the aggregated Irga6 protein in absence of IFNγ-induction was in the GTP-bound, active conformation was confirmed by immunofluorescence with the 10D7 monoclonal antibody specific for this conformation (Figure 35, for proof of 10D7 conformation specificity see (Papic 2007)). Irga6 expressed in absence of IFNγ was efficiently recognised by this antibody, while endogenous, IFN-induced Irga6 as well as Mifepristone-induced Irga6 localised normally due to co-induction with IFNγ were not detected (Figure 35). Only the residual aggregates present due to the expression of a double dose of Irga6 that exceeds the normalising capacity of the IFN-induced GMS proteins, were 10D7 positive (Figure 35). Thus, Irga6 is in the active, GTP-bound conformation in absence of IFNγ but in the inactive conformation in IFN-induced cells. In the latter case Irga6 is presumably bound to GDP due to the protein's higher affinity for GDP than for GTP (~7-15 fold, Figure 36, (Uthaiah 2003)) and the cellular concentration of the respective nucleotides (300 μM GTP and 100 mM GDP (Kleineke 1979)). Consistent with this idea, addition of aluminium fluoride, which traps Irga6 oligomers *in vitro* by stabilising the transition state of GTP-hydrolysis, trapped Irga6 homodimers in cells expressing Irga6 in absence but not in presence of IFNγ (Papic 2007).

IV.9. GTP-bound homomeric Irga6 and Irgb6 complexes dissociate *ex vivo*

Aggregate formation of Irga6 and Irgb6 protein in absence of IFNγ was not reflected in an altered running behaviour in analytical size exclusion chromatography (Figure 39 A-E). Furthermore, the aggregates were completely soluble in several non-ionic detergents (TritonX-100 (Figure 40), NP-40 and Thesit (Kaiser 2005; Papic 2007)). Thus, the GTP-dependent homomeric complexes formed *in vivo* (Figure 33 B, Figure 38 A) dissociated in cell lysates *ex vivo,* presumably due to hydrolysis of GTP (Figure 36). Irgb6 and Irgd behaved as a monomer in size exclusion chromatography (Figure 39 C-F). The elution profiles of the GMS proteins were consistent with the formation of dimeric complexes probably representing homodimers, as the elution profiles of the three GMS proteins did not overlap completely (Figure 39 G-I). Furthermore, at least for Irgm2, the complexes were formed in absence of other IFN-inducible proteins. Interactions with other, non-IFN-inducible factors cannot be formally excluded.

The apparent molecular weight range (50-200 kDa) displayed by cellular Irga6 in gel filtration independent of IFNγ (Figure 39 A-B) suggested that the protein was in equilibrium between the monomeric and the tetrameric state. The high molecular weight species of the protein were also observed for recombinant protein and were dependent on the presence of myristoyl modification (Papic 2007; Uthaiah 2002). Thus, the myristoyl group apparently induces major conformational changes that either strongly influence the apparent molecular weight of the protein or favour homomeric interaction.

IV.10. The three GMS proteins negatively regulate Irga6 and Irgb6

In view of the complex representation of the IRG proteins in the mouse (Figure 12), their extensive functional non-redundancy in pathogen resistance (Martens 2006) and their massive expression following IFN-induction (Bekpen 2005b; Boehm 1998; Martens 2004b), it was conceivable that IRG proteins themselves are the IFNγ-dependent regulators of Irga6 and Irgb6. Coexpression of the three GMS proteins, indeed, restored the wild type localisation of Irga6 and Irgb6 in over 80-90% of the cells (Figure 41-43). The GKS proteins tested (Irgd and Irgb6 or Irga6) did not mediate any normalising effect (Figure 42 and data not shown). Except for Irgm1 in the case of Irgb6, each of the three GMS proteins partially restored the wild type localisation of Irga6 and Irgb6, but only the coexpression of all three GMS proteins together mediated an efficient relocalisation (Figure 41, Figure 43). Thus, the GMS proteins cooperatively mediate normal localisation of Irga6 and Irgb6. There are two possible explanations for this cooperativity. Either the individual GMS proteins regulate Irga6 and Irgb6 at different subcellular compartments or the formation of heteromeric IRG protein complexes is necessary for efficient control of Irga6 and Irgb6 (Irgm2 and Irgm1 both mainly localise to the Golgi while Irgm3 is predominantly ER associated, (Figure 29-31). The role of Irgm1 in regulation of Irgb6 is difficult to interpret in absence of more data concerning the interactions of IRG proteins, as Irgm1 alone was inert but had an either beneficial or negative effect when coexpressed with Irgm2 or Irgm3, respectively (Figure 43). The formation of competing heteromeric interactions, however, is clearly one possible

explanation, supported by distinct interaction profiles of the four involved IRG proteins with other family members in Y2H (Figure 44).

The normalising effects of Irgm1, Irgm2 and Irgm3 were strictly nucleotide dependent, as the three GMS to GMN mutants failed to influence Irga6 and Irgb6 localisation (Figure 41-43). The P loop serine to asparagine mutant of Irgm3 (S98N), like Irga6(S83N) (Figure 36) and many other GTPases (Feig 1988; Nuoffer 1994; Praefcke 2004a; Stenmark 1994), was shown to be deficient in nucleotide binding (Taylor 1997). Thus, the analogous mutants of the other two GMS proteins are probably also nucleotide free (Irgm1(S90N), Irgm2(S78N)).

The beneficial effect of the GMS proteins was clearly dose dependent. Transfection of larger DNA amounts of GMS protein expression constructs had a stronger effect on Irga6 relocalisation (Figure 34). In the case of Irgb6, the beneficial effect of IFN-induction was much smaller (data not shown) than the effect of cotransfecting the three GMS proteins together with Irgb6-FLAG, which leads to a uniformly high expression of all the transfected proteins. Thus, the expression levels of all the involved IRG proteins have to be balanced to prevent Irga6 and Irgb6 aggregation. Consistent with these data, Irga6 and Irgb6 were aggregated in Irgm1-deficient mouse embryonic fibroblasts (Könen-Waisman unpublished data).

These findings have important implications for the interpretations of the susceptibility phenotypes of GMS-deficient mice. Since the regulation of Irga6 and Irg6 by the three GMS proteins is carefully balanced and the IRG proteins seem to function in a complex network of nucleotide-dependent homo- and heteromeric interactions that presumably compete with each other (Figure 44-46), the system might be put completely out of balance by depletion of one family member. The GMS proteins are probably especially vulnerable to such an effect considering their extraordinary importance in regulation of Irga6 and Irgb6 and the presence of only three GMS genes as opposed to 16 intact, immune-type GKS coding units, including six Irga6- and eight Irgb6-related sequence (Figure 12). The non-redundancy of IRG proteins in pathogen resistance observed in mice and cells deficient for single members of the IRG family might therefore actually reflect an unbalancing of the whole IRG resistance system. Thus, the observed susceptibility phenotypes may not necessarily be due to the non-redundant activity of single IRG members. In agreement with this, the two GMS-deficient mice analysed thus far displayed the broadest range of susceptibilities (Table 1). One could even postulate that GMS proteins are solely regulators in the IRG resistance system, while the GKS proteins constitute the – partially redundant (Martens 2005) - effector molecules. Consequently, all pathogen susceptibilities identified to date only for GMS-deficient mice would be attributable to certain GKS proteins not yet tested for involvement in resistance to these particular pathogens.

IV.11. Direct, nucleotide-dependent interactions of IRG proteins

Functional interactions of IRG family members take place *in vivo* since the three GMS GTPases were both necessary and sufficient to regulate the nucleotide status and the intracellular localisation of Irga6 and Irgb6 (Figure 35, 37, 38, 41-43). A Y2H analysis performed with the IRG proteins revealed that numerous of these interactions

are direct. Strong homomeric interactions were documented for Irga6 and Irgb6 (Figure 44) in line with the biochemical and cell biological data presented and discussed above. Furthermore, Irga6 interacted with Irgb6, which fits nicely with the observation that Irga6 colocalised with severe Irgb6 aggregates when the two proteins were expressed together in mouse fibroblasts in absence of IFNγ (data not shown). GMS-GMS protein interactions were also observed, namely Irgm1-Irgm3, Irgm2-Irgm3 and Irgm1 homomeric interaction. Furthermore, two of the GKS-GMS regulatory interactions observed in mammalian cells were reflected in Y2H: Irgb6 interacted with Irgm2 and Irga6 with Irgm3. Interestingly, all the interactions documented for Irga6 in Y2H were independent of the myristoyl group, since the N-terminal fusion with the Gal4 AD and BD obliterates the myristoylation motif. This is in line with the G-domains forming the interaction interface for Irga6 homomeric (Pawlowski unpublished data) as well as for regulatory GKS-GMS interactions (Figure 46).

All the observed interactions were strictly nucleotide dependent as they were completely abolished if one of the interaction partners carried the G1 motif serine to asparagine mutation interfering with nucleotide binding. As the Irga6(S83N) mutant loses affinity for both nucleotides, while Irga6(K82A) preserved wt affinity in both cases, GTP- and GDP-dependent processes could not be distinguished in the Y2H system (Figure 36). Mutants exclusively binding GDP were not identified thus far (Pawlowski unpublished data). In case of the Gx_4GKS to Gx_4GAS mutants of Irga6 and Irgb6, some of the interactions observed for the wild type proteins were preserved while others were lost, again suggesting that the two mutants possess an altered conformation that may hinder interaction with other proteins. Potential formation of homo-oligomers inside the yeast cells might also effectively reduce the free pool of these two mutants available for interaction with other proteins. Furthermore, the interactions observed in Y2H were independent of membrane association, as activation of the reporter genes requires translocation of the interacting proteins into the nucleus as well as binding to the respective promoters. Together with the data from mouse fibroblasts, this may suggest that interactions between IRG proteins are not restricted by conformational changes induced by membrane association in the uninfected cell but rather by spatial exclusion of certain interactions due to distinct subcellular localisations of the involved proteins.

One of the limitations of the yeast two-hybrid approach used is that heteromeric complexes formed by more than two proteins cannot be detected (McAlister-Henn 1999). This could be one of the reasons why Irgm1 showed only rather weak interactions in Y2H and Irgd non at all (Figure 44). Another likely explanation for these effects is the structural constraints imposed on the IRG proteins by N-terminal fusion to the Gal4 DNA binding and activating domains. Altered behaviour of tagged and fusion constructs of the IRG proteins has been observed previously in mammalian cells as well as with recombinant protein *in vitro* (Martens unpublished data (Martens 2004a; Uthaiah 2002)). Furthermore, hybrid proteins may not efficiently move into the nucleus and may not carry post-translational modifications necessary for proper function (McAlister-Henn 1999).

An independent Y2H study of the wild type IRG proteins based on the LexA DNA-binding domain and the B42 activation domain documented partially overlapping

but not identical interactions, including Irgd-Irgm2 and -Irgb6 interaction (Kaiser 2005). Interestingly, one of the newly identified IRG genes, Irgb10, was shown to interact with Irgm2, Irga6 and Irgd in that Y2H assay (termed cIGP9 in that study (Kaiser 2005)). Thus, different expression conditions highlight distinct interactions as reported in the literature (Van Criekinge 1999). Together this suggests that the nucleotide dependent, direct interactions documented by Y2H in this study probably represent a solid basis, while more interactions within the IRG family remain to be discovered.

IV.12. Direct, GDP-dependent interaction of Irga6 with Irgm3 occurs via the G-domain

As a representative of the documented direct, nucleotide-dependent regulatory GKS-GMS protein interactions, Irga6 interaction with Irgm3 was confirmed by coimmunoprecipitation from IFNγ-induced mouse fibroblasts. Irgm3 co-precipitated with Irga6 in absence of exogenous nucleotides (Figure 45). As the regulatory effect observed in fibroblasts as well as the interaction of Irga6 with Irgm3 seen in Y2H were nucleotide-dependent (for both partners in case of the Y2H), this presumably reflects the pre-existence of heterodimeric complexes *in vivo* containing trapped nucleotides. The amount of co-immunoprecipitated Irgm3 was markedly sub-stoichiometric relative to Irga6. This might indicate the transient nature of the Irga6-Irgm3 interaction *in vivo* with only a minor proportion of Irga6 bound to Irgm3 at any given time. Alternatively, the complex could be of limited stability, so that only a minor proportion of the Irgm3 associated with Irga6 inside the cells would still be present at the end of the immunoprecipitation procedure. Since depletion of bivalent cations by addition of EDTA did not prevent Irga6-Irgm3 interaction, nucleotide binding for Irgm3, as shown for Irga6 *in vitro* (Figure 36), must also be largely independent of Mg^{2+} (Figure 55). Addition of exogenous GTPγS to the reaction, which was shown to favour homomeric interactions of cellular Irga6 *ex vivo* in a similar setup (Papic 2007), abolished the co-precipitation of Irgm3. This suggests that both partners are in the GDP-bound state in the Irga6-Irgm3 interaction. The amount of co-immunoprecipitated Irgm3 was, however, noticeably decreased upon addition of exogenous GDP. This could be either due to the fact that one of the interaction partners, presumably Irgm3, has to be in the GTP-bound state or to competing hetero- and homomeric interactions with other IRG proteins present in the cell lysate.

The nucleotide status of the Irga6-Irgm3 interaction was further clarified in pull down experiments employing purified GST-Irga6 to capture cellular Irgm3. A weak but specific nucleotide independent affinity of Irgm3 for Irga6 was observed that was not detected with GST protein alone (Figure 46). The absence of an efficient Irga6-Irgm3 interaction without exogenous nucleotide in pull down as opposed to co-immunoprecipitation supports the presence of preformed cellular Irgm3-Irga6 complexes with trapped nucleotide in the latter case. Efficient Irgm3 pull down from IFNγ-induced cells as well as from cells ectopically expressing Irgm3 in absence of IFN was obtained only in the presence of GDP. Thus, the Irgm3-Irga6 interaction is independent of other IFN-induced factors and probably requires both proteins to be GDP bound, though it cannot be formally excluded that endogenous GTP was trapped in the cellular Irgm3

protein. It has been reported that Irgm3 immunoprecipitated from cells predominantly coprecipitated GTP (Taylor 1997). However, considering the high concentrations of exogenous GDP present, it seems unlikely that Irgm3 could retain GTP. Furthermore, exogenous GTPγS as well as a mixture of GDP and GTPγS reflecting physiological concentrations (Kleineke 1979) both did not promote Irga6-Irgm3 interaction. The GDP-dependent precipitation of Irga6 was abolished when the Irga6(S83N) mutant was used. Thus, Irgm3 and Irga6 interact in cells and cell lysates whereby Irga6 and most likely also Irgm3 are GDP bound.

Several non-Ras GTPases have recently been shown to accelerate GTP hydrolysis by G-domain-G-domain dimerisation (Egea 2004; Focia 2004; Ghosh 2004; Scrima 2006; Smith 2002; Sun 2002), the two G-domains functioning effectively as GAP proteins for each other. Such mutually activating GTPase pairs may be heterodimeric as for members of the signal recognition particle (SRP) family (Focia 2004) or homodimeric as shown for hGBP1 (Ghosh 2006) and Irga6 (Pawlowski unpublished data). Mant-labelled GTP interferes with oligomerisation of Irga6 *in vitro* by sterically hindering the formation of the essential G-domain-G-domain interaction interface that normally involves the bound nucleotides (Pawlowski unpublished data). The G-domain-G-domain interaction interface was shown here to be essential for the Irga6-Irgm3 interaction as well, since addition of mGDP to the pull down reaction reduced the interaction to the basal level (Figure 46).

Thus, Irgm3 may negatively regulate Irga6 via GDP-dependent, heterodimeric G-domain-G-domain interaction by successful competition with other Irga6 molecules for this interaction interface and by hindering GTP uptake, thus acting in this sense as guanine nucleotide dissociation inhibitor (GDI). The same principles may apply for the other GMS proteins and Irgb6.

IV.13. Regulatory IRG interactions in *Toxoplasma* infection

All IRG proteins analysed so far, namely Irga6, Irgb6, Irgd, Irgm1 and Irgm3, were shown to contribute to the IFN-mediated resistance to the protozoan parasite *Toxoplasma gondii* whereby the two GMS proteins play a much more important role than the tested GKS proteins (reviewed in (Martens 2006; Taylor 2007), Martens, Parvanova unpublished data). The mediated resistance was shown to be cell autonomous for Irgm1, Irgm3 and Irga6 (Butcher 2005; Halonen 2001; Ling 2006; Martens 2005). The mechanism of the anti-*Toxoplasma* activity conferred by the IRG proteins is poorly understood but involves disruption of the membrane of both the parasitophorous vacuole and the parasite by vesiculation (Ling 2006; Martens 2005). Both Irga6 and Irgm3 as well as subsequent autophagic events have been implicated in this process (Ling 2006; Martens 2005). Five IRG proteins (Irga6, Irgb6, Irgd, Irgm2 and Irgm3) were previously shown to accumulate at the parasitophorous vacuole in *Toxoplasma gondii* infected IFNγ-induced primary astrocytes (Martens 2005). For Irgm3, the vacuolar localisation was confirmed in primary macrophages (Ling 2006). Irgm1 association with the *Toxoplasma* PVM could not be detected (Butcher 2005; Martens 2005).

In this study, the accumulation of Irga6, Irgb6, Irgd, Irgm2 and Irgm3 at the *Toxoplasma* PVM was generalised to mouse fibroblasts and was shown to be dependent

on IFNγ for all but one of these IRG proteins (Figure 47-53). Despite massive aggregation in the cell, ectopically expressed Irgb6 translocated to the PVM in absence of IFN (Figure 51), though somewhat less efficiently than in IFNγ-induced cells. Thus, accumulation of Irgb6 in the active, GTP-bound state in homooligomers in absence of the negative regulation of the GMS proteins still allows localisation to the vacuole. In contrast, Irga6 aggregated due to the lack of negative regulation by the GMS proteins was incapable of associating with the PV (Figure 47-48). Thus, activation of Irga6 prior to infection prevents PV association of the protein. This paradoxical difference between Irga6 and Irgb6 allows several interpretations.

First, Irga6 might be completely in the GTP-bound, oligomeric state following ectopic expression while a part of the Irgb6 protein might remain nucleotide free or GDP bound. This might be due to the larger cytosolic pool of Irgb6 and the therefore lower local concentration of the protein. Alternatively, Irgb6 could possess different biochemical properties than Irga6 (e.g. lower affinity for GTP or higher GTP-hydrolysis rate). However, the similar behaviour of the two proteins in most analyses rather suggests the contrary. Second, Irgb6 could be a pioneer in colonisation of the parasitophorous vacuole directly recognising features of and binding to the PVM (see also chapter IV.15), while other family members might require the presence of Irgb6 at the vacuole to efficiently accumulate there. The second hypothesis is consistent with the observation that cotransfection of Irgb6, Irgd, Irgm1, Irgm2 and Irgm3 but not of the three GMS proteins alone efficiently restored the vacuolar localisation of Irga6 expressed in absence of IFNγ (Figure 47-48). Furthermore, the percentage of Irga6 positive vacuoles was reduced by roughly one third relative to IFNγ-induced cells in the latter case, while the average intensity of Irga6 signal at the PVM was four fold reduced (Figure 48). Thus, efficient association of Irga6 (and presumably also other IRGs) with the PVM seems to require Irgb6 at the vacuole probably as initiator of an activating interaction promoting Irga6 oligomerisation. Coexpression of Irgd with Irga6 and the three GMS proteins also resulted in Irga6 amounts on the PVM similar to those following IFN-induction, suggesting that Irgd might be able to fulfil a similar function for Irga6. Thus, the presence of GMS proteins is sufficient to allow vacuolar association of Irga6 but other GKS proteins seem to be necessary for its efficient accumulation at the PVM. Alternative interpretations of that data are discussed in the next chapter (IV.15). As observed for Irga6 regulation in uninfected cells, the beneficial effect of the three GMS proteins on Irga6 was completely abolished for the three GMN mutants and is thus nucleotide dependent.

The Irga6 protein accumulating at the PVM was shown to be in the active, GTP-bound conformation by use of the conformation-specific 10D7 antibody (Figure 49). Furthermore, the association of Irga6 with the PVM required nucleotide binding, as Irga6(S83N) was incapable of doing so even in the presence of IFNγ (Figure 50). The lack of vacuolar association of Irga6(S83N) was observed in both Irga6-deficient as well as in wild type mouse embryonic fibroblasts. In the latter, the endogenous wild type Irga6 localised normally to the PVM. Thus, consistent with the documented inability of Irga6(S83N) to interact with wild type Irga6 in Y2H (Figure 44, confirmed in coimmunoprecipitation (Papic 2007)), the mutant did not function as a dominant negative. The Irga6(K82A) mutant failed to localise to the parasitophorous vacuole in

IFN-induced Irga6-deficient MEFs and was strongly impaired in doing so in IFN-induced wild type MEFs (Figure 50). Furthermore, this mutant acted as a dominant negative, almost completely preventing endogenous Irga6 from reaching the vacuole (Figure 50). Irga6(S83N) is not able to localise to the PVM, presumably due to its inability to bind GTP. In contrast, both wild type Irga6 and Irga6(K82A) were not able to relocalise to the PVM in absence of IFNγ and the latter also not in presence of IFN, probably due to the binding of GTP leading to activation and homooligomer formation prior to infection. Thus, Irga6 has to be in the inactive (presumably) GDP-bound state prior to infection and then needs to bind GTP in order to be able to form oligomers at the PVM. This process is regulated by competing activating (e.g. Irga6-Irga6) and inhibiting interactions (e.g. Irga6-Irgm3) of IRG family members with each other. That vacuolar localisation of Irga6(K82A) is not promoted by IFNγ induction suggests that this mutant cannot be regulated by the GMS proteins. On the one hand, this could be due to the altered conformation of the mutant protein, indicated by its changed interaction properties in Y2H as well as the anomalous binding of mant-GTP and the slow GTP-dependent oligomerisation *in vitro* (Figure 36, Figure 44). On the other hand, the inability of the mutant to hydrolyse GTP (Figure 36), trapping the protein in GTP-dependent G-domain-G-domain homodimers that cannot be resolved, might be the reason. As the activating self-interaction as well as the inhibitory Irga6-Irgm3 interaction relies on the same interface that includes the bound nucleotide, Irgm3 might only be able to compete with the Irga6 self-interaction upon dissociation following GTP hydrolysis. That Irga6(K82A) forms severe aggregates even in presence of IFN supports this latter hypothesis (Figure 37). As the endogenous Irga6 protein was found in aggregates together with Irga6(K82A) and the two forms of the protein interacted with each other in Y2H (Figure 44), Irga6(K82A) presumably exerts its dominant negative function by trapping wild type Irga6 protein in such inert complexes locked in the GTP-bound state.

Together, these results suggest that Irga6 has to be held in the inactive, GDP-bound conformation by interaction with the GMS proteins prior to infection and that Irga6 is activated upon *T. gondii* infection, accumulating at the PVM in the GTP-bound oligomeric form. The biochemical data as well as the Y2H analysis and the requirement of other GKS proteins for efficient Irga6 accumulation at the PV suggest that homooligomers as well as heterooligomers with other GKS proteins are probably formed at the vacuole. GTP-dependent interactions with other GKS proteins already present at the PV might be necessary to initiate efficient Irga6 homooligomerisation at the vacuole. Irgb6 is the prime candidate for such a pioneering oligomerisation-initiator, as the protein is able to translocate to the PV in absence of other IRG proteins (Figure 51). Irgb6 localisation to the PVM was nucleotide dependent. Similar to the respective Irga6 mutant, Irga6(S70N) did not accumulate at the vacuole (Figure 51). Irgb6(K69A), however, was still able to accumulate at the PVM rather efficiently in presence of IFNγ, but in contrast to the wild type protein did not do so in absence of IFNγ (Figure 51). Thus, Irgb6(K69A) cannot reach the vacuole autonomously. In view of the preserved interaction of this mutant with wild type Irgb6 and Irga6 documented in Y2H, Irgb6(K69A) might reach the vacuole by interaction with these proteins.

Irgd required the presence of IFNγ or the other five IRG proteins (Irgb6 Irga6, Irgm1, Irgm2 and Irgm3) for localisation to the PV, whereas the three GMS proteins

were not sufficient (Figure 52). Vacuolar localisation of both Irgm2 and Irgm3 depends on the presence of the five other IRG proteins and on nucleotide binding since the GMN mutants did not accumulate at the PVM (Figure 53 and data not shown). In case of Irgd, Irgm2 and Irgm3 the minimal requirements for vacuolar localisation were not analysed.

Taken together, these data indicate that the complicated network of nucleotide-dependent interactions observed in Y2H is of functional relevance for the IRG proteins in infected cells. Considering the complexity of the IRG system and the lack of data available concerning the molecular mechanisms of IRG function thus far, it is difficult to interpret many of the obtained results.

IV.14. Irgc – a protein in search of a function

Despite extensive analysis no phenotype could be documented for mice genetically deficient for Irgc r (Rhode 2007). Testis and sperm morphology as well as fertility were not impaired. In a Y2H screen, two potential interaction partners of Irgc, the elongation release factor 1 (eRF1) and a testis specific splice variant of the Golgi reassembly and stacking protein 2 (tsvGORASP2), were identified (Rhode 2007). The interactions were dependent on an intact nucleotide-binding site, but the functional implications of these potential interactions remain unclear.

Thus, the unexpected, IFNγ-dependent relocalisation of Irgc ectopically expressed in mouse fibroblasts to the *T. gondii* parasitophorous vacuole is, to date, the only documented 'function' of this protein (Figure 54). The six IRG proteins Irgm1, Irgm2, Irgm3, Irgd, Irga6 and Irgb6 were sufficient to mediate vacuolar localisation of Irgc, and so were the three GMS proteins alone, though with markedly lower efficiency in the latter case (Figure 54). The other IRG proteins probably mediate Irgc translocation to the PVM via direct nucleotide dependent interaction, as Irgc interacted with both Irga6 and Irgm3 in Y2H and intact nucleotide binding sites were required on both sides of the interaction (Figure 44). As the G-domains display the highest sequence conservation within the IRG family, it is conceivable that Irgc interaction with other IRG proteins occurs via the G-domain-G domain interface involved in Irga6 homomeric (Pawlowski unpublished data) and Irga6-Irgm3 heteromeric interaction (Figure 46, see also chapter IV.12). Whether the Irgc translocation to the *Toxoplasma* PVM reflects an *in vivo* function of Irgc in resistance or only documents conserved interaction properties within the IRG family remains to be determined by infection of Irgc-deficient mice with *Toxoplasma gondii.* Interestingly, both virulent and avirulent *T. gondii* strains were reported to disseminate efficiently and rapidly (within 4-5 days) to the testis and to replicate there in high numbers following intraperitoneal infection of BALB/c and C57BL/6 mice (Hitziger 2005). Despite the obvious need for immunity against *Toxoplasma* in the testis, several facts argue against Irgc functioning as a host resistance factor in this immunoprivileged region. First, Irgc is not inducible by interferons but is developmentally regulated paralleling sexual maturity (Rhode 2007). Second, Irgc is only expressed in haploid spermatids (Rhode 2007) and thus would leave the majority of the testis tissue unprotected. Third, the Irgc gene is highly conserved throughout the mammals (Figure 20) arguing for a conserved function rather than a function in innate immunity. Fourth, Irgc translocation to the *T. gondii* PVM was dependent on the

presence of other, IFN-inducible IRG proteins that are clearly absent from higher primates and maybe other animals like the pig (Table 7). Then again, one could argue that the GBP family of IFN-inducible GTPases which are conserved in humans and have recently been implicated in *Toxoplasma* resistance in mouse (Degrandi 2007), might facilitate Irgc vacuolar localisation in species where immune type IRG proteins are absent. A conclusion about Irgc function awaits further experimental evidence.

IV.15. Virulent *Toxoplasma gondii* inhibit IRG proteins

While at least 6 IRG proteins accumulated efficiently at the parasitophorous vacuoles containing *Toxoplasma gondii* of the avirulent ME49 strain, the PVs generated by the virulent RH strain were largely devoid of IRG proteins (Figure 55). The most drastic effect was observed for Irgb6 with a nearly 10-fold reduction in percentage of detectibly positive PVs, while the reduction for the other IRGs (Irga6, Irgd, Irgm2 and Irgm3) was much less pronounced (Figure 55 and data not shown). There was, however, a much stronger effect on the quantity of Irga6 protein accumulating at the PVM. Only the few remaining Irgb6 positive vacuoles (8%) acquired normal amounts of Irga6 protein, while most of the Irga6 positive vacuoles acquired only small amounts of protein (Figure 55). Thus, virulent *Toxoplasma* possess a yet unidentified mechanism to efficiently inhibit the accumulation of IRG proteins at the PVM. As the effect was most drastic for Irgb6, the pathogen presumably targets this IRG protein and the effects on the other IRG proteins are secondary. This hypothesis is supported by fact that Irgb6 is the only IRG protein that was able to target the PVM independent of other IRGs (Figure 51). Furthermore, a striking inclusion relationship was observed for IRG protein accumulation at the PV of avirulent *T. gondii:* all of the Irgm3 positive vacuoles were also positive for Irgm2, all of which were positive for Irgd, followed by Irga6 and Irgb6 (Könen-Waisman unpublished data. These results, together with the restoration of efficient Irga6 PVM localisation in cells expressing Irga6 and the three GMS proteins by the presence of Irgb6 (Figure 48, see also chapter IV.13), strongly argue for a pioneer role of Irgb6 in vacuolar targeting. The presence of Irgb6 at the parasitophorous vacuole would then promote association of other IRG proteins presumably by GTP-dependent heteromeric interactions.

Surprising in this context was, however, the observation that also the presence of Irgd enhanced the amount of Irga6 accumulation at the PVM in cells coexpressing Irga6 and the three GMS proteins (Figure 48). A possible explanation for this effect is that the addition of Irgd in this setup indirectly releases Irga6 from a too strong inhibition by the GMS proteins by competing with Irga6 for the interaction with these proteins. As the inhibitory effect of the GMS proteins was dose dependent in uninfected cells (Figure 41), it is conceivable that overexpression of the 3 GMS by transient transfection in cells expressing normal amounts of Irga6 following Mifepristone induction might lock the majority of the Irga6 protein in the inactive form. Consequently, transient transfection of the 3 GMS proteins into IFNγ-induced cells should inhibit Irga6 accumulation at the PVM. This could indeed be the case, since Bernstein-Hanley and colleagues (Bernstein-Hanley 2006) observed that absence as well as overexpression of Irgm3 reduced cellular resistance to *Chlamydia trachomatis*. The same argumentation made for Irgd could of

course also apply to Irgb6; however, the additional data listed above strongly support the pioneer role of Irgb6 in vacuolar colonisation.

Recent experiments generalised the IRG protein accumulation at the PVM to other avirulent strains, and the inhibition of Irgb6 and Irga6 accumulation to other virulent strains of *Toxoplasma* (Khaminets unpublished data). The inhibition of Irgb6 accumulation at the parasitophorous vacuole by virulent *Toxoplasma gondii* was shown to function locally in coinfection with virulent and avirulent strains (Khaminets unpublished data) but the mechanism remains unknown. Masking of potential target structures on the PVM and modulation of the nucleotide status of Irgb6 either directly or indirectly by unbalancing the complicated regulatory network of IRG protein interactions is, however, conceivable. Two rhoptry kinases, ROP16 and ROP18, have been identified as important virulence factors, accounting at least for part of the tremendous virulence differences of *Toxoplasma* strains (El Hajj 2006; Saeij 2006; Saeij 2007; Taylor 2006), and are therefore potential candidates for inhibitors of the IRG proteins.

Very recently, mGBP1, -2, -3, -6, -7 and -9 have also been shown to localise to the PV of avirulent *T. gondii,* while mGBP5 could not be detected at the vacuole. As for the IRG proteins, mGBP1 and mGBP2 accumulation at the PVM was inhibited by the virulent BK strain (Degrandi 2007). This adds another six large, IFN-inducible GTPases potentially cooperating in resistance to the already crowded vacuole.

IV.16. Model of IRG function in uninfected and infected cells

In view of the complex pattern of activating as well as inhibitory interactions between IRG family members, it is difficult to draw a conclusive picture about the molecular mechanisms governing individual IRG proteins in uninfected and *T. gondii* infected cells. Based on the data generated in this study, however, it is for the first time possible to formulate a model of how the IRG resistance system might function as an entity (Figure 58). In view of the complexity of the situation this model is primarily based on the data obtained for Irga6, Irgb6 and the three GMS proteins.

In this model, GMS proteins negatively regulate self-activating GKS proteins by formation of GDP-dependent G-domain-G-domain interactions at endomembranes in absence of infection, thus preventing nucleotide exchange and premature activation. Upon infection with avirulent strains of *Toxoplasma gondii,* the inhibition by the GMS proteins is released and the GKS proteins translocate to the membrane of most but not all parasitophorous vacuoles, where they form GTP-bound, homo- and heterooligomers via the same G-domain-G-domain interface, promoting vacuolar destruction. The presence of GTP-bound, active GKS proteins at the PVM promotes the further accumulation of these proteins due to activating heteromeric interactions. It is not clear what triggers the translocation to the vacuole and how the IRG proteins get there, but microtubules are not involved (Khaminets unpublished data). The simplest possibility is diffusion from the GDP-bound, cytosolic pool of the GKS GTPases followed by spontaneous activation on the PV membrane that is not efficiently protected by GMS proteins. A specific pathogen derived activation signal could, however, also be involved. Furthermore, the GKS proteins could travel to the vacuole in dimeric complexes with the GMS proteins, possibly explaining their presence at the vacuole. The absence of the GDI-like activity of

the GMS proteins, as is the case in ectopic expression of GKS proteins in cells not induced with IFNγ, results in premature activation and GTP-dependent oligomerisation of GKS proteins on endomembranes, thus hindering or even preventing translocation to the parasitophorous vacuole upon infection (not depicted in the model). Virulent *Toxoplasma* strains possess a virulence mechanism effectively preventing the accumulation of IRG proteins on most but not all parasitophorous vacuoles.

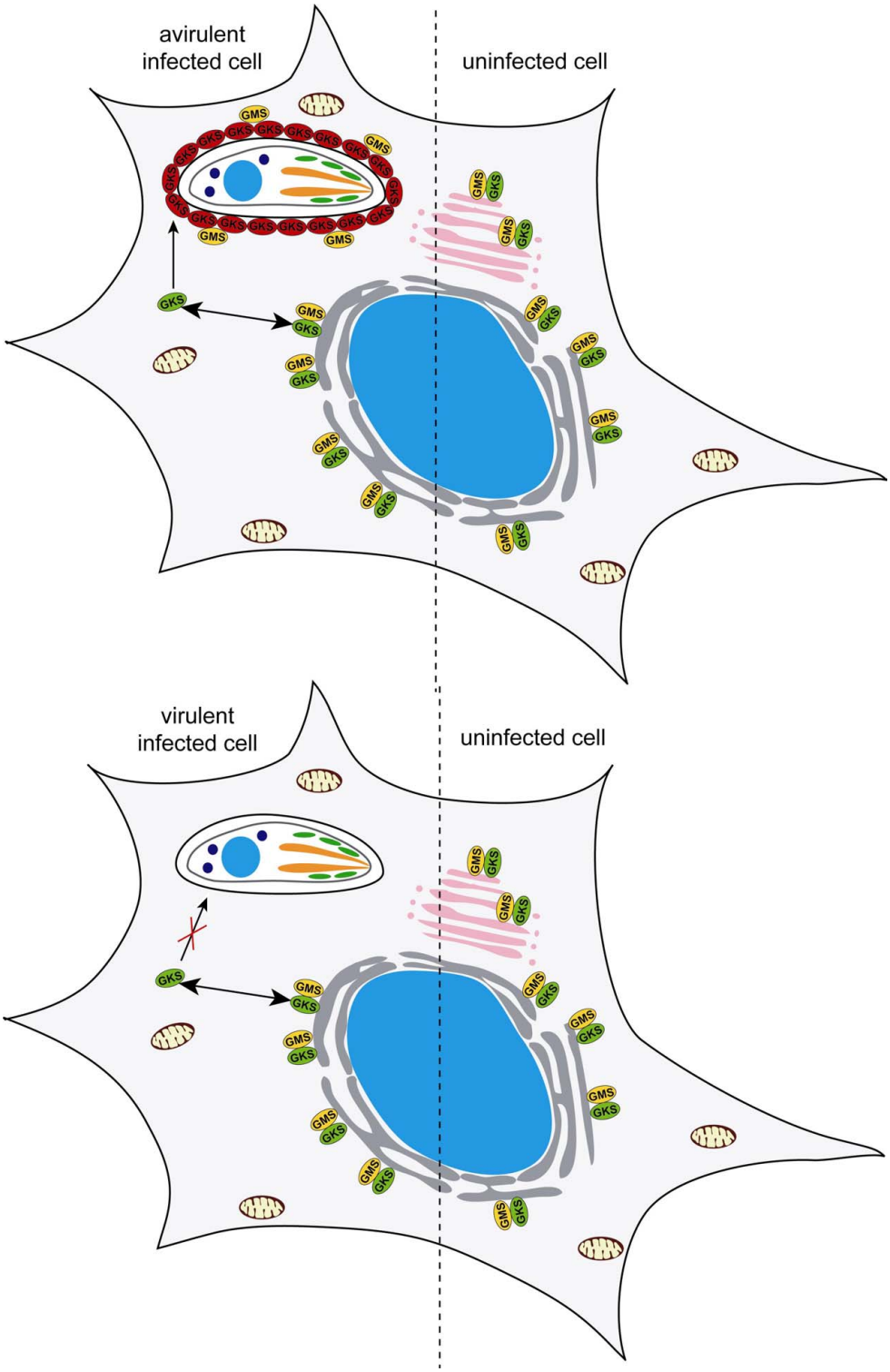
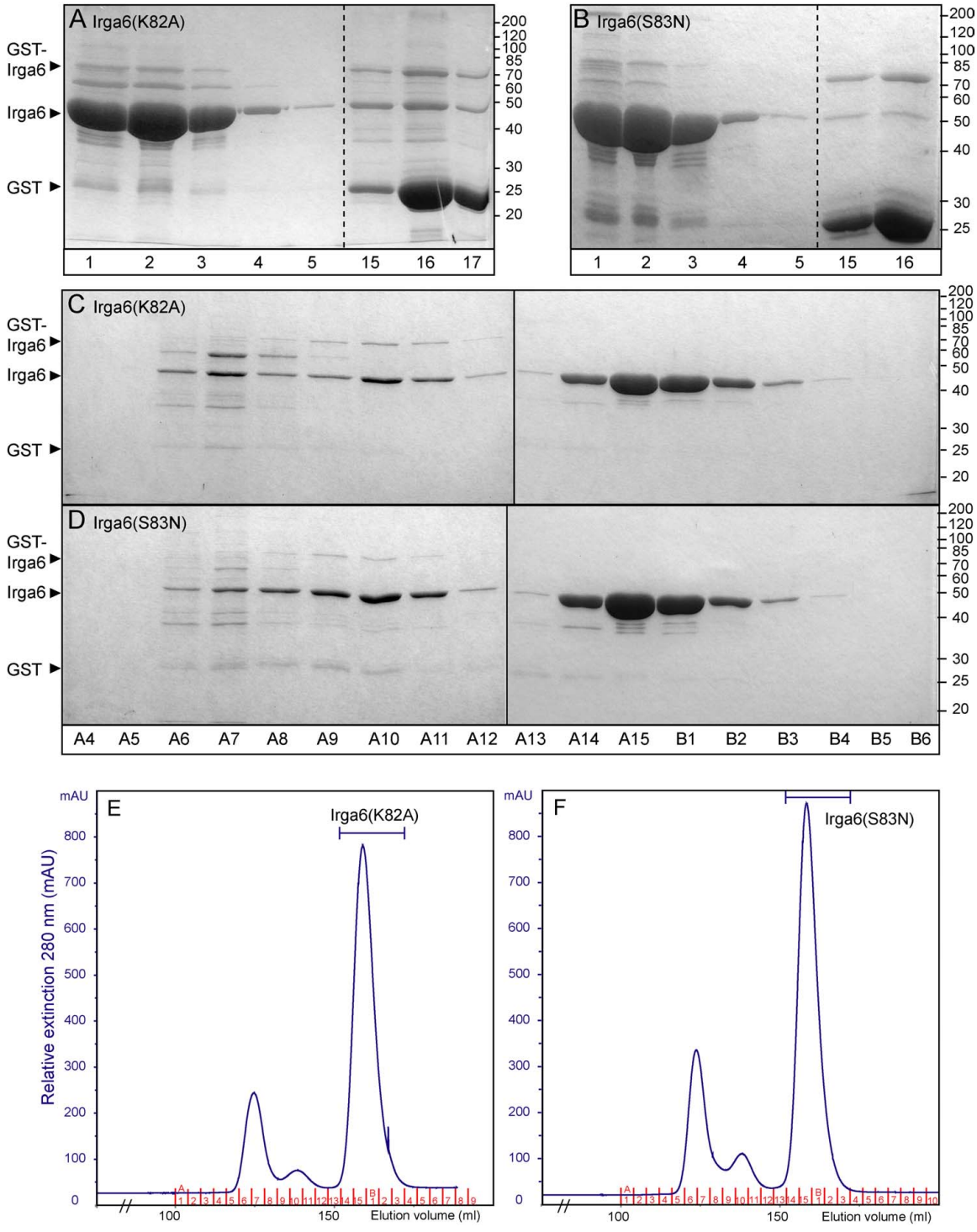


Figure 58 Model of IRG function in uninfected and infected cells. Regulatory interactions of IRG proteins governing the resting localisation in uninfected cells and the translocation to the PVM of avirulent but not virulent *T. gondii* strains (see text for details). Red ellipse: GTP-bound self-activating GKS proteins, green ellipse: GDP-bound self-activating GKS proteins, yellow ellipse: GMS proteins bound to GDP at the endomembranes and in unknown nucleotide status at the *Toxoplasma gondii* PVM

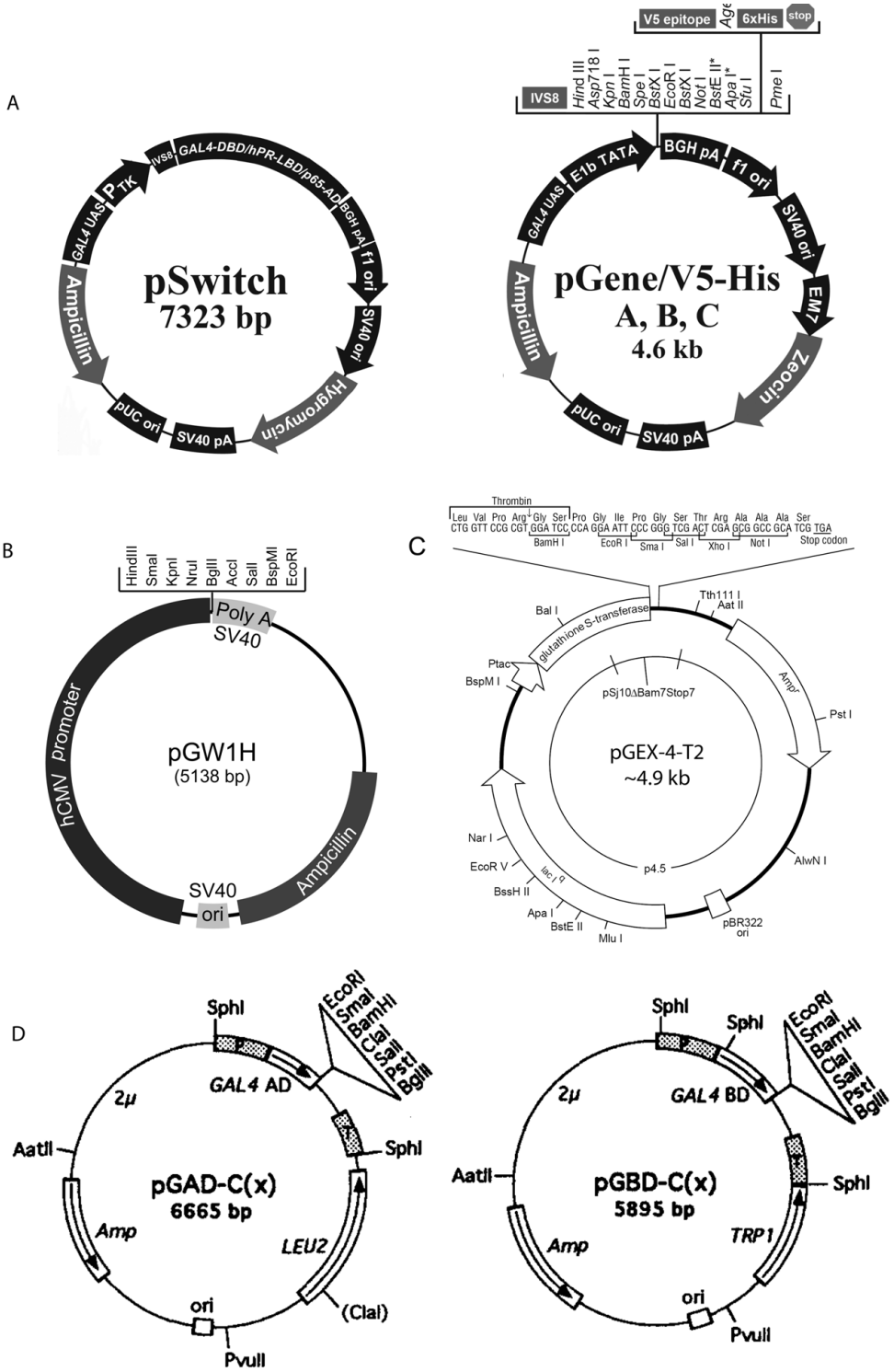
V. Appendix

V.1. Purification of recombinant Irga6(S83N) and (K82A) protein



Purification of recombinant Irga6(S83N) and Irga6(K82A) protein by glutathione Sepharose affinity and size exclusion chromatography. (A-B) Coomassie stained SDS-PAGE gel of fractions eluted from a glutathione Sepharose affinity column (C-F) Size exclusion chromatography of Irga6(K82A) and Irga6(S83N) protein containing fractions (1-3) from (A) and (B) respectively. Fractions A14-B3 were pooled and used for biochemical analyses (C-D) Coomassie stained SDS-PAGE gel of eluted fractions. (E-F) Elution profiles. For purification details see material and method section. (mAU: milli-absorbance unit).

V.2. Expression vectors



Plasmid maps of expression vectors. (A) Mammalian Mifepristone-inducible expression vectors of the GeneSwitch system (Invitrogen). (B) Mammalian expression vector used for transient transfection studies. (C) Bacterial expression vector used for generation of recombinant IRG protein. (D) Yeast expression vectors used for the yeast two-hybrid assay. (For details see material and methods section; hPRD-LBD: human progesterone receptor Ligand binding domain, TK: thymidine kinase, hCMV: human cytomegalovirus, pA/PolyA: polyadenylation signal, UAS: upstream activating sequences, Amp: ampicillin resistance gene, ori: bacterial origin of replication, T: transcription termination, P: promoter, AD. activation domain, BD: DNA binding domain, 2μ: yeast origin of replication; BGH: bovine growth hormone; SV40: Simianes Virus 40; TATA: TATA box; E1b: adenovirus E1b protein; ISV8: intron-8 polythymidine sequence; EM7: *E. coli* EM7 promoter; f1: bacteriophage f1; pUC: plasmid University of California; LEU2: 3-isopropylmalate dehydrogenase gene, TRP1: N-(5 -phosphoribosyl)-anthranilate isomerase gene, involved in leucine and tryptophan biosynthesis respectively)

V.3. DNA sequence of the expressed murine IRGs

Open reading frames of the mouse IRGs expressed using the vectors depicted above (V.2, see chapter II.1.6 for cloning strategy). A Kozak sequence (underlined) was introduced before the start codons and the sequences were flanked by SalI sites (in capitals) (GTCGACcaccatg....stopGTCGAC) (Kozak 1987).

Irgc ORF

```
atggcaacttccaggttgcccgccgtgcctgaggagaccaccatcctcatggccaaggaagagctggaggccctgcgcactgcttttga
gtctggcgacatccctcaagccgcctctcgccttcgggagctgctggccaactcagagaccacccggctggaagtgggcgtcacgggtg
agtcgggagccggcaagtcctccctcatcaatgccctacgcggcctgggggccgaggatcctggcgcagctctcactggggtcgtggag
accaccatgcagccttcgccctacccgcacccgcagtttcccgacgtgaccctgtgggacctgccgggggccggttctccaggctgctc
agcagacaagtatctgaagcaggtggatttcggccgctatgacttcttcttgctcgtctccccccgtcgctgtggcgccgtggagtccc
gcctcgcttctgagatcctgcgccaggggaagaagttttactttgtgcgcaccaaggtggacgaggatctggcggccacccgcagccag
aggccctcgggtttcagcgaggctgcagtcctccaggagatccgagatcactgcacggagcggctgcgggtagcgggtgtgaatgatcc
ccgcatcttcctggtgtccaacctgtcgccaacccgctatgacttcccgatgctcgtgaccacctgggagcacgacctgcccgcccacc
gtcgccacgccggtctgctgtccctgcctgacatctcgctggaggctctgcagaagaagaaggacatgctacaagagcaggtgcttaag
actgccttggtatctggtgtcatccaggccctgccggtccccggactggcagccgcctacgacgacgccttgcttatccgctcactgcg
gggctaccaccgcagcttcggcctagacgacgactcgctggccaagctggccgagcaggtgggcaaacaggcaggggacctgcgctccg
tcatccgctcccccctggccaacgaggtctcaccagagactgtcctgagactctactcgcagtcctcagatggtgccatgcgggtggcc
cgtgccttgagaggggcatccctgttttcggcacgctggtggccgggggtatcagcttcggcacggtctacaccatgctccagggctg
tctcaatgagatggctgaggacgcccaacgcgtccgcatcaaagccctggaggaagatgagccccaggggggtgaggtgagcttggagg
cggctggtgacaatttagtggaaaagcggagcactggggaaggaaccagcgaggaagccccgctgtccacccgcaggaagctcggcctc
ctcctcaagtatattcttgacagctggaagaggcgcgacttgtcagaagacaaataa
```

Irgm3 ORF

```
atggatttagtcacaaagttgccacaaaatatctggaagacttttacacttttcataaacatggcaaattatctcaagcgcctcatcag
cccgtggtctaaatctatgactgctggtgagtcactttattccagccagaactcatcctctccagaagtcattgaagatattggtaagg
cggtgacagagggaaatttacagaaagtgataggtatagtcaaagatgaaattcagagtaagtcaagatacagagtaaagattgctgtg
actggggactctggcaatggcatgtcatctttcatcaatgcccttaggttcattggacatgaggaggaggagtcagctcccactggggt
ggtgaggaccactaagaaaccagcctgctactcctctgactcccactttccctatgtggagctgtgggacctgcctggcttaggggcca
cagcccagagtgtggagagctacctggaggagatgcagatcagcacatttgaccttatcatcatcgtagcttctgagcagttcagctca
aatcatgtgaagctggccataaccatgcagaggatgagaaagaggttctatgtcgtctggaccaagctggacagggacctcagcacaag
taccttccctgaaccccagctactgcagagtatccaaaggaatatccgggagaatctccagcaggctcaggtgagggacccccccctat
tcctgatatcctgttttagcccgtcttttcacgacttcccagagcttagaaacacactgcaaaaagacatcttcagcatcaggtacaga
gatcccttagagatcatttctcaagtctgcgacaagtgcatcagcaataaggcgttctctctgaaggaagatcagatgctcatgaaaga
cctggaggcagctgtcagctccgaggatgacactgccaacctggagaggggtctccagacctaccagaagctctttggtgtggatgatg
ggtcacttcagcaggtagctcggagtacagggagactggagatgggctccagggctctgcagttccaggacttgatcaagatggacagg
agactggagctgatgatgtgttttgccgtgaacaagttcctcaggcttctcgaaagctcatggtggtatggcttgtggaacgtcgtcac
ccgctacttcagacaccagaggcacaagctcgtcattgaaatagtggctgagaacaccaagacctccctgaggaaagctctaaaggact
ctgtcctccctcccgaaattcactga
```

Irga6 ORF

```
atgggtcagctgttctcttcacctaagagtgatgagaataatgatttgccctccagctttactggttattttaagaaatttaatacggg
aagaaaaatcatttctcaagagatcctcaatttgattgaattaaggatgagaaaagggaatattcagttgacaaactctgcaatcagtg
atgcattaaaagaaatcgatagtagtgtgctcaatgttgctgtcaccggggagacgggatcagggaagtccagcttcatcaataccctg
agaggcattgggaatgaagaagaaggtgcagctaaaactggggtggtggaggtaaccatggaaagacatccatacaaacaccccaatat
acccaatgtggtttttgggacctgcctgggattggaagcacaaatttcccaccaaacacttacctggagaaaatgaagttctatgagt
acgatttcttcattattatttcggccacacgcttcaagaaaaatgatatagacattgccaaagcaatcagcatgatgaagaaggaattc
tacttcgtgagaaccaaggtggactctgacataacaaatgaagcagatggcaaacctcaaacctttgacaaagaaaaggtcctgcagga
catccgccttaactgtgtgaacacctttagggagaatggcattgctgagccaccaatcttcctgctctctaacaaaaatgtttgtcact
atgacttccccgtcctgatggacaagctgataagtgacctccctatctacaagagacacaattttatggtctccttacccaatatcaca
gattcagtcattgaaaagaagcggcaatttctgaagcagaggatttggctggaaggatttgctgctgacctagtgaatatcatcccttc
tctgacctttctcttggacagtgatttggagactctgaagaaaagcatgaaattctaccgcactgtgtttggagtggatgaaacatctt
tgcagagattagctagggactgggaaatagaggtggatcaggtggaggccatgataaaatctcctgctgtgttcaaacctacagatgaa
gaaacaatacaagaaaggctttcaagatatattcaggagttctgtttggctaatgggtacttacttcctaaaaatagttttcttaaaga
aatattttacctgaaatattatttccttgacatggtgactgaggatgctaaaactcttcttaaagagatatgtttaagaaactag
```

Irgd ORF

```
atggatcagttcatctcagccttcctgaagggtgcttcagaaaatagtttccaacaattagctaaggagtttctgcctcagtactctgc
attaatcagtaaggcaggaggcatgctctctccagaaaccctcactggtattcacaaagccctccaggagggaaatctctctgatgtca
tgatccagattcagaaagcaattagtgctgcagagaatgctatcctggaggtggctgtgatcgggcagtctgggactggcaagtccagt
tttataaatgcccttcgagggctgggccacgaagcagatgaatccgctgatgttgggactgtggagaccactatgtgtaaaacccccta
tcaacatccaaaatatcccaaagtgatcttctgggacctgcctgggactgggacacccaatttccacgcagacgcttatctagaccaag
tgggatttgccaactatgacttcttcatcatcatttcttcttcccgcttcagcctcaatgatgctctcctggctcagaaaataaaggat
gctgggaagaagttctactttgttagaaccaaggtggatagtgacttatataatgaacagaaagccaaacccatagctttcaagaagga
gaaagtccttcagcagattcgagactactgtgtgactaatctcatcaaaactggggtgactgaaccatgcatcttcctgatctccaact
tggatctgggtgcgtttgatttccccaagctggaggagactctgctaaaggagctccctgggcacaagcgtcatatgtttgccctgctc
ttgcccaatatctcggatgcttccattgagctaaagaaacattttcttcgggagaagatctggctggaggccttgaagtcagcagctgt
gagcttcatcccccttcatgactttctttaagggctttgatttgcctgaacaagaacagtgcttgaaggattaccgaagctattttggct
tggatgatcaatcgattaaagagattgctgaaaaattgggcgcacccttggcagacatcaaggggaacttaagtgcttggatttctgg
tcccttgtaaaggataacagcataatagcgcaagctacgagtgctgctgaagctttttgtgctgtgaaaggaggccctgaatcttctgc
cttccaggctttaaaagtctattataggcgcacacagttccttaacatagtggttgacgatgctaaacatctcttgagaaagatagaaa
cagtaaatgttgcctag
```

Irgm1 ORF

```
atgaaaccatcacacagttcctgcgaggctgctccactactccccaacatggcagagacccattatgctcccctgagctcagccttccc
ctttgtcacgtcataccaaacaggctccagcaggttacctgaggtcagtaggagcaccgaaagagctttaagagaaggaaaactactgg
aactggtctacggaatcaaggagactgtggcaacattgtcccagattccagtgagcatctttgtgactggggactctggcaatggcatg
tcatctttcatcaatgcacttcgagtcatcggccatgatgaagatgcctcggctcccactggggtggtgaggaccacgaagacgcggac
tgagtactcttcatcccactttcccaatgtggtgctgtgggacttacctggattgggggccacagcccaaaccgtagaggactatgtgg
aagagatgaaatttagcacatgtgacttattcatcatcattgcctctgagcagttcagctcgaatcatgtgaagctgtccaaaattatc
cagagcatgggaaagaggttctatattgtctggaccaagctggacagggacctcagcaccagtgtcctatcagaggtccggctcctaca
gaatatccaggagaatatccgagagaatctgcagaaggagaaagtgaagtaccccccgtgttcctggtatccagtctagacccttac
tatatgacttcccgaagcttagggacacacttcataaagatctctccaacatcaggtgctgtgaacccttaaagacccttttatggcact
tatgagaagatcgttggtgataaagtagcagtctggaagcagagaatagccaacgagtccttgaagaattctctcggtgtcagagatga
tgacaacatgggcgagtgtctgaaagtgtaccgactgatatttggtgtagatgacgaatcagttcagcaggtagcccagagtatgggga
cagtagtcatggagtacaaggacaacatgaagtcccaaaacttttatactctccgcagagaggactggaaactgaggctgatgacatgt
gcaattgtgaatgcattcttccgtttgttgagatttctcccatgcgtatgctgctgtttaagacgcttgagacataaacgcatgctttt
cttagttgcccaggacaccaagaacatcctagagaaaatcctgagggactccatcttccctccgcagatctag
```

Irgb ORF

```
atggcttgggcctccagctttgatgcattctttaagaattttaaaagggaaagcaaaatcatctctgaatatgacatcaccttgattat
gacttacatagaggaaaataagctacagaaagctgtttctgtaattgaaaaggtactgagagacatcgagagtgctcctctgcacatag
ctgtgacaggggaaacaggcgcagggaagtccactttcatcaataccctgaggggggtggggcatgaagaaaaaggtgcagcccccact
ggcgcaatagagacaaccatgaagagaactccatacccacacccaaagcttcccaacgtgacaatatgggacctgcctggcattgggac
cactaacttcacaccacaaaactatctgacagaaatgaagtttggtgagtatgacttcttcattatcatctcagctacacgtttcaaag
aaaatgatgcacaactggccaaagccattgcacagatggggatgaatttctactttgtcagaaccaagatagacagcgacttagataat
gaacagaagtttaagcctaagagtttcaataaggaggaagtcctcaagaatattaaggattactgctctaatcatcttcaggagtctct
cgacagtgagcctccagtcttcctagtctctaacgttgatatatcgaagtatgacttcccaaagctggaaactaaactcctacaggatc
tcccagcccacaagcgtcacgtattctcactgtctttgcaaagtcttactgaggccaccattaactacaagagagattccctgaagcaa
aaagtcttcctagaagccatgaaggctggagcattagccaccattccacttggtggcatgatcagtgatatcttagagaatctggatga
aacattcaatctctacaggtcttactttgggctggatgatgcttcactggaaaacattgcccaggatttgaacatgtctgtggatgact
tcaaggtacaccttcgatttccccatttgtttgcagaacacaatgatgagtccttagaagacaagctatttaaatatatcaaacacatt
tcttcagttactggtgggccagtcgctgcagtcacttactatcgcatggcttattatttgcagaatctctttcttgatactgcagccaa
tgatgccatagctcttctgaatagtaaagcacttttttgagaagaaggtgggaccatatatatctgagccccccgagtactgggaagctt
ga
```

V.4. Abbreviation of species names

AC	*Anolis carolinensis*	green anole
AM	*Acropora millepora*	coral
AT	*Ambystoma tigrinum tigrinum*	eastern tiger salamander
BF	*Branchiostoma floridae*	lancelet fish
BT	*Bos taurus*	cattle
CF	*Canis familiaris*	dog
CJ	*Callithrix jacchus*	common marmoset
CP	*Cavia porcellus*	Guinea pig
DN	*Dasypus novemcinctus*	armadillo
DR	*Danio rerio*	zebra fish
EC	*Equus caballus*	horse
EE	*Erinaceus europaeus*	western European hedgehog
ET	*Echinops telfairi*	lesser hedgehog tenrec
FC	*Felis catus*	cat
HS	*Homo sapiens*	human
LA	*Loxodonta africana*	African elephant
LE	*Leucoraja erinacea*	little skate
MD	*Monodelphis domestica*	opossum
ML	*Myotis lucifugus*	little brown bat
MM	*Mus musculus*	mouse
MN	*Macaca mulatta*	Rhesus monkey
MU	*Microcebus murinus*	grey mouse lemur
OA	*Ornithorynchus anatinus*	platypus
OC	*Oryctolagus cuniculus*	rabbit
OG	*Otolemur garnettii*	bushbaby
PP	*Pongo pygmaeus*	orang-utan
PT	*Pan troglodytes*	chimpanzee
RN	*Rattus norvegicus*	rat
SA	*Sorex araneus*	common shrew
SS	*Sus scrofa*	pig
ST	*Spermophilus tridecemlineatus*	thirteen-lined ground squirrel
XT	*Xenopus tropicalis*	South African clawed frog

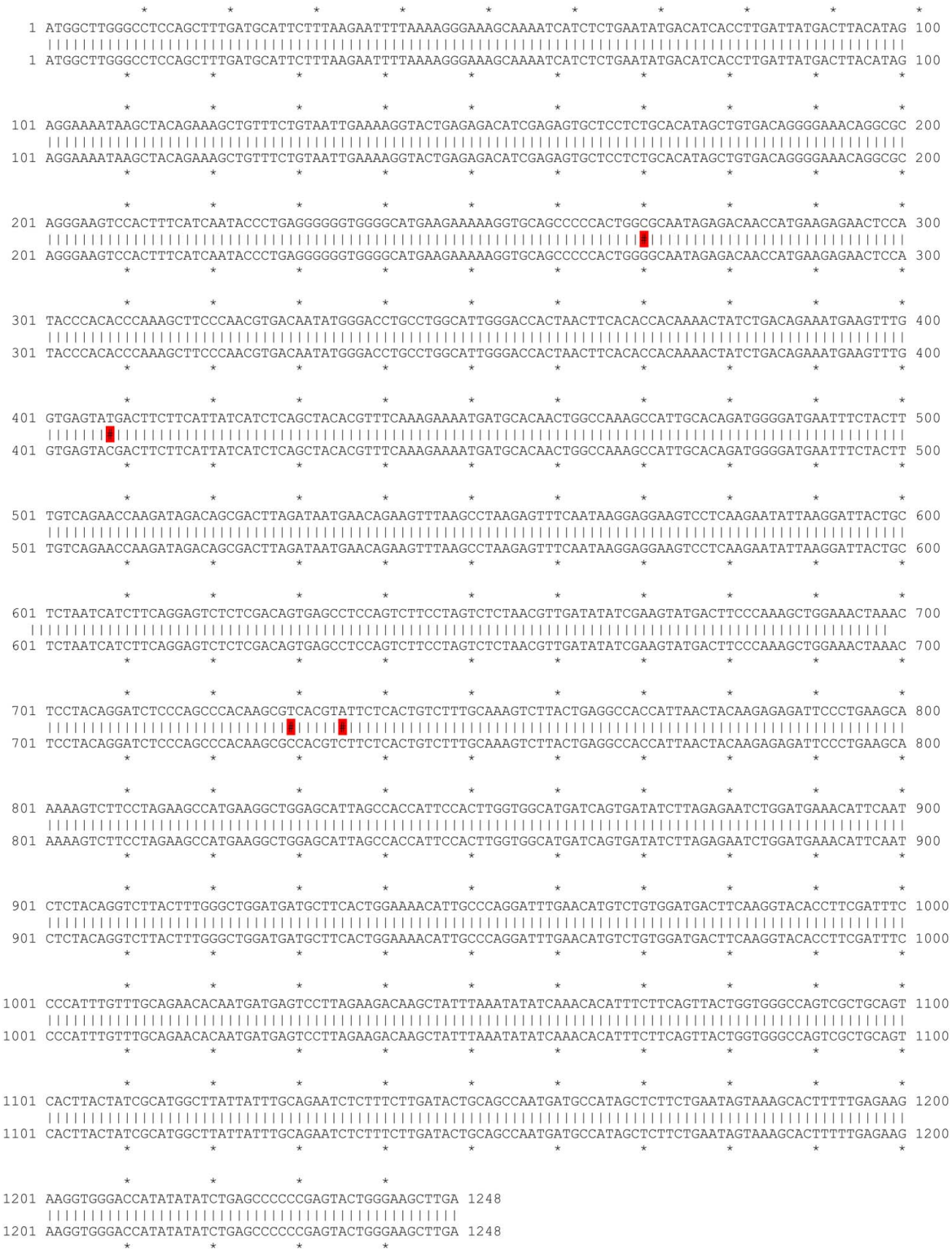
List of abbreviations of species names used in the phylogenetic analyses.

V.5. Sequence sources of mouse and human IRG genes

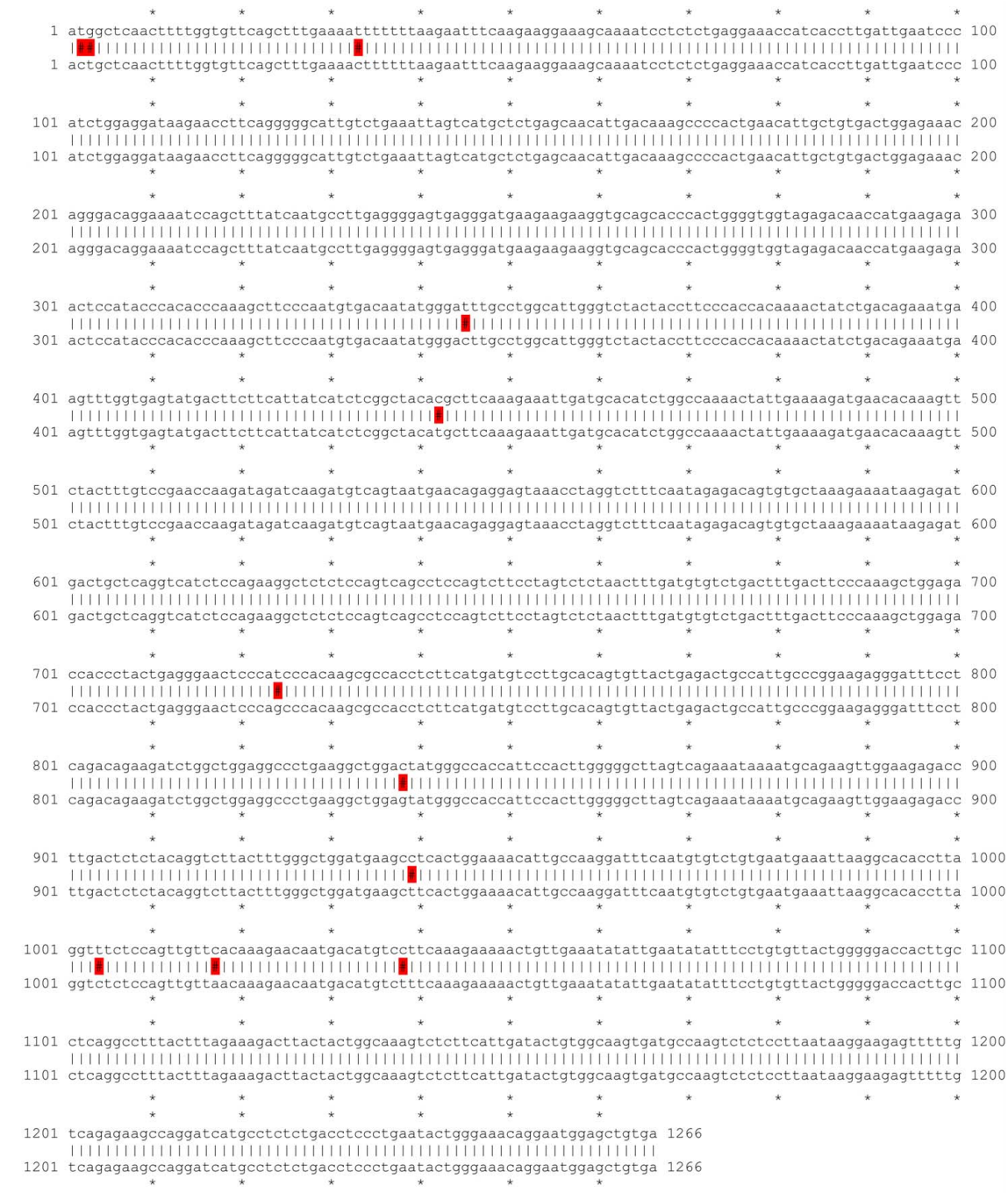
Gene name	Genesymbol/ID	Synonyms	Genomic sequences /Accession no.	cDNA or EST sequence Accession numbers	Notes
Mouse					
Irga1	*Irga1* MGI:1795294 MGI:1653512		AC132320 AC102225	BI658674 (NMRI, 5'EST, nearly 100%) BG915086 (NMRI, 5'EST; not 100%)	
Irga2	*Irga2* MGI:915200 MGI:1257137 MGI:1257136		AC132320 AC102225 XM_140378	AA968296 (C57BL/6, 5'EST, 100%) AA968378 (C57BL/6, 3'EST, 100%)	Inducible by IFN-γ.
Irga3	*Irga3* New gene		AC132320 XM_140379 (C57BL/6J)	BY751179 (NOD, EST, not 100%)	Inducible by IFN-γ.
Irga4	*Irga4* New gene		AC132320 XM_140380 (*Irgb4/Irgb5* tandem)	BY750970 (NOD, EST, nearly 100%) BU696309 (C57BL/6, EST, nearly 100%)	Inducible by IFN-γ.
Irga5Ψ	*Irga5* New gene		AC132320	None	A transcript is inducible by IFN-γ but the coding sequence of the gene is disrupted repeatedly.
Irga6	*Irga6* MGI:1926259 MGI:2147195 MGI:2147350	IIGP, IIGP1, Iigp1	AC135638	AJ007971 (C57BL/6, 100% correct) AF194871 (C57BL/6, also NM_021792, 100%) BC004649 (C57BL/6, cDNA 100%,)	(Boehm 1998) (MGI:1889878); (Zerrahn 2002) Inducible by IFN-γ.
Irga7	*Irga7* New gene		NT_039674 (C57BL/6J) XM_487533 (C57BL/6, 100%)	None known	
Irga8	*Irga8* MGI:953940 (C57BL/6) MGI:2384767 MGI:1489193 (CZECHII)	MGC:28198 BC023105	AC135638	BC023105/NM_145357 (CZECHII, not 100%) BB637466 (C57BL/6J, 5'EST, not 100%) BF163606 (CZECHII, not 100%, not full length) BE198503 (C57BL/6, 3'EST, 100%) BE198089 (C57BL/6, 3'EST, 100%) BX520309 (C57BL/6, 3'EST, 100%)	In C57BL/6 a non-canonical guanine after bp 849 (= aa 204) puts the sequence out of frame just before Helix H4; the reading frame is complete in BC023105 (CZECHII, *Mus musculus musculus*). Inducible by IFN-γ.
Irgb1	*Irgb1* MGI:1519766		AL645849	BC022776 tandem *Irgb2/Irgb1* (CZECHII, not 100%) AK145236 tandem *Irgb2/Irgb1* (Jyg, not 100%) BF144722 (CZECHII, EST, not 100%, starts with 3' end of *Irgb2*)	The *Irgb2/Irgb1* gene pair is almost certainly transcribed in tandem. The protein has not yet been described. Inducible by IFN-γ.
Irgb2	*Irgb2* MGI:1518599		AL645849	BC022776 tandem *Irgb2/Irgb1* (CZECHII, not 100%) AK145236 tandem *Irgb2/ Irgb1* (Jyg, not 100%) BF144934 (CZECHII, 5' *Irgb2* cDNA, not 100%)	See note above, *Irgb1*.
Irgb3	*Irgb3* MGI:1553791 (FVB/N)		AL627237 AL669850 (unordered) AF060196 (129/SvJ, genomic, 1 bp difference)	BF539106 (FVB/N, 3'EST, not 100%)	The genomic sequence of *Irgb3* is followed after 950 bp by a retroposon corresponding to the proteasome regulator PA28b (Li 1999, MGI:1331589). The presence or absence of this retroposon unambiguously distinguishes *Irgb3* from *Irgb4*.
Irgb4	*Irgb4* MGI:1795392 MGI:3041173	9930111J21Rik	AL627237 AL669850 (unordered)	BC066104 (C57BL/6, *Irgb5/Irgb4* tandem, 100%) BI655221 (NMRI, EST, not 100%)	See note above for *Irgb3*. *Irgb4* is probably normally expressed as a distinct 3' exon in a tandem transcript downstream of *Irgb5*.
Irgb5	*Irgb5* MGI:3041173 MGI:2401562	9930111J21Rik	AL627237 AL645688 AL669850 (unordered)	BC066104 (C57BL/6, *Irgb5/Irgb4* tandem; not 100% at 5' end) AK037088 (C57BL/6, cDNA, = NM_173434, 100%, unknown 5' end) (protein = BAC29698= Q8CB10)	*Irgb5* is probably normally expressed as a separate 5' exon in a tandem transcript upstream of *Irgb4*. However AK037088 does not splice into *Irgb4*. Thus *Irgb5* can exist as a single p47 unit or as a tandem with *Irgb4*. The reference number MGI:2401562 refers to several ESTs belonging to *Irgb5* and *Irgb9*. Inducible by IFN-γ.
Irgb6	*Irgb6* MGI:98734 MGD-MRK-15077	TGTP, Mg21, Gtp2	AL627237 AL645688 AL669850 (unordered)	L38444 (C57BL/6, 100%); NM_011579 (NOD, 2 aa difference); U15636 (C.D2-Idh-1/Pep-3, 2 aa difference); BC085259 (NMRI, cDNA, 100%); BC034256 (CECHII, cDNA, not 100%); AK163978 & BE632518 (C57BL/6, Irgb6*)	Carlow 1995, Lafuse 1995 Inducible by IFN-γ.
Irgb7Ψ	*Irgb7* New gene		AL645688 AL669850 (unordered)	None known	Pseudogene: STOP codon before G-domain. Not inducible by IFN-γ, no known transcript.
Irgb8	*Irgb8*		AL645849	AK144287 tandem *Irgb9/Irgb8* (C57BL/6)	The *Irgb9/Irgb8* gene pair is almost certainly transcribed in tandem. So similar

	MGI:1672892			AK165747 tandem *Irgb9/Irgb8 (*C57BL/6) BG974191 (NMRI, 3' EST, not 100%)	to *Irgb1*, *b3* and *b4* that non-identical EST sequences are hard to disentangle.
Irgb9	*Irgb9* MGI:2401562 New gene		AL645849 XM_204704 (C57BL/6, full length, 100%)	AK144287 tandem *Irgb9/Irgb8* (C57BL/6) AK165747 tandem *Irgb9/Irgb8* (C57BL/6) BB630182 (EST, short)	See note above, *Irgb8.* The reference number MGI:2401562 refers to several ESTs belonging to *Irgb5* and *Irgb9*.
Irgb10	*Irgb10* MGI:1282384		AL928857	AI122314 (C57BL/6, short EST, not 100%)	Short, terminates before end of G domain in S6. Inducible by IFN-γ.
Irgc	*Irgc* New gene	CINEMA	AC073810 (RP23-57J6) GENSCAN00000140134	BB615720 (C57BL/6 cDNA, 99%) 36 ESTs, none full length (e.g. CA464745 5'mRNA, 100% except of first two bp)	An *Irgc*-related sequence has recently been named HGTP-47 (MacMicking 2004). This sequence (NP_950178=NM_199013= AK089224, NOD) contains 4 frameshifts relative to the C57BL/6 genomic sequence leading to a largely incorrect protein sequence. The reference numbers MGI:2685948 and MGI:2685320 both relate to this error sequence.
Irgd	*Irgd* MGI:99448 MGD-MRK-16217	IRG-47, IRG47, Ifi47, 47kDa, Iigp4	AL645688 AL669850 (unordered)	M63630 (B6D2F1, =NM_008330, 100% correct)	(Gilly 1992). This is the first report of a p47 GTPase and has given its name (IRG-47) to the whole family. Inducible by IFN-γ.
Irgm1	*Irgm1* MGI:107567 MGD-MRK-36139	LRG-47, LRG47, Ifi1, Iigp3	AL645849	U19119 (BALB/c, =NM_008326, 100% correct)	(Sorace 1995). Two 5' splice variants exist. See notes human IRGM below. Inducible by IFN-γ.
Irgm2	*Irgm2* MGI:1926262 MGI:2144195	GTPI Iigp2	AL928857	AJ007972 (C57BL/6; 100%) NM_019440 (CZECHII, = BC005419, not 100%)	(Boehm 1998) MGI:1889878. Two 5' splice variants exist. Inducible by IFN-γ.
Irgm3	*Irgm3* MGI:107729 MGD-MRK-36305 MGI:2144580	IGTP Igtp	AL928857	U53219 (C57BL/6, cDNA, 100%) NM_018738 (NOD, cDNA, not 100%)	(Taylor 1996) MGI:82341 Inducible by IFN-γ.
Irgq	*Irgq* MGI:2667176	FKSG27	AC073810	AF322649 (C57BL/6, mRNA, = NM_153134)	
Human					
IRGC	UniGene Hs.515444 R30953_1 GeneID: 56269	CINEMA human IIGP5, cinema1	AC005622 HChr.19 cosmid	BC066939 (cDNA, 100%) NM_019612 (cDNA, 100%)	
IRGM	UniGene Hs.519680 GeneID: 345611 MIM: 608212	human LRG-47-like protein (LRG47, LRG-47), IFI1	AC010441 Chr.5 XM_293893 (splice variant a, 100%)	BC038360 (splice variant c, 3'EST); BC038539 (short EST); BI764111 (short EST) Sequences have been confirmed by RT-PCR (Bekpen Ph.D: thesis)	5 different 3' splice variants (a-e) (Bekpen 2005) The orthology of *Irgm1* with human *IRGM* implied by use of the name *LRG47* or IFI1 for the human gene is incorrect. The use of LRG47 as a synonym or alias for human IRGM is therefore not recommended.
IRGQ	UniGene Hs.546476 GeneID: 126298	Homo sapiens FKSG27, Irgq1	AC006276	AF322648 (=NM_001007561 mRNA, 100%)	

V.6. Nucleotide alignment of mouse Irga6 and Irga6* (C57BL/6)



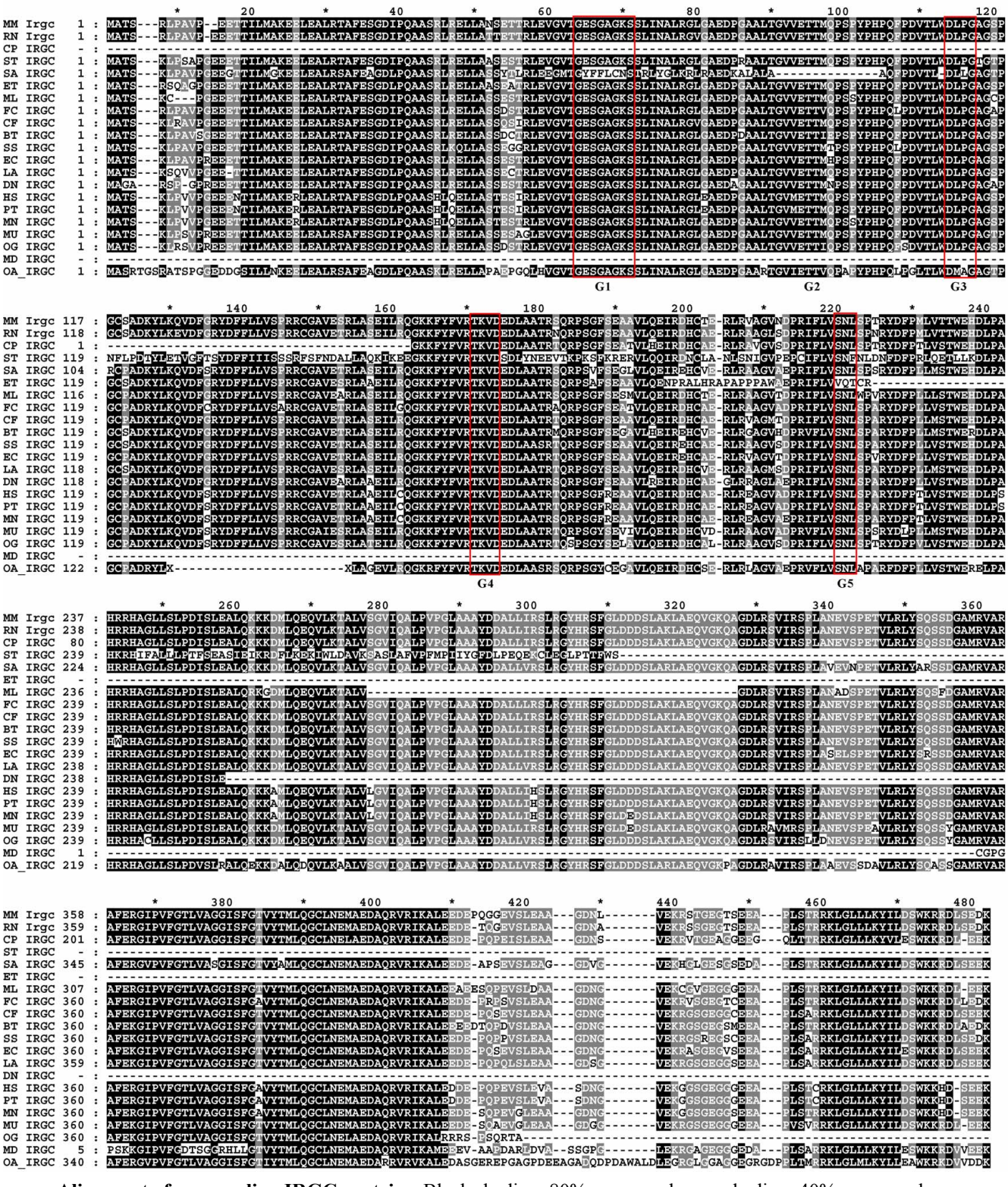
V.7. Protein and nucleotide alignment of mouse Irgb3 and Irgb4



Protein alignment of murine Irgb3 and Irgb4

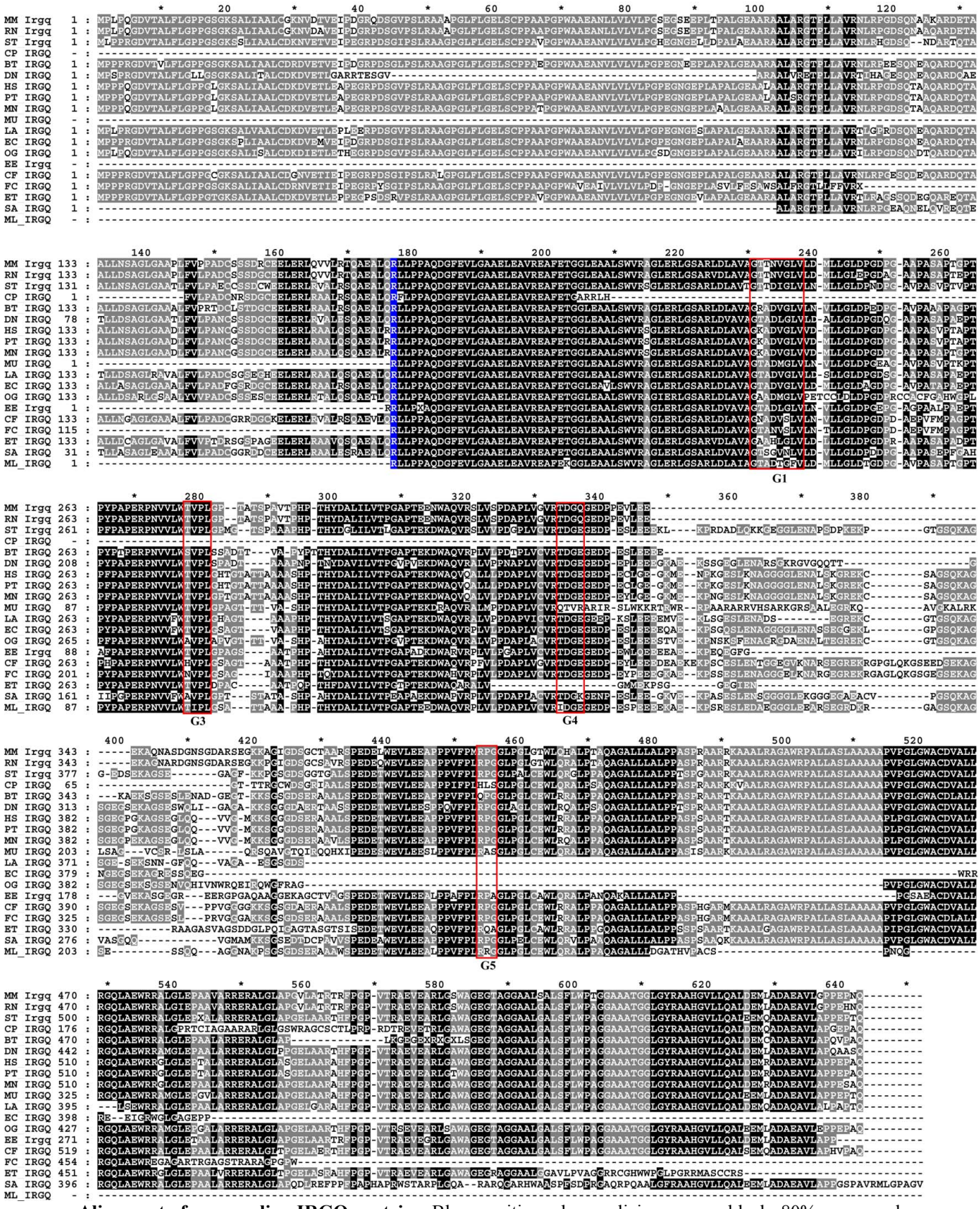
```
MAQLLVFSFENFFKNFKKESKILSEETITLIESHLEDKNLQGALSEISHALSNIDKAPLNIAVTGETGTGKSSFINALRGVRDEEEGAAP
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TAQLLVFSFENFFKNFKKESKILSEETITLIESHLEDKNLQGALSEISHALSNIDKAPLNIAVTGETGTGKSSFINALRGVRDEEEGAAP

TGVVETTMKRTPYPHPKLPNVTIWDLPGIGSTTFPPQNYLTEMKFGEYDFFIIISATRFKEIDAHLAKTIEKMNTKFYFVRTKIDQDVSN
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGVVETTMKRTPYPHPKLPNVTIWDLPGIGSTTFPPQNYLTEMKFGEYDFFIIISATCFKEIDAHLAKTIEKMNTKFYFVRTKIDQDVSN

EQRSKPRSFNRDSVLKKIRDDCSGHLQKALSSQPPVFLVSNFDVSDFDFPKLETTLLRELPAHKRHLFMMSLHSVTETAIARKRDFLRQK
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
EQRSKPRSFNRDSVLKKIRDDCSGHLQKALSSQPPVFLVSNFDVSDFDFPKLETTLLRELPAHKRHLFMMSLHSVTETAIARKRDFLRQK

IWLEALKAGLWATIPLGGLVRNKMQKLEETLTLYRSYFGLDEASLENIAKDFNVSVNEIKAHLRFLQLFTKNNDMSFKEKLLKYIEYISC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
IWLEALKAGVWATIPLGGLVRNKMQKLEETLTLYRSYFGLDEASLENIAKDFNVSVNEIKAHLRSLQLLTKNNDMSFKEKLLKYIEYISC

VTGGPLASGLYFRKTYYWQSLFIDTVASDAKSLLNKEEFLSEKPGSCLSDLPEYWETGMEL*
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
VTGGPLASGLYFRKTYYWQSLFIDTVASDAKSLLNKEEFLSEKPGSCLSDLPEYWETGMEL*
```

V.8. Alignment of mammalian IRGC proteins



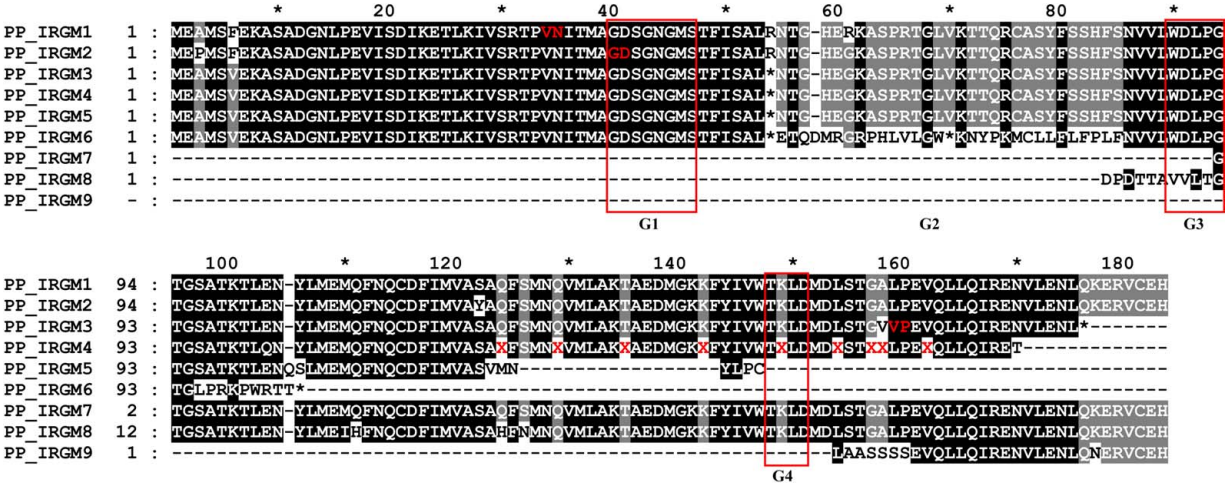
Alignment of mammalian IRGC proteins. Black shading: 80% conserved, grey shading: 40% conserved, red boxes: nucleotide-binding motifs

V.9. Alignment of mammalian IRGQ proteins



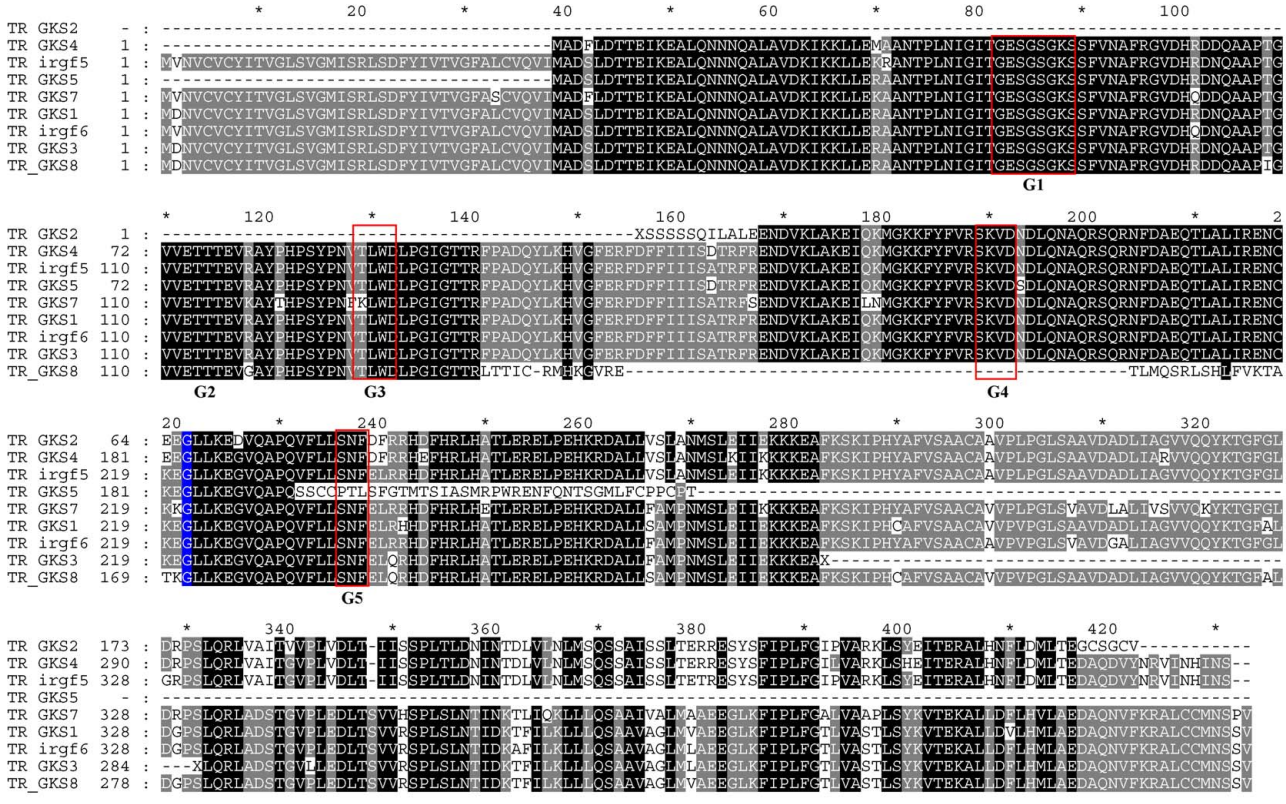
Alignment of mammalian IRGQ proteins. Blue: position where splicing occurs, black: 80% conserved, grey: 40% conserved, red: GTPase motifs. Note that the IRGQ proteins from different species contain insertion of variable length between G4 and G5 motif.

V.10. Alignment of orang-utan IRGM sequences *(Pongo pygmaeus)*



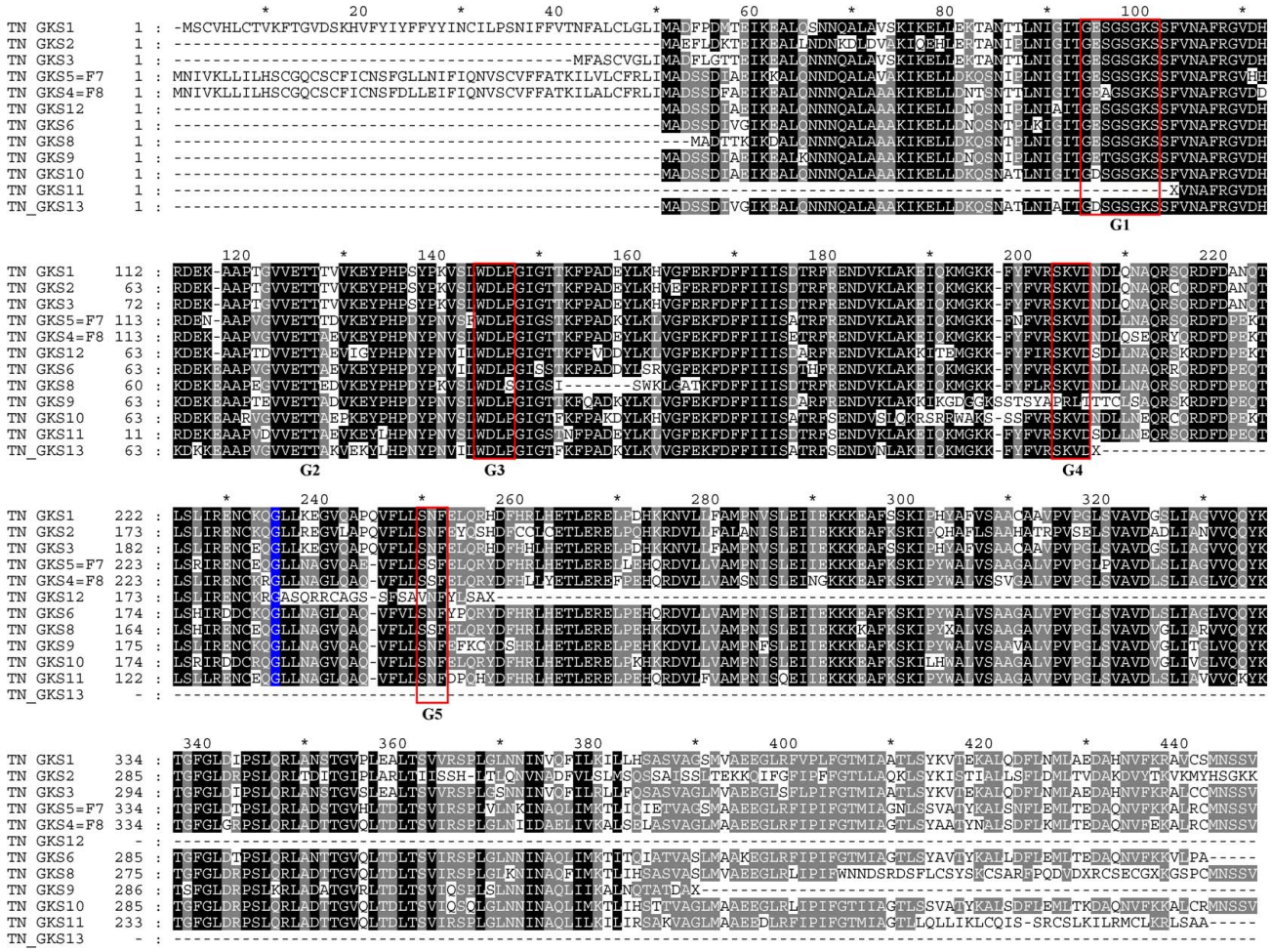
Alignment of IRGM sequences from orang-utan *(Pongo pygmaeus)*. Note that the IRGM4-9 sequences are disrupted by frame shifts and/or premature stop codons. Black shading: 70% conserved, grey shading: 40% conserved; * premature stop codon; red highlighting: position of frame shifts. Red boxes mark the GTPase motifs. The alignment was used to generate a consensus sequence (see appendix V.13).

V.11. Alignment of *Takifugu rubripes* IRG proteins.



Alignment of *Takifugu rubripes* (TR) IRG proteins. Blue: position of intron, black: 90% conserved, grey: 50% conserved, red: GTPase motifs

V.12. Alignment of *Takifugu rubripes* IRG proteins.



Alignment of *Tetraodon nigroviridis* (TN) IRG proteins. Blue: position of intron, black: 85% conserved, grey: 50% conserved, red: GTPase motifs.

V.13. IRG protein sequences

The position of introns is indicate by yellow highlighting

>MM_Irgq

MPLPQGDVTALFLGPPGSGKSALIAALCGKNVDTVEIPDGRQDSGVPSLRAAAPGLFLGELSCPPAAPGPWAAEANLLVLVLPGSEGSE
EPLTPALGEAARAALARGTPLLAVRNLRPGDSQNAAKARDETAALLNSAGLGAAPLFVPPADCSSSDRCEELERLQVVLRTQAEALQRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARLDLAVAGTTNVGLVLDMLLGLDPGDPGAAPASAPTGPTPYPAP
ERPNVVLWTVPLGPTATSPAVTPHPTHYDALILVTPGAPTEENWAQVRSLVSPDAPLVGVRTDGQGEDPPEVLEEEKAQNASDGNSGDA
RSEGKKAGIGDSGCTAARSPEDELWEVLEEAPPPVFPMRPGGLPGLGTWLQHALPTAQAGALLLALPPASPRAARRKAAALRAGAWRPA
LLASLAAAAAPVPGLGWACDVALLRGQLAEWRRALGLEPAAVARRERALGLAPGVLATRTRFPGPVTRAEVEARLGSWAGEGTAGGAAL
SALSFLWPTGGAAATGGLGYRAAHGVLLQALDEMLADAEAVLGPPEPNQ*

>MM_Irgc

MATSRLPAVPEETTILMAKEELEALRTAFESGDIPQAASRLRELLANSETTRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGVVE
TTMQPSPYPHPQFPDVTLWDLPGAGSPGCSADKYLKQVDFGRYDFFLLVSPRRCGAVESRLASEILRQGKKFYFVRTKVDEDLAATRSQ
RPSGFSEAAVLQEIRDHCTERLRVAGVNDPRIFLVSNLSPTRYDFPMLVTTWEHDLPAHRRHAGLLSLPDISLEALQKKKDMLQEQVLK
TALVSGVIQALPVPGLAAAYDDALLIRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMRVA
RAFERGIPVFGTLVAGGISFGTVYTMLQGCLNEMAEDAQRVRIKALEEDEPQGGEVSLEAAGDNLVEKRSTGEGTSEEAPLSTRRKLGL
LLKYILDSWKRRDLSEDK*

>MM_Irga1

MGQLFSLLKNKCQFLVSSVAEYFKKFKKIVIIILQEVTTSIELDMKKENFQEANSAICDALKEIDSSLVNVAVTGETGSGKSSFINTLR
GIGHEEEGAAKTGVVEATMERHPYKHPNMPNVVFWDLPGIGSTKFPPKTYLEKMKFYEYDFFIIISATCFKKNDIDLAKAISMMKKEFY
FVRTKVDTDLRNEEDFKPQTFDKEKVLQDIRLNCVNTFKENGIAEPPIFLISNENVCHYDFPVLMDKLISDLPDYKRHNFMLSLPNITD
SVIETKRQSLKQRHWLQGFAGVLLSYLH*

>MM_Irga2

MGQLFSSRRSEDQDLSSSFIEYLKECEKGINIIPHEIITSIEINMKKGNIQEVNSTVRDMLREIDNTPLNVALTGETGSGKSSFINTLR
GIGHEEGGAAHTGVTDKTKERHPYEHPKMPNVVFWDLPGTGSEDFQPKTYLEKMKFYEYDFFIIISATRFKKNDIDLAKAIGIMKKEFY
FVRTQVDSDLRNEEDFKPQTFDREKVLQDIRLNCVNTFRENGIAEPPIFLISNKNVCHYDFPVLMDKLISDLPVFKRQNFMFSLPNITD
SVIEKKRNFLRWKTWLEGFADGLLSFFLESDLETLEKSMKFYRTVFGVDDASLQRLARAWEIDQVDQVRAMIKSPAVFTPTDEETIQER
LSRYNQEFCLANGYLLPKNHCREILYLKLYFLDMVTEDAKTLLKEICLRN*

>MM_Irga3

MGQLFSHIPKDEDKGNLESSFTEYFRNYKQETKIISEETTRSIELCLKRGDFQRANSVISDALKNIDNTPINIAVTGESGAGKSSLINA
LREVKAEEESAAEVGVTETTMKVSSYKHPKVKNLTLWDLPGIGTMKFQPKDYLEKVEFKKYDFFIIVSSSRFTKLELDLAKATRIMKKN
YYFVRSKVDCDLDNEKKSKPRNFNRENTLNQVRNSYLDTFRESKIDEPQVFLISNHDLSDYDFPVLMDTLLKDLPAEKRQNFLLSLPNI
TEAAIQKKYNSTKQIIWLQATKDGLLATVPVVGILKDLDKERLKKRLDYYRDLFGVDDESLMFMAKDAQVPVELLIKNLKSPNLLKCKE
ETLEELLLNCVEKFASANGGLLAAGLYFRKTYYLQFHFLDTVAEDAKVLLKAAQTHFAHSF*

>MM_Irga4

MGQLLSDTSKTEDNEDLVSSFNEYFKNIKTEKIISQETIDLIKLYLNKGNIHGANSLISDALRNIDNAPINIAVTGESGAGKSSLINAL
IGIGPEEEGAAEVGVIETTMKRTSYKHPKIETLTLWDLPGIGTQKFPPKTYLEEVKFKEYDFFIIVSATRFTKLELDLAKAITNMKKNY
YFVRTKVDIDVENERKSKPRTFEREKALKQIQSYSVKIFNDNNMAVPPIFLISNYDLSDYDFPFLVDTLIKELHVQKRHNFMLSLPNFT
DQAIDRKYKATQQFIWLEAFKIGVVAIFPVLGNLRNKDMKKIKNTLNYYQKIFGVDDESLELVAKDFQVPVEQVKKTMKTPHLLKKYRE
ETFRNDFKKLVSTFGRLLAVGLYFPAIYYLQLHILDTVTEDAKVLLRWKYSKPRSNSTYP

>MM_Irga5_pseudo

MGQLFSGTSKSEALCSSFTEYFQKFKVENKIISQEISTLIELYLTLGDVQQANNAITYALR*LARTPQNVALIGESGRGKYSFINVFRG
LDMKRKMATVGVVETTMNRTPYRNPNIPNVIIWDLPGIGTTNFPPKHYLKKMQFYVMYDFFIIVSATCFRKNDIDLSKAVVMIKKDFLL
RTRTKEDIDIENENE*

>MM_Irga6

MGQLFSSPKSDENNDLPSSFTGYFKKFNTGRKIISQEILNLIELRMRKGNIQLTNSAISDALKEIDSSVLNVAVTGETGSGKSSFINTL
RGIGNEEEGAAKTGVVEVTMERHPYKHPNIPNVVFWDLPGIGSTNFPPNTYLEKMKFYEYDFFIIISATRFKKNDIDIAKAISMMKKEF
YFVRTKVDSDITNEADGKPQTFDKEKVLQDIRLNCVNTFRENGIAEPPIFLLSNKNVCHYDFPVLMDKLISDLPIYKRHNFMVSLPNIT
DSVIEKKRQFLKQRIWLEGFAADLVNIIPSLTFLLDSDLETLKKSMKFYRTVFGVDETSLQRLARDWEIEVDQVEAMIKSPAVFKPTDE
ETIQERLSRYIQEFCLANGYLLPKNSFLKEIFYLKYYFLDMVTEDAKTLLKEICLRN*

>MM_Irga7

MDQLLSDTSKNEDNDDLVSSFNAYFKNIKTENKIISQETIDLIELHLNKGNIHGANSLIREALKNIDNAPINIAVTGESGVGKSSFINA
LIGTGPEEEGAAEVGVIETTMKRNFYKHPKIETLTLWDLPGIGTQKFPPKTYLEEVKFKEYDFFIIVSSTRFTKHELDLAKAIGIMKKN
YYFVRTKVDIDLENERKSKPRTFDREKTLKQIQSYAMNTFSDNNMAIPPIFMVSNYDLSKYDFPVMMDTLIKDLHAEKRHNFMLSLPGI
TEAAIDRKHKATQQIVWLEAFNVGLLANFPVTGILGDNDVKKLEKSLNYYRKIFGVDDESLELVAKDFQVPVEQVKEIMKSPHLLKTNG
KETLGEKLLKYLEKFETATGGLLAVGLYFRKTYYLQLHFLDTVTEDAKVLLRWKYSKPRSNSTYP*

>MM_Irga8

MGQLFSNMPKDEDKGNLESSFTEYFRNYKQETKIISEETTRSIELCLKKGDIQRANSIISDALKNIDNAPINIAVTGESGAGKSSLINA
LREIKAEEESAAEVGVTETTMKVYSYKHPKVKNLTLWDLPGIGTKKFPPKTYLETVEFKKYDFFIIVSAIRFTNHEIELAKAIRIMKKN
YYFVRSKVDFDLYNEEKSKPRNFNRENTLNQVRNYYLDTFRESKIDEPQVFLISNHDLSDYDFPVLMDTLLKDLPAEKRHNFLLSLPNI
TEAAIQKKYNSPKQYIWLQAMEDGLLATVPAVGILKDLDKERLKRSLDYYRDLFGVDDESLMFMAKDAQVPFELLKIKLKSPYLLELEE
ETLGGLILNCVEKFASANGGLLATGLYFRKTYYLQFHFLDTVAEDAKVLLKEAY*

>MM_Irgb1

QHPPLNTATCQTSTGRTSQITAQLLEFNFKNFFKNFKKESKILSEETITLIESHLENKNLKEALTVISHALRNIDKAPLNIAVTGETGT
GKSSFINALRGISSEEKDAAPTGVIETTMKRTPYPHPKLPNVTIWDLPGIGSTNFPPQNYLTEMKFGEYDFFIIISATRFKEIDAHLAK
AIAKMNIKFYFVRTKIDQDISNEQRSKPKSFNRDSVLKKIKDECLGLLQKVLSSQPPIFLVSNFDVSDFDFPKLETTLLKELPAHKRHL
FMMSLHSVTETTIARKRDFLRQKIWLEALKAGLWATIPLGGLVRDKMQKLEETLTLYRSYFGLDEASLENIAKDFNVSVNEIKAHLRSL
QLLTKNNDMSFKEKLLKYIEYISCVTGGPLASGLYFSKTYYWQSLFIDTVASDAKSLLNKEEFLSEKPGSCLSDLPEYWETGMEL*

>MM_Irgb2

MGQTSSSTSPPKEDPPLTFQVKTKVLSQELIASIESSLEDGNLQETVSAISSALGDIEKVPLNIAVMGETGAGKSSLINALQGVGDDEE
GAAASTGVVHTTTERTPYTYTKFPSVTLWDLPSIGSTAFQPHDYLKKIEFEEYDFFIIVSAIRIKQSDIELAKAIVQMNRGLYFVRTKT
DSDLENEKLCNPMRFNRENILKSIRICLSSNLKERFQQEPPVFLVSNFDVSDFDFPKLESTLLSQLPAYKHQIFMSTLQVVINAIVDRK
RDMLKQKIWKESIMPRAWATIPSRGLTQKDMEMLQQTLNDYRSSFGLNEASLENIAEDLNVTLEELKANIKSPHLFSDEPDTSLTEKLL
KYIGNPYFSKVFHLQNYFIDTVASDAKIILSKEELFTEQVSSFNSKASPYREESVGKVFPVSPGSTFLFHFFEMFQSDSDKLCHVHVLL
LLTSWGLSGETVT*

>MM_Irgb3

QHPPLHTATCQPSSSRPSPLMAQLLVFSFENFFKNFKKESKILSEETITLIESHLEDKNLQGALSEISHALSNIDKAPLNIAVTGETGT
GKSSFINALRGVRDEEEGAAPTGVVETTMKRTPYPHPKLPNVTIWDLPGIGSTTFPPQNYLTEMKFGEYDFFIIISATRFKEIDAHLAK
TIEKMNTKFYFVRTKIDQDVSNEQRSKPRSFNRDSVLKKIRDDCSGHLQKALSSQPPVFLVSNFDVSDFDFPKLETTLLRELPSHKRHL
FMMSLHSVTETAIARKRDFLRQKIWLEALKAGLWATIPLGGLVRNKMQKLEETLTLYRSYFGLDEASLENIAKDFNVSVNEIKAHLRFL
QLFTKNNDMSFKEKLLKYIEYISCVTGGPLASGLYFRKTYYWQSLFIDTVASDAKSLLNKEEFLSEKPGSCLSDLPEYWETGMEL*

>Mouse_Irgb4

QHPPLHTATCQPSSSRPSRLTAQLLVFSFENFFKNFKKESKILSEETITLIESHLEDKNLQGALSEISHALSNIDKAPLNIAVTGETGT
GKSSFINALRGVRDEEEGAAPTGVVETTMKRTPYPHPKLPNVTIWDLPGIGSTTFPPQNYLTEMKFGEYDFFIIISATCFKEIDAHLAK
TIEKMNTKFYFVRTKIDQDVSNEQRSKPRSFNRDSVLKKIRDDCSGHLQKALSSQPPVFLVSNFDVSDFDFPKLETTLLRELPAHKRHL

FMMSLHSVTETAIARKRDFLRQKIWLEALKAGVWATIPLGGLVRNKMQKLEETLTLYRSYFGLDEASLENIAKDFNVSVNEIKAHLRSL
QLLTKNNDMSFKEKLLKYIEYISCVTGGPLASGLYFRKTYYWQSLFIDTVASDAKSLLNKEEFLSEKPGSCLSDLPEYWETGMEL*

>MM_Irgb5

MGQTSSSTPPPKEDPDLTSSFGTNLQNFKMKTKILSQELIAFIESSLEDGNLQETVSAISSALGGIEKAPLNIAVMGETGAGKSSLINA
LQGVGDDEEGAAASTGVVHTTTERTPYTYTKFPSVTLWDLPGIGSTAFQPHDYLKKIEFEEYDFFIIVSSGRFKHNDAELAKAIVQMNR
SFYFVRTHTDLDLMVVKRSNPRRFNRENTLKQIRHTISSMLKEVTHQEPPVFLVSNFDVSDFDFPKLESTLLSQLPAYKHHMFMLTLPI
VTDSTIDRKRDMLKQKVWKESTMPRAWATIPSLGLTQKDMEMLQQTLNDYRSSFGLDEASLENIAEDLNVTLEELKANIKSPHLLSDEP
DTSLTEKLLKYIGNPYFSKVFHLQNYFIDTVASDVKIILSKEELFTEQVSSFNSKASLYREESVGKVFPVGPGSTFLFHFIEMFQSDSD
ELCHVHVLLLLTSGGLSSETVT*

>MM_Irgb6

MAWASSFDAFFKNFKRESKIISEYDITLIMTYIEENKLQKAVSVIEKVLRDIESAPLHIAVTGETGAGKSTFINTLRGVGHEEKGAAPT
GAIETTMKRTPYPHPKLPNVTIWDLPGIGTTNFTPQNYLTEMKFGEYDFFIIISATRFKENDAQLAKAIAQMGMNFYFVRTKIDSDLDN
EQKFKPKSFNKEEVLKNIKDYCSNHLQESLDSEPPVFLVSNVDISKYDFPKLETKLLQDLPAHKRHVFSLSLQSLTEATINYKRDSLKQ
KVFLEAMKAGALATIPLGGMISDILENLDETFNLYRSYFGLDDASLENIAQDLNMSVDDFKVHLRFPHLFAEHNDESLEDKLFKYIKHI
SSVTGGPVAAVTYYRMAYYLQNLFLDTAANDAIALLNSKALFEKKVGPYISEPPEYWEA*

>MM_Irgb7_pseudo

PFWFVPPLGTIDICQDWVKLPLLHPLQRRILLLTFQMKTKILSQELITFIELYLEDGNL*ETVSAISSALGDIEKVPLNIAVMGETGAG
KSSLINALQGTGADEDGVTAPVGVVYTTIEKKSYPYAKFPSAILWELPAIGFHHFQPHDYLKKIKFEEYDFIIVSAGRIKHSDVELAKA
IVQMNRGLYFNRTKTDIDLKNEKLYNPMRFNRENTLKSLQICISSNLKECFHQEPPVFLVSNFDVSDFDFPKLESTLLSQLPAYKHQIF
MRTLQVVINAIVDWKRDMLKQKVWKESTTPRAWATIPSLGLTQKDMEMLQQTLNDYRSSFGLDEASLKNIAEDLNVTLEELKANIKSPH
LLSDEPDTSLTEKLLKYIGNPYFSKVFHLQNYFIDTVASDVKIILSKEELFTEQVSSFNSKASPYREESVGEVFPVGPGSTFLFHFFEM
FQSDSDKLCHVHVLLLLTSWGLSGETVT

>MM_Irgb8

QHPPLHTATCQPSSSRPSPLMAQLLVISFENFFKNFKKESKILSEETITLIESHLEDKNLQGALSEISHALSNIDKAPLNIAVTGETGT
GKSSFINALRGVRGEEEGAAPTGVVETTMKRTPYPHPKLPNVTIWDLPGIGSTNFQPQNYLTEMKFGEYDFFIIISATRFKEIDAHLAK
AIAKMNTKFYFVRTKIDQDVSNEQRSKPKSFNRDSVLKKIRDDCSGHLQKVLSSQPPVFLVSNFDVSDFDFPKLENTLLRELPAHKRHL
FMMSLHSVTETAIDRKRDFLRQRIWLEALKAGVWTTIPLGGLVRDKMQKLEETLTLYRSYFGLDEASLENIAKDFNVSVNEIKAHLRSL
QLLTKNNDMSFKEKLLKYIEYISCVTGGPLASGLYFRKTYYWQSLFIDTVASDAKSLLNKEEFLSEKPGSCLSDLPEYWETGMEL*

>MM_Irgb9

MGQTSSSTLPPKDDPDFIASFGTNLQNFKMKTKILSQELIAFIESSLEDGNLRETVSAISSALGGIEKAPLNIAVMGETGAGKSSLINA
LQGVGDDEEGAAASTGVVHTTTERTPYTYTKFPSVTLWDLPGIGSTAFQPHDYLKKIEFEEYDFFIIVSSGRFKHNDAELAKAIVQMNR
SFYFVRTHTDLDLMVVKLSDPRKFNKENILEQIRNSISNILKEVTHQEPPVFLVSNFDVSDFDFPNLESTLLSQLPAYKHHMFMLTLPI
VTDSTIDRKRDMLKQKIWKESIMPRAWATIPSRGLTQKDMEMLQQTLNDYRSSFGLDEASLENIAEDLNVTLEELKANIKSPHLLSDEP
DTSLTEKLLKYIGNPYFSKVFHLQNYFIDTVASDVKIILSKEELFTEQVSSFNSKASPYWEESVGKVFPVGPGSTFLFHFFEMFQSDSD
KLCHVHVLLLLTSWGLSGETVT*

>MM_Irgb10

MGQSSSKPDAKAHNMASSLTEFFKNFKMESKIISKETIDSIQSCIQEGDIQKVISIINAALTDIEKAPLNIAVTGETGAGKSTFINALR
GIGHEESESAESGAVETTKDRKKYTHPKFPNVTIWDLPGVGTTNFKPEEYLKKMKFQEYDFFLIISSARFRDNEAQLAEAIKKMKKKFY
FVRTKIDSDLWNEKKAKPSSYNREKILEVIRSDCVKNLQNANAASTRVFLVSSFEVAQFDFPSLESTLLEELPAHKRHIFVQCLPTITE
PAIDRRRDVLKQTIWLEALKAGASATIPMMSFFNDDIEEFEKILSHYRACFGLDDESLENMAKEWSMSVEELESTIKSPHLLSSEPNES
VADKLVKTMEKIFAVTGGFVATGLYFRKSYYMQNYFLDTVTEDAKVLLKKKVFLQDSVDSE*

>MM_Irgd

MDQFISAFLKGASENSFQQLAKEFLPQYSALISKAGGMLSPETLTGIHKALQEGNLSDVMIQIQKAISAAENAILEVAVIGQSGTGKSS
FINALRGLGHEADESADVGTVETTMCKTPYQHPKYPKVIFWDLPGTGTPNFHADAYLDQVGFANYDFFIIISSSRFSLNDALLAQKIKD
AGKKFYFVRTKVDSDLYNEQKAKPIAFKKEKVLQQIRDYCVTNLIKTGVTEPCIFLISNLDLGAFDFPKLEETLLKELPGHKRHMFALL
LPNISDASIELKKHFLREKIWLEALKSAAVSFIPFMTFFKGFDLPEQEQCLKDYRSYFGLDDQSIKEIAEKLGAPLADIKGELKCLDFW
SLVKDNSIIAQATSAAEAFCAVKGGPESSAFQALKVYYRRTQFLNIVVDDAKHLLRKIETVNVA*

>MM_Irgm1

MKPSHSSCEAAPLLPNMAETHYAPLSSAFPFVTSYQTGSSRLPEVSRSTERALREGKLLELVYGIKETVATLSQIPVSIFVTGDSGNGM
SSFINALRVIGHDEDASAPTGVVRTTKTRTEYSSSHFPNVVLWDLPGLGATAQTVEDYVEEMKFSTCDLFIIIASEQFSSNHVKLSKII
QSMGKRFYIVWTKLDRDLSTSVLSEVRLLQNIQENIRENLQKEKVKYPPVFLVSSLDPLLYDFPKLRDTLHKDLSNIRCCEPLKTLYGT
YEKIVGDKVAVWKQRIANESLKNSLGVRDDDNMGECLKVYRLIFGVDDESVQQVAQSMGTVVMEYKDNMKSQNFYTLRREDWKLRLMTC
AIVNAFFRLLRFLPCVCCCLRRLRHKRMLFLVAQDTKNILEKILRDSIFPPQI*

>MM_Irgm2

MEEAVESPEVKEFEYFSDAVFIPKDGNTLSVGVIKRIETAVKEGEVVKVVSIVKEIIQNVSRNKIKIAVTGDSGNGMSSFINALRLIGH
EEKDSAPTGVVRTTQKPTCYFSSHFPYVELWDLPGLGATAQSVESYLEEMQISIYDLIIIVASEQFSLNHVKLAITMQRMRKRFYVVWT
KLDRDLSTSTFPEPQLLQSIQRNIRDSLQKEKVKEHPMFLVSVFKPESHDFPKLRETLQKDLPVIKYHGLVETLYQVCEKTVNERVESI
KKSIDEDNLHTEFGISDPGNAIEIRKAFQKTFGLDDISLHLVALEMKNKHFNTSMESQETQRYQQDDWVLARLYRTGTRVGSIGFDYMK
CCFTSHHSRCKQQKDILDETAAKAKEVLLKILRLSIPHP*

>MM_Irgm3

MDLVTKLPQNIWKTFTLFINMANYLKRLISPWSKSMTAGESLYSSQNSSSPEVIEDIGKAVTEGNLQKVIGIVKDEIQSKSRYRVKIAV
TGDSGNGMSSFINALRFIGHEEEDSAPTGVVRTTKKPACYSSDSHFPYVELWDLPGLGATAQSVESYLEEMQISTFDLIIIVASEQFSS
NHVKLAITMQRMRKRFYVVWTKLDRDLSTSTFPEPQLLQSIQRNIRENLQQAQVRDPPLFLISCFSPSFHDFPELRNTLQKDIFSIRYR

DPLEIISQVCDKCISNKAFSLKEDQMLMKDLEAAVSSEDDTANLERGLQTYQKLFGVDDGSLQQVARSTGRLEMGSRALQFQDLIKMDR
RLELMMCFAVNKFLRLLESSWWYGLWNVVTRYFRHQRHKLVIEIVAENTKTSLRKALKDSVLPPEIH*

>RN_Irgq

MPLPQGDVTALFLGPPGSGKSALIAALCGKNVDAVEIPDGRPDSGVPSLRAAAPGLFLGELSCPPAAPGPWAAEANLLVLVLPGSEGSE
EPLTPALGEAARAALARGTPLLAVRNLRPGDSQNAAQARDETAALLDSAGLGAAPLFVLPADCSSSDGCEELERLQVVLRTQAEALQRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARVDLAVAGTTNVGLVLDMLLGLEPGDAGAAPASAPTEPTPYPAP
ERPNVVLWTVPLGPTATSPAVTPHPTHYDALILVTPGAPTEENWAQVRSLVSPDAPLVGVRTDGQGEDPPEVLEEEKAGNARDGNSGDA
RSEGKKPGIGDSGCSAVRSPEDEQWEVLEEAPPPVFPLRPGGLPGLGTWLQRALPTAQAGALLLALPPASPRAARRKAAALRAGAWRPA
LLASLAAAAAPVPGLGWACDVALLRGQLAEWRRALGIEPAAVARRERALGLAPGVLATRTRFPGPVTRAEVEARLGSWAGEGTAGGAAL
GALSFLWPAGGAAATGGLGYRAAHGVLLQALDEMLADAEAVLGPPEHNQ*

>RN_Irgc

MATSRLPAVPEEETTILMAKEELEALRTAFESGDIPQAASRLRELLATTETTRLEVGVTGESGAGKSSLINALRGVGAEDPGAALTGVV
ETTMQPSPYPHPQFPDVTLWDLPGAGSPGCSADKYLKEVDFGRYDFFLLVSPRRCGAVETRLASEILRQGKKFYFVRTKVDEDLAATRN
QRPSGFSEAAVLQEIRDHCAERLRAAGLSDPRIFLVSNLSPNRYDFPMLVTTWEHDLPAHRRHAGLLSLPDISLEALQKKKDMLQEQVL
KTALVSGVIQALPVPGLAAAYDDALLIRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMRV
ARAFERGIPVFGTLVAGGISFGTVYTMLQGCLNEMAEDAQRVRIKALEEDETQGEVSLEAAGDNAVEKRSSGEGTSEEAPLSTRRKLGL
LLKYILDSWKRRDLSEDK*

>RN_Irgm1

MPETSTHNAPLNLSFPSVPSYQIGCSSLPEISRSTERALKEGKLPELVYGVKETVATLSQIPVSIFVTGDSGNGMSSFINALRIIGHEE
DASAPTGVVRTTQTRAEYSSSHFPNVVLWDLPGLGATAQTVENYIEEMKFSTCDLFIIVASEQFSSNHVKLAKIIQSMGKRFYVIWTKL
DRDLSTSVLSEVRLIQNIQENIRENLQKEGVKEVPIFLVSNLDPLLHDFPELRNTLQTDLSNIRCCEPLKTLYVIYEKIIGDKVANWNQ
IIANGRLKSSLGVRDDDDMGECLKRYRLIFGIDDESLQQIAHGMGTVVMEYKANIKSQDFHTLRRADWKLRLMTCTTVNALFCLFKFLP
CLCHCFKRMRHKRMLLLVAKDTKNILKKILMDAVSPPQI*

>RN_Irgm2

MEEAVGLPEDKQFACLSDAVFISKDNSILSVEVIKSIQAAVAGGNGVEVVSIVKEIVQKVSRTTMKIAVTGDSGNGMSSFVNALRLIGH
EEEDSAPTGVVRTTQKPACYSSSHFPYVELWDLPGIGTTAQSMESYLDEMQFSAYDLIIIIASEQFSSNHVKLAEAMQRMRKKFYVVWT
KLDRDISTSTFPEPQLLQSIQKNIRENLQKGKVKEPPIFLVSIMKPLLHDFERLRETLRKDLSDIKYHGLLETLYQICENTINERVESI
KKIIDENNLQREFGILTPDNLTETRKVFQEIFGVDDQSLSQVSRSMEKPDTHYKASIESQEIQGYQQDGWPLVWLHRPVIQFFSTGLDR
VPCCFYSPHHRYTQQKGVLDETAGKTKNFLWKILKDSISHLQKT*

>RN_Irgm3

MAKPLKPPLFKSITAGESSYSSQNSSSPEVIEKVGKAVAEGDLQKVIYTVKEEMQSKSRYTVKIAVTGDSGNGMSSFVNALRLIGHEEE
DSAPTGVVRTTQKPACYSSFHFPYVELWDLPGTGVTAQSMESYLDEMQFSAYDLIIIIASEQFSSNHVKLAEAMQRMRKRFYVVWTKLD
RDISTSTFPEPQLLQSIQKNIRENLQKAQVRDPPIFLVSCFSPSFHDFLDLRETLRKDIHNIRYRDPLETLSQVCDKCINNKALSLKED
LMFTKHLEAAVSPPYDIADLERSLDTYQKLFGVDNESLRRVAQSTGRPEMSTRALQFQDLIKMDRRLRLMMCFVVNILLRVLGSPWWFG
LWDVVTRYFRHQRQKRIIEIVAKNTKTSLRRALEDYTLPPEILCEGSGVPSSGIQAASGSFCIEP*

>RN_Irgd

MDQFITAFLKGASEKNFQQLAMEFLPQYSALISKSGGMLSPETLSAIHYALQEGRLSDVMNQIQEAISAAENAVLEVAVIGESGTGKSS
FINALRGLGHEEAESADVGTVETTMYKTPYQHPKYPNVIFWDLPGTGTPNFHTDTYLDRVGFANYDFFIIISSSRFSVNDALLAQKIKD
AGKKFYFVRTKVDSDLYSEERTRPRTFRKEQVLQRIRDYCLSNLTDIGVSEPRIFLISNFDLDAFDFPKLEETLLKELPGHKRHMFALL
LPNISDASIELKKHFLQEKICLEALKSGAMSFIPFMPFISGFDLPQQEQCLKDYRSYFGLDDKSIEEIAERLETPLEDIKGQLKCLDFW
SFVKDDSIIARARSAGEAFCSVKGGLGSSVVQALKVYYMRTQFLNVVVEDAKHLLRKMETVNIA*

>RN_Irga5

MGQLFSGTAKSEALYSSFSEYFKKFKAENKIISQETITLIELYLILEDLPQANNEITSALRKLANTPLNVAVIGESGTGKSSFINVFRG
VGHEDETAAPIGVVETTMRRTPYRHPNIPNVVIWDLPGIGTTNFPPKDYLEKMKFCEYDFFIIISATRFRKNDIDLAKAVSMMKKDFYF
VRTKMDIDLENEMECKDTFSRETFLKHIRSHCVTMFKKNNLHVPPIFLISNRNVSDYDFPILKAMLQNKLSTHTYHNIMVSLPNITEAA
IERKHTFIQQFIWLEAFKDGVLMTIPVVDTLKDSDVEKLKMSLNHYRVLFGVDDATLQFMAKDSQVPVEQLRKIIKSPYLLETKKRKAL
EGMLLKYMEKSASANGGLLATGLYFRKSFYLQLLFLDTVAEDAKVLLRETHSRN*

>RN_Irga11

MGQLFSLTTNEQGEDLPSSFAKYFKKFKTGHKIISEEIITSVELSMTKGNIQMANSAISEAFREIDSTPLNVAVTGESGAGKSSFINAL
RGIGHEEEGAAEIGVVETTMWRHKYQHPSMPNVVIWDLPGIGTTNFPPKTYLEKMKFYEYDFFIIISATRFKKNDIDLAKAISMMKKDF
YFVRTKVDSDLRNEENTKPRSFDREKVLQNIRLNCVKHFKENGMDEPPIFLISNIDLSDYDFPILMDKLISDLPVYKRHTFMLSLPNIT
DSTIEMKRQCLKQRIWLEAFAADLLRILPSLTFLLDSDLETLKKCLKFYRTVFGVDDAALQSLAKDWQMPLVELEAMMKSPIVFKPTDE
ETIHERLSRYYHDYCSANGHLFTDDRDLREISYLKYYFLDIVTEDAKTLLKEICVRNKLVSN*

>RN_Irga12

MGQWFSSKNEQHQDLASSFKEYFKKFKTGHKIISEEIITSVELSMTKGNIQMANSTISEALRDIDGTPLNVAVTGESGAGKSSFINALR
GIGHEEEGAAEIGVVETTAERWPYKHPSMPNVVIWDLPGIGTTNFPPKTYLEKMKFYEYDFFIIISATRFKKTDTDLAKAISMMKKDFY
FVRTKVDSDLRNEENTKPRSFDREKVLQNIRLNCVKHFKENGMDEPPIFLISNIDLSDYDFPILMDKLISDLPVYKRHTFMLSLPNITD
STIEMKRQCLKQRIWLEAFAADLLRILPSLTFLLDSDLETLKKCLKFYRTVFGVDDAALQSLAKDWQMPLVELEAMMKSPIVFKPTDEE
TIHERLSRYYHDYCSANGHLFTDDRDLREISYLKYYFLDIVTEDAKTLLKEICVRNKLVSN*

>RN_Irga13

MGQWFSSKNEQHQDLASSFKEYFKKFKTGHKIISEEIITSVELSMTKGNIQMANSAISEALREIDGTPLNVAVTGESGAGKSSFINALR
GIGHEEEGAAKIGVVETTAERWPYKHPSMPNVVIWDLPGIGTTTFPTKTYLEKMKFYEYDFFIIISATRFKKNDIDLAKAISMMKKDFY
FVRTKIDSDLRNEEEFKPRSFDREKVLQNIRFNCVKHFKENGIDEPPIFLISNRNLSDYDFPILMDKLISDLPVYKRHTFMLSLPNITD

SAIEKKRQSLKQKIWLEAFAADLLSIIPSLTFLLDSDLETLKKCLKFYRTVFGVDDAALQSLAKDWQMPLVELEAKMKSPIVFKPTDEE
TIHERLSRYYRDFCLANGYLVTQNLYLREIFYLKFYFLDIVTEDSKTLLKEICLRNKLVSN*

>RN_Irga14p

MGQLFSDTSKNEDNEGDLVSSFKAYFKKLIQKQKSFLQKLSIxIELHLSKGDFLGANDLISDALKNIDNIPINIAVTGESGAGKSSFIN
ALIGIRPEEEGAAAVGVVETTMKRTPYQHPKIKTLTLWDLPGIGTQKFPPKTYLEKVKFKEYDFFIIASATRFTKLELDLAIAIRILKK
NYYFVRTKVDIDLNHALLTEKRPxQIRSYSVNTFSNNYMDVPQIFLISNYDLSDYDFPILVDTLNKNLPAQKRYNFMLSLPKITEAAID
RKHKASQQFIWLEAFKVGILATFPVVGILRDNDVEKLTESLNHYRQLFGVDDESLELMARDFQVPVGQLKEKIKSPHLLKTNREETLGK
KLLKYLEKFASANGGLLATGLYFRKTFYLQLHFLDTVAEDAKVLLQWTYSKQ*

>RN_Irga15

MGQLFSDTSKSEDNGGDLVSSSNAYFKKINTKTKIISPETIRLIELHLSKGNILGASDLISDALKNIESIPINIAVTGESGAGKSSFIN
ALRGIRPEEEGAAEVGVVETTMERTPYQHPKIKTLTLWDLPGIGTQKFPPKTYLEKVKFEEYDFFIIASATRFTKLELDLAKAIRIMKK
NYYFVRTKVDFDLENEKRSKPRTFDREKTLKKIRGCTMKTFRENNMDVPQIFLISSYNLSDYDFPVLMDTLIKDIPAQKRYNFMLSLPK
ITEAAIDRKHKAMQEFVWLEAFKTGALATIPALGILRDNDVEKLRQKLNNYRQLFGVDDESLEFMAKDFQVPVAQLKEILKSPHLLKTD
REETLQDKLLKYLEIFASANGGLLATGLYFRKTYYLQLHFLDTVAEDAKVLLQWKYSKH*

>RN_Irga16p

MKTGNLESSLQICFTAYFRKYKPETKIISEEITRLTELCLKRGDLQGANSVISDALKNIDIAPINIAVTGVSGAGKSSLINALREVKDE
GEGAAEVGVAESTMKTDSYEQPQNxKSLTLWDLPGIETQKFQQKKIIWKKxEFKKYDFFIIVSSIQIAKHEVDLAKAIGIMKKNYYVVR
TKVDSDLERGEIHRPHSFNRENTLNQI*GDCLDTSRDNEIDEPQLFLISDHNLSDYDFPVLMDTLIKDLPAEKRHNFLPSLPNITEAAI
QTKKYNSTKQFIWLEAMKDGVLATVPVVGILNDLDMEGLxRYRDLLGAEDESLAFLAKDAQVPLTYxEKRKLKSPYLLEIKKEEALxSI
ALELFGEIYFSxINDGFLATGLYFGKTYFLQTYFLDTVTEEPKVLLKEAYSKNIAQTQLAHSCRQLRDQRVNSRESLDKFML*

>RN_Irgb10

MGQSSSKPDAKAHNMASSFNEFFKSFKMESKILSEETINSIQSCVQEGDIQKGISIINAALADIEKAPLNIAVTGETGAGKSTFINALR
GVGHEESESAKIGAVETTMDKFPKFPNVTIWDLPGVGTCNFKPEEYLKKLRFQEYDFFLIISATRFRENDAQLAKAIKKMKKNFYFVRT
KIDSDLWNQKKCKPKSYNKEKILEEIRKDCVEKLQNARVASARVFLVSSVEVAQFDFPELESTLLEELPAHKRHVFMQCLPSITERAID
RRRDALRQKIWLEALKYGASATIPMMCFFNDDIEELEKILTHYRGSFGLDDESLKNMASEWSMSVEELKSFINSPHLLSCEMNESVSDK
MVKIMEKIFAVTGGLIATGLYFRKSYYMQNYFLDTVSEDAKILLKKKVFLQGSEDSE*

>RN_Irgb13

QHPPGHTATCKSSSSRSSPLTAQLLSLGLKIFFKSFKKESKILSEETVTLIESHLEDKNLQGALSTISHALRNIDKAPLNIAVTGETGA
GKSSFINALRGVRDDEECAAPTGVVEKTKERTPYPHPKLANVTIWDLPGIGSTTFPPQNYLAEMKFGEYDFFVIISATRFKETDAHLAK
AIAKMNTKFYFVRTKIDQDLRNEEKSKPKVFNRDGVLKKIRDDCSQHLQKDLSSEPPIFLVSNFDVSDFDFPKLETTLLSELPAHKRHI
FMLSLHNVTETAIDRKRDFLKQKIWLEALKAGAWTTIPFGGLVHDKKQTLEDTLNLYRSYFGLDEASLEKIANNFNVSVDEIKAHIKSL
HLLTENKDMSFGEKLLKYIEYISSFTGGPLASGLYFRKTYYWKSLFIDTVASDAKALLNKEAFLSEKPGLRVSDHTEYWEAGMEL*

>RN_Irgb14

MGQTSSSTTPPKEDPDLTSSFGTNLQNFEMKTKILSQELITFIESSLEGGNLRETVSAISDALSDIEKAPLNIAVIGETGAGKSSLINA
LQGVGADKEGTAAPTGVVHTTSERTPYTYTKFPCVTLWDLPGIGSPAFQPHDYLKKIKFEEYDFFIIVSSGRFKHNDAELAKAIVQMNR
SFYFVRTHIDLDLMVVKLSAPKRFDKENILEEILNSISSILKEVTYQEPPVFLVSNFNVSDFDFRKLETTLLEELPAYKRHIFMLTLPT
VTESTIDRKRDMLKQKIWKESIMPRAWASIPFRGLTQNDIEMLEQTLNDYRSSFGLDEASLENIAGDLNVTLEELKANIKSPHLLSYEP
DISLRDKLLKYISHPYFSKVFHLQNYFIDAVASDVKLILSKEELLTNKVRSFNSNVSRQLF*

>CP_IRGC

XXGKKFYFVRTKVDEDLAATRTQRPSGFSEATVLHEIRDHCAERLRAVGVSDPRIFLVSNLSPTRYDFPTLVSTWEHDLPAHRRHAGLL
SLPDISLEALQKKKDMLQEQVLKTALVSGVIQALPVPGLAAAYDDALLIRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLA
NEVSPETVLRLYSQSSDGAMRVARAFERGIPVFGTLVAGGISFGTVYTMLQGCLNELAEDAQRVRIKALEEDEPQPEISLEAAGDNSVE
KRVTGEAGGEEGQLTTRRKLGLLLKYVLESWKKRDLEEK*

>CP_IRGQ

FVLPADCNRSDGCEELERLRAALRSQAEALQRFLPPAQDGFEVLGAAELEAVREAFETGARRLHXXXGTTTRGCWDSGRIAALSPEDET
WEVLEEAPPPIFPLHLSGLPGLCEWLQRALPSAQAGALLLALPPASPRAARKKVAALRAGAWRPALLASLAAAAAPVPGLGWACDVALL
RGQLAEWRRALGPRTCIAGAARARLGLGSWRAGCSCTLPRPRDTRxEVETRLGAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAA
HGVLLQALDEMQADAEAVLAPGEPAQ*

>CP_GMS1

MENEPLSIAVTGESGAGMSTFINALRGTWDEEEGVAPTGPMATTLQATAYTLQSARGTLWDLPGLGRPNFPLQEYLKEIPDCDLFFIIF
ATRLKCNDIELAKAIAQIKKKVYFIQTKIDNALVSCQKARPLTFDKDKVLQEIRNCGLAQLQEARVVADKIFLVSSLDVCAYDFPELQS
TLVRDLPAHKRHAFMQRLSSVTEAAVNGKRDSLKQKIWLEALKAGAWAAVPLVGLFSDSERKKLEDTLSLYRSHFGVDDESLEKMAQDL
HVSLEDLKANLQSPDLLSAESDESFWEKTKQMVEMVLLVTGGPIATGLQFTKTFYLQNYFLDIVASDVKALLAKEDLFSASVGSAGLQK
EGN*

>CP_GMS2

MEEDLVSHRTPLSASFTSGVSYHRDGGVPLEVSRDLEKAVKEGQLLEVVAIARKVVEMASRAPVSLAVMGDSGNGMSSFVNALRGIGHE
DAASAPTGVVRTTLTPARYTSPSFPNVFLWDLPGMGASDQSLDHYLRELQHSQYDLFLLIASEQFSLHHVRLAKTIQGMGKRFYVIWTK
VDRDLSTTPLSRGLLLRNIQENILETLQKEGVHKPPIFLVSSLDPDLHDFPDLRKKLRIDIFNIRCSGPLEAMSYACKETINEKVASLK
TRVLQNCPEDTLGSCGADDLEQYLRVYRGCFGIDDDSLLQVAWSTGRVVSEYRTLLRSQDLCGLRRADCRLRLATCSVMKMLLRLLRWV
PWLGPRAVRWFAQVTHKRMLHLVAQDTKGILKKILEDSTCPA*

>CP_GMS2b

MEEDLVSHRTPLSASFTSGVSYHRGGGVPLEVSRDLEKAVKEGQLLEVVAIARKVVEMASRATVSLAVTGDSGNGMSSFVNALRGIGHE
DVASAPTGVVRTTLTPARYTSPSFPEVFLWDLPGMGASDQSLEHYLRELQHSQYDLFLLIASEQFSLHHVRLAKTIQGMGKRFYVIWTK

VDRDLSTTPLSRGLLLRNIQENILETLLKEGVHKPPIFLVSSLDPDLHDFPDLRKELRIDIFNIRCSGPLEAMSCACKETINEKVASLK
TRVLQNCPEDTLGSCGADDLEHYLRVYRGCFGIDDDSLLQVAWSTGRVVSEYRTLLRSQDLCGLRRADCRLRLXXX

>CP_GKS1

XSGRAT*RLQWIYEEILSGAKNAALEVAVIGASGTGKSSFINALRGLGHEEEGAAEVGVVETTMEKTPYKHPQYPNVTFWDLPGTGTPT
FTPDTYLEAVGFATFDFFIIISSSRFTCSDGLLAQKIQEAGKNFYFVRSKVDCDLDNERRAKPKSFQRERVLQEIRDYCLANLRNMGVT
DPPIFLVSNFELHGFDFPGLQRTLLGELPAHKRQVFALMLPALSDASIELKRSILKEKIWLEALKLAAVAFIPFGSIFKGFDLPEQEQC
LQLYQKYFGLDDESIEEIAKKLHTSVQDIKGELRCLDSSALLQDDSKAAMAMYCAEKFCSVTGGPISSTMHFVKAYSIRLKILDAVAQD
AKVLLHKTLRAPF*QNEVEALCPAS

>CP_GKS2

MRHFLCLPASPPPCILAPALSPPLWLPLHLCFLLQVLFLSCLSCLKSCQKSLVVLRPVLLFGCCFPVTLFSLLFTDLPESAAMDALISD
FLKNLTQKNFQQLAADFMSQSAAFISTAGGVIPPGTLSKIEVVLKEGNLRAAVDIIEEILSEVENAALEVAVIGESGTGKSSFINALRG
LGHEEEGAAEVGVVETTMKKTPYKHPKYPNVTFWDLPGTGTPTFTPDTYLEAVGFATFDCFIIISSSRFTCNDALLARKIQEAGKNFYF
VRSKVDCDFDNERRAKPQSFKRERVLQLIRDYCLANLYDIGVPDPRIFLVSNFELHDFDFPDLQRTLLGELPAHKRQAFAVMLPTLSDA
CIELKRGFLKEKIWLEAVKSSALAFIPLMPIFKGFDLSEQEACLKLYRKYFGLDDKSIAETAMKLGTSVQDIKGYTRCLDFWALVKDDS
TAAKAMRCAESFCSVNGGVTSVVAQFLKACFLRSKFLDTVADDAKLLLRKTINAHI*

>CP_GKS3

MGQASSSSTNPDTGENGDLASSFDEYFKNITMETKILTNKDQDMIKLYLQQGDVQKAASVINAVLKDIENAPLSIAVTGEGGSGKSTLI
NALRGVAHEDEGAAATGLTETTTEGTEYRHPKFPNVSIWDLPGVGTTKFSPEKYLKKVNFADYDLFLIVSCTRFKNNDAHLAKAIAKMK
KKFYFVRTKIDIDLSNEERAKPRNFNKEKLLEKMRNDIVTQLKAAGVSAAQIFLISSLDVGDYDFPEMERTLLRDLPAHKRHVFRMSLP
SLTEPAINQKTDSLKQKIWLEALKVGATATVPLVGLFSDSERKKLEDTLGLYRSHFGLDDESLEKMPQDXXXXX

>CP_GKS4

MRHFLCLPASPPPCILAPALSPPLWLPLHLCFVLHVLFLSCLSCLKSCQKSLEVLRPVLLFGCCFPVTLFSLLFTDLPESAAMDALISD
FLKNLTQKNFQQLAADFMSQSAAFISNAGGVIPPETLSKIEAVLKEGNLRAAVDIIEEILSEAKNAALEVAVIGGSGTGKSSFINALRG
LGHEEEGAAEVGVVEITKKKTPYKHPKYPNVTFWDLPGTGTPTFTPDTYLEAVGFATFDCFIIISSSRFTCNDALLAGKIQEAGKNFYF
VRSKVDSDLYNERRGKPQSFQRERVLQLIRDNCLANLYDIGVPDPRIFLVSNFELHDFDFPGLQRTLLGELPAHKRQAFAVMLPTLSDA
CIELKRGFLKQKIWLDAVKSSALAFIPFMPILKGFDLSEQEACLKLYRKYFGLDDKSIAETAMKLGTSVQDIKGYTRCLDFWALVKDDS
TAAKAMRCAESFCSVNGGVTSVVAQFLKACFLRSKFLDTVADDAKLLLHKIINAHI*

>CP_GKS5

MDALISDFMKNLTQKNFQQLAADFMSQSAAFISNAGGVIKPETLSKIEAVLKEGNLRAAVDIIEEILSEAKNAALEVAVIGGSGTGKSS
FINALRGLGHEDEGAAATGLTETTTEGTEYRHPKFPNVSIWDLPGVGTTKFSPEKYLKKVNFADYDLFLIGSCTRFKNNDAHLAKAIAK
MKKKFYFVRTKIDIDLSNEERAKPRNFNKEKLLEKMRNDIVTQLKAAGVSAAQIFLISSFDVGDYDFPEMERTLLRDLPAHKRHVFRMS
LPSLTEPAINQKTDSLKQKIWLEALKVGATATVPLVGLFSDSERKKLEDTLGLYRSHFGLDDESLERWPRICTCPWRTSRQTFXXX

>CP_GKS6

MGQASSSSTNPATGENGDLASSFDEYFKNITMETKILTNKDQDMIKLYLQQGDVQKAASVINAVLKDIENAPLSIAVTGEGGSGKSTLI
NALRGVAHEDEGAAATGLTETTTEGTEYRHPKFPNVSIWDLPGVGTTKFSPEKYLKKVNFADYDLFLIVSCTRFKNNDAHLAKAIAKMK
KKFYFVRTKIDIDLSNEERAKPRNFNKEKLLEKLRNDIVTQLKAVGVSAAQIFLISSFDVGDYDFPEMEKTLLRDLPAHKRHVFRMSLP
SLTEPAIXXXESFSEKMMQIVQTVLSVTGGPIATGLYFTKAFYIHNYFLDTVANDANVLLNKEDLFGASAGSVEGYQE*

>CP_GKS7

XEAFSVGNLQAVVDAIQEILSAAENAVLEVAVIGESGTGKSSFINALRGLGHEEEGAAKGGVVETTMKKTPYKHPQYPNVTFWDLPGVG
TPTFTPDTYLEAVGFATFDCFIIISSSRFTCNDALLARKIQEAGKSFYFVRSKVDS*LYNERRGKPQSFQRERVLQLIRDNCLANLYDI
GVPDPRIFLVSNFELHDFDFPGLQRTLLGELPAHKRQAFAVMLPTLSDACIELKRGFLKQKIWLDAVKSSALAFIPFMPILKGFDLSEQ
EACLKLYRKYFGLDDKSIAETAMKLGTSVQDIKGYTRCLDFWAPGEG*

>ST_Irgq

MLPPRGDVTALFLGPPGSGKSSLIAALCDKNVETVEIPEGRPDSGIPSLRAAGPGLFLGELSCPPAVPGPWAAEANVLVLVLPGHEGNG
ELLDPALAEAARAALARGTPLLAVRNLRHGDSQNDARTQTAALLNSAGLGAATLFVLPAECCSSDCWEELERLRVALRSQAEALQRLLP
PAQDGFEVLGAAELEAVREAFETGGLEAALSWVRSGLERLGSARLDLAVTGTTDIGLVLNMLLGLDPNDPGAVPASVPTVPTPFPAPER
PNVVLWTVPLGPMGTSPAAAPHPTHYDGLILVTLGAPTEKDWAQVRSLVVPDGPLVCVRTDGEGEDPESLEEEKLKPRDADLQKKGEGG
LENAPSDPKEKPGTGSQKAGGEDSEKAGSEGAGFKKPGSGDSGGTGALSPEDETWEVLEEAPPPVFPLRPGGLPALCEWLQRGLPPAQA
GALLLALPPTSPGAARRKAAALRAGAWRPALLASLAAAAAPVPGLGWACDVALLRGQLAEWRRALGLEPxALARRERALGLAPGELAAR
THFPGPVTRAEVEARLGAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHGVLLQALEEMQADAEAVLAPPEPTQ*

>ST_IRGC

MATSKLPSAPGEEETTILMAKEELEALRTAFESGDIPQAASRLRELLAASESTRLEVGVTGESGAGKSSLINALRGLGAEDPRAALTGV
VETTMQPSPYPHPQFPDVTLWDLPGTGTPNFLPDTYLETVGFTSYDFFIIISSSRFSFNDALLAQKIKEEGKKFYFVRTKVDSDLYNEE
VTKPKSFKRERVLQQIRDNCLANLSNIGVPEPCIFLVSNFNLDNFDFPRLQETLLKDLPAHKRHIFALLLPTFSEASIEIKRDFLKEKI
WLDAVKSASLAFVPFMPIIYGFDLPEQEKCLEGLPTTFWS*

>ST_GKS1

MGQASSSTTPNKEAQDFTFSCDKFFKTFKMESKILSPETIASIQSHLDEGNIQKTVSAINDALKNIQNAPLNIAVTGESGAGKSTFINA
LRGVGHEEKDAAATGVVETTMERTRYQHPKLPNVIIWDLPGIGTTNFQPRKYLKKMMFGEYDFFIIISSTRFKENDAHLAKAIAKMNKQ
FYFVRTKIDSDIYNQKICTPKSFNRDKLLQKIRDDCLKHLKDNNINGAQVFLVSSVHVSDYDFPNLETTLLKELPAHKRYIFMQCLPSV
TEAAIDRKRDSLKQMVWLEALKAGASATIPMMGLINDNDIQKLKETLTLYRSYFGLDDASLENMAKDFQVSMQELKANIKSPHLLSVDR
DESLGEKILTYIEKVCSVTGGFLATGLYFRKVFYLQTYFLDTVVSDAKVLLKREVLFTDSEISEQSFKNGVSEAESP*

>ST_GMS1

MSHNFCASLSVTILTPCNSISVGYSTLINMAEAPLSPNTMFSSSHTSVMLDHKCSSILSTDIIRDIEKAWKEGKLLEVVSVVRKTETVS
RMPVSIAVTGESGNGMSSFINALRRLKHEDEDSAPIGVVKTTQTRGSYSSFHFPNVVLWDLPGMDTTAQSLENYLNEMQFSQYDLFIII
ASEQFSMNHVKLAKAIQGMGKRFYVVWTKLDRDFSTSALLNEKILQNIQEKIKENLQNEGVKEPPIFLVSNFDPSLHDFPKLRNTLRID
ISKIRYHSAQKIILSICEDIINENVTSLLNKIDTGGFQDVLGIQDPNNLGDCLKSYRFFFGVDDKSLKQVAQSIGKSVEEYTNIVKSQD
AQTYHQESSTLSWIGNITVFYFCIGLSYIPYYGNSVAKYLNYMEQRRLLESVAKDTKTILRKVLEDFINLDEVH*

>OO_GIS

MAEFSRSPPTLLSASLTSVMSHQDWGVLSKDEAVKIEKDLEDGNLLKVVSDMRRPLELVSQMPVNIAVTGESGNGISTFINALREIGHE
EEASAPTGVVTTTETRAAYSSPLFPKAVLWDLPGTGVATDTLQEYHVAMQFSQYDLFIIIASQQFSMNHVMLAKTIADMGKKFYIVWTK
LDLDLSSSTLPGCELQQTIKNHIMRSLQKQRVCEPPIFLVSSLKPSSYNFPGLRDTLQTGLSQVRCDGPLQTPHACL*

>OO_GKS1

MGASFSAELSKECQDLESSFKDYIRNFREENKILSQETILSIKSRLSRGDIQGAHSIISGILENIDKIPLNIAVTGESGSGKSSLVNSL
RGVGHEEEDAAPTGVEETTIMRTPYKHPKFPNVTIWDLPGIGTTNFQPKDYLEKVKFGEYDFFIIVSATRFKKNDLDLAKVIKAMKKNF
YFVRTKVDLDLQNEQEFKPTTYVRDKVLEEIRNKSLKEFKDNNIETQIFLISNKNLSEFDFPILMETLLKDLPAQKRHAFTLSLPNITE
AAIDRKRDSLKQIVWLEAFKAGISAIVPAVGIIKDNDVKKLKASLHQYQFHFGVDDTSLQSLAKDLQVPVEELKAIIKSPYLFDTEKEE
TTGEMALKFLEISSSVAFPPLAAGLYFMKVFYLQFHFLDIVTSDAKVLLKKS*

>EE_GKS2

MxSGSQLWQVFKVICTENKILSQESIGLxSVRICRRANIQDTASVIRYALSDIEKAPIDIAVTGETRAGNFSFINALRGVSPEEEGAAE
TGVVKTTVERVPYKHPKFPSVTVWDLPGIGTTRFPPHNYLQEMKFQEYDIFFIISAKHFTPNDTQLCVAIKKMRKNFYFVRTNVDSDLD
KERVRKCRPLSPEEKKHVLQKIKTDCVTNLQKTKVTNSPVFLVSSFEVSGYDFPDLQTTLLRELPAYKRHIFMLSLPAVVEATMDRKSN
SLKQKIWLDALKIRVPDPIPxAVSYFSPEEVDSLRETLTLYRSYFQLDDASLEQVAKDLHVSIGELKANLKXXX

>EE_GKS1

MGQSFSMPSRTMNHDLASSFGDF*KNFKTESNIISQETIALIQSHLLxGNIQKATSVISGALSDIEKAPIDIAVTGETGAGKSSFINTP
WGVGPEEEGAAEIGVVETTVERVPYKHSKLPRVTLWDLPSIGTTCFPPHNYLQEMKF*ECDFFLIVCGTHFKHSDAQLATAIRNMKKNF
CFVPMKVDSELLGSQKAQLSMFNEEETLQKMCFECVRQLEQAQVKESQVILVSSYDLGDYDFPKPEATLLxGLDRESLDSMAKNLHVPE
EMLERNLQSLHLLSIKNHKSLGEQLLRLLEIISSISGGPVADGLCFRKFYLQCYFLYTMVKDAKCFLNKEEIIKHSETLKRACQSCGVG
VRMWKMRQLASDVFSGVACHWPGKRVTFQAGP*

>EE_GKS3

MGQSCSTPPSDTKNYDLASSFGKFFKDICTENKILSQETIALIQSHLSEGNIQDTASVIRDALSDIEKGPIDxSVTGETxGAGKSVIIN
GVRGVRPEKIVPAETGGGEDNVDGKSQAKNP*S*NVVSLRPRLNLRPIFRPPAY*LXX

>EE_Irgq

RRLLPxAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARLDVAVAGTADLGLVLNVLLGLDPGEPGAGPAALPAEPTAF
PAPERPNVVLWTVPLGPAGSAATPHPAHYDALILVTPGAPADKDWARVRPLVLPGAPLVCVRTDGEGEDPEWLQEEEEAEKPEQEGFGG
VEKASGEGREERGPGAQAAGGEKAGCTVAGSPEDETWEVLEEALPPAFPLRPAGLPGLCAWLQRALPANQAKALLLALPPxxxPGSAEA
CDVALLRGQLAEWRRALGLETAALARRERALGLAPGELAARTRFPGPVTRAEVEGRLGAWAGEGTAGGAALGALSFLWPAGGAAATGGL
GYRAAHGVLLQALDEMLADAEAVLAPP

>SA_GKS1

SKVLKSLDECQALMETFEVGSLPAVAAKLQATLQALENVQLDIGITGSPGSGKSTLVNALRGLGDEDMDSAQTSVVETKVAPTPYPHRQ
YLNVVIWDLPGIGPASFQADKYLQQVLQKPYDLLLLLSAGCFPDSLAQLARRLGERGTHFCLVRSKVDVDVAASRSRRPSTFSEDAVLS
QIRNDCAKRLEGEGDRLLGGGGAECRVAQDGTGERGGPQVGIPVMAESNSQGPGASPLLLRFFFLGFFFLSHIRRCFWLYAQELLLVLL
GSIQVGRTPVQCLTCCTICSGLSEAFLLFFLF*GLGDVVLQQSSGRSGGHSW*

>SA_IRGC

MATSKLPAVPGEEGTTILMGKEELEALRSAFEAGDLPQAASRLRELLASSYTLRLEEGMTGYFFLCNSTRLYGLKRLRAEDKALALAA*
xxxxxxxxxxxxxQFPDVTL*DLLGAGTPRCPADKYLKQVDFSRYDFFLLVSPRRCGAVETRLASEILRQGKKFYFVRTKVDEDLAATR
SQRPSVFSEGLVLQEIREHCVERLRAAGVAEPRIFLVSNLSPSRYDFPLLMSTWEHDLPAHRRHAGLLSLPDISLEALQKKKDMLQEQV
LKTALVSGVIQALPVPGLAAAYDDALLVRSLRGYHRSFGLDDDSLARLAEQVGKQAGDLRSVIRSPLAVEVNPETVLRLYARSSDGAMR
VARAFERGVPVFGTLVASGISFGTVYAMLQGCLNEMAEDAQRVRIKALEEDEAPSEVSLEAGGDVGVEKHGLGESGSEDAPLSTRRKLG
LLLKYILDSWKKRDLSEEK*

>SA_IRGQ

ALARGTPLLAVRNLRPGEAQNELQVREQTETLLASAGLEAAALFVLPADCGGRDDCEELERLRAALESRAEALQRLLPPAQDGFEVLGA
AELEAVREAFETGGLEAALSWVRAGLERLGSARLDVAVAGTSGVNLVLDVLLGLDPDDPGAAPASEPFGAHIIPGPERPNVVFWAVPLG
PTSTATASHPAHYDALILVTPEAPAEKDWAPVRPLVLPDAPLACVRTDGKGENPESLEEGKVEKPASESLENSGGGGLEKGGGEGAEAC
VPGSQKAGVASGQQVGMAMKKSGSEDTDCPAVVSPEDEAWEVLEEAPPPVFPLRPGGLPELCEWLQRVLPAAQAGALLLALPPASPSAA
QKKAAALRAGAWRPALLASLAAAAAPVPGLGWACDVALLRGQLAEWRRALGLEPAALARRERALGLAPQDLREFPPFPAP*HAPRWSTA
RPLGQARARQGARHWAASPFSDPRGAQRPQAALGFRAAHGVLLQALEEMLADAEAVLAPPGSPAVRMLGPAGV*

>ET_IRGQ

MPPPRGDVTALFLGPPGTGKSALIAALCDKDVETLEPPEGPSDSRVPSLRAAGPGLFLGELSCPPAVPGPWAAEANVLVLVLPGPEGNG
EVLAPALGEAARAALARGTPLLAVRTLRAGSSQDEGQAREQTAALLDCAGLGAVALFVVPTDRSGSPAGEELERLRAAVQSQAEALQRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARVDLAVAGAAHLGLVLDLLLGLDPGDPRAAPASAPADPTPYPAP
ERPNVVLWTVPLDPACAAVxTPQPTHFDAVILVTPGTPTEKDWAQARALVXXXXXGMMEKPSGGEGIENRAAGASVAGSDDGLPQIGAG
TASGTSISEDETWEVLEEAQPPVFPLRQAGLPGLCAWLRRALPPGQAGALLLALPPSSPSAARTKAAALGAGAWRPALLASLAAAAAPI
PGLGWACDVALLRGQLAEWRRGLGLEPAALVRRERALGLTPGELASRAHFPGPVTRAEVEARLGAWAGEGRAGGAALGGAVLPVAGGRR
CGHWWPGLPGRRMASCCRS*

>ET_IRGC

MATSRSQAGPGEEETTILMAKEELEALRTAFESGDLPQAASRLRELLAASEATRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGV
VETTMQPSPYPHPQFPDVTLWDLPGAGSPGCSADKYLKQVDFGRYDFFLLVSPRRCGAVESRLAAEILRQGKKFYFVRTKVDEDLAATR
SQRPSAFSEAAVLQENPRALHRAPAPPPAWAEPRIFLVVQTCRXXX

>ET_GKS1

MGQLFSSDTDETENLASSMKTYFTNFNMINKIISEETIKLITSKLEAGHTDSMGVIINSVLKKVTNATLNIAVTGESGSGKSSFINALR
GVKPGDEQSADTGVVETTMEPTKYQHPKFPTVIFWDLPGIGTTSFPPQEYLKKVKFGEYDFFIIISSTRFKDTDVQLAKAIKDLGKHFY
FVRSKIDIDLQSEQKQKPKTFDQEKFLEQIRNNCLQDFRQIKMEKPPIFLISSDEVFEYDVPILMNTILNDLSTQKRHIFMLSLPNVTE
EVIESKRDSLKQKVWLEAFKNALWATLPFGITSDYDKKKVENSLNEYRALFGVDDASLQKLATYSKVPVEQLKSVIESPRLLAKEKMNQ
HQKDFYNGWKNSVQXXX

>ET_GKS2

MGQSSSKTPADPKNQNWASSVDEFVKDFKMESQMFSPETIDFIKSNLETGNMQGAVSVISNVMKEIENATIHIAVTGESGSGKSSFINA
LRGVGHEDLDAAACGVVETTMERKPYRHPKIPHVTLWDLPGIGTTTFQPQEYLKKMNFNEYDFFIIVCAQRFKLNDSQLATVIRKMKKN
FYLVRTKVDHDVYNQRRAQPKNFCKDKTLQMIREDCLRNLRAIGVNDAGVFLISSFEVSDYDFPKLETTLLEELPAYKRHIFTQCLPCV
TDAAIERRKDYLRQKIGLEALQCGALATIPVMGCISDSDVQQLEATLTFL*

>ET_GKS3

MGQSSSNTPADPKNKTWASSVEEFFKDFKIESQMLSSEMIKLIKSNIEKGDVQGTANVISAAMKEIENAAVNIAVTGESGSGKSSFINA
LRGVGHEDLDAAACGVVETTMERKSYCHAKIPNVTLWDLPGIGTTKFQPQEYLKKVNFNEYDFFIIICATRFKEYDSQLATVIKKMKKN
FYFVRSKVDSDLKNQRRAQPKNFCKDKTLQMIREDCLRNLRAIGVNDARVFLISSFEVSDYDSQSWRPPFWRSFQLINVTSSRNAFLVL
RMLPLNG*KDYLRQKIWLEALKCGALATIPVMGYISDSDVEQLKATLTFYRSYFGLDDASLEHIAKDLHLSVEELKANLKSPDLLSVAK
DDDLLADKLWKMLEKLCFIGGGLLATGLYFRKTFYLQIHFLDTVVDDAKFLLQKAVIKEQH*

>ET_GKS4

MDQFILSFITKNNFQQLSSEFLSYYFTIISKAGGILSQETLSDIQAALQKGNLEDIVKKIQGALVQAENAPLDVAVTGQSGTGKSSFIN
ALRGLSHEEEGTALVGIVETTMEKTPYQHPKYPKVTFWDLPGTGTPNFCADTYLEKMGFTNYDFFIIISSTRFTYNDALLAQKIKSMGK
NFYFVRTKVDNDLYNENISKPTSFNEKKVLQQIRDXXXKVERHVLNLSAQ*

>ET_GLS

MSLLFLLLQVTPLFCNMAAYPLSPEDLSSVSSGLFETGATSIEKDVMDGNLPNVIAKAKKTWKNMNNNVTLNIAVTGASGNGLSTFINA
LRNVGHDEETSAPTGVVRTTFTRASYSASCFPNVVLWDLPGMGASAQSLENYMAEMHFSQYDFVIIVASEQFSMNHAVLAKTIQKLGKK
FYIVWNKLDMDLSKNVLSKEKAVKMIREYILETLWMEQVNEPPIFLVSSFDLSLHGFRELKQTLKKDLLLIRYQDPLQNLLHTCEAIIN
DKQTFLQKKVDTLSFQDIHGIEDPDNSEACLKVYKVHFGVDDESILRVAQIMESDVLDYTANMKSQDLQTLSNVDWLIYCMNFNTTSYL
STIFSYIPVFGVPITNYIRWTKHRYFLEVVAKDTKTILRKILKDSIN*

>ET_GMS1

XPQ*ILFLTGDSGNGMSTFMSALRGIGHEEEDSAPTGVVKTTYTHASYSSLHFPNTLLWDLPGMETTPQNLEKYIMEMEFSRYDLFIII
ASEQFSMNHVMLAKTIKSMRKKFYIVWTKLDTDLSTSVLGEEQLLKNRRVNILENLQKVQVCEPPLFMVSSLDPLFYDFPKLRDTLQKD
LIHIRCHGPLQKLSDTCEKILNDKVVTLQEKIDAHSFQDIWDPDDLMECLAAYQLLFGMNDESIQQVAQKMGTATMEYTALMKSQNPWT
FSRIDCKMACLTCRVFQAFLNALRCIPLLGNPITNYFRRMKHKHILQIVAKDSKVILRKILKDSTIPG*

>ML_IRGC

MATSKCPGEEETTILMAKEELEALRTAFESGDIPQAASRLRELLASSESTRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGVVET
TMQPSSYPHPQFPDVTLWDLPGAGCPGCPADKYLKQVDFGRYDFFLLVSPRRCGAVEARLASEILRQGKKFYFVRTKVDEDLAATRSQR
PSGFSESMVLQEIRDHCTERLRAAGVTDPRIFLVSNLWPVRYDFPLLLSTWEHDLPAHRRHAGLLSLPDISLEALQRKGDMLQEQVLKT
ALVxxxGDLRSVIRSPLANAxDSPETVLRLYSQSFDGAMRVARAFERGIPVFGTLVAGGISFGTVYTMLQGCLNEMAEDAQRVRIKALE
EAEESQPEVSLDAAGDNGVEKCGVGEGGGEEAPLSTRRKLGLLLKYILDSWKKRDLEEK*

>ML_GKS1

MANQIFQFNISELSKDTGALKETFEAGNLPAVAAKLQATLHSLEDVRLDIGVTGAIGSGKSTFVNAIRGLGDEDPKSAGTGVVEMTVEP
TPYPHPKYFNIIIWDLPGIDAPTFQADEYLQRVLVDRYDFFIIITPDSFTARHAQLARGLRQQGKAFYFVRSKVDVDLAASSSRRPSTF
SEQRVLRQILGDCWQRWKVECPLWGARG*

>ML_IRGQ

RLLPPAQDGFEVLGAAELEAVREAFEKGGLEAALSWVRAGLERLGSARLDLAIAGTADTGFVLDMLLGLDTGDPGAVPASAPTGPTPYP
APERPNVVLWTIPLGSATTAAAPHPTHYDALILVTPGAPTEEDWAQVRPLVLPDAPLVCVRIDGEGEDPESPEEEEKAEKPSRESLEDA
EGGGLEEARSEGRDKRGAGSQKAGSESSQQAGGNAKPSGSGDSERAAAWSPEDETWEVLEEALPPVFPLRRGGLPGLCEWLQRALPPAQ
AGALLLDGATHVPACSPNQG

>CF_IRGQ

MPPPRGDVTALFLGPPGCGKSALIAALCDGNVETIEIPEGRPDSGIPSLRALGPGLFLGELSCPPAAPGPWAAEANVLVLVLPGPEGNG
EPLAPALGEAARAALARGTPLLAVRNLRPGESQDEAQARDQTAALLNGAGLGAAALFVLPADCGRRDGCKELERLRVALRSQAEVLQRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARLDLAVAGNADVSLVLNVLLGLDPGDPDAEPVFMPAGPTPHPAP
ERPNVVLWHVPLGSAGTAAATPHPTHYDALILVTPGAPTEKDWAQVRPFVLPDAPLVGVRTDGEGEDPEYLEEEDEAEKEKPSCESLEN
TGGEGVKNARSEGREKRGPGLQKGSEEDSEKAGSGEGSEKAGSESVPPVGGGGKKSGSGDAERAAALSPEDETWEVLEEAPPPVFPLRP
GGLPGLCEWLRRALPPAQAGALLLALPPASPHGARMKAAALRAGAWRPALLASLAAAAAPVPGLGWACDVALLRGQLAEWRRALGLEPA
ALARRERALGLTPGELAERTHFPGPVTRAEVEARLGAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHGVLLQALSEMQADAEA
VLAPHVPAQ*

>CF_IRGC

MATSKLRAVPGEEETTILMAKEELEALRSAFESGDIPQAASRLRELLASSQSIRLEVGVTGESGAGKSSLINALRGVGAEDPGAALTGV
VETTMQPSPYPHPQFPDVTLWDLPGAGSPGCPADKYLKQVDFGRYDFFLLVSPRRCGAVETRLASEILRQGKKFYFVRTKVDEDLAATR
TQRPSGFSEAAVLQEIRDHCAERLRVAGMTDPRIFLVSNLSPARYDFPLLMSTWEHDLPAHRRHAGLLSLPDISLEALQKKKDMLQEQV
LKTALVSGVIQALPVPGLAAAYDDALLIRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMR
VARAFEKGIPVFGTLVAGGISFGTVYTMLQGCLNEMAEDAQRVRIKALEEDEPQSEVSLEAAGDNGVEKRGSGEGGCEEAPLSARRKLG
LLLKYILDSWKKRDLSEEK*

>CF_C-like

MASRVFQSNFTWSKILELWRDTSALKGAFEVGDLPTVATKLQATLHLPENARLDTGITGGTGSGKSTFVNAIRGLGDEDPRSAYTGVVE
MSVDPTPYAHPKYPNVVIWDLLGIDTPTFQAKKYLQQVLLDRYDFLLLITLESFTAHHTQLACEILQQGKRFYFIRSKVDVDMRPHAAG
APAPFQRKEHSARPGRTVGNGWKVSLGSGRWKRVGALGVVDSGGMARPFPFLFQLRPTVGKGSSKVTLRD*

>CF_Qgene

YNSGITCGSREYSSVYGCIQESGTRMYSFINAL*GHIHEEKNSATVGIFKTTMKKPPYQHPKDPSMTFWDLPGTGTPNFLLHTYLKMVG
FANYDFFIIVSSSWYNFNDVLLNQNFRELGKKFHFDRTKVDNDLYNEGKNKARSLKRERMLQQI*DSCLTTLSNIRVHEPGNFLVSKFD
LDDFDFPILEETLLKDLPVHKCYIFVILLPNLGEASIEMERTLFKERIWLDSLKSLAFTSIPFMACFSDFD*SQQEKCLNVYENYFGLD
EKSVKGIVRKLVMSVEEIKSFIKSLHFWLLLKNNCIAAKAMKCTEGHCFMNGSLSSNLFQLLKIYFLHLKFMNKVIDDARIFLHKISES
ISLRRCQTHSSETN*

>CF_IRGB11

MGQSPPSTPSNRNGGDLASSFDKFFKEFKLDSKIISQETISTIQSHLEKGDLQSAFSAINDALRDIDNAPLNIAVTGESGTGKSSFINA
LRGMGHDEEGAAPTGPVETTFLRKAYKHPKFPNVTFWDLPGIGTTSFQPQDYLEKMVFREYDFFIIICATRFKINDVQLATAIKKMKKN
FYFVRSKVDSDLYNLKRIKPREFNKDEILQKIRNDCVKHLMEANMSDAQVFLVSSFELSDYDFQSLETTLLRELPSHKRHIFMQYLPIV
TEATIDRKRDCLRQKVWLEAIKAGASASIPLVGYISDNDVETLKDTLTLYRSYFGLDDISLKTIAKDLNVSVEKLKANLMFPHLLSVEK
YDEPLGEKLLKYVEKFCSVSGGPIAAGIYFRKIYYLKNYFLDTVVSDAKVLLKKEEIFKDPVDSEQTYLHTNVGNENGKSDTSSS*

>CF_IRGB12

MGQSSSTPSHKTGGDLASSFGKFFKDFKLESKILSQEAITSIEKSLKEGNLQKAVSDINKALKDIDNAPLSIAVTGESGTGKSSFINAL
RGVGHDEEGAAPIGAVETTFDRTEYKHRKFPNVTLWDLPGVGTTTFHPQEYLEKMKFREYDFFIIISSTRFTINDAQLATAIRKMKKNF
YFVRSKVDSDLYNLKRTKPSDFNKDEILLKIRNDCITQLQNVKVCDPQVFLVSNLDLSSYDFQSLETTLLKELPAHKRHIFMQYLPNIT
ESAIDRKRDSLRQKVWLEAVKAGASATIPFMGLINDNEVEKLEETLHLYRSYFGLDDASLETIAKDLNVSVEKLKANLTSPHLLSVEKE
DESLGEKLLRYVEKFCSVSGGLIATGVYFRKIFYLQNYFLEAVVSDAKVLLNKEEIFKETVGSGQAYLLQDVGIENRKSDATSS*

>CF_GKS1

MGQSSSTSSHIKGDDLASSFGKFFKDFKIESQILSQETITLIKTHLKEGDLQKAASAISDALRDIDNAPLNIAVTGEPGTGKSSFINAL
RGMEHDEEGAAPTGLVETTLERISYKHPKFPMxTSWDLPGMMTTTFQPQMYLKKMKFCEYDVFIIISFTRFKISDMHLAAAIKKMKKNF
YFVRTKVDNDLYNIKMSKPSTFNKDEVLQGIRNDCVTQLQNANMSDTQVFLISSLDLFSYDFQSLETTLLKKLPAYKCHIFMQYLPNVT
EAAIDQKRDSVRQKVWLEAIKAGATATIPFMGLIGDKDVEKLGETLTLYRSYFGLDDASLETMAKDLNVSMEKLKANLKSPHLLTVEKE
DESLGKKLLRYVEKFCSVSGGLIATGLYFRKIFYLQNYFLETVASDAKVILNKEEIFKESLGSGKAYLLQDVGIENGKSDTTNS*

>CF_IRGD

MDKFMCDFLVGKNFQQLAINFIPHYTTLVNKAGGIIASENLDRIQAALKEAKLKDVADIIEESLVAAENAPLDVAVIGESGTGKSSFIN
ALRGLSYEEEGSASVGVVETTMKKTPYQHPKYPKVTFWDLPGTGTPNFHPHEYLEMVEFATYDFFIIISSSRFSLNDALLAQNIKEIGK
KFYFVRTKVDNDLYNEEKSKPMSFKRERVLQQIRDNCLANLSNIGVPEPCIFLVSNFDLDDFDFPRLEETLLKELPVHKRHIFALLLPN
LSYTSIEMKRAFFKEKIWLDALKSSALSFIPFMACFNGFDFPQQEKCLNLYQSHFGLDEKSVKGIAEKLDMSVEEIKSFTKSLDFWLLV
KDDSIAEKAMKCVECYCSVNGGLPSTIFQFFKIYFLHLKFINTVADDAKILLHKTLEILSHRR*

>CF_IRGM4

MAQPTQSLHTPSPTSFTSTVPYHKGGSILSESGAMNIEKALGEGKLLDMVSVVRETLETASSVPVSIAVTGDSGNGMSTFINALRKIGH
NEEDSAPTGVVRTTQIPTCYSFSDIPNVELWDLPGTGAATQNLETYLEEMQFSKYDLFIIIASEQFSMNLVKLVKSIQGQGKRFYIVWT
KLDRDLSTCVLSEEQLLRNIRENIRETLHKEGVCEPIIFLVSSFNPFLHDFPELRKSLHRDISNIGYRGHLENLTHTCEKVINGKVTTL
QGQIGSKSFQDILGIQNANDLGEFLNAYHRLFGVDDDSLQEVAQSMGKPKEEYKAIMKSQDLHTALAWDWALSWMNCNAASYLYSVLSY
IPILGTTGIHYLKWWSQGHLLEIVAEDTKTILKKILEDAII*

>CF_IRGM5

MTQPNHSLHIPLSTSFTSIVPYNMGWTVLPKATATNIEKALGDGKLLEVVSMIRETLETVSSAPVSIAVTGDSGNGMSSFINALREIGH
DEKDSAPTGVVRTTQVPTCYSSSHFPYMELWDLPGTGTGTQSLENYLEKIHFSQYDLFIIIASEQFSMNLVKLVKAIQRQGKRFYIVWT
KLDRDLSTRVLPEEQVLQNIWENIQETLQKVGVCEPIIFLVSSFEPLLHDFPELRDALNRDISDIRYCGPLENLSDTCEKIINDKVTSF
QEQIGSKTFQDILGIQDEDDLGQCLIAYHLFFGVDDKSLQQMAQSMGKPMEEYRAIMKSQDVHTVLTGDWALSCMNCKTASYLYSILSY
IPFLGDTVINYLRVWKHRHFLEIVAKDTRSIVKKILTDSII*

>CF_IRGM6

LHCFFPLLQVTPLLSDVTQPTHSLHTPLLTSSNYDMPYNMGWSSLSKETAINIEKALGGRKLLEVVPMVRETLERASSVPLRIAVTGDS
GNGMSSFINALRGIGHDEEDSAPTGVVKTTQIPTCYSYPHFPNVELWDLPGTGAGTQSLENYLEEMKFSWYDLFIIIASEQFSMNLVKL
AKAIQVLGKRFYIVWTKLDRDLSTSALLKERLLQNIQENIQENLQKERVFEPIIFLVSSFEPLLHDFPELRNTLNRDISDIRYCGPLKN
LSHTYEKVISDKVTMFRGKIASKSFDTLGIWNADDLGECLIAYHLFFGVDDESLQQIAQSMGKPMEEYRAIMKSRDLHTIIRGDWAVSC
MNCNTSSCLYTILRYIPLLGDFIINFLRKWKHRRLLEIVAEDTRTILKKILKDSII*

>CF_GKS2pseudo

MGQSSSTTSCHMVDGGLDSRVGKFFRELKLGSQIVPQEAINLIQSHLxRRGTFRKWFMISDALRKIENTSLNFAVIGESGAGKSSFINA
xKGMGQEVEVAATKLVETTLKRVKYxTQILSQCDIVKPIWHRDHYLFTTEYLKKMKYWEYNFFIIISSTCFTINDVQLATAIRKMKKNI
YFV*TKVDSDLHNKNxSQRKNFNKGKILQMICNGYVKHLMEANMSDAQIFVFSFELLxNYDFQSLEITLLxGKSQPTNLTSSCNITKSA

IDQKRDSLRQKIWLEAMKVGATATIP*GCCLISDNKVEKLVDTLHLYRSYFGLDNAFLxNMAKDFHVPVETLKANLMSPHLLSIEKEDE
SLEEKPLRYVEKICSVSGGLIATGLYFRKIFYLQNYFLDTVVNDVKVLLLKKDFLRILRALGKPICI*

>FC_IRGC

MATSRLPAVPGEEETTILMAKEELEALRTAFESGDIPQAASRLRELLASSDSTRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGV
VETTVQPSPYPHPQLPDVTLWDLPGAGAPGCPADRYLKQVDFCRYDFFLLVSARRCGAVETRLASEILGQGKKFYFVRTKVDEDLAATR
AQRPSGFSEATVLQEIRDHCAERLRAAGVADPRIFLVSNLSPARYDFPLLVSTWEHDLPAHRRHAGLLSLPDISLEALQKKKDMLQEQV
LKTALVSGVIQALPVPGLAAAYDDALLLRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMR
VARAFERGIPVFGTLVAGGISFGAVYTMLQGCLNEMAEDAQRVRIKALEEDEPRPSVSLEAAGDNGVEKRVSGEGTCEEAPLSTRRKLG
LLLKYILDSWKKRDLLEDK*

>FC_C_like

MAGRVLQSSLSQSKILELWRDTSALMGAFEAGGLAAVAARLQATLHSLENARLDIGITGGTGSGKSTFVNAIRGLGDEDPNSACTGVVE
MTVDPTPYPHPKYPNIVIWDLPGIGTSRFRTGRYLQRVLLERYDFFIIITSDSFTAHHAQLACEILQRGKRFYLVRSKVDVDIAASRSR
RPSTFSEERVLRQIREDCGRRLRGEVQALERRNRAVLLSSAAEGLKDPKVFLLSMFELGKYDFHLLEELMVKELESHKQHAFLLAVPNV
SKPILEKKAASLRQHIWLVATVACGVNPSPVPGVREVACDLYVLISSLEGYRRSLGLDEDSLVRLAEQRGQPLHKILEVVQGWKTKVTE
ALVVELLGQASRDASAFTQELLGVPILGALATCGISFATIYQMLRTALDEVVKDAQRVLTQAFLDDSDHELPDKCNQ*

>FC_IRGQ

MPPPRGDVTALFLGPPGSGKSALIAALCDRNVETIEIPEGRPYSGIPSLRAAGPGLFLGELSCPPAAPGPWAVEAIVLVLVLPDP*GNG
EPLASVLFESAWSALFRGTLLFFVRXXXXXRLLPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARLDLAVAGTANV
SLVLNTLLGLDPGDPDAEPVFMPAGPTPYPAPERPNVVLWNVPLGSAGIAAAPHPTQYDALILVTPGAPTEKDWAHVRPLVLPDAPLVC
VRTDGEGEDPEYPEEEGKAEKPSSESLENAGGGELKNARGEGREKRGAGLQKGSGEGSEKAGSGEGSEKAGSESLPRVGGGAKKSGSGD
SERAAALSPEDETWEVLEEAPPPVFPLRPGGLPGLCEWLRRALPPAQAGALLLALPPASPHGARMKAAALRAGAWRPALLASLAAAAAP
IPGLGWACDLALLRGQLAEWREGAGARTRGAGSTRARAGPGPWXX

>HS_IRGC

MATSKLPVVPGEEENTILMAKERLEALRTAFESGDLPQAASHLQELLASTESIRLEVGVTGESGAGKSSLINALRGLEAEDPGAALTGV
METTMQPSPYPHPQFPDVTLWDLPGAGSPGCPADKYLKQVDFSRYDFFLLVSPRRCGAVETRLAAEILCQGKKFYFVRTKVDEDLAATR
TQRPSGFREAAVLQEIRDHCAERLREAGVADPRIFLVSNLSPARYDFPTLVSTWEHDLPSHRRHAGLLSLPDISLEALQKKKAMLQEQV
LKTALVLGVIQALPVPGLAAAYDDALLIHSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMR
VARAFERGIPVFGTLVAGGISFGAVYTMLQGCLNEMAEDAQRVRIKALEDDEPQPEVSLEVASDNGVEKGGSGEGGGEEAPLSTCRKLG
LLLKYILDSWKKHDSEEK*

>HS_IRGQ

MPPPQGDVTALFLGPPGLGKSALIAALCDKDVETLEAPEGRPDSGVPSLRAAGPGLFLGELSCPPAAPGPWAAEANVLVLVLPGPEGNG
EPLAPALGEAALAALARGTPLLAVRNLRPGDSQTAAQARDQTAALLNSAGLGAADLFVLPANCGSSDGCEELERLRAALQSQAEALRRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRSGLERLGSARLDLAVAGKADVGLVVDMLLGLDPGDPGAAPASVPTAPTPFPAP
ERPNVVLWTVPLGHTGTATTAAAASHPTHYDALILVTPGAPTEKDWAQVQALLLPDAPLVCVRTDGEGEDPECLGEGKMENPKGESLKN
AGGGGLENALSKGREKCSAGSQKAGSGEGPGKAGSEGLQQVVGMKKSGGGDSERAAALSPEDETWEVLEEAPPPVFPLRPGGLPGLCEW
LRRALPPAQAGALLLALPPASPSAARTKAAALRAGAWRPALLASLAAAAAPLPGLGWACDVALLRGQLAEWRRGLGLEPTALARRERAL
GLASGELAARAHFPGPVTRAEVEARLGAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHGVLLQALDEMRADAEAVLAPPEPAQ
*

>HS_IRGM

MEAMNVEKASADGNLPEVISNIKETLKIVSRTPVNITMAGDSGNGMSTFISALRNTGHEGKASPPTELVKATQRCASYFSSHFSNVVLW
DLPGTGSATTTLENYLMEMQFNRYDFIMVASAQFSMNHVMLAKTAEDMGKKFYIVWTKLDMDLSTGALPEVQLLQIRENVLENLQKERV
CEY*

>PT_IRGC

MATSKLPVVPGEEENTILMAKERLEALRTAFESGDLPQAASHLQELLASTESIRLEVGVTGESGAGKSSLINALRGLEAEDPGAALTGV
METTMQPSPYPHPQFPDVTLWDLPGAGSPGCPADKYLKQVDFSRYDFFLLVSPRRCGAVETRLAAEILCQGKKFYFVRTKVDEDLAATR
TQRPSGFREAAVLQEIRDHCAERLREAGVADPRIFLVSNLSPARYDFPTLVSTWEHDLPSHRRHAGLLSLPDISLEALQKKKAMLQEQV
LKTALVLGVIQALPVPGLAAAYDDALLIHSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMR
VARAFERGIPVFGTLVAGGISFGAVYTMLQGCLNEMAEDAQRVRIKALEDDEPQPEVSLEVASDNGVEKGGSGEGGGEEAPLSTCRKLG
LLLKYILDSWKKHDSEEK*

>PT_IRGM

MEAMNVEKASADGNLPEVISNIKETLKIVSRTPVNITTAGHSGNGMSTFISALRNTGHEGKASPPTGLVKATQRCASYFSSHFSNVVLW
DLPGTGSATKTLENYLMEMQFNRYDFIMVASAQFSMNHVMLAKTAEDMGKKFYIVWTKLDLDLSTGALPEVQLLQIRENVLENLQKERV
CEF*

>PT_IRGQ

MPPPQGDVTALFLGPPGLGKSALIAALCDKDVETLEAPEGRPDSGVPSLRAAGPGLFLGELSCPPAAPGPWAAEANVLVLVLPGPEGNG
EPLAPALGEAALAALSRGTPLLAVRNLRPGDSQTAAQARDQTAALLNSAGLGAADLFVLPANCGSSDGCEELERLRAALQSQAEALRRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRSGLERLGSARLDLAVAGKADVGLVVDMLLGLDPGDPGAAPASVPTAPTPFPAP
ERPNVVLWTVPLGHTGTATTAAAASHPTHYDALILVTPGAPTEKDWAQVQALLLPDAPLVCVRTDGEGEDPECLGEGKMEKPKGESLKN
AGGGGLENALSKGREKCSAGSQKAGSGEGPGKAGSEGLQQVVGMKKSGGGDSERAAALSPEDETWEVLEEAPPPVFPLRPGGLPGLCEW
LRRALPPAQAGALLLALPPASPSAARTKAAALRAGAWRPALLASLAAAAAPLPGLGWACDVALLRGQLAEWRRGLGLEPTALARRERAL
GLASGELAARAHFPGPVTRAEVEARLGTWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHGVLLQALDEMRADAEAVLAPPEPAQ
*

>MN_IRGM

TEATNVEKALVGGNLPEVASAINDTLKIVSRTPVNIAMAGDSGSGMNTFISAL*NTGHEGKPSPPTGLVKATQRCAPYLSSHFPNMVLW
DLPGTGSATKTLENYLMEMQFNQYDFIMVASAQFSMNHVMLAKTTEDMGKKFYIVWTKLDMDLSTGALPEVQLLRIRENVLENLQKEQV
CEH*

>MN_IRGC

MATSKLPVVPGEEETTILMAKERLEALRTAFESGDLPQAASHLQELLASTESTRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGV
VETTMQPSSYPHPQFPDVTLWDLPGAGSPGCPADKYLKQVDFSRYDFFLLVSPRRCGAVETRLAAEILCQGKKFYFVRTKVDEDLAATR
SQRPSGFREAAVLQEIRDHCAERLREAGVAEPRIFLVSNLSPARYDFPTLVSTWEHDLPAHRRHAGLLSLPDISLEALQKKKAMLQEQV
LKTALVLGVIQALPVPGLAAAYDDALLIHSLRGYHRSFGLDEDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMR
VARAFERGIPVFGTLVAGGISFGAVYTMLQGCLNEMAEDAQRVRIKALEEDESQPEVGLEAAGDNGVEKGGSGEGGSEEAPLSTRRKLG
LLLKYILDSWKKHDSEEK*

>MN_IRGQ

MPPPQGDVTALFLGPPGLGKSALIAALCDKDVETLEAPEGRPDSGVPSLRAAGPGLFLGELSCPPATPGPWAAEANVLVLVLPGPEGNG
EPLAAALGEAARAALARGTPLLAVRNLRPGDSQTAAQARDQTAALLNSAGLGAADLFVLPANCGSSDGCEELERLRAALQSQAEALRRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRSGLERLGSARLDLAVAGKADVGLVVDMLLGLDPGDPGAAPASAPTGPTPFPAP
ERPNVVLWTVPLGPTGTATTAAAASHPTHYDALILVTPGAPTEKDWAQVQALVLPDAPLVCVRTDGEGEDPEYLGEGKMEKPNGESLKN
AGGGGLENALSKGREKCSAGSQKAGSGEGPEKAGSEGLQQVVGMKKSGGGDSERAAVLSPEDETWEVLEEAPPPVFPLRPGGLPGLCEW
LRQALPPAQAGALLLALPPASPSAARTKAAALRAGAWRPALLASLAAAAAPLPGLGWACDVALLRGQLAEWRRGLGLEPAALARRERAL
GLAPGELAARAHFPGPVTRAEVEARLGAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHGVLLQALDEMRADAEAVLAPPESAQ
*

>PP_IRGM_consensus

MEAMSVEKASADGNLPEVISDIKETLKIVSRTPVNITMAGDSGNGMSTFISALRNTGHEGKASPRTGLVKTTQRCASYFSSHFSNVVLW
DLPGTGSATKTLENYLMEMQFNQCDFIMVASAQFSMNQVMLAKTAEDMGKKFYIVWTKLDMDLSTGALPEVQLLQIRENVLENLQKERV
CEH*

>CJ_IRGC_consensus

MAASKLPAVPGKEETTVLMAKEELEALRTAFESGDIPQAASRLRELLASTHSTRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGV
IETTMQTSPYPHPQFPDVTLWDLPGAGCPGCPADKYLKQVDFSRYDFFLLVSSRRCGAVETRLAAEILYQGKKFYFVRTKVDEDLAATR
TQRPSGFSEAAVLQEIRDHCAERLRAAGVAEPRIFLVSNLSPARYDFPMLVSTWEHDLPAHRRHAGLLSLPDISLEALRKKKAMLQEQV
LKTALVSGVMQALPVPGLAAAYDDALLVRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANELSPETILRLYSQSSDGAMR
VARAFERGIPVFGTLVAGGISFGAVYTMLQGCLNEMAEDAQRVRIKALEEDKPQPEVSLETAGDNSMEKGGSGEGDGEESPLSARRKLG
LLLKYILDSWKKYSRKR

>CJ_GMS1

MNVERASADGDLPEVVSAIKESLKIVFRTPVNIAMAGDSGNSISTFISALQIAGHEAKASPPTGLVKATQRCASYFSSRFPNVVLWDLP
GAGSATKTLENYLMEM*FNQYDFIMVASAQFSMNHVILAKTIEDMGKKFYIVWTKLDMDLSTGALPEVQLL*IRENVLESLQREQVCEL
PIFMASSLEPLLHDFPKLRDTLQKTHPN*

>CJ_IRGQ_consensus

MPPPQGDVTALFLGPPGLGKSALITALCDKDVETLEAPEGRPDSGVPSLRAAGPGLFLGELSCPPAAPGPWAAEANLLVLVLPGPEGNG
EPLAPALGEAARAALARGTPLLAVRNLRPGDSQNATEARDQTAALLNSAGLGAADLFVLPANCGSCDGCEELERLRVALQSQAEALRRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRSGLERLGSARLDLAVAGKADVGLVLDMLLGLDPGDPGAAPASVPMGPTPFPAP
ERPNVVLWTVPLDPAGTATTPAATSHSSHYDALILVTPGAPTEKDWAQVQALVLPDAPLVCVRTDGEGEDPECVGEGKMEKPSGESLKN
AGGGGLENALSKGREKYSTGSQKAGSREGSGKAGSEGLQQVVGMKKSGGGDSERAAALSPEDEMWEVLEEAPPPVFPLRPGGLPGLCEW
MRRVLPPAQAGALLLALPPASPSAARTKAAALRAGAWRPALLASLAAAAAPLPGLGWACDIALLRGQLAEWRRGLGLEPAALARRERAL
GLAPGELAARARFPGPVTRAEVEARLGAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHGVLLHALDEMLADAEAVLAPPESAQ
*

>OG_IRGC

MATSKLRSVPREEETTILMAKEELEALRTAFESGDIPQAASRLRELLASSDSTRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGV
VETTIQPSPYPHPQFSDVTLWDLPGAGSPGCPADKYLKQVDFSRYDFFLLVSPRRCGAVESRLATEILRQGKKFYFVRTKVDEDLAATR
TQSPSGYSELAVLQEIRDHCALxRLRAAGVSDPRIFLVSNLSPTRYDFPVLVSTWEHDLPAHRRHACLLSLPDISLEALQKKKDMLQEQ
VLKTALVSGVIQALPVPGLAAAYDDALLVRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSLLDNEVSPETVLRLYSQSSYGAM
RVARAFEKGIPVFGTLVAGGISFGTVYTMLQGCLNELAEDAQRVRIKALRRRSPSQRTA*

>OG_GKS1

MGQHSSAASHDDQDLASSCNAYFKNFKMENKIISQETIDLMELHLKKGDIQGANSVINDALKEIDNVPLSIAVIGECGVGKSSLINALR
GVGNGDQDLAPTGVVGTTRERSPYKHPKFPNVTFWDLPAIGTSNYQQKDYLEKVKFGEYDFFIIVSAVCFKKNDIDLAQMIQIMKKNFY
FVRTKVDLDLEAEQVFKKTAFDREKVLQQIRNDYLNIFKQNKINEPPIFLISNRDLSEYDFPILIDTLIKDLPIQKHHIFMLSLPNVTE
AAIESKRDSLKQMIWLDALKAGALAPLRMVGITSNRDVEKLKTSLNQYRVLFGVDDASLERLAKDLQVPVKQLKATLKSPSLLQNKKEE
SIGEMLLKYLERVCSADGFLASGISFGKTYYLQLHFLNTVTEDAKVLLKETFKKRVESGTCGSKE*

>OG_IRGQ

MPLPQGDVTALFLGPPGSGKSALISALCDKDIETLETHEGRPDSGVPSLRAAGPGLFLGELSCPPAAPGPWAAEANVLVLVLPGSDGNG
EPLAPALGEAARAALARGTPLLAVRILRPGDSQNDTQARDQTAALLDSARLGSAALYVVPADCSSSESCEELERLRTALQSQAETLQRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGIERLGSARLDLAVAGAADMGLVPETCCLDLDPGDPRCCACFGAHWGPLPFP
APERPNVVLWAVPLAPVGTTTVASHPAHYDALILVTPGVPTEKDWAQVRALVPPDAPLACVRTDGEGEDPESLEEESTVEKPNSKSFEN
AGRGDAENALTEGRERCGTGSQKAGSGEGSEKSGSENVQHIVNWRQEIRQWGFRAGxxxxxPVPGLGWACDVALLRGQLAEWRRAMGLE
PGALARRERALGLAPGELAARTHFPGPVTRSEVEARLSAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHGVLLQALEEMLADA
EAVLEPPEPAQ*

>MU_GKS1

MDQFSSATSHDGEHQDLASSFSAYFNNFRVGSKIISKETTVSIESHLRNGNIVAAHSVIEDALKEIDNATLNIAVTGESGAGKSSFINA
LRELGPEDEGAAPTGVVETTMERIPYKHPKFPRVTLWDLPGIGSTNFQPKDYLQKVQLVEYDFFIIVSATRFKNNDIQLAKVIKSMKKN
FYFVRSKVDCDLRNEQESKPKSFNKEKVLQQMRNNCLRIFNENKIDEAPVFLISNRELSKYDFPILIDTLSKDLPAQKHHIFMLSLPNI
TDAAIESKRNSLEQHIWLEAFRDALMAGVPIISIFSDSEEEKLKEILNNYQVLFGVDDESLQHLAKHLQVPVEQLKAGLKSPGLLKKKK
ELTGNTLKGVMKKLFSVFGGVIRGVYYFGKIYYLRFNFLDTVANDAKVLLKETYSRKV*

>MU_GKS2

MGQLFSATSHDEQHQDLASSFTAYFKNFSVGSNIISQEAAESIKSHLTKGNIQAANSVISAALKEIDNAPLSIAVTGESGTGKSSFINA
LRGLGPEDEGAAPTGAVETTMERKPYTHPKFPCVTLWDLPGIGTMNFQPKDYLRKVNFVEYDFFIIVSATRFKDNDIELAKMIKTMKKN
FYFVRSKVDIDLRNEQESKPKSFTEEGVLQQIRNNCRYIFKENKIDEAPVFLISNRDLSAYDFPILIDTLLKDLPAEKHHIFSAFPAQ

>MU_IRGC

MATSKLPSVPREEETTILMAKEELEALRSAFESGDIPQAASRLRELLASSESAGLEVGVTxGESGAGKSSLINALRGLGAEDPGAALTG
VVETTMQPSPYPHPQFPDVTLWDLPGAGTPGCPADKYLKQVDFSRYDFFLLVSPRRCGAIESRLASEILRQGKKFYFVRTKVDEDLAAT
RTQRPSGYSEVIVLQEIRDHCVDRLRAAGVADPRVFLVSNLSPSRYDLPLLMSTWEHDLPAHRRHAGLLSLPDISLEALQKKKDMLQEQ
VLKTALVSGVIQALPVPGLAAAYDDALLIRSLRGYHRSFGLDEDSLAKLAEQVGKQAGDLRAVMRSPLANEVSPEAVLRLYSQSSYGAM
RVARAFEKGIPVFGTLVAGGISFGTVYTMLQGCLNEMAEDAQRVRIKALEEDESQAEVGLEAAGDGGVEKRGSGEGGGEEAPVSVRRKL
GLLLKYILDSWKKRDLEEK*

MU_GKS4

GLKAGLKAGLEAGLEAGLEAVLEAGLKATVKAVVKAAVNAAVPILDIFCDGEVEKLKEILSNYRVLFGVDDESLQRLAKDLQVPVEQLK
AGLISPGLLEKKAKSIEYILLTNLENFIPSNIFLLRGAFTFSKIYYLQLQFLDTVADDAKVLLKETYSTKKKLDFPIN*

>MU_GKS5

PAEKPHIFLLSLISITDAAIEGKRDSM*QILCLEALKAAAWATLPMVGIISDRDVEKLKGILNDYRVFFGVDDKSLKHLANDLQVPVVQ
LKARLKSPGLLEMKEESTGEMLLRYAENFSSATGGPLAAGLYFRKNYYLQFHFLDTVVNDAKVLLKEIYSRKV*

>MU_GKS6_corrected

MHPFIPAILSGKSFQSLAIDFVSPYSTLTSKTRGILALQTLTGIEETLKNGKVRDMVDKIEESLAEAENDSLHVAVTGQSGTGKSSVMN
ALRGLSHEGKGSACVGAVETTMKKVPYQHPKYPNVTFCDLPKTGTPSFLPDTSLEMVGLLTTYNFFILISSSWFSLNDALLAQKIKEMG
KKFYFIRSKVDNDLYNEKKPRSFRKERVLQQI*GNCLANLSDTGEPELCVFSVSDSDLDDFDFPRLEETLLKELPAHKCQTFVLQVPNW
SDTSTEMKREMLKEKIWLF*

>MU_GMS1

MLSQMTLMHPPNSPALLPLQVPPLLTDVTAFPLSPHTPLSASLTAVLPQCKDWSMLSKVEALSIEKAIAGGNLPELVSAVRETVKMVSR
TPVNVAVTGDSGNGMSSFINALRNIGHEEEASAPVGVLKTTQTHACYLSPHFPNVVLWDLPGTVSAAQSLENYATEMQFSRYDFFIIIA
SEQFSMNHVMLAKTVEDMGKQFYIVWTKLDMDLNTSALPKGQLRQIIRENILENLQKQRVCEPPIFLVSSFDPLSYDFPKLRDSLQMDL
MKNRRHELLQNLSHTCERAVNDKVSFLQKKIATESLQDACGISDADDLAACLKAYQLLFGVDDESLWQVAQRVGRTFADYTNITKSCDV
QGLSTRNWKLTCMTCTVFRAFLGLLRCIPWLGNLTIHYFRRKKQRCLLEIVAEDTKAILRKVLEDSIIPQ*

>MU_GMS2

MLSQMILMHPVPPHSPALLPLQVPPLLADVTAFPLPPHTPLSASLTAVLPQCKDWSMLSKAEATYINKAIADRNLPELVSAVRETVKMA
FRTPVNVAVTGNSGNGMSSFINALRNIGHEEEASAPVGVLKTTQTRACYLSPHFPNVELWDLPGTECAAQSLENYATEMQFSRYDFFII
IASEQFSMNHVMLAQTMENMRKKFYVVWTKLDIDLNTSALPEGQLRQIIRENILENLQKQQVCEPPIFLVSSFDPLSYDFPKLRDSLQM
DLMRIRCDELLQNLSHTCERAVNDKVSFLQKKISTESLQDACGISDADDLAACLKAYQLLFGVDDESLWQVAQRVGRTFADYTNIMKSH
DVEGLSTRNWKLTCMTCTVFRAFLGLLRCIPWLGNLIIGYFRGKKQRCLLEIVAEDTKAIIRKVLEDSIIPQ*

>MU_GMS3

MNIEKALADGNLPEVVSAIRESLKMASRTPVNVALTGNSGNDMSSFINPLRNIGHEEEASVPVAVLKTTQTRACYLSPHFPNVVLWDLP
GTECAAQSLENYSMEMQFSRYDFFIIIVSDSSA*NHVILAKTM*DMRKQFYVVWTKLDMDLNTSALPEGQLLQIIRENILENLQKQRVC
GPPIFLVSSFDPLSYDFPKLRDSLQMDLMKNRYHELLQNLSHMCERAVNDKVSFLQKKIATESLQDACGISDADDLAVCLKAYQLLFGV
DDESLWQVAQRVGRTFADYMSIMKSCDVQGLSTRNWKLTCMTCTVFRSFLGLLRCIPWLGNLIIRYFRQKKQRCLLEIVAEDTKAILRK
VLEDSIIPQ*

>MU_GMS4

MLSQMILMHPVPPHSPALLPLQVPPLLADVTAFPLSPHTLLSASLTAVLPQCKDWGMLSRVEAMSIEKAVAGGNLPELVSAVRETAKVA
FRTPVNVAVTGDSGNGMSSFINALRNVGHEEEASAPVGVLKTTQTRACYLSPHFPNVELWDLPGTECAAQSLENYATEMQFSRYDFFII
IASEQFNTNHVMLAKAVEDMGKQFYIVWTKLAMDLNTSALPEGQLQQIIRKNILENLQKKRVCDTPIFLVSSFDHLSYNFPKLRDIL*

>MU_GMS5

MLSKAEATYINKAIADRNLPELVSAIRESSKMASRTPVNVAVTGNSGNGMSSFINALRNIGHEEEASAPVGVLKTTQTRACYLSPHFPN
VELWDLPGT

>MU_GMS6

MNIEKAVAGGNLPEGVSAIRETVEIVSRTPVNVAVTGDSGNGMSSFINALQNIGHEGEDSAPHFQCPFLDGVV*

>MU_IRGQ

RLLPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARLDLAVAGTADMGLVLNVLLGLDPGEAGAVPASVPTKPTPFP
APERPNVVLWTVPLGPAGTTTVASHPTHYDALILVTPGAPTEKDRAQVRALMPPDAPLVCVRQTVRARIRSLWKKRTRWRRPAARARRV
HSARKGRSAALEGRKQAVGKALRKLSAGVCSRLSLAQRSQAVGTQIRQQH*IPEDESWEVLEESLPPVFPLRASGLPGLCEWLQRALPP
AQAGALLLALPPASISAARKKAAALRAGAWRPALLASLAAAAAPVPGLGWACDVALLRGQLAEWRRAMGLEPGVLARRERALGLAPGEL
AARAHFPGPVTRAEVEARLGAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHGVLLQALEEMLADAEAVLAPPEPTQ*

>LA_IRGC

MATSKSQVVPGEETTILMAKEELEALRTAFESGDLPQAASRLRELLASSECTRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGVV
ETTMQPSPYPHPQFPDVTLWDLPGAGSPGCSADKYLKQVDFGRYDFFLLVSPRRCGAVESRLASEILRQGKKFYFVRTKVDEDLAATRT
QRPSGYSEAAVLQEIRDHCVERLRAAGMSDPRIFLVSNLSPARYDFPLLMSTWEHDLPAHRRHAGLLSLPDISLEALQKKKDMLQEQVL
KTALVSGVIQALPVPGLAAAYDDALLVRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMRV
ARAFEKGIPVFGTLVAGGISFGTVYTMLQGCLNEMAEDAQRVRIKALEEDEPQPQLSLEAAGDSGVEKRGSGEGGSEEAPLSARRKLGL
LLKYILDSWKKRDLSEEK*

>LA_IRGQ

MPLPRGDVTALFLGPPGSGKSALVAALCDKDVETLEPLEERPDSGVPSLRAAGPGLFLGELSCPPAAPGPWAAEANVLVLVLPGPEGNG
ESLAPALGEAARAALARGTPLLAVRTLGPRDSQNEAQARDQTATLLDSAGLRAVALFVLPADCSGSEGHEELERLRAALQSQAEALQRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARLDLAVAGTADVGLVLDSLLGLDPGDSGAAPASAPAEPTPYPAP
ERPNVVFWTVPLGHAGTAAAPHPTHYDAVILVTSGAPTEKDWAQVRALVPPDAPVICVRTDGEGEEPKSLEEEEMVEKLSGESLENADS
EGREKCGTGSQKAGSGESEKSNNGFQQVAGAESGSGDSXXXLSEWRRALGLEPAALARRERALGLAPGELGARAHFPGPVTRAEVEARL
GAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHGVLLQALDEMQADAQAVLALPAPTQ

>LA_GKS1

MGHILSSTSYNEEHQNLASSMNAYFMNFKVENKIISEETINLITFNLEKGYIDSVGTVIDDVLKEIDSAPLNIAVTGESGVGKSSFINA
LRGVGYEGEGAAPTDVVETTMKATPYKHPRFPNVTLWDLPGIGTTNFPPQDYLRKVKFDEYDFFIIISATRFKENDAQLAKAIKGMKKN
FYFVRTKVDIDLRNERKCKPKNFKRGEILQKIRNNCLVSFRQLKMDEPQIFLVSTHDISKYDFPILIDTLLKDLPAQKRHMFMLSLPNV
TKEAIERKRGSLGQKIWLEALKAGAWATLPMVGIFNDNDIKKLEASLKEYRVLFGVDDITLPKFG*

>LA_GKS#2

XKVTFWDLPGSRTPNFLPETYLETVEFAKYDFFIIISSTRFTYSDALLAQKIKNDGKKFYFVRTKVDNDLYNEERSKPKSFKKERVLQQ
IRDHCLANLSHIGVPEPCIFLVSNFDLADFDFPRLEETLLMELPSHKRHTFALLIPCLSDSAIEMKRDFFKGKIWLDALSTAALALIPF
MPFIYGFDTPEQEKCLGLYRSHFGLDDKSIKEIAQKLGTPEQEIRSSIKSLDFQSLVKDDEIVAKAKKCVESYCSLKGGLPSSVFQFVK
VYLLRLKFLNTVADDAKIILQKTLEG*

>LA_GKS3_p

SSNLASLTTGRDQDLQHIGKITITKWRTTTQYWESWPNQVDIYFWGTQFNYDSAPLNIAMTWESGAGKSSFIKALRGVEYEEEHAAPTG
VVQTTVKATPYKYSQFPxLTFWDLPGIGTTNFQPQNHLWKVKFGEYDFFIVISATCFRKNDAHLAKAIKYMKKNFYFIRTKVDINLQNE
QKSKPKNFEREKALLQIQNNCLLRFRQIKMxEPQILVSSNDVSKYD*DLSAQKWHVFMFSLPNVTEEAIGRKRDSLRQKIWLKALKAGA
WATWPMVGILKDNDVKNLEASLQEYQVLFGVDDITLQNLAGNLQVPVEQLQAIIESPNLLEKKKDESIGERLLKYVKFCLANGGLLATG
LYFRKNFHLQLHFLNMMTNDAKTLLKITFSEKRLSSH*

>LA_GMS1

MKAQKDLENGNLLRSGLCCNAKLVRRVSTTGVNIAVTGESGNGMSTFINALRGIGHEEEASAPTGVVRTTQTPASYSCSHFPNVVLWDL
PGMGATSQSLKDYVVEMEFNQYDLFIIVASEHFSMNHVTLAKTIEEMGKKFYIVWTKLDMDLSTTIFTEEQLLKSIREHILENLQMMLV
CKSPIFLISSFEPLRHDFPKLRATLQVDISDIRCHGPLQDLFHTCEKIINDKATCLQEKIATDSFQNTPAIRDADGLAECLKAYQFSFG
VNDKSLQQVAQRMGKESLEYTAITMSQDVQALSRADWKLMCLTCRVMKAFLFLLSYIPSFSSPIVQCIRRMQHKRILMIVVEDTKTVLR
KVLKDSSISG*

>LA_GMS2

MAILTALLFLPLCRVTPLLTDMAEYSGSAHDPVSASFPSVVPYCTGWSISPEAGAMDIEKALADGNLLEVVSVVKNTLKTASRTPVKIA
VTGDFGNGMSSFINALRGIGHEEEASAPTGVVRTTQTPASYLSSNFPNVELWDQPGMGRSHRAWRAMCWKWGLASTNSSSSLHPNSST*

>BT_IRGC

MATSKLPAVSGEEETTILMAKEELEALRTAFESGDIPQAASRLRELLASSDCTRLEVGVTGESGAGKSSLINALRGLGAEDPDAALTGV
VETTIEPSPYPHPQFPDVTLWDLPGAGSPGCSADKYLKQVDFGRYDFFLLVSPRRCGAVETRLASEILRQGKKFYFVRTKVDEDLAATR
MQRPSGFSEGAVLHEIREHCVERLRGAGVHDPRVFLVSNLSPARYDFPLLMSTWERDLPAHRRHAGLLSLPDISLEALQKKKDMLQEQV
LKTALVSGVIQALPVPGLAAAYDDALLIRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMR
VARAFEKGIPVFGTLVAGGISFGTVYTMLQGCLNEMAEDAQRVRIKALEEEEDTQPDVSLEAAGDNGVEKRGSGEGSMEEAPLSTRRKL
GLLLKYILDSWKKRDLAEDK*

>BT_GKS1_D

MDPLLLNVIKKNDSKQLASEFLSGYKTLVSEVGGILSQVSLTRILKGFEKGQPKDVADEIQRALQSAENARQNVAVIGQSGSGKSSFIN
VLRGIGHEGAGSASVGVVPTTRKKTPYPHPKYPNVTFWDLPGTGTPESLPNPYQEVVGDDNYDYFIIISSSRFSSNDAFLAQKIQEKGK
KFYFVRTKVDSDLYNESKSKPRSFNKETVLQQIRDNCLINLSKIVSEPTVFLVSNFKSKEFDFPKLQETLLQDLPAEKRYTALLLLPNL
SESFIQLKRATIKEKLWLTAFRAAILAFIPLTPFCCGFELSEHERDLKQYQSHFGLDEESVSQIAQNLGTSEQEIYSLMKSTDFNSLVK
DDSIIANVKKCAEFVFSVTGLLQSSVFQFYKVYFLHLKFIDTVAEDAKRVLAKIEEMRSGRMK*

>BT_IRGQ

MPPPRGDVTVLFLGPPGSGKSALIAALCDRDVETVEIPDGRPDSGLPSLRAAGPGLFLGELSCPPAEPGPWAAEANVLVLVLPGPEGNE
EPLAPALGEAARAALARGTPLLAVRNLRPEESQNEAQARDQTAALLDSAGLGAAALFVPRTDCLSTDGCEELERLRAALRSQAEALQRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARLDLAVAGRADVGLVLNVLLGLDPEDPGAVPAAAPAGPTPYPTP
ERPNVVLWSVPLSSADTTVAPYPTTHYDALILVTPGAPTEKDWAQVRPLVLPDTPLVCVRTDGEGEDPESLEEEEKAEKSGSESLENAD
GEGTKKSGSGDSERAAALSPEDETWEVLEEAPPPVFPLQPGGLPGLCEWLRRALPPAQAGALLLALPPASPRAARTKAAALRAGAWRPA
LLASLAAAAAPVPGLGWACDLALLRGQLAEWRRALGLEPAALARRERALGLAPLKGXXXXXEGEXRXGXLSGEGTAGGAALGALSFLWP
AGGAAATGGLGYRAAHGVLLQALDEMCADAEAVLAPQVPAQ

>BT_GKS2

GRWDILSQLCCLSSSPEDGELGVRQPMASEVFQSCLSQCKILELSKDTRALKEAFEAGDLPAVAAKLQSTLHSLENVRLDVGVTGGMGS
GKTTFVNAIRGLGDEDPNSACTGVVEMTVDPTPYPHPKYPSVVFWDLPGVGTPAFRADKYFQRVQLFRYDFFLIITSESFTTDLAELAL

EILRRGKHFYCVRSKVDVDIAASRSRRPSSFSEERVLNQIRDDCAQRLEAQGLRDPKVFLLSMFELGKYDFHLLEESMVRDLESHKRHA
FLVALPNVSKP

>SS_Irgc

MATSKLPAVPGEEETTILMAKEELEALRTAFESGDIPQAASRLKQLLASSEGGRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGV
VETTMTPSPYPHPQLPDVTLWDLPGAGSPGCSADKYLKQVDFGRYDFFLLVSPRRCGAVETRLASEILRQGKKFYFVRTKVDEDLAASR
TQRPSGFSEAAVLQEIREHCAERLRAAGVSDPRIFLVSNLSPARYDFPLLVSTWEHDLPAHWRHAGLLSLPDISLEALQKKKDMLQEQV
LKTALVSGVIQALPVPGLAAAYDDALLIRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLANEVSPETVLRLYSQSSDGAMR
VARAFEKGIPVFGTLVAGGISFGTVYTMLQGCLNEMAEDAQRVRIKALEEDEPQPPVSLEAAGDNGVEKRGSREGSCEEAPLSARRKLG
LLLKYILDSWKKRDLSEDK*

>SS_C_like

MASEAFPSGLSQSKMLELLGDTRALKEAFEAGDLPAVAAMLQSTLHSLENVRLDIGVTGGTGSGKSTFVNAIRGLGDEDPTSACTGVVE
MTMAPTPYLHPKYPNVIIWDLPGIGAPAFQADKYVQRVLLDRYDFLLLLTSESFTAHHARLAREILQQGKRFYCVR

>EC_IRGC

MATSKLPAVPREEETTILMAKEELEALRTAFESGDIPQAASRLRELLASSESTRLEVGVTGESGAGKSSLINALRGLGAEDPGAALTGV
VETTMHPSPYPHPQFPDVTLWDLPGAGSPGCPADKYLKQVDFGRYDFFLLVSPRRCGAVETRLASEILRQGKKFYFVRTKVDEDLAATR
TQRPSGFSEAAVLQEIREHCAERLRVAGVTDPRIFLVSNLSPVRYDFPLLVSTWEHDLPAHRRHAGLLSLPDISLEALQKKKDMLQEQV
LKTALVSGVIQALPVPGLAAAYDDALLIRSLRGYHRSFGLDDDSLAKLAEQVGKQAGDLRSVIRSPLASELSPETVLRLYSRSSDGAMR
VARAFEKGIPVFGTLVAGGISFGTVYTMLQGCLNEMAEDAQRVRIKALEEDEPQSEVSLEAAGDNGVEKRASGEGVSEEAPLSARRKLG
LLLKYILESWKKRDLSEEK*

>EC_IRGQ

MPPPRGDVTALFLGPPGSGKSPLIAALCDKDVEMVEIPDGRPDSGIPSLRAAGPGLFLGELSCPPAAPGPWAAEADVLVLVLPGPEGNG
EPLAPALAEAARAALARGTPLLAVRNLRPGDSQNEAQARDQTAALLASAGLGAAALFVLPADFGSRDGCEELERLRAALRSQAEALQRL
LPPAQDGFEVLGAAELEAVREAFETGGLEAVLSWVRAGLERLGSARLDLAVAGTADVGLVLDMLLGLDAGDPGAVPATAPAEPTPYPAP
ERPNVVFWTVPLGSAGTVAAPHPTHYDALILVTSGAPTEKDWAQVRPLVLPDAPLVCVRTDGEGEDPESLEEEEQAEKPSGQSLENAGG
GGLENASSEGGEKLGPGSQKAGNGEGSEKAGRESSQEGWRRREEIGRWGLGAGEPP*

>EC_C_like

NKAFGARPGIWLPSSAEGCGEERAAQFEILPYLSCQWSSGPGDGEPGPGSPWQAVFQRSLSQARVSELLRDTGALKEAFEAGSLPAVAA
TLQATLHLLENVRPDIGIPGGTGSGKVTFVNAIWGLGDEDPNSACKGMVGMTVDPTPYPHPKYPNVIIWDLPGIGTPTFQADxYLQWAL
LDHYDFIIIMVTSEHFTANQARLTCEILQ*GKRPYFIRSKVDVDVAAxEHRHPVPSLGRVRSARSGTTVGSTQWVRRRLWGGWVGGPLR
DCPPLPLHLLVEAWSGKRGRPKSHCTTELGPTVTEVYSCHFLRAGTSQxVEGLRDPRVFLLCVIDLSKDDFHLLQESVEKELESHKRHA
FLVAPPSTSKPVLAKEASPLGHIWLVATVACGVNLSPMLGAQDMGCDLHMLIHSLEGCRHSFGLDDCPVKRAERTGRPLHKIWEAVQGQ
EPKVSQALVVKLLGQAPGDASAFTRKHLNVPVLGTLAACGTSFTTIYWMLSLSLDAAVKDAQDMLTQAFFNNSDHKLPEKPHHNPSPEP
RPAGGGSEELGGSRLYLSPGQVVSGDGKKETQAPAVLD*

>DN_GKS1

MDQSSSSKRSPTKHNDYNDLASSFDKFFKTTHIESQIISQETITLIKSQLKEGNIQGAVSAINDALIDIENAPMSIAVTGETGTGKSSF
INALRGLRHQGKGAAPVGVTETTTERESYQHPKFPNVTIWDLPGIGCTNFPPKEYLKKMKFLEYDFFIIISATRFRDHDAQLATAIEKM
RKKFYFVRAKIDSDLQNDRKCNPSTFNKERTLEKIRDDCHEKLLQAKIRVPQVFLISSFDVSDYDFPNLETTLLRELPAHKRYIFMLSL
PCVTEATIDRRRESLKQKVWLEALKAGISATMPLVGCISDSDVEMLEDTLNLYRSYFGLDNVSLENMAEDLDISIEQIKAIIKSPHLLS
VKKENESLGEKLLKCMEKTFSVIGTPIASGLYFRKTFYLHSHFLDTVANDAKALLNKEEIFISKVHSM*

>DN_GMS1

MTDMADSPVSPQSPLPANFNSMPPQIEALSIEKAWSDDEWxQRVISMFRETLRSVPCTPANIAVTGASGNGMSSFINALQGIGHEEAAS
APTGVVRPTQTNASYSSSCFPNVLLWYLPGMGATTQSLESYLLEMKFNQYDLFIIIAPEEFSMNHVRFAKTIEDIGKKFYVV*TKLDRD
LSMSxRSAPAEYPRKYPGKSPEGLGLxEPPIFLVSNLDPSLHDFPKLRDTL*KDFSHSRCDGLLQNLFHTCKKIINDKVTFLKEKISTR
SFQDTLDSPDAVNLEECFKFYRWYFGVDDESL*

>DN_GKS2

MEQLPSATPSDNFGSSFDEFPKNFKTENKILSEDNITLIKSYVERGNILKVISIIRDAMKDIENAPLSIAVIGEIGVGKSSFTNALMGF
GPEEKCAAKTGPAETTKERAPYKSLKFRNVTLWDLPGIGSTDFQSVQKYLKKMKFEEYDFFIIISAARFKETDAQLAKAIGEMRKTFYV
VRTKVDSDLQNEKKSKPNIFNKETILQKIRENYKENLHKAKVKAPQIFLVSNIDVSNYDFPQLKTTLLHELPTHKYHKLINFVSSFTEP
IIDQKRDSLKQKVWQDAWALKFFTQPFMIFISDNDVKELEEILSHYRACFGLDDVSLTDMAKYFRVSVKELKAHIKSPSLLSVEDEDKP
LRTKLLECIEKVCSNGGLLENALPFRRAYCLHIYFIDTVASDAKTLLRNVNLLVKKVPMSNNSGY*

>DN_GKS3

MGQLCFSKHSH*KPNGLASSFDKFFKxLHRKIISQETIIFIQLQLKKGDIQGAVSSISDTSTDIENAPKSIVVPGEAGMKxSFINAVRK
AELERKGVAPIRVVKTTCKTASSQHLMFPNVTFxWDLPGIGTTNF*AKQFLKKVKFGKYDFFIIISATHFKENDAQLATARTRKNEEKF
YFVQTKTDLDLQDEKKKMCKPSSFNKDRILQNIQNHCQENLQKVNVREPQVFLVSSTDVSDYxFPKMNTTLLEEFPAHKHYIFMLSLTS
FT*ATIDRKESLKQKVWLETLKAGASATILFMDFINDSDMEILEETLYSSYFGLDDVSPRNYGKGFGHFSGLKTCIKSLHFLSEE*R*V
LRGKTVKRWAKVCSVTGRPIATDLYFRKTLYLQYHFLDTVASDAKSLLKKEELFISKVDSMQV*

>DN_GKS4

QQLASNFVPYYSTLISKGGSMLSQETLTGIKKALK*GKLKVMIDMIQKSLHETENSPLDVAVIGESGxGKSSFINAL*GLSHEEESSAG
VGVVESTREKIPYQHPKYPNVTFWDLPGTGTPDFLPETYLGKVGFANYDFFIIISSSWFEYNDALLAQKIKEIGKHFYFVRTKVHSDLY
NEEKTKPSSLYKERVLQKTQDKCLANLNVIGVPAPCIFLISNFDLGDFDFPSLEETLLKEIPAHKCQPFVLLLPSLSEAAIELKRDFLK
EKIWLDALKSAALDFMPFMAIISGFDLPKQEKCLDLH*SYFGLDDKSVNEIAQYLDISMQEIKSSTKSLNF*LLVKDDSIVAKVMKCAE
LYCSVNGGLPSTIFQFFKTYFLHLKFLNTVADDAKFILHKTLESSLRE*

>DN_IRGC_K5

MAGARSPGPREEETTILMAKEELEALRTAFESGDIPQAASRLRELLASSESTRLEVGVTGESGAGKSSLINALRGLGAEDAGAALTGVV
ETTMNPSPYPHPQFPDVTLWDLPGAGAPGCPADKYLKQVDFGRYDFFLLVSPRRCGAVEARLAAEILRQGKKFYFVRTKVDEDLAATRS
QRPSGFSEAAVLREIRDHCAEGLRRAGLAEPRIFLVSNLSPARYDFPLLVSTWEHDLPAHRRHAGLLSLPDISLEXXX

>DN_IRGQ

MPSPRGDVTALFLGLLGSGKSALITALCDKDVETLGARRTESGVXXXARAALVRETPLLAVRTIHAGESQNEAQARDQAETLLDSAGLG
AATLFVLPANCSSSDGCEELERLRVALESQAEALQRLLPPAQDGFEVLGAAELEAVREAFETGGLEAALSWVRAGLERLGSARLDVAVA
GTADLGLVLDALLGLDPGDQGAAPASAPAEPTPYPAPERPNVVLWTVPLSPADTAAAPNPTNYDAVILVTPGVPVEKDWAQVRALVPPN
APLVCVRTDGEGEDPEPLEEEGKAEKSSGEGLENARSGKRGVGQQTTGSGEGSEKAGSESWQLIGAGAKKSGGGDAERTAASSPEDETW
EVLEESPPQVFPLRPGGLAGLCEWLRQALPSAQAGALLLALPPTSPRAARTKAAALRAGAWRPALLASLAAAAAPVPGLGWACDVALLR
GQLAEWRRAMGLEPAALARRERALGLPPGELAARTHFPGPVTRAEVEARLGAWAGEGTAGGAALGALSFLWPAGGAAATGGLGYRAAHG
VLLQALDEMLADAEAVLAPQAASQ

>MD_Irgc

CGPGPSKKGIPVFGDTSGGRHLLGTVYTMLQGCLNEMAEDAQRVRIKAMEEEVAAPDARLDVASSGPGLEKRGAGEGGGEDAPLSARRK
LGLLLKYILESWKKRDVVEEK*

>MD_C_like

MAAPKETFEEGTIPVASGKLRATLQDLENVRLDIAVTGGSGSGKSTFVNVIRGLGDEDQGSAKTGVVEIITDPTPYPHPKFPNVIIWDL
PGVETPLMEANKYLDKLLPSRYDFFLIMVSERFTTSHAQLACAIQDQGKPFYFVRSKVDIDMAASRQRRPTTFSEEGVLCEIRAKCCTE
LEDEGVMDPEVFLLSMFELGKYDFQLLIEIIEKDMDAQKHHAFLVALPNVSQAILEKKQATLKQHIWLVATVACSSNPQPVPGVRGVAC
DLSQLLRSLDTYRRSFGLDSTSLCKLAQQTGQPLRVMQSAILGPSSEVTEAQVIQLLAEATQGPTGVFAKELSSVPVLGVVASCGYSFA
VVYQMLRTYLTGAAKNAQRVMKRAFHSSL*

>MD_GKS1_D

MEHSLSSLSLKEKPNALASSIVNNYETETGGRGFLPQENITLLKSVLKEGRLEEAVAIIKEKIEDMDKALLNIAITGESGTGKSTFINA
FRGTSHEGDDAANTGVVETTTEIIPYEHPKFPNVKLWDLPGIGTPNFQPKLYLEQVNFNSYDFFFIVSSTRFRDNDANLAQEIRKMGKK
FYFVRTKIDSDLDNERKAKPRSFKEENVLQQIQENCLQNLRKVGIEDPQVFLISNFELASFDFPKLQDILVEELPAHKRHVFLLSLPNI
SEAAINQKKAALQDKIWLKALKSGLRSTVPFVGIIKEDSIAKLQKTLKNYQSLFGVDDASLRKIAQKSNRSLEDLKKLIKSPDLLTVKK
DESLSEKLVKYAEVVFSVSGGLCATGLYYRKIYNIQLHFLNTVADDAKILLSKIWEPPAKER*

>MD_GKS2

METSDTALHDANIQSLASNFMPGYRMISSTDGSLLTKDTLTWIQSALEEGRLGEVASRIQETLEASENIPLNIAVTGESGSGKSTFINA
I*GIGHEEKDSASTGVVETTKEATPYKHPKYPNVNFWDLPGIGIPDFHPETYLNKVNIDQYDFFIIFSASRFTVNHVKLVQEIRKMKKK
FYFVRSMVDKDLENERKAKPSTFSQEKVLQIIRDYCLKYFNDECMSEAPVFLVSNFDLSVYDFPKLEETLIKDLPACKRHAFLLALPDV
SKAVIDRKKAALRERIWLQALKTGALATIPVAGFFLNDLVTLNNHLKHYRKVFGMDDESLLQVARRLGRPVEEIKVPLKSLDLEALIKE
ESTIELLLKLVEGFCSANGIILSAGFHFGKTYYTQLLFLNTVADDAKILLQKNLGVPF*

>OA_IRGC

XLAGEVLRQGKRFYFVRTKVDEDLAASRSQRPSGYCEGAVLQEIRDHCSERLRLAGVAEPRVFLVSNLAPARFDFPLLVSTWERELPAH
RRHAGLLSLPDVSLRALQEKKDALQDQVLKAALVSGVIQALPVPGLAAAYDDALLVRSLRGYHRSFGLDDDSLARLAEQVGKPAGDLRA
VIRSPLAAEVSSDAVLRLYSQASSGAMRVARAFERGVPVFGTLVAGGISFGTIYTMLQGCLNEMAEDARRVRVKALEDASGEREPGAGP
DEEAGADQDPDAWALDLEGRGLGGAGGEGRGDPPLTMRRKLGLMLKYLLEAWKRKDVVDDK*

>OA_GKS1

MEPQASRPSGNQVPPLNPDAIREFTAALGSGNVPNVARKLMEAVKMAAYAKLNVAVTGEPGAGKSTFINAMRQLGDEDPGAAATGVVDT
TRDPTPYPHPKYPNITLWDLPGIGSPEFRAEGYLARVASDRYDFFILLASQRFTYNLAQLACAIQQQGKKFYLVRCKVDVDLEASRRRR
PSSFDPEQVLAEIRRDCQDQLSRQGLSGPRVFALSNFNRGLYDFPLLEETLEKELPGHKRQAFLLARPNDSLEVLEKKKAALEEQIWKL
ALAACTVNSVPVPGLPGLPAACEVAILTDSLSDYCRSFGLERETLEELARELGLPQEEVQGLIRSPLAQDITRGLVLQLLASASQSAQL
LFQYFRQAVPTFGTLAAAGLSFAAIYLTLRTFLQHVAEDARRIVLLAQRKRGQRDGEPPRRAPPAPPEPRGSAEALPPAASPPPPAPLP
GPSKSPS*

>OA_GKS2

MAAKAKPPTGPGASSGQGPEEKNFLQLSAEELGALREAFEDGNLSAAARKLQDTLQSLENVRLDVAVTGETGSGKSTFVNAIRGLGDEE
PGSAETGVVETTLDPTPYPHPKYPNVTIWDLPGIGTPAFQASRYLERVQFGRYDFFILIASERFTAHHAQLACEIQRLGKRFYFIRSKV
DVDLAASRQRRPATFSEEGVLNEIRTHCRDRLRAEGVAEPRVFLLSTF*LGKYDFHLLGETLERELEGHKRHALLLALPNVSRQILDKK
RASLHQHIWLISAISCGLDPVPVPKVPGLLLRPLTAGTGPTWLLSELRVGMSTTXXX

>OA_GKS4

MASRTGSRATSPGGEDDGSILLNKEELEALRSAFEAGDLPQAASKLRELLAPAEPGQLHVGVTGESGAGKSSLINALRGLGAEDPGAAR
TGVIETTVQPAPYPHPQLPGLTLWDMAGAGTPGCPADRYLXXX

>OA_GKS5

RKRFWDHFGAERSQEHLNITVMGESGAGKSTIINALWGVGHEEEGSALTGVVTTKEPTCYQHPNLDMNYWDPPGIGTPNFQPGSYLEKV
AFHRYDFFILVASERFTSNHALLAREIGKMGKRFYFVRAKVDMDLENCRIRKPRSHNEDRILWEI*

>OA_GKS6

IIASEPFKSSSVDLAREIQRMEKKFYFVRTMTVEDPASAKEGKPGSSNEESLPQLIRDDCLESLKKGSVSHPRVFLISSFDMGLYDFGD
LEDTLVRELPAHQRHTLRLALPSVTEEAIEKKKLTLQEKIGLEATKECLRAPFSGIFSPNNREELEKCLPHYRSLLGVDDASLAKLAER
EGREVRDLKAALKSYDMETVLKNSDALPSLVNSVKEKIKTSLSCVSFFGNLLSTIAISRTQIYTLQWHALDTVAEDAKLLLKKTQSTG*

>AC_GKS1

MKVKKSHVIFACFFTFRSRLDQMAAAPQEAAQNHQVNFEEYDIITEEDIEEIKDALEGGRMAEAASKIMENLQALENARLDIAVTGESG
SGKSSFINAIRGLGDEEEGSAPTGVVETTREPTPYPHPKHPNVTMWDLPGIGTPDFQSSTYLEQVNFSRYDFFFLIASERFKANHAMLA
NEIKKQGKHFYFVRSKVDADLEASKKRRPRSYNEEMVLTKIRENCQDCLKKEGVDDPHVFLLSSWELSKYDFMLLEKTLERELPNHKRH
AFLLALPNISLEILQKKKETLQKQIWKLATISCGVAAVPIPGLSVTCDVALLIKSLSTYRKNFGLDEESLIKLAEKVDKPVEEIKEAIK
SPLTKEISKDLVVKMLTKAGGGALMFMEYLASTVPIFGSMAAGGISFGTTYYMLWSFLNEVAGDAQNVLIKAFESDV*

>AC_GKS2

MASAPQEAVQNHQVNFEEYHIIAEKDIKVIKDALEGGRMSDAVSKIMENLQNIENARLDIAVTGECGSGKSSFINAIRGLGDEEEGSAP
TGVVETTREPIPYPHPKHPNVTMWDLPGIGTPDFQSSTYLKQVNFSRYDFFFLIASERFKANHAMLANKIKKQGKHFYFVRSKVDADLE
ASKKRRPRSYNEEMVLTKIRENCQDCLKKEGVDDPHVFLLSSWELSKYDFMLLEKTLERELPNHKRHAFLLALPNISIEILKKKKKVLQ
KQIWKLATISCVVATVPIPGLSITCDIALLIKSMSTYRKDFGLDEESLIQLAEKVGKPVEEIKEVIKSPLVKKISKDLVIKLLTKAGGG
ALMVLGYFVSMVPIFGSLAAGGISFGTTYYMLKKFIDELAEDAQNVLIKVSESGV*

>AC_GKS3

MSFCPLSLYSRTATMEIQEVDLREMGSLIQSRVILEVSNQVQSLLESVETTTLEIAVTGESGAGKSTFINALLGLNDGDPRAAPTGVTE
TTFSPMPYEHPRLPRVHIWDLPGTGTPRFQAETYLQQVGFERYDFFIIIASERFRENHVKLARAVAAMGKRFYFVRSKIDLDLQASRQR
RPTRFEEVHVLREIEADCIRQLKKEGLDSPKVFLISSFELHRFDFQCLEDTLAEELEGHKRHVLLLSFPSVTTEAVQKKKASLRRNIWK
KALMACFFSALPGLPFHLNIPMLLKTLDSYRRSFGLDDDSLGALALTTSKNPSQLKNQVSSTLARDLSENAVHVILSQAATYGKVAARL
LKDRVPFLDNLVAGGISFVAAYYLLHTALDNFAKDAERVLLAAYDLEDEFQKSVFYPPEPGFIFD*

>AC_GKS4

MGIALTKSLVKEELQRFKAAKEEGNIEDVKAKIMKDMELLNNTTLHIAITGDSGSGKSSLVNALREMTDEEKDSAPTGEIEMTKDKQSY
KHPKFPQVTLWDLPGIGTPNFKADEYLKQVNFRQYDFFIIASGERFTENDVKLACEIKEMKKKFYYVRTKMDVSISNNKKRKNFSEEKT
LQEIRNYCIQRLRGTGEHSPRVFLISRWHLNNYEFPLLVKALEEDLDALKRHVLIMALPTFSKEVIKKKENAMSDLIWKLSLLSCTVGA
IPIPGLSLACDIGILVSTMRKIFNCFGLDAESLRRLGERVGKPVDELKSARKKFQGVDTLTKEFVIDYLQKSVAWITVSTVELVLDFVP
VVGSVVGGGASFVTTYYMLKNFLSDAVEDAQNVLAKVCEKKE*

>AC_GKS5

MFLLLLRSNEPPKGSFLRSAISRTILKLDLENLTLALEKEDISDVVAKTQQQLDFLTNATLDIAITGRSGAGKSSLVNALRGMSDYDEG
SAETGVIQTTMEIKGYPHPTFPKVTLWDLPGIGTPEFKAKEYLRKVNFSKYDFFIIVSAGRFTTNDITLAHEIQKMKKQFYYVRTKIDI
NIDSEKRKQNFNEHETLEKIRNDCSENLKKAEGSSSRVFLVSRSDLSLYDFPSLHGTLEEDLDDLKRHAFIMASPTFSGEMVKKKKKSL
ESLIWMLALVSCFVGAVPVPGLSLACDISILMGAMVHFCKVFGLDDNSLHVLAKRVGKPVEELRSAIKNTPLPHTITTDLVLSLLSKSL
LCATLTVIELVCDFVPVLGSIVGGASSFVCTFFMLRSFLQSVEVDAANVRAKATK*

>AC_GKS6

MRSISKTIISLPHAEEDLIRNLAEKLYETFNAVSIYEVARELERSLELWGKTELHIAVIGETGAGKSSFINAMRGLKPGDVGAAQVGVT
ATTSNPICYENPNDPNVKFWDLPGIGAPEFRPNDYLEKVNFHHYDFFIILCSKRFRDIHIDLAQIIDNEKKKFYFVRTKVDEDLTNMER
DYPRTFNEANVLQNIRDDCNRYFQVYIRSFIPQVFLLFSCNLERFDFYLLIERLEGDLPTLKKLKFVLNLPNMSSDIIQQKKQLLKGEL
WKVSLVSAFINAVPLPDFPVAFNASFLQKYMADMYKKFGLDGDSLALLAWHINKPVEELKAQIKSSQEEYISTDSIIKKCFNIVMTFVV
FWNKYFPLSVFTSLANVGLSFISTSRMLYKFLDQVAEDAERVSKEALKHVEYDINRPLNRT*

>AC_GKS7

MGGTNSHPSSIQELEDVTFNIAVIGETGTGKSSLVNALRGMKTDVEGGAAVTDVIQVLNEPTAYVHPEYSDITLWDLPGIGTNEFKSEE
YVKNIDLNKYDFFIIVSEIRFTEDDQRLVHAIQKIKRKFVYVRTKVDKSLESERENPNFSEEGILKKIRNHYCENLTKAGEPSPRVFVI
SRWHLEMYDFPCLVNALKDELADFKREVSRWSPEKKRKRLDR*

>AC_GKS8

MRSVLKTIISLPNAKEDLIQNIAEKLCETFNAVSIYEVAHELERSLELWEKTELHIAVIGETGAGKSSFINAMRGLKPGDVGAAQVGVT
ATTSNPICYESPNDPNVKFWDLPGIGAPEFRSNDYLEKVNFHRYDFFIILCSKRFRDIHIDLAQIIDNEKKKFYFVRTKVDEDLTNMER
DYPRTFNEANVLQNIRDECNRYFQVYMGSFIPQVFLLSSCKLERFDFYLLIERLEGDLPILKKLKFILNLPNVSSDIIQQKKQLLEGEL
WKVSLVTACINAVPLPDFPVAFNASFLQKYMADMYKKFGLDGDSLALLAWHINKPVEELKAQIKSSQEEYISTDSIIKKCFNIVMTFVV
FWNKYFPLSVVTSLANVGLSFTSTSRMLYKFLDQVAEDAERVSKEALKHVEYDINRPLNRT*

>AC_GKS9

MAMEDFKAAVYEGRLTDAVSNVVGKPLQYFSEVPLNISVAGEPGSGKSSFINAMLGLHAGDPGAAETGIQTTTVDVKAFPHPHLLRSFL
DLPERAAFLVRYFADDs*xRFDFSS*LVPSVSRTIHADLVNEIQGMSKKFYFVRSNIDLDLEASKRQRPSDYNEEKILLRIKDDCHEGL
RREGVANPQVFLVSSYETSRFDFPLLWEKLKSDLLGLRRKAFLLNLPSIYLPVLNNKKMAMKKQILTRALWLWIFAAIPIPGLSYFPAR
KVHSWCYRNFGLDDPSLTDLSQLVGKTAATFKAVMKPLSFTSVVLWGFAELVRAVVIIGDYNHHRHFPLYGYLLSGGISLLSTYLILKK
FVSNATDNTQRVLTEALVCEEKKSI*

>AC_GKS10

MSAKGTEAKVSLTEMQIRDLQRAFEKGSFSDLSDRLREAHSALENLRLDVAITGEPAVGKSTFINAFRDVSADDDDAAPTGGSAATTTT
ATTPEPRPYTHPKYADVRLWDLPSIGGPGFQPRDYLQQVEFSRYDLFLVLSSGRFSALHGRLARLLQKAQADVYFVRSRVDQDVEARRR
RREGPAGFSEAAALKEIREACLQDLAAAGVKSPKVSEPQGRPRGPPGPAPPHAERAGKS*

>ATT_GKS1

MTDCKLHIAITGESGSGKSSFVNSLRGLMDDDVDSAQTGVIEMTACPTPYQFSKNPDVLLWDLPGIGTPLFQPKSYFQQVKFHQYDVYI
IIASERFKSNHAMLARGIANMGKRFYFVRTKVDLDLRASKRKLKSVYNEAQILQTIRQNTVDCLLTEGLDASQVFLISNFELDRHDFPA
LLKALENDIPAHVLNASKSQWNRGCRLQ*

>XT_Q1

VESTYCSCFRCPSYIGTWSPFNEPAFFSLPFAHRPLSLVQQGYEVIGDKEISELQSALERGGIPEMVRFIQNSLDTLLDTRIDFALEGG
DCSVRHSLLNALRGVPDEEDGSAHEKDPETPTEYPSLKYPKASLWDLPGLESQEFQVKDYLEHVNINHYDFIFLLLTHGTEATASFQSL
LSELKKTGKKVFLIQAYQEGEIQLEEESGQNEVFVLCPTNLPGPEFHRLLGNLEKQLSTHKSRAFMLSLPNLSTEVIQKKKDALAHEIW
KVALLSSLVACVPVPGVAIACDMSLLTARLDTYRQELGLDANSIASIAQQYGVSAATLTLEIHSPAGKGVTRKLVGKMLGSATGLGLEV
AGVILHRFPLLGPLATGGIAFHANYSVLSRCLEGMAEDAQRVLEKALKLTSTPMN*

>XT_Q2

MKSYAVYVPPHPFQVMETLGELSLSEEKTHPWGTEDPTPCLTEECCEFSVAVAGSPGVGKTVLIQALLGLTESHQHSNAASLHPTRPGG
PLLYIHPSHPGLSLWELPLEDGSPEKWISEADLLLLVTNGCFSPVHIDLARSALSQGKKLCLVRSQVDCDLHTLKRRMKEMFSREEALH
SLCTVTLQPFTHLIREGEIRLFLVSGYEPWKLDTAGLKAVIEETAALCGRYVYMGIWVSLFLEGAEQT*

>XT_Q1+2

CWQTVLIQALLGLTESHQHSNAASLHPTRPGGPLLYIHPSHPGLSLWELPLEDGSPEKWISEADLLLLVTNGCFSPVHIDLARSALSQG
KKLCLVRSQVDCDLHTLKRRMKEMFSREEALHSLCTVTLQPFTHLIREGEIRLFLVSGYEPWKLDTAGLKAVIEETAALCGRPLSLVQQ
GYEVIGDKEISELQSALERGGIPEMVRFIQNSLDTLLDTRIDFALEGGDCSVRHSLLNALRGVPDEEDGSAHEKDPETPTEYPSLKYPK
ASLWDLPGLESQEFQGK

>DR_irge1

MPEKEEDKNENLYIISSEFLDIMSNATDDPDSISEDMKEVIDAKPKEKTRKLKDKLTELENVTLNMAITGMTGAGKSSFVNALRGLRDD
DEGAASTGTTETTMKPNMYEHPFMPNVKIWDLPGIGSPKFRAKKYLKDVNFHMYDFFLIVTSERFRENDIELAKAINKSNKLFYFIRTK
IDNDVRAESNKRNFDERVLLDKIREDCKVNLLKLNISKIFLISSFHLERYDFQKLVNTLEEELPKNKRFALIQSLPVYSLETLTKKITY
FKKLIWLNAVGAGVGAFPPIPGVSLAVDYGIMKKFFKQVFMAFGLSNQALQVLSERVNKPVEVLNAAKTSRFKDGVTDRILIDMMSNPV
IAITKTLGTIMALLPGGALPAGGAAVASVHYLLNVGLKEMADDTRKVLVVSQLA

>DR_irge2

MKIQKQKQELSNSSKPDTHSHSTAKENVSLKSANTVQVEHIYEMPDVHLNSSAEYINEMECVIEQNKQLGNVTLHVAVTGSTGAGKSSF
INAIRGLTSDDENAAPTGVTETTLVPTMYRHPTMPNIELWDLPGTGSPKFKAKKYLKDVKLETFDFFIIISSERFKENDIMLANAIKER
KKLFYFLRSKIDNDIHAESHRKDFDEQKVLSHIRENCHRNLKDIDDPHAFLICSFELHKYDFQTFVDTLEKQLPDHKRDALILSLPIYS
SKILEEKIEIFMKQTWSAAVASGSVAVVPVPGLSMACDAAILLGFFTKCYYAFGLDEKSIDKLSVRVNNLSLKAIRRSPLVVAIGQKKL
TNKELSALTSKEAAVKFAWSMVPVVGSIKTAQMSYSTTLNLLRTGVQDLAETAS

>DR_irge3

METQDPAIAEAVQASGESTLEKATAKAKESFDQFMNVSLNIAVTGKTGSGKSSFINALRGLKDDDEGAAPTGVTETTMEPNMYEHPAMP
NVKIWDLPGIGSPNFKADKYLKDVKLKNYDFFIILNSERFMQNDVMLAKEIRKQKKNFYFVRSKIDNDISAEQRKKTFDEQRVLCTIRE
DCLKNLKQLGDPKVFLISSFDLEKYDFEELQNTLAEELPVHKRNALLQAWPVCSAASLEMKIKMFEGVIWAASLASAGIAVVPLPGLSA
ACDTGMVALFLTRCYFAFGLDDGSLARLSEKINKPLVGHLAKSKIASAIQEKALTRLQVSGTLVVLFSAEYVASLVPGVGSVAAAGLSF
GTTYYLLRSGLKELANVAREIRKEVLDSVR

>DR_irge4

MTDDSSADMNFSGALQRLGESDPNAAAVKAKEELDRLDSVTLNIAVTGEAGAGKSSFINALRDLSDEDENSAPTGLTETTKKATMYTHP
TKPNVRLWDLPGIGTPNFKANQYLKDVKFETYDFFIIISSERFKENDVYLAKEIQKKQKRFYFVRNKIDNDICSVANGKINEQQLLCAI
REDCYRNLKEVGNPKVFLISSFDLRKYDFNLVGTLESELSDQKGFALVQSVPVYSLAMLEKKKALLEKFIWLAALASSACTLVPNQFIS
LITDKAILIVYLIGCHYALGLNEKSLKQLSERTNKPVSLLKLAIKSPVSLAVLDRMRISPMAKPVKSLEDLLDSKNLAVNVQNTADAFR
NSHTNLTRALNEMIKDMRQVLQVAGLDE

>DR_irge5

KEEEDENENLYIVSSEFINIMSNATDDPDSISVDMKEVIDAKPNEKTTKLKDKLTELENVTLNMAITGMTGVGKSSFVNALRGLRDDDK
DAAFTGTTETTMKPNMYEHPFMPNVKIWDLPGIGSPKFRAKKYLKDVNFHMYDFFFIVTSERFRENDIELAKAIKKSNKLFYFIRTKID
NDVRAESYKRNFDEPMLLDKIREDCKVNLLKVRISKIFLISSFHLERYDFQKLVNTLEEELPKNKRFALIQSLPVYSLEALTKKITYFK
KLIWLNAVGAGVGAIAPIPGVSLAVEYVIMKKFFKQVFMAFGLSNQALEVLSGRVNKPVKVLKAAKTSRFKDGITEHILMDMISNPVIA
IAVTLGTIMALLPGGALPAGGTAVATVHYLLNVGLREMADDTRKVLAISQLA

>DR_irge6

MECVIEQNKQLGNVTLHVAVTGSTGAGKSSFINAVRGLTSDDENAAPTGVTETTLVPMMYKHPTMPNVELWDLPGTGSPKFKAKKYLKE
VKLETFDFFIIISSERFKENDIMLANAIKERKKLFYFLRSKIDNDIHAESHRKDFDEQKVLSHIREDCHRNLKDMDDPHVFLICSFELH
KYDFQTFVDTLEKQLPDHKRDALILSLPIYSSKILEEKIEIFMKQTWSAAVASGSVAVVPVPGLSMACDAAILLGFFIKCYYAFGLDEK
SIDKLSVRVNNPSLKAIRRSPLVVAIGQKKLTNKELSALTSKEAAVKFAWSMVPVVGSKKTAQMSYSTTLKLLRTGVQDLAETAREVLK
AAGVTGVY

>DR_irgf1

MATFEDYCVITQEDLDDIKDSISTQDLPSAVNTIKEYLKQQDLVELNIGVTGESGSGKSTFVNAFRGLGDEDEGSAETGPVETTMEPEV
YIHPKYHNVKVWDLPGIGTPNFKADEYLELVEFERYDFFIIIASDRFRECHTQLAKEIMRMGKKFYFVRSKIDASITAEKKKKNFDQKK
TLDSIRKDCINGLRKIGIEDPIVFLISGWELSKYDLNLLQDRMEKELPQHKRRVLMLALPNITLEINEKKKKALEENIRKVAFLSACVA
LFPLPGLSISADIAIIAEELRKYYSAFGLDDPSLQKLCERSGKTVEELKSLMKSPLHHGINPSSILTLLGAASVLISEDAVELLVSFIP
IIGSVVAGGLSYLTVSGMLKKALNEIAEDARNVLMASLETEV

>DR_irgf2

VDALEHLYEIKVEDKLKEIKEILYTQDLPTAFGTISNYFKETSLVLNIGVTGESGSGKSTFVNAFRGLGDEDEGSAKTSSVVTTAEPEV
YFHPKYENVKLWDLPGIGTPNFKADKYLELVEFERYDFFIIIASDRFRECHTQLAKGIMRMGKKFYFVRSKIDASITAEKKKKNFDQKK
TLDSIREDCENGLRKIGIEYPVVFLISGWDLGKYDLNLLQEMMEKEILKCKRILLKSALLNVKQEVIEQRKDTLKRNIERVTEQSVAIT
DVHLPGLSISVNVDIIAEELTKYYSEFGLDDQSLQKLCERSGKTIEELKSLMKSPLCYGINTSLIINLLEAEVPKIENEYFLSFMPFIG
TEIKKIKSSVAVSSMLKTALNVIAEDIRNVI

>DR_irgf3

MDILEDYDIITQNDLEEIKESISTEDLPTAVSRIREYLRKQDLVELNVGVTGESGSGKSTFVNAFRGLGDEDEGSAETGVVETTMEPKA
YNHPKIQHVKVWDLPGIGTPNFKADEYLQQVEFERFDFFIIIASDRFKECHTHLAKEIMRMGKKFYFVRSKIDASITAEKRKKNFDLKK
TLDVIREDCVNGLRKIGIEDPVVFLISNFELGKYDLNLLEEKMEEELPQHKRRVLLLALPNITQEINEKKKEALGQNIGKVAILSACVA
AVPIPGLSVAVDLVIVKREIEIYYSTFGLDDPSLQMLCERSGKTIEEFKSLMKSPLRGGINPASLLSLVGAVSVVGAESTVEYILSLVP
ILGTVVAGGLSYLTVSTMLRRALNDIAEDARNVLNASLETEV

>DR_irgf4

MSNISQKVVLLFAEQEELVDLRKAISTQDLPTAINTIKECLRKQDLVELNIGVTGESGSGKSTFVNAFRGLGNEEKGSAETGFEETTME
PKDYIHPNFKNVRLWDLPGIGTPNFKAKDYLKLVKFERYDFFIIISSDRFKEHHSLLAEEIVRLRKTFYFVRSKIDQSIDSEKYKKTFD
QEKMLDNIRDKCKSELSKIVKDPAVFLISCNELNKYDFQLLQERMETELPLHKRRVLMLALPNVSLDVIKKKKEVLEKDIAKVAFISAT
VSAVPIPGLSVAVDVMIIKEETEKYFRGFNLDDESLQRLCDVSGKSLEEIKSLMKSPLKAGIGSYSILALLSSATLVLGGMSVLAAESA
LEYFLSTIPLIGSVAAAAMSYKTITLMLKKTLNDLAKDAETVFKALLETEV

>DR_irgg1

MFFSRLCMPAKVQEDHLGTIRDVFAGESPETIPHRLISLLEVFDRFKIDIAVTGDSGAGKSSLINAILGLKPDDKGAAQTGAIETTKQA
TMYQQSNLPHIRLWDLPGMGTPSFASKSYVKMMNFDLYDMFMVVISERVRENNMLLVDEIDKRKKPFYFIRTKIDNDVKSQRRKSKFSE
TQALEQMRQDCEKYLKEKKLDPHIFLVSTHDTHNYEFQKFISTFKDEVFKIRAEEFSGFLDKMLHGGWLKAR

>DR_irgq1

MLHGGWLKARYATQHVQQTEKLETEDITKLQNMYKSTGFGAAKVSAVLEALSHFQLDVAVLGETGSGVSTLVNALVGLENEESSGAGAS
ISNPALSPVYPDVRFWDISGIEAVMDYSVFEMKQAMKCYDFYIIIVSDWEKVRHVKLAKEVEKLRKHYLLVQTKVDSCLQTQGDLCCEE
TEILDGLRAQFTQELQREKLSEQQMFLINSQDRSAFDFVSLESALSSDLNTIRTSAFAYYIARTVKENL

>DR_irgq2

MADVIKGLNLLETLKESIEKNNISDIRDALEDMLISRINIAIAGERNAEKATFINSLRGLSQEDEGAAQNPPSAAPEELAVFTNPKHPD
FRLWDLPPISSDANFKPEDYIERFKATRYNAIILTSTDRPSANSVAVWKEVRSLQKETVYFVLLASVKDTEKSLEAKKAASLDVLKAEG
VPLPKVFLVQPSALEKLDFLTFLEVMRGDLPEIRAHALLLALPTFSSSLVTQKKDAFKALVWAAASLSGGVSAIPVPLVSSMVDATVGV
RILVKAQISLCLDDESLQRLARQRGLDPAKLKALRTCALSVEVSKSEVKRRLAEAEKDTSTATTRLVELAIPRQARSVSRSFTVMLQAL
NNAIDDMGADAEKVVAMVTGERQ

>DR_irgq3

MAIQCTHRICSYLTNSLFFRFVVSTALRSMKINQDDLDQISKLSQTRDFTDNPSKLQAILGALDHFRLDVGVLGETGCGSSSLINALLG
LKNSNETAALTGVTETTKEAVEYALPDSHNIRFWDLPGLGKIGDLNSLSANAFSSSEGQQVASVLALCDGYIHILVSPLRVRLRTIQLL
QQASSMGKECYLVISMVDLIEDKAVEEVRQWTEKVLSKLDIQQSLFLVSANYPETLDLAKLKGMLKAAIPSHKKVALARYVSKQLDEDV
FWKRSDSCKFM

>TN_GKS1

MADFPDMTEIKEALQSNNQALAVSKIKELLEKTANTTLNIGITGESGSGKSSFVNAFRGVDHRDEKAAPTGVVETTTVVKEYPHPSYPK
VSLWDLPGIGTTKFPADEYLKHVGFERFDFFIIISDTRFRENDVKLAKEIQKMGKKFYFVRSKVDNDLQNAQRSQRDFDANQTLSLIRE
NCKQGLLKEGVQAPQVFLLSNFELQRHDFHRLHETLERELPDHKKNVLLFAMPNVSLEIIEKKKEAFSSKIPHYAFVSAACAAVPVPGL
SVAVDGSLIAGVVQQYKTGFGLDIPSLQRLANSTGVPLEALTSVVRSPLGLNNINVQFILKILLHSASVAGSMVAEEGLRFVPLFGTMI
AATLSYKVTEKALQDFLNMLAEDAHNVFKRAVCSMNSSV*

TN_GKS2

MAEFLDKTEIKEALLNDNKDLDVAKIQEHLERTANIPLNIGITGESGSGKSSFVNAFRGVDHRDEKAAPTGVVETTTVVKEYPHPSYPK
VSLWDLPGIGTTKFPADEYLKHVEFERFDFFIIISDTRFRENDVKLAKEIQKMGKKFYFVRSKVDNDLQNAQRCQRDFDANQTLSLIRE
NCKQGLLREGVLAPQVFLLSNFEYQSHDFCCLCETLERELPQHKRDVLLFALANISLEIIEKKKEAFKSKIPQHAFLSAAHATRPVSEL
SVAVDADLIANVVQQYKTGFGLDRPSLQRLTDITGIPLARLTIISSHLTLQNVNADFVLSLMSQSSAISSLTEKKQIFGFIPFFGTLLA
QKLSYKISTIALLSFLDMLTVDAKDVYTKVKMYHSGKK*

TN_GKS3

MFASCVGLIMADFLGTTEIKEALQNNNQALAVSKIKELLEKTANTTLNIGITGESGSGKSSFVNAFRGVDHRDEKAAPTGVVETTTVVK
EYPHPSYPKVSLWDLPGIGTTKFPADEYLKHVGFERFDFFIIISDTRFRENDVKLAKEIQKMGKKFYFVRSKVDNDLQNAQRSQRDFDA
NQTLSLIRENCEQGLLKEGVQAPQVFLLSNFELQRHDFHHLHETLERELPDHKKNVLLFAMPNVSLEIIEKKKEAFSSKIPHYAFVSAA
CAAVPVPGLSVAVDGSLIAGVVQQYKTGFGLDIPSLQRLANSTGVSLEALTSVVRSPLGSNNINVQFILRLLFQSASVAGLMVAEEGLS
FLPIFGTMIAATLSYKVTEKALQDFLNMLAEDAHNVFKRALCCMNSSV*

TN_GKS4 = irgf8

MNIVKLLILHSCGQCSCFICNSFDLLEIFIQNVSCVFFATKILALCFRLIMADSSDFAEIKEALQNNNQALAAAKIKELLDNTSNTTLN
IGITGEAGSGKSSFVNAFRGVDDRDEKAAPVGVVETTAEVKEYPHPNYPNVSLWDLPGIGTTKFPADEYLKLVGFEKFDFFIIISETRF
RENDVKLAKEIQKMGKKFYFVRSKVDNDLQSEQRYQRDFDPEKTLSLIRENCKRGLLNAGLQAQVFLLSSFELQRYDFHLLYETLEREF
PEHQRDVLLVAMSNISLEINGKKKEAFKSKIPYWALVSSVGALVPVPGLSVAVDLSLIAGLVQQYKTGFGLGRPSLQRLADTTGVQLTD
LTSVIRSPLGLNIIDAELIVKALSELASVAGLMAAEEGLRFIPIFGTMIAGTLSYAATYNALSDFLKMLTEDAQNVFEKALRCMNSSV*

TN_GKS5 = irgf7

MNIVKLLILHSCGQCSCFICNSFGLLNIFIQNVSCVFFATKILVLCFRLIMADSSDIAEIKKALQNNDQALAVAKIKELLDKQSNIPLN
IGITGESGSGKSSFVNAFRGVHHRDENAAPVGVVETTTDVKEYPHPDYPNVSFWDLPGIGSTKFPADKYLKLVGFEKFDFFIIISATRF
RENDVKLAKEIQKMGKKFNFVRSKVDNDLLNAQRSQRDFDPEKTLSRIRENCEQGLLNAGVQAEVFLLSSFELQRYDFHRLHETLEREL
LEHQRDVLLVAMPNISLEIIEKKKEAFKSKIPYWALVSAAGALVPVPGLPVAVDLSLIAGVVQQYKTGFGLDTPSLQRLADSTGVHLTD
LTSVIRSPLVLNKINAQLIMKTLIQIETVAGSMAAEEGLRFIPIFGTMIAGNLSSVATYKALSNFLEMLTEDAQNVFKKALRCMNSSV*

>TN_GKS6

MADSSDIVGIKEALQNNNQALAAAKIKELLDNQSNTPLKIGITGESGSGKSSFVNAFRGVDHRDEKEAAPVGVVETTAEVKEYPHPDYP
NVILWDLPGISSTKFPADDYLSRVGFEKFDFFIIISDTHFRENDVKLAKEIQKMGKKFYFVRSKVDNDLLNAQRRQRDFDPEQTLSHIR
DDCKQGLLNAGLQAQVFVLSNFYPQRYDFHRLHETLERELPEHQRDVLLVAMPNISLEIIEKKKEAFKSKIPYWALVSAAGALVPVPGL
SVAVDLSLIAGLVQQYKTGFGLDTPSLQRLANTTGVQLTDLTSVIRSPLGLNNINAQLIMKTITQIATVASLMAAKEGLRFIPIFGTMI
AGTLSYAVTYKALLDFLEMLTEDAQNVFKKVLPA*

>TN_GKS8

MADTTKIKDALQNNNQALAAAKIKELLDKQSNTPLNIGITGESGSGKSSFVNAFRGVDHKDEKEAAPEGVVETTEDVKEYPHPDYPKVS
LWDLSGIGSISWKLGATKFDFFIIISDTRFRENDVKLAKEIQKMGKKFYFLRSKVDNDLLNAQRSQRDFDPEQTLSHIRENCEQGLLNA
GVQAQVFLLSSFELQRYDFHRLHETLERELPEHKKDVLLVAMPNISLEIIEKKKKAFKSKIPY*ALVSAAGAVVPVPGLSVAVDVGLIA
RVVQQYKTGFGLDRPSLQRLADTTGVQLTDLTSVIRSPLGLKNINAQFIMKTLIHSASVASLMVAEEGLRLIPIFWNNDSRDSFLCSYS
KCSARFPQDVD*RCSECG*KGSPCMNSSV*

TN_GKS9

MADSSDIAEIKEALKNNNQALAAAKIKELLDNQSNIPLNIGITGETGSGKSSFVNAFRGVDHKDEKEAAPTEVVETTADVKEYPHPDYP
NVSLWDLPGIGTTKFQADKYLKLVGFEKFDFFIIISDARFRENDVKLAKKIKGDGGKSSTSYAPRLTTTCLSAQRSKRDFDPEQTLSLI
RENCKQGLLNAGVQAQVFLLSNFEFKCYDSHRLHETLERELPEHKKDVLLVAMPNFSLEIIEKKKEAFKSKIPYWALVSAAVALVPVPG
LSVAVDVGLITGLVQQYKTSFGLDRPSLKRLADATGVRLTDLTSVIQSPLSLNNINAQLIIKALNQTATDAXX

TN_GKS10

MADSSDIAEIKEALQNNNQALAAAKIKELLDKQSNATLNIGITGDSGSGKSSFVNAFRGVDHRDEKEAARVGVVETTAEPKEYPHPDYP
NVSLWDLPGIGTFKFPAKDYLKHVGFEKFDFFIIISATRFSENDVSLQKRSRRWAKSSSFVRSKVDNDLLNEQRCQRDFDPEKTLSRIR
DDCRQGLLNAGVQAQVFLLSNFELQRYDFHRLHETLERELPKHKRDVLLVAMPNISLEIIEKKKEAFKSKILHWALVSAAGALVPVPGL
SVAVDVGLIVGLVQQYKTGFGLDRPSLQRLADTTGVQLTDLTSVIQSQLGLNNINAQLIMKTLIHSTTVAGLMAAEEGLRLIPIFGTII
AGTLSSVATYKALSDFLEMLTKDAQNVFKKALCRMNSSV*

TN_GKS11

XXXVNAFRGVDHRDEKEAAPVDVVETTAEVKEYLHPNYPNVSLWDLPGIGSTNFPADEYLKLVGFEKFDFFIIISDTRFSENDVKLAKE
IQKMGKKFYFVRSKVDSDLLNEQRSQRDFDPEQTLSLLRENCEQGLLNAGLQAQVFLLSNFDPQHYDFHRLHETLERELPEHQRDVLFV
AMPNISQEIIEKKKEAFKSKIPYWALVSAAGAVVPVPGLSVAVDLSLIAVVVQKYKTGFGLDRPSLQRLADTTGVQLTDLTSVIRSPLG
LNNINAQLILKILIRSAKVAGLMAAEEDLRFIPIFGTMIAGTLLQLLIKLCQISSRCSLKILRMCLKRLSAA*

>TN_GKS12

MADSSDIAEIKEALQNNNQALAAAKIKELLDNQSNIPLNIAITGESGSGKSSFVNAFRGVDHKDEKAAPTDVVETTAEVIGYPHPNYPN
VILWDLPGIGTTKFPVDDYLKLVGFEKFDFFIIISDARFRENDVKLAKKITEMGKKFYFIRSKVDSDLLNAQRSKRDFDPEKTLSLIRE
NCKRGASQRRCAGSSFSAVNFYLSAXX

>TN_GKS13

DIILFIQNVSCVFFATKILALCLRLIMADSSDIVGIKEALQNNNQALAAAKIKELLDKQSNATLNIAITGDSGSGKSSFVNAFRGVDHK
DKKEAAPVGVVETTAKVEKYLHPNYPNVILWDLPGIGTFKFPADKYLKVVGFEKFDFFIIISATRFSENDVNLAKEIQKMGKKFYFVRS
KVDXXX

>TR_irgf6

MVNVCVCYITVGLSVGMISRLSDFYIVTVGFALCVQVIMADSLDTTEIKEALQNNNQALAVDKIKKLLERAANTPLNIGITGESGSGKS
SFVNAFRGVDHQDNQAAPTGVVETTTEVRAYPHPSYPNVTLWDLPGIGTTRFPADQYLKHVGFERFDFFIIISATRFRENDVKLAKEIQ
KMGKKFYFVRSKVDNDLQNAQRSQRNFDAEQTLALIRENCKEGLLKEGVQAPQVFLLSNFELRRHDFHRLHATLERELPEHKRDALLFA
MPNMSLEIIEKKKEAFKSKIPHYAFVSAACAAVPVPGLSVAVDGALIAGVVQQYKTGFGLDGPSLQRLADSTGVPLEDLTSVVRSPLSL
NTIDKAFILKLLLQSAAVAGLMLAEEGLKFIPLFGTLVASTLSYKVTEKALLDFLHMLAEDAQNVFKRALCCMNSSV*

>TR_irgf5

MVNVCVCYITVGLSVGMISRLSDFYIVTVGFALCVQVIMADSLDTTEIKEALQNNNQALAVDKIKKLLEKRANTPLNIGITGESGSGKS
SFVNAFRGVDHRDNQAAPTGVVETTTEVRAYPHPSYPNVTLWDLPGIGTTRFPADQYLKHVGFERFDFFIIISATRFRENDVKLAKEIQ
KMGKKFYFVRSKVDNDLQNAQRSQRNFDAEQTLALIRENCKEGLLKEGVQAPQVFLLSNFELRRHDFHRLHATLERELPEHKRDALLVS
LANMSLEIIKKKKEAFKSKIPHYAFVSAACAAVPLPGLSAAVDADLIAGVVQQYKTGFGLGRPSLQRLVAITGVPLVDLTIISSPLTLD
NINTDLVLNLMSQSSAISSLTETRESYSFIPLFGIPVARKLSYEITERALHNFLDMLTEDAQDVYNRVINHINS*

>TR_GKS1

MDNVCVCYITVGLSVGMISRLSDFYIVTVGFALCVQVIMADSLDTTEIKEALQNNNQALAVDKIKKLLERAANTPLNIGITGESGSGKS
SFVNAFRGVDHRDNQAAPTGVVETTTEVRAYPHPSYPNVTLWDLPGIGTTRFPADQYLKHVGFERFDFFIIISATRFRENDVKLAKEIQ
KMGKKFYFVRSKVDNDLQNAQRSQRNFDAEQTLALIRENCKEGLLKEGVQAPQVFLLSNFELRHHDFHRLHATLERELPEHKRDALLSA
MPNMSLEIIEKKKEAFKSKIPHCAFVSAACAVVPVPGLSAAVDADLIAGVVQQYKTGFALDGPSLQRLADSTGVPLEDLTSVVRSPLSL
NTIDKTFILKLLLQSAAVAGLMVAEEGLKFIPLFGTLVASTLSYKVTEKALLDVLHMLAEDAQNVFKRALCCMNSSV*

>TR_GKS2

XSSSSSQILALEENDVKLAKEIQKMGKKFYFVRSKVDNDLQNAQRSQRNFDAEQTLALIRENCEEGLLKEDVQAPQVFLLSNFDFRRHD
FHRLHATLERELPEHKRDALLVSLANMSLEIIEKKKEAFKSKIPHYAFVSAACAAVPLPGLSAAVDADLIAGVVQQYKTGFGLDRPSLQ
RLVAITVVPLVDLTIISSPLTLDNINTDLVLNLMSQSSAISSLTERRESYSFIPLFGIPVARKLSYEITERALHNFLDMLTEGCSGCV*

>TR_GKS3

MDNVCVCYITVGLSVGMISRLSDFYIVTVGFALCVQVIMADSLDTTEIKEALQNNNQALAVDKIKKLLERAANTPLNIGITGESGSGKS
SFVNAFRGVDHRDNQAAPTGVVETTTEVRAYPHPSYPNVTLWDLPGIGTTRFPADQYLKHVGFERFDFFIIISATRFRENDVKLAKEIQ
KMGKKFYFVRSKVDNDLQNAQRSQRNFDAEQTLALIRENCKEGLLKEGVQAPQVFLLSNFELQRHDFHRLHATLERELPEHKRDALLFA

MPNMSLEIIEKKKEAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXLQRLADSTGVLLEDLTSVVRSPLSLNTIDKTFILKLLLQSAAVAGLMLAEEG
LKFIPLFGTLVASTLSYKVTEKALLDFLHMLAEDAQNVFKRALCCMNSSV*

>TR_GKS4

MADFLDTTEIKEALQNNNQALAVDKIKKLLEMAANTPLNIGITGESGSGKSSFVNAFRGVDHRDDQAAPTGVVETTTEVRAYPHPSYPN
VTLWDLPGIGTTRFPADQYLKHVGFERFDFFIIISDTRFRENDVKLAKEIQKMGKKFYFVRSKVDNDLQNAQRSQRNFDAEQTLALIRE
NCEEGLLKEGVQAPQVFLLSNFDFRRHEFHRLHATLERELPEHKRDALLVSLANMSLKIIKKKKEAFKSKIPHYAFVSAACAAVPLPGL
SAAVDADLIARVVQQYKTGFGLDRPSLQRLVAITGVPLVDLTIISSPLTLDNINTDLVLNLMSQSSAISSLTERRESYSFIPLFGILVA
RKLSHEITERALHNFLDMLTEDAQDVYNRVINHINS*

>TR_GKS5

MADSLDTTEIKEALQNNNQALAVDKIKKLLERAANTPLNIGITGESGSGKSSFVNAFRGVDHRDNQAAPTGVVETTTEVRAYPHPSYPN
VTLWDLPGIGTTRFPADQYLKHVGFERFDFFIIISDTRFRENDVKLAKEIQKMGKKFYFVRSKVDSDLQNAQRSQRNFDAEQTLALIRE
NCKEGLLKEGVQAPQSSCCPTLSFGTMTSIASMRPWRENFQNTSGMLFCPPCPT*

>TR_GKS7

MVNVCVCYITVGLSVGMISRLSDFYIVTVGFASCVQVIMADFLDTTEIKEALQNNNQALAVDKIKKLLEKAANTPLNIGITGESGSGKS
SFVNAFRGVDHQDDQAAPTGVVETTTEVKAYTHPSYPNFKLWDLPGIGTTRFPADQYLKHVGFERFDFFIIISATRFSENDVKLAKEIL
NMGKKFYFVRSKVDNDLQNAQRSQRNFDAEQTLALIRENCKKGLLKEGVQAPQVFLLSNFELRRHDFHRLHETLERELPEHKRDALLFA
MPNMSLEIIKKKKEAFKSKIPHYAFVSAACAVVPLPGLSVAVDLALIVSVVQKYKTGFGLDRPSLQRLADSTGVPLEDLTSVVHSPLSL
NTINKTLIQKLLLQSAAIVALMAAEEGLKFIPLFGALVAAPLSYKVTEKALLDFLHVLAEDAQNVFKRALCCMNSPV*

>TR_GKS8

MDNVCVCYITVGLSVGMISRLSDFYIVTVGFALCVQVIMADSLDTTEIKEALQNNNQALAVDKIKKLLERAANTPLNIGITGESGSGKS
SFVNAFRGVDHRDDQAAPIGVVETTTEVGAYPHPSYPNVTLWDLPGIGTTRLTTICRMHKGVRETLMQSRLSHLFVKTATKGLLKEGVQ
APQVFLLSNFELQRHDFHRLHATLERELPEHKRDALLSAMPNMSLEIIEKKKEAFKSKIPHCAFVSAACAVVPVPGLSAAVDADLIAGV
VQQYKTGFALDGPSLQRLADSTGVPLEDLTSVVRSPLSLNTIDKTFILKLLLQSAAVAGLMVAEEGLKFIPLFGTLVASTLSYKVTEKA
LLDFLHMLAEDAQNVFKRALCCMNSSV*

>AM_GKS1

FGTREDDRLKGDTGARHRAADDINVSKTEEAREQRWPKVEVHIALTGDSGAGKSSFINAIRDLREDDEGAATVDVTQCTKEPTAYDHPA
FPNAKFWDLPGIGTPSYPDMETYTKKVELEKYDAFLIFTATRFTENSLKLAVKIKSMKRKFIFIRTKIDNSARAESRKQSFDEQGMLTK
IRCKCVERLGDLLSCEEDVFLISSHHPNKWDFSRLTEAIFDAVIMLQEQESLTLRVLQDLVITS*

>LE_GKS1

MAQSMLPKYFSDTEMRSLQTGYSNGDVVSAMLRIKRVEDSGNVPINIAVLGDGGAGKSTFINTMRGVRSGDQGAAPVGGYEASVNPVGY
PYPSLPNVQLWDLPGSNSLGFEMSRYLKQVQFESYDFYIIVSQSRFRESDGELSKKIQQQGKCFYYIRSKIDNDAFSMQMQGTDFGEGQ
RQIRQDCLKHFHRVSVEPPAIFLISGLEVTGYDFPKLQSALASDLPKIKSTAFRLAIPRMMQEIQRPRRQILMWCIILWAFLSGALGVL
RLLTLPLLTATLCTVSGWIYLRRQLGV

>SqA_GKS1

MAGVLSHLFYSTTDVKHLAATYRHGGMAALQSEIEAKAIQFKHVKLHVAVMGEAGAGKSSFINALRGLGVNDKGAAPTDVVECTKEVTP
YLHPTLPNVTYLDFPGIGTERFPIKKFLKQTNFSQFDFGIIVSDARFTDNDA

>SqA_GKS2

EGRRTILRDCVSNFQSVGVTPPAIFLISSFDLDKYDFPDIRSTLVSNLPSIKSNVFLLSLPKIMLEIIEPKVRMLKKRVWLMAVLAGAL
GAVPVPGLSFITGIVLTVAGLIYLQKQVGLNDKSLQSLASQXQxKPTSALKTEMNRRLPSKIPPVFTRVLLGIPIVACMIAGVNHSFSP
LTLSIFGTVSVLL

>SqA_GKS3

FSSVNKDLEDDLPTIKKSVFILALPNLTLEIVEKKRLELRKRVWMLATLSGAVGAVPVPGVSLAADIGIVIGGIIHFRKCLGLDDASLQ
RLANKAGKPVEDLKAVVKTPLVGEITPDLIARLSWGLAAVTISALEIALDIIPVVGSIFGAGSSFLMTYKLLSDALDDLAENAQRVVKA
AFGTDGDGLHQTSIQ*VNAVANILKGSNPERCRNSTLQYSLFAKLFPGILLIH

>BF_GKS1

KELQEMLSNLSAATGAARIKILKNIQEYAAENTEAWKNQKVNLGILGDPGAGKSTFINSIRGLKPKAPGAAKVGLRHTTTVVTGYPHPA
RPDNLIFVDFPGVLLKKGTGQARDFDIQQYLDEFGEKMQQCHVFLVFTSGHIQHNAVQIGMEARKMGKKVLFVRSQFDLDVGKRKKDDP
DYFTGKTKADLMEELRQDYIHVLKEVGWEGQVDPKDVFIISGVLENVLEGSWDIPKFRKAMIGNLSALQKMVVISTCRDFSKATIKERG
DIYREYVWAVAMAATAGTFVPFAGAASVPGK*

>BF_GKS2

SIRREIRILEMNEEEQGMLRNLAAAEGANKAEILKKIQEYAAEKVEAWKRQKVYVGIVGDPGAGKSTFINSIRGLSPTDDGAAEVGLTH
TTTTTTVYPHPERPDNLIFVDFPGVLLKKGTEEERDFDIQQYLAKFKGKMQQCHVFLVFSSGRIQHNAVQIGMKAREMGKKVLFVRSKF
DSDLPNMRKDKPQYFKDKTEADLMEELCQDYIKVLKEVGWKGEVNPRDVFIISGHLDYVLKGSWDIPKFRKAMIGNLSALQKMVVISTC
RDFSKATIKERGDIYREYVWAVAMAATAGTFVPFAGAASIPGK*

>BF_GKS3

MLRNLADPGSANKAEILKKIQEYAAEKVEAWKRQKVYVGIVGEPGAGKSTFINSIRGLKPKAPGAAEVGLGHTTSVVTEYPHPARPDNL
IFVDFPGVLLKKGTGQARDFDIQQYLNEFGEKMQECHVFLVFSSGRIQHNAVQIGMKARDMEKKVLFVRSQFDKDLADKQNDDPEYFDD
KTEADLMEELRQDYIKVLKKVGWEGEVNPREVFIISGRLKNVLEGSWDIPKFRKAMIEGLDALQKMIVINTCRDFSKATIKERGDVYRG
QVWWVAAAATAGTLVPYVGGAAVPGEKCHSHSGREREKF*

>BF_GKS4

MNEEEQRMLKNLAAAGGANKAEILKEIQQYAAENMEAWKNQKVNIGIVGDPGAGKSTFINSIRGMKPKAPGAAGVGLTHTTTEAIDYPH
PERPNSLVFVDFPGVLLKKGTRGKRDFDIQQYLNEFGEKMQQCHVFLVFSSGRIQHNAVQIGMEAREMGKKVLFVRSQFDKDLEDKHND
DDEYFDDKTEADLMEELRHNYIKVLKDVGWEGEVNPRDVFIISGRLKHVMKGFWDIPTFRKTMIEELGTLHKMVVINTCRDYSTATIKE
RGDIYRSKVWTVAMAASTGSFVPYAGAAAIPGR*

>BF_GKS5

MDEEEQGMLRNLAAAGGANKAEILKKIQQYAAENMEAWKNQKVNIGIVGDPGAGKSTFINSIRGLKPTGLPVQRRWGLHTLPK*TTDYP
HPTRPDSLVFVDFPGVLLGRGSGDERDFDIQQYLAKFGEKMDQCHVFLVFSSGRIQHNAVQIGMKAREKGKKVLFVRSQFDKDLADKKN
DDPEYFDDKTEADLMEELRQDYINVLKRVGWEGEVNPREVFIISGRLRNVLEGSWDIPKFRKAMIDGLGALQKMVVINTCRDFSKATIK
ERGDIYRGKVWTVAMATTAGSFVPYAGAAAIPGR*

>BF_GKS6

MVCYIAVSCFQFDDLGPDERKQFRKFADKSVKDQTASEMLQALWEEVGQYIDQEAWMGQAYVRIGLVGLSGAGKSTFINSLRGLRPTDP
GAADVGVKETTTIPTEYPHPEHKHVILVDFPGNVFKLKEGLMTPVGFDTNEYIKKNGKKMEECNVFLVFTSGRVHDNAVWIAKKCKDMG
KKLLFVRSKFDSDVTNTQEDKPTYFKNGQDEGESHLLNEVKADFVSKLNEQGFGQVDEKDVFIIGGKYDKVQMGEWDTPLLKQAILKQL
DIQQQMLFITTCQDFSPTMARAKAQVYRSRAWKVALGTAPAGAIPFVGAGVTLGV*

>BF_GKS7

MIRDEDGDNTDFDKIAHRMTLKLQDSEHATVNIGIVGEAGAGKSTFINSFRGIKPGEEGAAEVSAFRHTTNDVTRYPVPDNQNIVLMDF
PGVIFRNTGTRLDMEEEFNTKSYLDLYGVEMEECDVFLVFVTCRVSNNIIWIAKEVGKMNKKVLFVRSKIDVDLANESRDNPRRFPEGT
SSTTIEVRRFVEELRKVTTNELERLSYAEVKETMVFVICGLPDDVASGTYDMTNLRKAIYNTLSPDKKGVLINGLVEFATDMVHEKAEY
LRSREVIAAAVANTVISATPIPGLGLALDIGTFFLLLCFLSFMSTAP*

>BF_GKS8

MLHERAADGGGNITAQEALDIAAASESEPDKVKIGIVGDAGAGKSTFINSFRGLSPDDVGAAKVSAFGHATTESESYDVPGKAVVLTDF
PGVLFKPKMEASSTDIDGTEKVIFNTSSYLDPNKAKMQECDFFLIFMPNRPGNNVVWIAKEVRKMGKRLLFVRSQADEDIERARHDNPK
DFPTNIDNQIAERRVMLKFKQSTKVTFEALGYGKVDEDDIFVICGLKEPVARGDYDMGALRIAMLNSLSTYKQGVLIKNIQDFSMATLM
KKGEVMRKIVWGVAAGAAAISAVPLPGIGVAMDMGKWLCIYTYMLDEVMLSSKLST*

>BF_GKS9

MGQAYVRIGIVGVSGAGKSTYINSFRGLRASDPGAAAVGVIETTTEAEEYRHPKHDHVILVDFPGALFKLEGGHRRSVTFDMKEYTRKF
EGKMKECNVLLVFTSVRVHDNAVWIAAKAREIGKKVLFVRSMLDVDIYNKQRDDPAYFTGGQEEGEKRLLQLHRQDYVTMLETMGYGRV
APEEVFIISGMLEHIQRGSWDAPALKEAMLKQLHIQQQILFITICQDFSPTMARAKAKIYKSRAWKVALRVAAAGVIPFAGGSINAGV*

>BF_GKS10

MRLKSKFTSVFFLLLNIASVLHVFCTFFLSHRRECPLEVDQLTISEKHLSQEELDELEMIRDEDGDNTDFEKIAHRMTLKLQDSEHATV
NIGIVGEAGAGKSTFINSFRGIKPGEEGAAEVSAFRHTTNEVTRYPVPDNQNIVLMDFPGVIFRNTGTRLDMEEEFNTKSYLDLYGSEM
EECDVFLVFVTCRVSNNIIWIAKEVGKMNKKVLFVRSKIDVDLANESRDNPRRFPEGTSSTTIEVRRFVEELRKVTTKELERLSYAEVK
ETMVFVICGLPDDVASGTYDMTNLRKAIYNTLSPDKKGVLINGLVEFATDMVHEKAEYLRSREVIVAAVANTVISATPIPGLGLALDIG
TLCSSLFFSFFMPTAP*

>BF_GKS11

LEEHMSQAEIDELKKACDGSDDFEDMAKRLTSTFQSAKYATVNIGIVGDAGAGKSTFINSFRGLRPKEEEAAEVSAIRHTTNKVTRYPV
PDNQNIVLLDFPGVIFRSTGTGVAEEFDTKSYLDLHGAEMKNCDLFLIFVTGRVSNNIIWIAKEARKMEKNVLFVRSKIDIDLANESRD
HPTRFPKGTSYTTLETQHFVEELRQATAEELKRLDYGEVKETRVFVISGVLECMAEGQYDMTNLRKTIYNSVSPDKKAVLITSLSDFAT
DMVHEKAKYLRTREAIVAAMCNAVISAAPVPGIGIVTDIGKVLNFSTFLTVACYLMILLLMIETY*

>BF_GKS12

MWVFFFLQYSVEIRRDDLGPEEQEIFRRFADQVQGVRGDLPPEEMLQMLRRELGTDDREDAWMGQAYVRIGIVGSSGAGKSTFINSFRG
LEASDPGAADVGTTETTRETAEYPHPEHDHVILVDFPGALFKLQEENWHPVTFNMEEYRVRFGGKMKECNVFLVFTSERVHDNAVWIAK
VAKDMGKKVLFVRSKFDRDLEDKQRDKRSYFAGGQKEGEERLLQEHREDYVTKLDTLGYGRVDIRDVFVISGILEHVQTGHWDAAALKE
AILKQLDIQQKMLLMTTTTDYSPTMVRVKAEIYRSRAWKVALTVAAGGLIPLAGTFINAGVYRILTDHLSAKVGRELESPWLRLDDFTF
QKYDKAAHFRRVLHCRYYGDDDGVVQERLWPLPDLGEKAG*

>BF_GKS13

MGQAYVRIGIVGSSGPGKSTFINSFRGLKAEDKGAAPVGTKETTKETAEYPHPEHDHVILVDFPGALFKLQGSDWQPVRFNMKEYTRKF
GDKMKECNVFLVFTSDRVHDNAVWIAAKAKEMKKKVLFVRSKFDRDLEDVRRDKPSFFAEKDGEERLLQEHRDDYVTKLGTLGYGRVDI
GDVFVISGILEHIEGGRWDAAALKEAILKQLDIQQKMLLMTTTTDYSPTMVRVKAEIYRSRAWKVALGVAAGGAIPWAGTFVNAGMYRI
RMDHLSKLVRQ*

>BF_GKS14

MFADKVKNANDQSSAEMIQTLHEEVGQHVDEEAWMGQAYVRIGIVGVTGAGKSTFINSFRGLEAGDPGAAAVDTIETTADTAEYPHPEH
DHVILVDFPGALFKLEDGGRRRSAKFDMKEYTRKFGEKMRECNVFLVFTSDRVHDNAVWIAAKAREIGKKVLFVRSKFDQDVANKRRDD
KTYFAGGQEEGEERLLQFHRQDYVTKLETMGYGRVAPEDVFIISGIVENIQRGRWDATAL

VI. References

Abascal, F., R. Zardoya, and D. Posada. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*. 21:2104-5.

Accola, M.A., B. Huang, A. Al Masri, and M.A. McNiven. 2002. The antiviral dynamin family member, MxA, tubulates lipids and localizes to the smooth endoplasmic reticulum. *J Biol Chem*. 277:21829-35.

Advani, A.S., H.K. Dressman, M. Quiroz, G.A. Taylor, and A.M. Pendergast. 2004. Elevated expression of a subset of interferon inducible genes in primary bone marrow cells expressing p185 Bcr-Abl versus p210 Bcr-Abl by DNA microarray analysis. *Leuk Res*. 28:285-94.

Aebi, M., J. Fah, N. Hurt, C.E. Samuel, D. Thomis, L. Bazzigher, J. Pavlovic, O. Haller, and P. Staeheli. 1989. cDNA structures and regulation of two interferon-induced human Mx proteins. *Mol Cell Biol*. 9:5062-72.

Ajzenberg, D., A.L. Banuls, M. Tibayrenc, and M.L. Darde. 2002. Microsatellite analysis of Toxoplasma gondii shows considerable polymorphism structured into two main clonal groups. *Int J Parasitol*. 32:27-38.

Alexander, D.L., J. Mital, G.E. Ward, P. Bradley, and J.C. Boothroyd. 2005. Identification of the moving junction complex of Toxoplasma gondii: a collaboration between distinct secretory organelles. *PLoS Pathog*. 1:e17.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J Mol Biol*. 215:403-10.

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25:3389-402.

Ames, J.B., T. Tanaka, L. Stryer, and M. Ikura. 1994. Secondary structure of myristoylated recoverin determined by three-dimensional heteronuclear NMR: implications for the calcium-myristoyl switch. *Biochemistry*. 33:10743-53.

Anderson, S.L., J.M. Carton, J. Lou, L. Xing, and B.Y. Rubin. 1999. Interferon-induced guanylate binding protein-1 (GBP-1) mediates an antiviral effect against vesicular stomatitis virus and encephalomyocarditis virus. *Virology*. 256:8-14.

Andrews, S.C., and J.R. Guest. 1988. Nucleotide sequence of the gene encoding the GMP reductase of Escherichia coli K12. *Biochem J*. 255:35-43.

Angata, T., E.H. Margulies, E.D. Green, and A. Varki. 2004. Large-scale sequencing of the CD33-related Siglec gene cluster in five mammalian species reveals rapid evolution by multiple mechanisms. *Proc Natl Acad Sci U S A*. 101:13251-6.

Araujo, F.G., D.M. Williams, F.C. Grumet, and J.S. Remington. 1976. Strain-dependent differences in murine susceptibility to toxoplasma. *Infect Immun*. 13:1528-30.

Arnheiter, H., and E. Meier. 1990. Mx proteins: antiviral proteins by chance or by necessity? *New Biol*. 2:851-7.

Atkinson, P.G., J.M. Blackwell, and C.H. Barton. 1997. Nramp1 locus encodes a 65 kDa interferon-gamma-inducible protein in murine macrophages. *Biochem J*. 325 (Pt 3):779-86.

Bachmann, B.J. 1983. Linkage map of Escherichia coli K-12, edition 7. *Microbiol Rev*. 47:180-230.

Bachmann, B.J. 1990. Linkage map of Escherichia coli K-12, edition 8. *Microbiol Rev*. 54:130-97.

Bafica, A., C.G. Feng, H.C. Santiago, J. Aliberti, A. Cheever, K.E. Thomas, G.A. Taylor, S.N. Vogel, and A. Sher. 2007. The IFN-Inducible GTPase LRG47 (Irgm1) Negatively Regulates TLR4-Triggered Proinflammatory Cytokine Production and Prevents Endotoxemia. *J Immunol*. 179:5514-22.

Bancroft, G.J. 1993. The role of natural killer cells in innate resistance to infection. *Curr Opin Immunol*. 5:503-10.

Barltrop, J.A., T.C. Owen, A.H. Cory, and J.G. Cory. 1991. 5-(3-carboxymethoxyphenyl)-2-(4,5-dimethylthiazolyl)-3-(4-sulfophenyl)tetrazolium, inner salt (MTS) and related analogs of 3-(4,5-dimethylthiazolyl)-2,5-diphenyltetrazolium bromide (MTT) reducing to purple water-soluble formazans As cell-viability indicators. *Bioorganic & Medicinal Chemistry Letters*. 1:611-614.

Barten, R., M. Torkar, A. Haude, J. Trowsdale, and M.J. Wilson. 2001. Divergent and convergent evolution of NK-cell receptors. *Trends Immunol*. 22:52-7.

Barton, C.H., T.E. Biggs, S.T. Baker, H. Bowen, and P.G. Atkinson. 1999. Nramp1: a link between intracellular iron transport and innate resistance to intracellular pathogens. *J Leukoc Biol*. 66:757-62.

Bazer, F.W., T.E. Spencer, and T.L. Ott. 1997. Interferon tau: a novel pregnancy recognition signal. *Am J Reprod Immunol*. 37:412-20.

Bekpen, C. 2005a. Evolutionary and functional studies on p47 GTPases involved in cell autonomous immunity. PhD thesis. Institute for Genetics, University of Cologne, Cologne.

Bekpen, C., J.P. Hunn, C. Rohde, I. Parvanova, L. Guethlein, D.M. Dunn, E. Glowalla, M. Leptin, and J.C. Howard. 2005b. The interferon-inducible p47 (IRG) GTPases in vertebrates: loss of the cell autonomous resistance mechanism in the human lineage. *Genome Biol*. 6:R92.

Bernstein-Hanley, I., J. Coers, Z.R. Balsara, G.A. Taylor, M.N. Starnbach, and W.F. Dietrich. 2006. The p47 GTPases Igtp and Irgb10 map to the Chlamydia trachomatis susceptibility locus Ctrq-3 and mediate cellular resistance in mice. *Proc Natl Acad Sci U S A*.

Bertani, G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *J Bacteriol*. 62:293-300.

Beutler, B. 2004. Innate immunity: an overview. *Mol Immunol*. 40:845-59.

Binder, E.M., and K. Kim. 2004. Location, location, location: trafficking and function of secreted proteases of Toxoplasma and Plasmodium. *Traffic*. 5:914-24.

Bininda-Emonds, O.R., M. Cardillo, K.E. Jones, R.D. MacPhee, R.M. Beck, R. Grenyer, S.A. Price, R.A. Vos, J.L. Gittleman, and A. Purvis. 2007. The delayed rise of present-day mammals. *Nature*. 446:507-12.

Binns, D.D., B. Barylko, N. Grichine, M.A. Atkinson, M.K. Helms, D.M. Jameson, J.F. Eccleston, and J.P. Albanesi. 1999. Correlation between self-association modes and GTPase activation of dynamin. *J Protein Chem*. 18:277-90.

Black, M.W., and J.C. Boothroyd. 2000. Lytic cycle of Toxoplasma gondii. *Microbiol Mol Biol Rev*. 64:607-23.

Blackston, C.R., J.P. Dubey, E. Dotson, C. Su, P. Thulliez, D. Sibley, and T. Lehmann. 2001. High-resolution typing of Toxoplasma gondii using microsatellite loci. *J Parasitol*. 87:1472-5.

Bodaghi, B., O. Goureau, D. Zipeto, L. Laurent, J.L. Virelizier, and S. Michelson. 1999. Role of IFN-gamma-induced indoleamine 2,3 dioxygenase and inducible nitric oxide synthase in the replication of human cytomegalovirus in retinal pigment epithelial cells. *J Immunol*. 162:957-64.

Boehm, U., L. Guethlein, T. Klamp, K. Ozbek, A. Schaub, A. Futterer, K. Pfeffer, and J.C. Howard. 1998. Two families of GTPases dominate the complex cellular response to IFN-gamma. *J Immunol*. 161:6715-23.

Boehm, U., T. Klamp, M. Groot, and J.C. Howard. 1997. Cellular responses to interferon-gamma. *Annu Rev Immunol*. 15:749-95.

Bogdan, C. 2001. Nitric oxide and the immune response. *Nat Immunol*. 2:907-16.

Bonhomme, A., G.T. Maine, A. Beorchia, H. Burlet, D. Aubert, I. Villena, J. Hunt, L. Chovan, L. Howard, S. Brojanac, M. Sheu, J. Tyner, M. Pluot, and J.M. Pinon. 1998. Quantitative immunolocalization of a P29 protein (GRA7), a new antigen of toxoplasma gondii. *J Histochem Cytochem*. 46:1411-22.

Boothroyd, J.C., and M.E. Grigg. 2002. Population biology of Toxoplasma gondii and its relevance to human infection: do different strains cause different disease? *Curr Opin Microbiol*. 5:438-42.

Bordier, C. 1981. Phase separation of integral membrane proteins in Triton X-114 solution. *J Biol Chem*. 256:1604-7.

Borghans, J.A., J.B. Beltman, and R.J. De Boer. 2004. MHC polymorphism under host-pathogen coevolution. *Immunogenetics*. 55:732-9.

Bourne, H.R., D.A. Sanders, and F. McCormick. 1990. The GTPase superfamily: a conserved switch for diverse cell functions. *Nature*. 348:125-32.

Bourne, H.R., D.A. Sanders, and F. McCormick. 1991. The GTPase superfamily: conserved structure and molecular mechanism. *Nature*. 349:117-27.

Bradley, D.J., B.A. Taylor, J. Blackwell, E.P. Evans, and J. Freeman. 1979. Regulation of Leishmania populations within the host. III. Mapping of the locus controlling susceptibility to visceral leishmaniasis in the mouse. *Clin Exp Immunol*. 37:7-14.

Bradley, P.J., C. Ward, S.J. Cheng, D.L. Alexander, S. Coller, G.H. Coombs, J.D. Dunn, D.J. Ferguson, S.J. Sanderson, J.M. Wastling, and J.C. Boothroyd. 2005. Proteomic analysis of rhoptry organelles reveals many novel constituents for host-parasite interactions in Toxoplasma gondii. *J Biol Chem*. 280:34245-58.

Burnet, F.M. 1959. The Clonal Selection Theory of Acquired Immunity. Cambridge University Press, Cambridge, UK.

Butcher, B.A., R.I. Greene, S.C. Henry, K.L. Annecharico, J.B. Weinberg, E.Y. Denkers, A. Sher, and G.A. Taylor. 2005. p47 GTPases regulate Toxoplasma gondii survival in activated macrophages. *Infect Immun*. 73:3278-86.

Cao, H., F. Garcia, and M.A. McNiven. 1998. Differential distribution of dynamin isoforms in mammalian cells. *Mol Biol Cell*. 9:2595-609.

Carey, K.L., A.M. Jongco, K. Kim, and G.E. Ward. 2004. The Toxoplasma gondii rhoptry protein ROP4 is secreted into the parasitophorous vacuole and becomes phosphorylated in infected cells. *Eukaryot Cell*. 3:1320-30.

Carlow, D.A., J. Marth, I. Clark-Lewis, and H.S. Teh. 1995. Isolation of a gene encoding a developmentally regulated T cell-specific protein with a guanine nucleotide triphosphate-binding motif. *J Immunol*. 154:1724-34.

Carlow, D.A., S.J. Teh, and H.S. Teh. 1998. Specific antiviral activity demonstrated by TGTP, a member of a new family of interferon-induced GTPases. *J Immunol*. 161:2348-55.

Carruthers, V.B., O.K. Giddings, and L.D. Sibley. 1999. Secretion of micronemal proteins is associated with toxoplasma invasion of host cells. *Cell Microbiol*. 1:225-35.

Carruthers, V.B., and L.D. Sibley. 1997. Sequential protein secretion from three distinct organelles of Toxoplasma gondii accompanies invasion of human fibroblasts. *Eur J Cell Biol*. 73:114-23.

Carter, C.C., V.Y. Gorbacheva, and D.J. Vestal. 2005. Inhibition of VSV and EMCV replication by the interferon-induced GTPase, mGBP-2: differential requirement for wild-type GTP binding domain. *Arch Virol*. 150:1213-20.

Casey, P.J. 1995. Protein lipidation in cell signaling. *Science*. 268:221-5.

Causier, B., and B. Davies. 2002. Analysing protein-protein interactions with the yeast two-hybrid system. *Plant Mol Biol*. 50:855-70.

Chebath, J., P. Benech, M. Revel, and M. Vigneron. 1987. Constitutive expression of (2'-5') oligo A synthetase confers resistance to picornavirus infection. *Nature*. 330:587-8.

Chebath, J., G. Merlin, R. Metz, P. Benech, and M. Revel. 1983. Interferon-induced 56,000 Mr protein and its mRNA in human cells: molecular cloning and partial sequence of the cDNA. *Nucleic Acids Res*. 11:1213-26.

Chee, A.V., P. Lopez, P.P. Pandolfi, and B. Roizman. 2003. Promyelocytic leukemia protein mediates interferon-based anti-herpes simplex virus 1 effects. *J Virol*. 77:7101-5.

Chelbi-Alix, M.K., F. Quignon, L. Pelicano, M.H. Koken, and H. de The. 1998. Resistance to virus infection conferred by the interferon-induced promyelocytic leukemia protein. *J Virol*. 72:1043-51.

Chen, S.Y., S.Y. Huff, C.C. Lai, C.J. Der, and S. Powers. 1994. Ras-15A protein shares highly similar dominant-negative biological properties with Ras-17N and forms a stable, guanine-nucleotide resistant complex with CDC25 exchange factor. *Oncogene*. 9:2691-8.

Cheng, Y.S., M.F. Becker-Manley, T.P. Chow, and D.C. Horan. 1985. Affinity purification of an interferon-induced human guanylate-binding protein and its characterization. *J Biol Chem*. 260:15834-9.

Cheng, Y.S., M.F. Becker-Manley, T.D. Nguyen, W.F. DeGrado, and G.J. Jonak. 1986. Nonidentical induction of the guanylate binding protein and the 56K protein by type I and type II interferons. *J Interferon Res*. 6:417-27.

Cheng, Y.S., R.J. Colonno, and F.H. Yin. 1983. Interferon induction of fibroblast proteins with guanylate binding activity. *J Biol Chem*. 258:7746-50.

Cheng, Y.S., C.E. Patterson, and P. Staeheli. 1991. Interferon-induced guanylate-binding proteins lack an N(T)KXD consensus motif and bind GMP in addition to GDP and GTP. *Mol Cell Biol*. 11:4717-25.

Chieux, V., W. Chehadeh, J. Harvey, O. Haller, P. Wattre, and D. Hober. 2001. Inhibition of coxsackievirus B4 replication in stably transfected cells expressing human MxA protein. *Virology*. 283:84-92.

Clarke, S. 1992. Protein isoprenylation and methylation at carboxyl-terminal cysteine residues. *Annu Rev Biochem*. 61:355-86.

Clarke, S., J.P. Vogel, R.J. Deschenes, and J. Stock. 1988. Posttranslational modification of the Ha-ras oncogene protein: evidence for a third class of protein carboxyl methyltransferases. *Proc Natl Acad Sci U S A*. 85:4643-7.

Collart, F.R., and E. Huberman. 1988. Cloning and sequence analysis of the human and Chinese hamster inosine-5'-monophosphate dehydrogenase cDNAs. *J Biol Chem*. 263:15769-72.

Collazo, C.M., G.S. Yap, S. Hieny, P. Caspar, C.G. Feng, G.A. Taylor, and A. Sher. 2002. The function of gamma interferon-inducible GTP-binding protein IGTP in host resistance to Toxoplasma gondii is Stat1 dependent and requires expression in both hematopoietic and nonhematopoietic cellular compartments. *Infect Immun*. 70:6933-9.

Collazo, C.M., G.S. Yap, G.D. Sempowski, K.C. Lusby, L. Tessarollo, G.F. Woude, A. Sher, and G.A. Taylor. 2001. Inactivation of LRG-47 and IRG-47 reveals a family of interferon gamma-inducible genes with essential, pathogen-specific roles in resistance to infection. *J Exp Med*. 194:181-8.

Combeau, C., and M.F. Carlier. 1988. Probing the mechanism of ATP hydrolysis on F-actin using vanadate and the structural analogs of phosphate BeF-3 and A1F-4. *J Biol Chem*. 263:17429-36.

Conklin, D.C., F.J. Grant, M.W. Rixon, and W. Kindsvogel. 2002. Interferon-epsilon. U.S.Patent 6329175.

Cooper, M.D., and M.N. Alder. 2006. The evolution of adaptive immune systems. *Cell*. 124:815-22.

Cory, A.H., T.C. Owen, J.A. Barltrop, and J.G. Cory. 1991. Use of an aqueous soluble tetrazolium/formazan assay for cell growth assays in culture. *Cancer Commun*. 3:207-12.

Costa, R.H., V.V. Kalinichenko, A.X. Holterman, and X. Wang. 2003. Transcription factors in liver development, differentiation, and regeneration. *Hepatology*. 38:1331-47.

da Silva, A.M., and C. Klein. 1990. A rapid posttranslational myristylation of a 68-kD protein in D. discoideum. *J Cell Biol*. 111:401-7.

Dalton, D.K., S. Pitts-Meek, S. Keshav, I.S. Figari, A. Bradley, and T.A. Stewart. 1993. Multiple defects of immune cell function in mice with disrupted interferon-gamma genes. *Science*. 259:1739-42.

Darnell, J.E., Jr., I.M. Kerr, and G.R. Stark. 1994. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*. 264:1415-21.

Daumke, O., M. Weyand, P.P. Chakrabarti, I.R. Vetter, and A. Wittinghofer. 2004. The GTPase-activating protein Rap1GAP uses a catalytic asparagine. *Nature*. 429:197-201.

de St. Groth, F., R.G. Webster, and A. Datyner. 1963. Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. *Biochem. Biophys. Acta*. 71:377–391.

De Zoysa, M., H.S. Kang, Y.B. Song, Y. Jee, Y.D. Lee, and J. Lee. 2007. First report of invertebrate Mx: cloning, characterization and expression analysis of Mx cDNA in disk abalone (Haliotis discus discus). *Fish Shellfish Immunol*. 23:86-96.

Degols, G., P. Eldin, and N. Mechti. 2007. ISG20, an actor of the innate immune response. *Biochimie*. 89:831-5.

Degrandi, D., C. Konermann, C. Beuter-Gunia, A. Kresse, J. Wurthner, S. Kurig, S. Beer, and K. Pfeffer. 2007. Extensive Characterization of IFN-Induced GTPases mGBP1 to mGBP10 Involved in Host Defense. *J Immunol*. 179:7729-40.

Delarbre, C., C. Jaulin, P. Kourilsky, and G. Gachelin. 1992. Evolution of the major histocompatibility complex: a hundred-fold amplification of MHC class I genes in the African pigmy mouse Nannomys setulosus. *Immunogenetics*. 37:29-38.

Der, S.D., A. Zhou, B.R. Williams, and R.H. Silverman. 1998. Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. *Proc Natl Acad Sci U S A*. 95:15623-8.

Deretic, V., S. Singh, S. Master, J. Harris, E. Roberts, G. Kyei, A. Davis, S. de Haro, J. Naylor, H.H. Lee, and I. Vergne. 2006. Mycobacterium tuberculosis inhibition of phagolysosome biogenesis and autophagy as a host defence mechanism. *Cell Microbiol*. 8:719-27.

Dever, T.E., M.J. Glynias, and W.C. Merrick. 1987. GTP-binding domain: three consensus sequence elements with distinct spacing. *Proc Natl Acad Sci U S A*. 84:1814-8.

Di Paolo, C., H.P. Hefti, M. Meli, H. Landis, and J. Pavlovic. 1999. Intramolecular backfolding of the carboxyl-terminal end of MxA protein is a prerequisite for its oligomerization. *J Biol Chem*. 274:32071-8.

Dobrowolski, J.M., and L.D. Sibley. 1996. Toxoplasma invasion of mammalian cells is powered by the actin cytoskeleton of the parasite. *Cell*. 84:933-9.

Doffinger, R., S. Dupuis, C. Picard, C. Fieschi, J. Feinberg, G. Barcenas-Morales, and J.L. Casanova. 2002. Inherited disorders of IL-12- and IFNgamma-mediated immunity: a molecular genetics update. *Mol Immunol*. 38:903-9.

Dreiding, P., P. Staeheli, and O. Haller. 1985. Interferon-induced protein Mx accumulates in nuclei of mouse cells expressing resistance to influenza viruses. *Virology*. 140:192-6.

Dubremetz, J.F., A. Achbarou, D. Bermudes, and K.A. Joiner. 1993. Kinetics and pattern of organelle exocytosis during Toxoplasma gondii/host-cell interaction. *Parasitol Res*. 79:402-8.

Dufner, A., and T.W. Mak. 2006. CARD tricks: controlling the interactions of CARD6 with RICK and microtubules. *Cell Cycle*. 5:797-800.

Durfee, T., K. Becherer, P.L. Chen, S.H. Yeh, Y. Yang, A.E. Kilburn, W.H. Lee, and S.J. Elledge. 1993. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev*. 7:555-69.

Egea, P.F., S.O. Shan, J. Napetschnig, D.F. Savage, P. Walter, and R.M. Stroud. 2004. Substrate twinning activates the signal recognition particle and its receptor. *Nature*. 427:215-21.

Ehrt, S., D. Schnappinger, S. Bekiranov, J. Drenkow, S. Shi, T.R. Gingeras, T. Gaasterland, G. Schoolnik, and C. Nathan. 2001. Reprogramming of the macrophage transcriptome in response to interferon-gamma and Mycobacterium tuberculosis: signaling roles of nitric oxide synthase-2 and phagocyte oxidase. *J Exp Med*. 194:1123-40.

El Hajj, H., E. Demey, J. Poncet, M. Lebrun, B. Wu, N. Galeotti, M.N. Fourmaux, O. Mercereau-Puijalon, H. Vial, G. Labesse, and J.F. Dubremetz. 2006. The ROP2 family of Toxoplasma gondii rhoptry proteins: proteomic and genomic characterization and molecular modeling. *Proteomics*. 6:5773-84.

Engelhardt, O.G., H. Sirma, P.P. Pandolfi, and O. Haller. 2004. Mx1 GTPase accumulates in distinct nuclear domains and inhibits influenza A virus in cells that lack promyelocytic leukaemia protein nuclear bodies. *J Gen Virol*. 85:2315-26.

Engelhardt, O.G., E. Ullrich, G. Kochs, and O. Haller. 2001. Interferon-induced antiviral Mx1 GTPase is associated with components of the SUMO-1 system and promyelocytic leukemia protein nuclear bodies. *Exp Cell Res*. 271:286-95.

Espert, L., G. Degols, C. Gongora, D. Blondel, B.R. Williams, R.H. Silverman, and N. Mechti. 2003. ISG20, a new interferon-induced RNase specific for single-stranded RNA, defines an alternative antiviral pathway against RNA genomic viruses. *J Biol Chem*. 278:16151-8.

Estojak, J., R. Brent, and E.A. Golemis. 1995. Correlation of two-hybrid affinity data with in vitro measurements. *Mol Cell Biol*. 15:5820-9.

Everett, R.D., and M.K. Chelbi-Alix. 2007. PML and PML nuclear bodies: implications in antiviral defence. *Biochimie*. 89:819-30.

Farazi, T.A., G. Waksman, and J.I. Gordon. 2001. The biology and enzymology of protein N-myristoylation. *J Biol Chem*. 276:39501-4.

Fazaeli, A., P.E. Carter, M.L. Darde, and T.H. Pennington. 2000. Molecular typing of Toxoplasma gondii strains by GRA6 gene sequence analysis. *Int J Parasitol*. 30:637-42.

Feig, L.A., and G.M. Cooper. 1988. Inhibition of NIH 3T3 cell proliferation by a mutant ras protein with preferential affinity for GDP. *Mol Cell Biol*. 8:3235-43.

Fellenberg, F., T.B. Hartmann, R. Dummer, D. Usener, D. Schadendorf, and S. Eichmuller. 2004. GBP-5 splicing variants: New guanylate-binding proteins with tumor-associated expression and antigenicity. *J Invest Dermatol*. 122:1510-7.

Feng, C.G., C.M. Collazo-Custodio, M. Eckhaus, S. Hieny, Y. Belkaid, K. Elkins, D. Jankovic, G.A. Taylor, and A. Sher. 2004. Mice deficient in LRG-47 display increased susceptibility to mycobacterial infection associated with the induction of lymphopenia. *J Immunol*. 172:1163-8.

Feuerstein, J., H.R. Kalbitzer, J. John, R.S. Goody, and A. Wittinghofer. 1987. Characterisation of the metal-ion-GDP complex at the active sites of transforming and nontransforming p21 proteins by observation of the 17O-Mn superhyperfine coupling and by kinetic methods. *Eur J Biochem*. 162:49-55.

Fischer, H.G., S. Stachelhaus, M. Sahm, H.E. Meyer, and G. Reichmann. 1998. GRA7, an excretory 29 kDa Toxoplasma gondii dense granule antigen released by infected host cells. *Mol Biochem Parasitol*. 91:251-62.

Flajnik, M.F., and M. Kasahara. 2001. Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. *Immunity*. 15:351-62.

Focia, P.J., I.V. Shepotinovskaya, J.A. Seidler, and D.M. Freymann. 2004. Heterodimeric GTPase core of the SRP targeting complex. *Science*. 303:373-7.

Ford, B., K. Skowronek, S. Boykevisch, D. Bar-Sagi, and N. Nassar. 2005. Structure of the G60A mutant of Ras: implications for the dominant negative effect. *J Biol Chem*. 280:25697-705.

Frech, M., I. Schlichting, A. Wittinghofer, and P. Chardin. 1990. Guanine nucleotide binding properties of the mammalian RalA protein produced in Escherichia coli. *J Biol Chem*. 265:6353-9.

Galtier, N., M. Gouy, and C. Gautier. 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci*. 12:543-8.

Gao, G., X. Guo, and S.P. Goff. 2002. Inhibition of retroviral RNA production by ZAP, a CCCH-type zinc finger protein. *Science*. 297:1703-6.

Garcia-Mata, R., Z. Bebok, E.J. Sorscher, and E.S. Sztul. 1999. Characterization and dynamics of aggresome formation by a cytosolic GFP-chimera. *J Cell Biol*. 146:1239-54.

Garcia-Mata, R., Y.S. Gao, and E. Sztul. 2002. Hassles with taking out the garbage: aggravating aggresomes. *Traffic*. 3:388-96.

Garcia, M.A., E.F. Meurs, and M. Esteban. 2007. The dsRNA protein kinase PKR: virus and cell control. *Biochimie*. 89:799-811.

Gavrilescu, L.C., B.A. Butcher, L. Del Rio, G.A. Taylor, and E.Y. Denkers. 2004. STAT1 is essential for antimicrobial effector function but dispensable for gamma interferon production during Toxoplasma gondii infection. *Infect Immun*. 72:1257-64.

Gessani, S., and F. Belardelli. 1998. IFN-gamma expression in macrophages and its possible biological significance. *Cytokine Growth Factor Rev*. 9:117-23.

Ghosh, A., G.J. Praefcke, L. Renault, A. Wittinghofer, and C. Herrmann. 2006. How guanylate-binding proteins achieve assembly-stimulated processive cleavage of GTP to GMP. *Nature*. 440:101-4.

Ghosh, A., R. Uthaiah, J. Howard, C. Herrmann, and E. Wolf. 2004. Crystal Structure of IIGP1; A Paradigm for Interferon-Inducible p47 Resistance GTPases. *Mol Cell*. 15:727-39.

Gietz, R.D., R.H. Schiestl, A.R. Willems, and R.A. Woods. 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*. 11:355-60.

Gilbert, L.A., S. Ravindran, J.M. Turetzky, J.C. Boothroyd, and P.J. Bradley. 2007. Toxoplasma gondii targets a protein phosphatase 2C to the nuclei of infected host cells. *Eukaryot Cell*. 6:73-83.

Gill, S.C., and P.H. von Hippel. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem*. 182:319-26.

Gilly, M., M.A. Damore, and R. Wall. 1996. A promoter ISRE and dual 5' YY1 motifs control IFN-gamma induction of the IRG-47 G-protein gene. *Gene*. 179:237-44.

Gilly, M., and R. Wall. 1992. The IRG-47 gene is IFN-gamma induced in B cells and encodes a protein with GTP-binding motifs. *J Immunol*. 148:3275-81.

Giniger, E., S.M. Varnum, and M. Ptashne. 1985. Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell*. 40:767-74.

Goldman, M., R.K. Carver, and A.J. Sulzer. 1958. Reproduction of Toxoplasma gondii by internal budding. *J Parasitol*. 44:161-71.

Gorbacheva, V.Y., D. Lindner, G.C. Sen, and D.J. Vestal. 2002. The interferon (IFN)-induced GTPase, mGBP-2. Role in IFN-gamma-induced murine fibroblast proliferation. *J Biol Chem*. 277:6080-7.

Graham, F.L., and A.J. van der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology*. 52:456-67.

Grigg, M.E., S. Bonnefoy, A.B. Hehl, Y. Suzuki, and J.C. Boothroyd. 2001. Success and virulence in Toxoplasma as the result of sexual recombination between two distinct ancestries. *Science*. 294:161-5.

Gros, P., E. Skamene, and A. Forget. 1981. Genetic control of natural resistance to Mycobacterium bovis (BCG) in mice. *J Immunol*. 127:2417-21.

GSC. Genome Sequencing Centre.

Gubbels, M.J., C. Li, and B. Striepen. 2003. High-throughput growth assay for Toxoplasma gondii using yellow fluorescent protein. *Antimicrob Agents Chemother*. 47:309-16.

Guenzi, E., K. Topolt, E. Cornali, C. Lubeseder-Martellato, A. Jorg, K. Matzen, C. Zietz, E. Kremmer, F. Nappi, M. Schwemmle, C. Hohenadl, G. Barillari, E. Tschachler, P. Monini, B. Ensoli, and M. Sturzl. 2001. The helical domain of GBP-1 mediates the inhibition of endothelial cell proliferation by inflammatory cytokines. *Embo J*. 20:5568-77.

Guenzi, E., K. Topolt, C. Lubeseder-Martellato, A. Jorg, E. Naschberger, R. Benelli, A. Albini, and M. Sturzl. 2003. The guanylate binding protein-1 GTPase controls the invasive and angiogenic capability of endothelial cells through inhibition of MMP-1 expression. *Embo J*. 22:3772-82.

Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*. 52:696-704.

Guo, X., J. Ma, J. Sun, and G. Gao. 2007. The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. *Proc Natl Acad Sci U S A*. 104:151-6.

Guo, Z.G., U. Gross, and A.M. Johnson. 1997. Toxoplasma gondii virulence markers identified by random amplified polymorphic DNA polymerase chain reaction. *Parasitol Res*. 83:458-63.

Gutierrez, M.G., S.S. Master, S.B. Singh, G.A. Taylor, M.I. Colombo, and V. Deretic. 2004. Autophagy is a defense mechanism inhibiting BCG and Mycobacterium tuberculosis survival in infected macrophages. *Cell*. 119:753-66.

Hackam, D.J., O.D. Rotstein, W. Zhang, S. Gruenheid, P. Gros, and S. Grinstein. 1998. Host resistance to intracellular infection: mutation of natural resistance-associated macrophage protein 1 (Nramp1) impairs phagosomal acidification. *J Exp Med*. 188:351-64.

Hakansson, S., A.J. Charron, and L.D. Sibley. 2001. Toxoplasma evacuoles: a two-step process of secretion and fusion forms the parasitophorous vacuole. *Embo J*. 20:3132-44.

Hall, A., and A.J. Self. 1986. The effect of Mg2+ on the guanine nucleotide exchange rate of p21N-ras. *J Biol Chem*. 261:10963-5.

Haller, O., M. Acklin, and P. Staeheli. 1987. Influenza virus resistance of wild mice: wild-type and mutant Mx alleles occur at comparable frequencies. *J Interferon Res*. 7:647-56.

Haller, O., H. Arnheiter, I. Gresser, and J. Lindenmann. 1979. Genetically determined, interferon-dependent resistance to influenza virus in mice. *J Exp Med*. 149:601-12.

Haller, O., H. Arnheiter, M.A. Horisberger, I. Gresser, and J. Lindenmann. 1980. Interaction between interferon and host genes in antiviral defense. *Ann N Y Acad Sci*. 350:558-65.

Haller, O., M. Frese, and G. Kochs. 1998. Mx proteins: mediators of innate resistance to RNA viruses. *Rev Sci Tech*. 17:220-30.

Haller, O., and G. Kochs. 2002. Interferon-induced mx proteins: dynamin-like GTPases with antiviral activity. *Traffic*. 3:710-7.

Haller, O., G. Kochs, and F. Weber. 2007a. Interferon, Mx, and viral countermeasures. *Cytokine Growth Factor Rev*. 18:425-33.

Haller, O., P. Staeheli, and G. Kochs. 2007b. Interferon-induced Mx proteins in antiviral host defense. *Biochimie*. 89:812-8.

Halonen, S.K., G.A. Taylor, and L.M. Weiss. 2001. Gamma interferon-induced inhibition of Toxoplasma gondii in astrocytes is mediated by IGTP. *Infect Immun*. 69:5573-6.

Han, B.H., D.J. Park, R.W. Lim, J.H. Im, and H.D. Kim. 1998. Cloning, expression, and characterization of a novel guanylate-binding protein, GBP3 in murine erythroid progenitor cells. *Biochim Biophys Acta*. 1384:373-86.

Harlow, E., and D. Lane. 1988. Antibodies. A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA. 726 pp.

Harris, R.S., K.N. Bishop, A.M. Sheehy, H.M. Craig, S.K. Petersen-Mahrt, I.N. Watt, M.S. Neuberger, and M.H. Malim. 2003. DNA deamination mediates innate immunity to retroviral infection. *Cell*. 113:803-9.

Harris, R.S., and M.T. Liddament. 2004. Retroviral restriction by APOBEC proteins. *Nat Rev Immunol*. 4:868-77.

Hauptmann, R., and P. Swetly. 1985. A novel class of human type I interferons. *Nucleic Acids Res*. 13:4739-49.

Hefti, H.P., M. Frese, H. Landis, C. Di Paolo, A. Aguzzi, O. Haller, and J. Pavlovic. 1999. Human MxA protein protects mice lacking a functional alpha/beta interferon system against La crosse virus and other lethal viral infections. *J Virol*. 73:6984-91.

Henriquez, F.L., M.B. Nickdel, R. McLeod, R.E. Lyons, K. Lyons, J.F. Dubremetz, M.E. Grigg, B.U. Samuel, and C.W. Roberts. 2005. Toxoplasma gondii dense granule protein 3 (GRA3) is a type I transmembrane protein that possesses a cytoplasmic dilysine (KKXX) endoplasmic reticulum (ER) retrieval motif. *Parasitology*. 131:169-79.

Herrmann, C., and N. Nassar. 1996. Ras and its effectors. *Prog Biophys Mol Biol*. 66:1-41.

Higgins, D.G. 1994. CLUSTAL V: multiple alignment of DNA and protein sequences. *Methods Mol Biol*. 25:307-18.

Hillig, R.C., M. Hanzal-Bayer, M. Linari, J. Becker, A. Wittinghofer, and L. Renault. 2000. Structural and biochemical properties show ARL3-GDP as a distinct GTP binding protein. *Structure*. 8:1239-45.

Hinshaw, J.E. 2000. Dynamin and its role in membrane fission. *Annu Rev Cell Dev Biol*. 16:483-519.

Hinshaw, J.E., and S.L. Schmid. 1995. Dynamin self-assembles into rings suggesting a mechanism for coated vesicle budding. *Nature*. 374:190-2.

Hitziger, N., I. Dellacasa, B. Albiger, and A. Barragan. 2005. Dissemination of Toxoplasma gondii to immunoprivileged organs and role of Toll/interleukin-1 receptor signalling for host resistance assessed by in vivo bioluminescence imaging. *Cell Microbiol*. 7:837-48.

Holzinger, D., C. Jorns, S. Stertz, S. Boisson-Dupuis, R. Thimme, M. Weidmann, J.L. Casanova, O. Haller, and G. Kochs. 2007. Induction of MxA gene expression by influenza A virus requires type I or type III interferon signaling. *J Virol*. 81:7776-85.

Hood, L., J.H. Campbell, and S.C. Elgin. 1975. The organization, expression, and evolution of antibody genes and other multigene families. *Annu Rev Genet*. 9:305-53.

Howe, D.K., and L.D. Sibley. 1995. Toxoplasma gondii comprises three clonal lineages: correlation of parasite genotype with human disease. *J Infect Dis*. 172:1561-6.

Huang, S., W. Hendriks, A. Althage, S. Hemmi, H. Bluethmann, R. Kamijo, J. Vilcek, R.M. Zinkernagel, and M. Aguet. 1993. Immune response in mice that lack the interferon-gamma receptor. *Science*. 259:1742-5.

Huang, T., J. Pavlovic, P. Staeheli, and M. Krystal. 1992. Overexpression of the influenza virus polymerase can titrate out inhibition by the murine Mx1 protein. *J Virol*. 66:4154-60.

Imai, S., M. Okumoto, M. Iwai, S. Haga, N. Mori, N. Miyashita, K. Moriwaki, J. Hilgers, and N.H. Sarkar. 1994. Distribution of mouse mammary tumor virus in Asian wild mice. *J Virol*. 68:3437-42.

Isaacs, A., and J. Lindenmann. 1957a. Virus interference. I. The interferon. *Proc R Soc Lond B Biol Sci*. 147:258-67.

Isaacs, A., J. Lindenmann, and R.C. Valentine. 1957b. Virus interference. II. Some properties of interferon. *Proc R Soc Lond B Biol Sci*. 147:268-73.

Itsui, Y., N. Sakamoto, M. Kurosaki, N. Kanazawa, Y. Tanabe, T. Koyama, Y. Takeda, M. Nakagawa, S. Kakinuma, Y. Sekine, S. Maekawa, N. Enomoto, and M. Watanabe. 2006. Expressional screening of interferon-stimulated genes for antiviral activity against hepatitis C virus replication. *J Viral Hepat*. 13:690-700.

Jackson, S.H., J.I. Gallin, and S.M. Holland. 1995. The p47phox mouse knock-out model of chronic granulomatous disease. *J Exp Med*. 182:751-8.

Jacobs, M., A. Samarina, S. Grivennikov, T. Botha, N. Allie, C. Fremond, D. Togbe, V. Vasseur, S. Rose, F. Erard, A. Monteiro, V. Quesniaux, and B. Ryffel. 2007. Reactivation of tuberculosis by tumor necrosis factor neutralization. *Eur Cytokine Netw*. 18:5-13.

Jainchill, J.L., S.A. Aaronson, and G.J. Todaro. 1969. Murine sarcoma and leukemia viruses: assay using clonal lines of contact-inhibited mouse cells. *J Virol*. 4:549-53.

James, P., J. Halladay, and E.A. Craig. 1996. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*. 144:1425-36.

Janeway, C.A., Jr., and R. Medzhitov. 2002. Innate immune recognition. *Annu Rev Immunol*. 20:197-216.

Janssen, R., A. Van Wengen, E. Verhard, T. De Boer, T. Zomerdijk, T.H. Ottenhoff, and J.T. Van Dissel. 2002. Divergent role for TNF-alpha in IFN-gamma-induced killing of Toxoplasma gondii and Salmonella typhimurium contributes to selective susceptibility of patients with partial IFN-gamma receptor 1 deficiency. *J Immunol*. 169:3900-7.

Janzen, C., G. Kochs, and O. Haller. 2000. A monomeric GTPase-negative MxA mutant with antiviral activity. *J Virol*. 74:8202-6.

Jayan, G.C., and J.L. Casey. 2002. Increased RNA editing and inhibition of hepatitis delta virus replication by high-level expression of ADAR1 and ADAR2. *J Virol*. 76:3819-27.

Jin, H.K., A. Takada, Y. Kon, O. Haller, and T. Watanabe. 1999. Identification of the murine Mx2 gene: interferon-induced expression of the Mx2 protein from the feral mouse gene confers resistance to vesicular stomatitis virus. *J Virol*. 73:4925-30.

Jin, H.K., T. Yamashita, K. Ochiai, O. Haller, and T. Watanabe. 1998. Characterization and expression of the Mx1 gene in wild mouse species. *Biochem Genet*. 36:311-22.

Johannes, L., R. Kambadur, H. Lee-Hellmich, C.A. Hodgkinson, H. Arnheiter, and E. Meier. 1997. Antiviral determinants of rat Mx GTPases map to the carboxy-terminal half. *J Virol*. 71:9792-5.

John, J., M. Frech, and A. Wittinghofer. 1988. Biochemical properties of Ha-ras encoded p21 mutants and mechanism of the autophosphorylation reaction. *J Biol Chem*. 263:11792-9.

Johnson, D.R., R.S. Bhatnagar, L.J. Knoll, and J.I. Gordon. 1994. Genetic and biochemical studies of protein N-myristoylation. *Annu Rev Biochem*. 63:869-914.

Johnston, J.A., C.L. Ward, and R.R. Kopito. 1998. Aggresomes: a cellular response to misfolded proteins. *J Cell Biol*. 143:1883-98.

Jones, D.T., W.R. Taylor, and J.M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci*. 8:275-82.

Jones, T.C., S. Yeh, and J.G. Hirsch. 1972. The interaction between Toxoplasma gondii and mammalian cells. I. Mechanism of entry and intracellular fate of the parasite. *J Exp Med*. 136:1157-72.

Jouanguy, E., F. Altare, S. Lamhamedi-Cherradi, and J.L. Casanova. 1997. Infections in IFNGR-1-deficient children. *J Interferon Cytokine Res*. 17:583-7.

Jouanguy, E., R. Doffinger, S. Dupuis, A. Pallier, F. Altare, and J.L. Casanova. 1999a. IL-12 and IFN-gamma in host defense against mycobacteria and salmonella in mice and men. *Curr Opin Immunol*. 11:346-51.

Jouanguy, E., S. Lamhamedi-Cherradi, D. Lammas, S.E. Dorman, M.C. Fondaneche, S. Dupuis, R. Doffinger, F. Altare, J. Girdlestone, J.F. Emile, H. Ducoulombier, D. Edgar, J. Clarke, V.A. Oxelius, M. Brai, V. Novelli, K. Heyne, A. Fischer, S.M. Holland, D.S. Kumararatne, R.D. Schreiber, and J.L. Casanova. 1999b. A human IFNGR1 small deletion hotspot associated with dominant susceptibility to mycobacterial infection. *Nat Genet*. 21:370-8.

Kaiser, F. 2005. Molekulare Charakterisierung und funktionelle Analyse der Interferon-induzierten 47 kDa GTPase IIGP. PhD thesis. Institut für Biotechnologie. Technischen Universität Berlin, Berlin.

Kaiser, F., S.H. Kaufmann, and J. Zerrahn. 2004. IIGP, a member of the IFN inducible and microbial defense mediating 47 kDa GTPase family, interacts with the microtubule binding protein hook3. *J Cell Sci*. Pt.

Kapur, S., B. Marcotte, and A. Marette. 1999. Mechanism of adipose tissue iNOS induction in endotoxemia. *Am J Physiol*. 276:E635-41.

Khalturin, K., Z. Panzer, M.D. Cooper, and T.C. Bosch. 2004. Recognition strategies in the innate immune system of ancestral chordates. *Mol Immunol*. 41:1077-87.

King, M.C., G. Raposo, and M.A. Lemmon. 2004. Inhibition of nuclear import and cell-cycle progression by mutated forms of the dynamin-like GTPase MxB. *Proc Natl Acad Sci U S A*. 101:8957-62.

Klamp, T., U. Boehm, D. Schenk, K. Pfeffer, and J.C. Howard. 2003. A giant GTPase, very large inducible GTPase-1, is inducible by IFNs. *J Immunol*. 171:1255-65.

Klein, J., and F. Figueroa. 1986. Evolution of the major histocompatibility complex. *Crit Rev Immunol*. 6:295-386.

Kleineke, J., C. Duls, and H.D. Soling. 1979. Subcellular compartmentation of guanine nucleotides and functional relationships between the adenine and guanine nucleotide systems in isolated hepatocytes. *FEBS Lett*. 107:198-202.

Ko, J.H., H.K. Jin, A. Asano, A. Takada, A. Ninomiya, H. Kida, H. Hokiyama, M. Ohara, M. Tsuzuki, M. Nishibori, M. Mizutani, and T. Watanabe. 2002. Polymorphisms and the differential antiviral activity of the chicken Mx gene. *Genome Res*. 12:595-601.

Kochs, G., M. Haener, U. Aebi, and O. Haller. 2002a. Self-assembly of human MxA GTPase into highly ordered dynamin-like oligomers. *J Biol Chem*. 277:14172-6.

Kochs, G., and O. Haller. 1999a. GTP-bound human MxA protein interacts with the nucleocapsids of Thogoto virus (Orthomyxoviridae). *J Biol Chem*. 274:4370-6.

Kochs, G., and O. Haller. 1999b. Interferon-induced human MxA GTPase blocks nuclear import of Thogoto virus nucleocapsids. *Proc Natl Acad Sci U S A*. 96:2082-6.

Kochs, G., C. Janzen, H. Hohenberg, and O. Haller. 2002b. Antivirally active MxA protein sequesters La Crosse virus nucleocapsid protein into perinuclear complexes. *Proc Natl Acad Sci U S A*. 99:3153-8.

Kochs, G., M. Trost, C. Janzen, and O. Haller. 1998. MxA GTPase: oligomerization and GTP-dependent interaction with viral RNP target structures. *Methods*. 15:255-63.

Koga, R., S. Hamano, H. Kuwata, K. Atarashi, M. Ogawa, H. Hisaeda, M. Yamamoto, S. Akira, K. Himeno, M. Matsumoto, and K. Takeda. 2006. TLR-dependent induction of IFN-beta mediates host defense against Trypanosoma cruzi. *J Immunol*. 177:7059-66.

Kotenko, S.V., G. Gallagher, V.V. Baurin, A. Lewis-Antes, M. Shen, N.K. Shah, J.A. Langer, F. Sheikh, H. Dickensheets, and R.P. Donnelly. 2003. IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex. *Nat Immunol*. 4:69-77.

Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res*. 15:8125-48.

Krell, T., J. Maclean, D.J. Boam, A. Cooper, M. Resmini, K. Brocklehurst, S.M. Kelly, N.C. Price, A.J. Lapthorn, and J.R. Coggins. 2001. Biochemical and X-ray crystallographic studies on shikimate kinase: the important structural role of the P-loop lysine. *Protein Sci*. 10:1137-49.

Kruchten, A.E., and M.A. McNiven. 2006. Dynamin as a mover and pincher during cell migration and invasion. *J Cell Sci*. 119:1683-90.

Kumar, S., and S.B. Hedges. 1998. A molecular timescale for vertebrate evolution. *Nature*. 392:917-20.

Kumar, S., C. Mitnik, G. Valente, and G. Floyd-Smith. 2000. Expansion and molecular evolution of the interferon-induced 2'-5' oligoadenylate synthetase gene family. *Mol Biol Evol*. 17:738-50.

Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform*. 5:150-63.

Kunzelmann, S., G.J. Praefcke, and C. Herrmann. 2006. Transient kinetic investigation of GTP hydrolysis catalyzed by interferon-gamma induced human guanylate binding protein 1 (hGBP1). *J Biol Chem*.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 227:680-5.

LaFleur, D.W., B. Nardelli, T. Tsareva, D. Mather, P. Feng, M. Semenuk, K. Taylor, M. Buergin, D. Chinchilla, V. Roshke, G. Chen, S.M. Ruben, P.M. Pitha, T.A. Coleman, and P.A. Moore. 2001. Interferon-kappa, a novel type I interferon expressed in human keratinocytes. *J Biol Chem*. 276:39765-71.

Lafuse, W.P., D. Brown, L. Castle, and B.S. Zwilling. 1995. Cloning and characterization of a novel cDNA that is IFN-gamma-induced in mouse peritoneal macrophages and encodes a putative GTP-binding protein. *J Leukoc Biol*. 57:477-83.

Landis, H., A. Simon-Jodicke, A. Kloti, C. Di Paolo, J.J. Schnorr, S. Schneider-Schaulies, H.P. Hefti, and J. Pavlovic. 1998. Human MxA protein confers resistance to Semliki Forest virus and inhibits the amplification of a Semliki Forest virus-based replicon in the absence of viral structural proteins. *J Virol*. 72:1516-22.

Lane, K.T., and L.S. Beese. 2006. Thematic review series: lipid posttranslational modifications. Structural biology of protein farnesyltransferase and geranylgeranyltransferase type I. *J Lipid Res*. 47:681-99.

Lapaque, N., O. Takeuchi, F. Corrales, S. Akira, I. Moriyon, J.C. Howard, and J.P. Gorvel. 2006. Differential inductions of TNF-alpha and IGTP, IIGP by structurally diverse classic and non-classic lipopolysaccharides. *Cell Microbiol*. 8:401-13.

Le Bon, A., and D.F. Tough. 2002. Links between innate and adaptive immunity via type I interferon. *Curr Opin Immunol*. 14:432-6.

Lee, E., R. Taussig, and A.G. Gilman. 1992. The G226A mutant of Gs alpha highlights the requirement for dissociation of G protein subunits. *J Biol Chem*. 267:1212-8.

Lee, S.B., and M. Esteban. 1993. The interferon-induced double-stranded RNA-activated human p68 protein kinase inhibits the replication of vaccinia virus. *Virology*. 193:1037-41.

Lefevre, F., M. Guillomot, S. D'Andrea, S. Battegay, and C. La Bonnardiere. 1998. Interferon-delta: the first member of a novel type I interferon family. *Biochimie*. 80:779-88.

Lehmann, T., C.R. Blackston, S.F. Parmley, J.S. Remington, and J.P. Dubey. 2000. Strain typing of Toxoplasma gondii: comparison of antigen-coding and housekeeping genes. *J Parasitol*. 86:960-71.

Leipe, D.D., Y.I. Wolf, E.V. Koonin, and L. Aravind. 2002. Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol*. 317:41-72.

Leister, D. 2004. Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. *Trends Genet*. 20:116-22.

Leong, J.C., G.D. Trobridge, C.H. Kim, M. Johnson, and B. Simon. 1998. Interferon-inducible Mx proteins in fish. *Immunol Rev*. 166:349-63.

Levine, N.D. 1988. Progress in taxonomy of the Apicomplexan protozoa. *J Protozool*. 35:518-20.

Lieberman, L.A., and C.A. Hunter. 2002. The role of cytokines and their signaling pathways in the regulation of immunity to Toxoplasma gondii. *Int Rev Immunol*. 21:373-403.

Lillie, J.W., and M.R. Green. 1989. Transcription activation by the adenovirus E1a protein. *Nature*. 338:39-44.

Lindenmann, J. 1962. Resistance of mice to mouse-adapted influenza A virus. *Virology*. 16:203-4.

Lindenmann, J. 1963. [Relations between virus and host at the cellular level]. *Arch Gesamte Virusforsch*. 13:35-40.

Lindenmann, J. 1964. Inheritance Of Resistance To Influenza Virus In Mice. *Proc Soc Exp Biol Med*. 116:506-9.

Lindenmann, J., E. Deuel, S. Fanconi, and O. Haller. 1978. Inborn resistance of mice to myxoviruses: macrophages express phenotype in vitro. *J Exp Med*. 147:531-40.

Ling, J., W. Pi, R. Bollag, S. Zeng, M. Keskintepe, H. Saliman, S. Krantz, B. Whitney, and D. Tuan. 2002. The solitary long terminal repeats of ERV-9 endogenous retrovirus are conserved during primate evolution and possess enhancer activities in embryonic and hematopoietic cells. *J Virol*. 76:2410-23.

Ling, Y.M., M.H. Shaw, C. Ayala, I. Coppens, G.A. Taylor, D.J. Ferguson, and G.S. Yap. 2006. Vacuolar and plasma membrane stripping and autophagic elimination of Toxoplasma gondii in primed effector macrophages. *J Exp Med*. 203:2063-71.

Love, K.R., A. Catic, C. Schlieker, and H.L. Ploegh. 2007. Mechanisms, biology and inhibitors of deubiquitinating enzymes. *Nat Chem Biol*. 3:697-705.

Low, H.H., and J. Lowe. 2006. A bacterial dynamin-like protein. *Nature*. 444:766-9.

Luan, Z., Y. Zhang, A. Liu, Y. Man, L. Cheng, and G. Hu. 2002. A novel GTP-binding protein hGBP3 interacts with NIK/HGK. *FEBS Lett*. 530:233-8.

Lubeseder-Martellato, C., E. Guenzi, A. Jorg, K. Topolt, E. Naschberger, E. Kremmer, C. Zietz, E. Tschachler, P. Hutzler, M. Schwemmle, K. Matzen, T. Grimm, B. Ensoli, and M. Sturzl. 2002. Guanylate-binding protein-1 expression is selectively induced by inflammatory cytokines and is an activation marker of endothelial cells during inflammatory diseases. *Am J Pathol*. 161:1749-59.

MacKenzie, C.R., R. Langen, O. Takikawa, and W. Daubener. 1999. Inhibition of indoleamine 2,3-dioxygenase in human macrophages inhibits interferon-gamma-induced bacteriostasis but does not abrogate toxoplasmastasis. *Eur J Immunol*. 29:3254-61.

MacMicking, J., Q.W. Xie, and C. Nathan. 1997. Nitric oxide and macrophage function. *Annu Rev Immunol*. 15:323-50.

MacMicking, J.D., G.A. Taylor, and J.D. McKinney. 2003. Immune control of tuberculosis by IFN-gamma-inducible LRG-47. *Science*. 302:654-9.

Mangeat, B., P. Turelli, G. Caron, M. Friedli, L. Perrin, and D. Trono. 2003. Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. *Nature*. 424:99-103.

Mantovani, R. 1999. The molecular biology of the CCAAT-binding factor NF-Y. *Gene*. 239:15-27.

Marck, C. 1988. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Res*. 16:1829-36.

Marks, B., M.H. Stowell, Y. Vallis, I.G. Mills, A. Gibson, C.R. Hopkins, and H.T. McMahon. 2001. GTPase activity of dynamin and resulting conformation change are essential for endocytosis. *Nature*. 410:231-5.

Marmorstein, R., M. Carey, M. Ptashne, and S.C. Harrison. 1992. DNA recognition by GAL4: structure of a protein-DNA complex. *Nature*. 356:408-14.

Martal, J.L., N.M. Chene, L.P. Huynh, R.M. L'Haridon, P.B. Reinaud, M.W. Guillomot, M.A. Charlier, and S.Y. Charpigny. 1998. IFN-tau: a novel subtype I IFN1. Structural characteristics, non-ubiquitous expression, structure-function relationships, a pregnancy hormonal embryonic signal and cross-species therapeutic potentialities. *Biochimie*. 80:755-77.

Martens, S. 2004a. Cell-biology of interferon inducible GTPases. PhD thesis. Institute for Genetics, University of Cologne, Cologne.

Martens, S., and J. Howard. 2006. The Interferon-Inducible GTPases. *Annu Rev Cell Dev Biol*.

Martens, S., I. Parvanova, J. Zerrahn, G. Griffiths, G. Schell, G. Reichmann, and J.C. Howard. 2005. Disruption of Toxoplasma gondii Parasitophorous Vacuoles by the Mouse p47-Resistance GTPases. *PLoS Pathog*. 1:e24.

Martens, S., K. Sabel, R. Lange, R. Uthaiah, E. Wolf, and J.C. Howard. 2004b. Mechanisms Regulating the Positioning of Mouse p47 Resistance GTPases LRG-47 and IIGP1 on Cellular Membranes: Retargeting

to Plasma Membrane Induced by Phagocytosis. *J Immunol*. 173:2594-606.
Mashimo, T., P. Glaser, M. Lucas, D. Simon-Chazottes, P.E. Ceccaldi, X. Montagutelli, P. Despres, and J.L. Guenet. 2003. Structural and functional genomics and evolutionary relationships in the cluster of genes encoding murine 2',5'-oligoadenylate synthetases. *Genomics*. 82:537-52.
Maurer-Stroh, S., B. Eisenhaber, and F. Eisenhaber. 2002. N-terminal N-myristoylation of proteins: refinement of the sequence motif and its taxon-specific differences. *J Mol Biol*. 317:523-40.
Maurer-Stroh, S., M. Gouda, M. Novatchkova, A. Schleiffer, G. Schneider, F.L. Sirota, M. Wildpaner, N. Hayashi, and F. Eisenhaber. 2004. MYRbase: analysis of genome-wide glycine myristoylation enlarges the functional spectrum of eukaryotic myristoylated proteins. *Genome Biol*. 5:R21.
McAlister-Henn, L., N. Gibson, and E. Panisko. 1999. Applications of the yeast two-hybrid system. *Methods*. 19:330-7.
Medzhitov, R., and C.A. Janeway, Jr. 1997. Innate immunity: impact on the adaptive immune response. *Curr Opin Immunol*. 9:4-9.
Meier, E., J. Fah, M.S. Grob, R. End, P. Staeheli, and O. Haller. 1988. A family of interferon-induced Mx-related mRNAs encodes cytoplasmic and nuclear proteins in rat cells. *J Virol*. 62:2386-93.
Meier, E., G. Kunz, O. Haller, and H. Arnheiter. 1990. Activity of rat Mx proteins against a rhabdovirus. *J Virol*. 64:6263-9.
Melen, K., and I. Julkunen. 1997. Nuclear cotransport mechanism of cytoplasmic human MxB protein. *J Biol Chem*. 272:32353-9.
Melen, K., P. Keskinen, T. Ronni, T. Sareneva, K. Lounatmaa, and I. Julkunen. 1996. Human MxB protein, an interferon-alpha-inducible GTPase, contains a nuclear targeting signal and is localized in the heterochromatin region beneath the nuclear envelope. *J Biol Chem*. 271:23478-86.
Melen, K., T. Ronni, B. Broni, R.M. Krug, C.H. von Bonsdorff, and I. Julkunen. 1992. Interferon-induced Mx proteins form oligomers and contain a putative leucine zipper. *J Biol Chem*. 267:25898-907.
Meurs, E.F., Y. Watanabe, S. Kadereit, G.N. Barber, M.G. Katze, K. Chong, B.R. Williams, and A.G. Hovanessian. 1992. Constitutive expression of human double-stranded RNA-activated p68 kinase in murine cells mediates phosphorylation of eukaryotic initiation factor 2 and partial resistance to encephalomyocarditis virus growth. *J Virol*. 66:5805-14.
Miura, T.A., J.O. Carlson, B.J. Beaty, R.A. Bowen, and K.E. Olson. 2001. Expression of human MxA protein in mosquito cells interferes with LaCrosse virus replication. *J Virol*. 75:3001-3.
Miyairi, I., V.R. Tatireddigari, O.S. Mahdi, L.A. Rose, R.J. Belland, L. Lu, R.W. Williams, and G.I. Byrne. 2007. The p47 GTPases Iigp2 and Irgb10 regulate innate immunity and inflammation to murine Chlamydia psittaci infection. *J Immunol*. 179:1814-24.
Modiano, N., Y.E. Lu, and P. Cresswell. 2005. Golgi targeting of human guanylate-binding protein-1 requires nucleotide binding, isoprenylation, and an IFN-gamma-inducible cofactor. *Proc Natl Acad Sci U S A*. 102:8680-5.
Mogensen, K.E., M. Lewerenz, J. Reboul, G. Lutfalla, and G. Uze. 1999. The type I interferon receptor: structure, function, and evolution of a family business. *J Interferon Cytokine Res*. 19:1069-98.
Mordue, D.G., N. Desai, M. Dustin, and L.D. Sibley. 1999a. Invasion by Toxoplasma gondii establishes a moving junction that selectively excludes host cell plasma membrane proteins on the basis of their membrane anchoring. *J Exp Med*. 190:1783-92.
Mordue, D.G., S. Hakansson, I. Niesman, and L.D. Sibley. 1999b. Toxoplasma gondii resides in a vacuole that avoids fusion with host cell endocytic and exocytic vesicular trafficking pathways. *Exp Parasitol*. 92:87-99.
Mordue, D.G., and L.D. Sibley. 1997. Intracellular fate of vacuoles containing Toxoplasma gondii is determined at the time of formation and depends on the mechanism of entry. *J Immunol*. 159:4452-9.
Mosmann, T. 1983. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods*. 65:55-63.
Mosmann, T.R., and T.A. Fong. 1989. Specific assays for cytokine production by T cells. *J Immunol Methods*. 116:151-8.
Muhlberg, A.B., D.E. Warnock, and S.L. Schmid. 1997. Domain structure and intramolecular regulation of dynamin GTPase. *Embo J*. 16:6676-83.
Munder, M., M. Mallo, K. Eichmann, and M. Modolell. 1998. Murine macrophages secrete interferon gamma upon combined stimulation with interleukin (IL)-12 and IL-18: A novel pathway of autocrine macrophage activation. *J Exp Med*. 187:2103-8.
Murray, H.W., A. Szuro-Sudol, D. Wellner, M.J. Oca, A.M. Granger, D.M. Libby, C.D. Rothermel, and B.Y. Rubin. 1989. Role of tryptophan degradation in respiratory burst-independent antimicrobial activity of gamma interferon-stimulated human macrophages. *Infect Immun*. 57:845-9.
Nakaar, V., H.M. Ngo, E.P. Aaronson, I. Coppens, T.T. Stedman, and K.A. Joiner. 2003. Pleiotropic effect due to targeted depletion of secretory rhoptry protein ROP2 in Toxoplasma gondii. *J Cell Sci*.

116:2311-20.

Nakajima, E., T. Morozumi, K. Tsukamoto, T. Watanabe, G. Plastow, and T. Mitsuhashi. 2007. A naturally occurring variant of porcine Mx1 associated with increased susceptibility to influenza virus in vitro. *Biochem Genet*. 45:11-24.

Nakayama, M., K. Yazaki, A. Kusano, K. Nagata, N. Hanai, and A. Ishihama. 1993. Structure of mouse Mx1 protein. Molecular assembly and GTP-dependent conformational change. *J Biol Chem*. 268:15033-8.

Nantais, D.E., M. Schwemmle, J.T. Stickney, D.J. Vestal, and J.E. Buss. 1996. Prenylation of an interferon-gamma-induced GTP-binding protein: the human guanylate binding protein, huGBP1. *J Leukoc Biol*. 60:423-31.

Nelson, D.E., D.P. Virok, H. Wood, C. Roshick, R.M. Johnson, W.M. Whitmire, D.D. Crane, O. Steele-Mortimer, L. Kari, G. McClarty, and H.D. Caldwell. 2005. Chlamydial IFN-gamma immune evasion is linked to host infection tropism. *Proc Natl Acad Sci U S A*. 102:10658-63.

Neun, R., M.F. Richter, P. Staeheli, and M. Schwemmle. 1996. GTPase properties of the interferon-induced human guanylate-binding protein 2. *FEBS Lett*. 390:69-72.

Nguyen, T.T., Y. Hu, D.P. Widney, R.A. Mar, and J.B. Smith. 2002. Murine GBP-5, a new member of the murine guanylate-binding protein family, is coordinately regulated with other GBPs in vivo and in vitro. *J Interferon Cytokine Res*. 22:899-909.

Nicholas, K.B., N.H.B. Jr., and D.W.I. Deerfield. 1997. GeneDoc: Analysis and Visualization of Genetic Variation. *EMBNEW.NEWS 4:14*. http://www.psc.edu/biomed/genedoc.

Ning, Z., A.J. Cox, and J.C. Mullikin. 2001. SSAHA: a fast search method for large DNA databases. *Genome Res*. 11:1725-9.

Nisole, S., J.P. Stoye, and A. Saib. 2005. TRIM family proteins: retroviral restriction and antiviral defence. *Nat Rev Microbiol*. 3:799-808.

Noel, L., T.L. Moores, E.A. van Der Biezen, M. Parniske, M.J. Daniels, J.E. Parker, and J.D. Jones. 1999. Pronounced intraspecific haplotype divergence at the RPP5 complex disease resistance locus of Arabidopsis. *Plant Cell*. 11:2099-112.

Nuoffer, C., H.W. Davidson, J. Matteson, J. Meinkoth, and W.E. Balch. 1994. A GDP-bound of rab1 inhibits protein export from the endoplasmic reticulum and transport between Golgi compartments. *J Cell Biol*. 125:225-37.

Ohteki, T., T. Fukao, K. Suzue, C. Maki, M. Ito, M. Nakamura, and S. Koyasu. 1999. Interleukin 12-dependent interferon gamma production by CD8alpha+ lymphoid dendritic cells. *J Exp Med*. 189:1981-6.

Olszewski, M.A., J. Gray, and D.J. Vestal. 2006. In silico genomic analysis of the human and murine guanylate-binding protein (GBP) gene clusters. *J Interferon Cytokine Res*. 26:328-52.

Orange, J.S., B. Wang, C. Terhorst, and C.A. Biron. 1995. Requirement for natural killer cell-produced interferon gamma in defense against murine cytomegalovirus infection and enhancement of this defense pathway by interleukin 12 administration. *J Exp Med*. 182:1045-56.

Oritani, K., K.L. Medina, Y. Tomiyama, J. Ishikawa, Y. Okajima, M. Ogawa, T. Yokota, K. Aoyama, I. Takahashi, P.W. Kincade, and Y. Matsuzawa. 2000. Limitin: An interferon-like cytokine that preferentially influences B-lymphocyte precursors. *Nat Med*. 6:659-66.

Osterlund, P.I., T.E. Pietila, V. Veckman, S.V. Kotenko, and I. Julkunen. 2007. IFN Regulatory Factor Family Members Differentially Regulate the Expression of Type III IFN (IFN-{lambda}) Genes. *J Immunol*. 179:3434-42.

Ottenhoff, T.H., F.A. Verreck, E.G. Lichtenauer-Kaligis, M.A. Hoeve, O. Sanal, and J.T. van Dissel. 2002. Genetics, cytokines and human infectious disease: lessons from weakly pathogenic mycobacteria and salmonellae. *Nat Genet*. 32:97-105.

Pai, E.F., W. Kabsch, U. Krengel, K.C. Holmes, J. John, and A. Wittinghofer. 1989. Structure of the guanine-nucleotide-binding domain of the Ha-ras oncogene product p21 in the triphosphate conformation. *Nature*. 341:209-14.

Pai, E.F., U. Krengel, G.A. Petsko, R.S. Goody, W. Kabsch, and A. Wittinghofer. 1990. Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 A resolution: implications for the mechanism of GTP hydrolysis. *Embo J*. 9:2351-9.

Paine, P.L. 1975. Nucleocytoplasmic movement of fluorescent tracers microinjected into living salivary gland cells. *J Cell Biol*. 66:652-7.

Palm, M., M. Leroy, A. Thomas, A. Linden, and D. Desmecht. 2007. Differential anti-influenza activity among allelic variants at the Sus scrofa Mx1 locus. *J Interferon Cytokine Res*. 27:147-55.

Pantoja, L.G., R.D. Miller, J.A. Ramirez, R.E. Molestina, and J.T. Summersgill. 2000. Inhibition of Chlamydia pneumoniae replication in human aortic smooth muscle cells by gamma interferon-induced indoleamine 2, 3-dioxygenase activity. *Infect Immun*. 68:6478-81.

Papic, N. 2007. Biochemical analysis of the immunity-related GTPase Irga6 in vivo and in vitro; the role of the myristoyl group. PhD thesis. Institute for Genetics, University of Cologne, Cologne.

Parvanova, I. 2005. Analysis of the role of the p47 GTPase IIGP1 in resistance against intracellular Pathogens. PhD thesis. Institute for Genetics, University of Cologne, Cologne.

Pavlovic, J., O. Haller, and P. Staeheli. 1992. Human and mouse Mx proteins inhibit different steps of the influenza virus multiplication cycle. *J Virol*. 66:2564-9.

Pavlovic, J., T. Zurcher, O. Haller, and P. Staeheli. 1990. Resistance to influenza virus and vesicular stomatitis virus conferred by expression of human MxA protein. *J Virol*. 64:3370-5.

Pestka, S., C.D. Krause, and M.R. Walter. 2004. Interferons, interferon-like cytokines, and their receptors. *Immunol Rev*. 202:8-32.

Pestka, S., J.A. Langer, K.C. Zoon, and C.E. Samuel. 1987. Interferons and their actions. *Annu Rev Biochem*. 56:727-77.

Pfefferkorn, E.R. 1984. Interferon gamma blocks the growth of Toxoplasma gondii in human fibroblasts by inducing the host cells to degrade tryptophan. *Proc Natl Acad Sci U S A*. 81:908-12.

Pfefferkorn, E.R., M. Eckel, and S. Rebhun. 1986. Interferon-gamma suppresses the growth of Toxoplasma gondii in human fibroblasts through starvation for tryptophan. *Mol Biochem Parasitol*. 20:215-24.

Pfefferkorn, E.R., and L.C. Pfefferkorn. 1977. Specific labeling of intracellular Toxoplasma gondii with uracil. *J Protozool*. 24:449-53.

Pickart, C.M., and D. Fushman. 2004. Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol*. 8:610-6.

Pitossi, F., A. Blank, A. Schroder, A. Schwarz, P. Hussi, M. Schwemmle, J. Pavlovic, and P. Staeheli. 1993. A functional GTP-binding motif is necessary for antiviral activity of Mx proteins. *J Virol*. 67:6726-32.

Plant, J., and A.A. Glynn. 1976. Genetics of resistance to infection with Salmonella typhimurium in mice. *J Infect Dis*. 133:72-8.

Ponten, A., C. Sick, M. Weeber, O. Haller, and G. Kochs. 1997. Dominant-negative mutants of human MxA protein: domains in the carboxy-terminal moiety are important for oligomerization and antiviral activity. *J Virol*. 71:2591-9.

Powers, S., K. O'Neill, and M. Wigler. 1989. Dominant yeast and mammalian RAS mutants that interfere with the CDC25-dependent activation of wild-type RAS in Saccharomyces cerevisiae. *Mol Cell Biol*. 9:390-5.

Praefcke, G.J., M. Geyer, M. Schwemmle, H. Robert Kalbitzer, and C. Herrmann. 1999. Nucleotide-binding characteristics of human guanylate-binding protein 1 (hGBP1) and identification of the third GTP-binding motif. *J Mol Biol*. 292:321-32.

Praefcke, G.J., S. Kloep, U. Benscheid, H. Lilie, B. Prakash, and C. Herrmann. 2004a. Identification of residues in the human guanylate-binding protein 1 critical for nucleotide binding and cooperative GTP hydrolysis. *J Mol Biol*. 344:257-69.

Praefcke, G.J., and H.T. McMahon. 2004b. The dynamin superfamily: universal membrane tubulation and fission molecules? *Nat Rev Mol Cell Biol*. 5:133-47.

Prakash, B., G.J. Praefcke, L. Renault, A. Wittinghofer, and C. Herrmann. 2000a. Structure of human guanylate-binding protein 1 representing a unique class of GTP-binding proteins. *Nature*. 403:567-71.

Prakash, B., L. Renault, G.J. Praefcke, C. Herrmann, and A. Wittinghofer. 2000b. Triphosphate structure of guanylate-binding protein 1 and implications for nucleotide binding and GTPase mechanism. *Embo J*. 19:4555-64.

Rajala, R.V., R.S. Datla, T.N. Moyana, R. Kakkar, S.A. Carlsen, and R.K. Sharma. 2000. N-myristoyltransferase. *Mol Cell Biochem*. 204:135-55.

Reeves, R.H., B.F. O'Hara, W.J. Pavan, J.D. Gearhart, and O. Haller. 1988. Genetic mapping of the Mx influenza virus resistance gene within the region of mouse chromosome 16 that is homologous to human chromosome 21. *J Virol*. 62:4372-5.

Regad, T., A. Saib, V. Lallemand-Breitenbach, P.P. Pandolfi, H. de The, and M.K. Chelbi-Alix. 2001. PML mediates the interferon-induced antiviral state against a complex retrovirus via its association with the viral transactivator. *Embo J*. 20:3495-505.

Reichelt, M., S. Stertz, J. Krijnse-Locker, O. Haller, and G. Kochs. 2004. Missorting of LaCrosse virus nucleocapsid protein by the interferon-induced MxA GTPase involves smooth ER membranes. *Traffic*. 5:772-84.

Rhode, C. 2007. Genetic and functional studies on the conserved IRG (immunity-related GTPase) protein Irgc (CINEMA). PhD thesis. Institute for Genetics, University of Cologne, Cologne.

Richter, M.F., M. Schwemmle, C. Herrmann, A. Wittinghofer, and P. Staeheli. 1995. Interferon-induced MxA protein. GTP binding and GTP hydrolysis properties. *J Biol Chem*. 270:13512-7.

Robertsen, B., J. Zou, C. Secombes, and J.A. Leong. 2006. Molecular and expression analysis of an

interferon-gamma-inducible guanylate-binding protein from rainbow trout (Oncorhynchus mykiss). *Dev Comp Immunol*. 30:1023-33.

Rojas, R.J., R.J. Kimple, K.L. Rossman, D.P. Siderovski, and J. Sondek. 2003. Established and emerging fluorescence-based assays for G-protein function: Ras-superfamily GTPases. *Comb Chem High Throughput Screen*. 6:409-18.

Sabin, A.B. 1941. Toxoplasmic encephalitis in children. *J Am Med Assoc*. 116:801–807.

Saeij, J.P., J.P. Boyle, and J.C. Boothroyd. 2005. Differences among the three major strains of Toxoplasma gondii and their specific interactions with the infected host. *Trends Parasitol*. 21:476-81.

Saeij, J.P., J.P. Boyle, S. Coller, S. Taylor, L.D. Sibley, E.T. Brooke-Powell, J.W. Ajioka, and J.C. Boothroyd. 2006. Polymorphic secreted kinases are key virulence factors in toxoplasmosis. *Science*. 314:1780-3.

Saeij, J.P., S. Coller, J.P. Boyle, M.E. Jerome, M.W. White, and J.C. Boothroyd. 2007. Toxoplasma co-opts host gene expression by injection of a polymorphic kinase homologue. *Nature*. 445:324-7.

Samarajiwa, S.A., W. Wilson, and P.J. Hertzog. 2006. Type I Interferons: Genetics and Structure. *In* The Interferons. M. Dr. Anthony, editor. 1-34.

Samuel, C.E. 2001. Antiviral actions of interferons. *Clin Microbiol Rev*. 14:778-809, table of contents.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 74:5463-7.

Santiago, H.C., C.G. Feng, A. Bafica, E. Roffe, R.M. Arantes, A. Cheever, G. Taylor, L.Q. Vieira, J. Aliberti, R.T. Gazzinelli, and A. Sher. 2005. Mice deficient in LRG-47 display enhanced susceptibility to Trypanosoma cruzi infection associated with defective hemopoiesis and intracellular control of parasite growth. *J Immunol*. 175:8165-72.

Schafer, D.A. 2002. Coupling actin dynamics and membrane dynamics during endocytosis. *Curr Opin Cell Biol*. 14:76-81.

Schafer, D.A. 2004. Regulating actin dynamics at membranes: a focus on dynamin. *Traffic*. 5:463-9.

Schafer, W.R., and J. Rine. 1992. Protein prenylation: genes, enzymes, targets, and functions. *Annu Rev Genet*. 26:209-37.

Scharton-Kersten, T.M., T.A. Wynn, E.Y. Denkers, S. Bala, E. Grunvald, S. Hieny, R.T. Gazzinelli, and A. Sher. 1996. In the absence of endogenous IFN-gamma, mice develop unimpaired IL-12 responses to Toxoplasma gondii while failing to control acute infection. *J Immunol*. 157:4045-54.

Schlesinger, D.H., G. Goldstein, and H.D. Niall. 1975. The complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. *Biochemistry*. 14:2214-8.

Schrem, H., J. Klempnauer, and J. Borlak. 2002. Liver-enriched transcription factors in liver function and development. Part I: the hepatocyte nuclear factor network and liver-specific gene expression. *Pharmacol Rev*. 54:129-58.

Schroder, K., P.J. Hertzog, T. Ravasi, and D.A. Hume. 2004. Interferon-gamma: an overview of signals, mechanisms and functions. *J Leukoc Biol*. 75:163-89.

Schumacher, B., and P. Staeheli. 1998. Domains mediating intramolecular folding and oligomerization of MxA GTPase. *J Biol Chem*. 273:28365-70.

Schwemmle, M., M.F. Richter, C. Herrmann, N. Nassar, and P. Staeheli. 1995. Unexpected structural requirements for GTPase activity of the interferon-induced MxA protein. *J Biol Chem*. 270:13518-23.

Schwemmle, M., and P. Staeheli. 1994. The interferon-induced 67-kDa guanylate-binding protein (hGBP1) is a GTPase that converts GTP to GMP. *J Biol Chem*. 269:11299-305.

Scrima, A., and A. Wittinghofer. 2006. Dimerisation-dependent GTPase reaction of MnmE: how potassium acts as GTPase-activating element. *Embo J*. 25:2940-51.

Sever, S., A.B. Muhlberg, and S.L. Schmid. 1999. Impairment of dynamin's GAP domain stimulates receptor-mediated endocytosis. *Nature*. 398:481-6.

Seyama, T., J.H. Ko, M. Ohe, N. Sasaoka, A. Okada, H. Gomi, A. Yoneda, J. Ueda, M. Nishibori, S. Okamoto, Y. Maeda, and T. Watanabe. 2006. Population research of genetic polymorphism at amino acid position 631 in chicken Mx protein with differential antiviral activity. *Biochem Genet*. 44:437-48.

Shan, S.O., and P. Walter. 2005. Molecular crosstalk between the nucleotide specificity determinant of the SRP GTPase and the SRP receptor. *Biochemistry*. 44:6214-22.

Sheehy, A.M., N.C. Gaddis, J.D. Choi, and M.H. Malim. 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature*. 418:646-50.

Sheppard, P., W. Kindsvogel, W. Xu, K. Henderson, S. Schlutsmeyer, T.E. Whitmore, R. Kuestner, U. Garrigues, C. Birks, J. Roraback, C. Ostrander, D. Dong, J. Shin, S. Presnell, B. Fox, B. Haldeman, E. Cooper, D. Taft, T. Gilbert, F.J. Grant, M. Tackett, W. Krivan, G. McKnight, C. Clegg, D. Foster, and K.M. Klucher. 2003. IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nat Immunol*. 4:63-8.

Shoshani, J., C.P. Groves, E.L. Simons, and G.F. Gunnell. 1996. Primate phylogeny: morphological vs. molecular results. *Mol Phylogenet Evol*. 5:102-54.

Shtrichman, R., and C.E. Samuel. 2001. The role of gamma interferon in antimicrobial immunity. *Curr Opin Microbiol*. 4:251-9.
Sibley, L.D. 2003. Toxoplasma gondii: perfecting an intracellular life style. *Traffic*. 4:581-6.
Sibley, L.D. 2004. Intracellular parasite invasion strategies. *Science*. 304:248-53.
Sibley, L.D., and J.C. Boothroyd. 1992. Virulent strains of Toxoplasma gondii comprise a single clonal lineage. *Nature*. 359:82-5.
Sibley, L.D., I.R. Niesman, S.F. Parmley, and M.F. Cesbron-Delauw. 1995. Regulated secretion of multi-lamellar vesicles leads to formation of a tubulo-vesicular network in host-cell vacuoles occupied by Toxoplasma gondii. *J Cell Sci*. 108 (Pt 4):1669-77.
Sigal, I.S., J.B. Gibbs, J.S. D'Alonzo, G.L. Temeles, B.S. Wolanski, S.H. Socher, and E.M. Scolnick. 1986. Mutant ras-encoded proteins with altered nucleotide binding exert dominant biological effects. *Proc Natl Acad Sci U S A*. 83:952-6.
Simon, A., J. Fah, O. Haller, and P. Staeheli. 1991. Interferon-regulated Mx genes are not responsive to interleukin-1, tumor necrosis factor, and other cytokines. *J Virol*. 65:968-71.
Sims, T.A., J. Hay, and I.C. Talbot. 1989. Immunoelectron microscopy for Toxoplasma antigen in association with intact tissue cysts in mouse brain. *Ann Trop Med Parasitol*. 83:639-41.
Sinai, A.P., and K.A. Joiner. 2001. The Toxoplasma gondii protein ROP2 mediates host organelle association with the parasitophorous vacuole membrane. *J Cell Biol*. 154:95-108.
Sinai, A.P., P. Webster, and K.A. Joiner. 1997. Association of host cell endoplasmic reticulum and mitochondria with the Toxoplasma gondii parasitophorous vacuole membrane: a high affinity interaction. *J Cell Sci*. 110 (Pt 17):2117-28.
Singh, G., and J. Lykke-Andersen. 2003. New insights into the formation of active nonsense-mediated decay complexes. *Trends Biochem Sci*. 28:464-6.
Singh, S.B., A.S. Davis, G.A. Taylor, and V. Deretic. 2006. Human IRGM Induces Autophagy to Eliminate Intracellular Mycobacteria. *Science*.
Sintchak, M.D., and E. Nimmesgern. 2000. The structure of inosine 5'-monophosphate dehydrogenase and the design of novel inhibitors. *Immunopharmacology*. 47:163-84.
Smith, M.D., A. Hiltbrunner, F. Kessler, and D.J. Schnell. 2002. The targeting of the atToc159 preprotein receptor to the chloroplast outer membrane is mediated by its GTPase domain and is regulated by GTP. *J Cell Biol*. 159:833-43.
Song, B.D., and S.L. Schmid. 2003. A molecular motor or a regulator? Dynamin's in a class of its own. *Biochemistry*. 42:1369-76.
Sorace, J.M., R.J. Johnson, D.L. Howard, and B.E. Drysdale. 1995. Identification of an endotoxin and IFN-inducible cDNA: possible identification of a novel protein family. *J Leukoc Biol*. 58:477-84.
Sprang, S.R. 1997. G proteins, effectors and GAPs: structure and mechanism. *Curr Opin Struct Biol*. 7:849-56.
Staeheli, P., P. Danielson, O. Haller, and J.G. Sutcliffe. 1986a. Transcriptional activation of the mouse Mx gene by type I interferon. *Mol Cell Biol*. 6:4770-4.
Staeheli, P., R. Grob, E. Meier, J.G. Sutcliffe, and O. Haller. 1988a. Influenza virus-susceptible mice carry Mx genes with a large deletion or a nonsense mutation. *Mol Cell Biol*. 8:4518-23.
Staeheli, P., and O. Haller. 1985. Interferon-induced human protein with homology to protein Mx of influenza virus-resistant mice. *Mol Cell Biol*. 5:2150-3.
Staeheli, P., O. Haller, W. Boll, J. Lindenmann, and C. Weissmann. 1986b. Mx protein: constitutive expression in 3T3 cells transformed with cloned Mx cDNA confers selective resistance to influenza virus. *Cell*. 44:147-58.
Staeheli, P., M.A. Horisberger, and O. Haller. 1984. Mx-dependent resistance to influenza viruses is induced by mouse interferons alpha and beta but not gamma. *Virology*. 132:456-61.
Staeheli, P., F. Pitossi, and J. Pavlovic. 1993. Mx proteins: GTPases with antiviral activity. *Trends Cell Biol*. 3:268-72.
Staeheli, P., D. Pravtcheva, L.G. Lundin, M. Acklin, F. Ruddle, J. Lindenmann, and O. Haller. 1986c. Interferon-regulated influenza virus resistance gene Mx is localized on mouse chromosome 16. *J Virol*. 58:967-9.
Staeheli, P., and J.G. Sutcliffe. 1988b. Identification of a second interferon-regulated murine Mx gene. *Mol Cell Biol*. 8:4524-8.
Stark, G.R., I.M. Kerr, B.R. Williams, R.H. Silverman, and R.D. Schreiber. 1998. How cells respond to interferons. *Annu Rev Biochem*. 67:227-64.
Stenmark, H., R.G. Parton, O. Steele-Mortimer, A. Lutcke, J. Gruenberg, and M. Zerial. 1994. Inhibition of rab5 GTPase activity stimulates membrane fusion in endocytosis. *Embo J*. 13:1287-96.
Stertz, S., M. Reichelt, J. Krijnse-Locker, J. Mackenzie, J.C. Simpson, O. Haller, and G. Kochs. 2006. Interferon-induced, antiviral human MxA protein localizes to a distinct subcompartment of the smooth

endoplasmic reticulum. *J Interferon Cytokine Res*. 26:650-60.

Stickney, J.T., and J.E. Buss. 2000. Murine guanylate-binding protein: incomplete geranylgeranyl isoprenoid modification of an interferon-gamma-inducible guanosine triphosphate-binding protein. *Mol Biol Cell*. 11:2191-200.

Stojdl, D.F., N. Abraham, S. Knowles, R. Marius, A. Brasey, B.D. Lichty, E.G. Brown, N. Sonenberg, and J.C. Bell. 2000. The murine double-stranded RNA-dependent protein kinase PKR is required for resistance to vesicular stomatitis virus. *J Virol*. 74:9580-5.

Stotz, A., P.P. Muller, and P. Linder. 1993. Regulation of the ADE2 gene from Saccharomyces cerevisiae. *Curr Genet*. 24:472-80.

Stowell, M.H., B. Marks, P. Wigge, and H.T. McMahon. 1999. Nucleotide-dependent conformational changes in dynamin: evidence for a mechanochemical molecular spring. *Nat Cell Biol*. 1:27-32.

Stranden, A.M., P. Staeheli, and J. Pavlovic. 1993. Function of the mouse Mx1 protein is inhibited by overexpression of the PB2 protein of influenza virus. *Virology*. 197:642-51.

Stremlau, M., C.M. Owens, M.J. Perron, M. Kiessling, P. Autissier, and J. Sodroski. 2004. The cytoplasmic body component TRIM5alpha restricts HIV-1 infection in Old World monkeys. *Nature*. 427:848-53.

Stremlau, M., M. Perron, M. Lee, Y. Li, B. Song, H. Javanbakht, F. Diaz-Griffero, D.J. Anderson, W.I. Sundquist, and J. Sodroski. 2006. Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5alpha restriction factor. *Proc Natl Acad Sci U S A*. 103:5514-9.

Studier, F.W., and B.A. Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*. 189:113-30.

Su, C., D. Evans, R.H. Cole, J.C. Kissinger, J.W. Ajioka, and L.D. Sibley. 2003. Recent expansion of Toxoplasma through enhanced oral transmission. *Science*. 299:414-6.

Sun, Y.J., F. Forouhar, H.M. Li Hm, S.L. Tu, Y.H. Yeh, S. Kao, H.L. Shr, C.C. Chou, C. Chen, and C.D. Hsiao. 2002. Crystal structure of pea Toc34, a novel GTPase of the chloroplast protein translocon. *Nat Struct Biol*. 9:95-100.

Suss-Toby, E., J. Zimmerberg, and G.E. Ward. 1996. Toxoplasma invasion: the parasitophorous vacuole is formed from host cell plasma membrane and pinches off via a fission pore. *Proc Natl Acad Sci U S A*. 93:8413-8.

Suzuki, Y., M.A. Orellana, R.D. Schreiber, and J.S. Remington. 1988. Interferon-gamma: the major mediator of resistance against Toxoplasma gondii. *Science*. 240:516-8.

Sweitzer, S.M., and J.E. Hinshaw. 1998. Dynamin undergoes a GTP-dependent conformational change causing vesiculation. *Cell*. 93:1021-9.

Takaoka, A., Y. Mitani, H. Suemori, M. Sato, T. Yokochi, S. Noguchi, N. Tanaka, and T. Taniguchi. 2000. Cross talk between interferon-gamma and -alpha/beta signaling components in caveolar membrane domains. *Science*. 288:2357-60.

Taylor, G.A. 2007. IRG proteins: key mediators of interferon-regulated host resistance to intracellular pathogens. *Cell Microbiol*. 9:1099-107.

Taylor, G.A., C.M. Collazo, G.S. Yap, K. Nguyen, T.A. Gregorio, L.S. Taylor, B. Eagleson, L. Secrest, E.A. Southon, S.W. Reid, L. Tessarollo, M. Bray, D.W. McVicar, K.L. Komschlies, H.A. Young, C.A. Biron, A. Sher, and G.F. Vande Woude. 2000. Pathogen-specific loss of host resistance in mice lacking the IFN-gamma-inducible gene IGTP. *Proc Natl Acad Sci U S A*. 97:751-5.

Taylor, G.A., C.G. Feng, and A. Sher. 2004. p47 GTPases: regulators of immunity to intracellular pathogens. *Nat Rev Immunol*. 4:100-9.

Taylor, G.A., M. Jeffers, D.A. Largaespada, N.A. Jenkins, N.G. Copeland, and G.F. Woude. 1996. Identification of a novel GTPase, the inducibly expressed GTPase, that accumulates in response to interferon gamma. *J Biol Chem*. 271:20399-405.

Taylor, G.A., R. Stauber, S. Rulong, E. Hudson, V. Pei, G.N. Pavlakis, J.H. Resau, and G.F. Vande Woude. 1997. The inducibly expressed GTPase localizes to the endoplasmic reticulum, independently of GTP binding. *J Biol Chem*. 272:10639-45.

Taylor, S., A. Barragan, C. Su, B. Fux, S.J. Fentress, K. Tang, W.L. Beatty, H.E. Hajj, M. Jerome, M.S. Behnke, M. White, J.C. Wootton, and L.D. Sibley. 2006. A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen Toxoplasma gondii. *Science*. 314:1776-80.

Terskikh, A.V., M.C. Easterday, L. Li, L. Hood, H.I. Kornblum, D.H. Geschwind, and I.L. Weissman. 2001. From hematopoiesis to neuropoiesis: evidence of overlapping genetic programs. *Proc Natl Acad Sci U S A*. 98:7934-9.

Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 22:4673-80.

Tripal, P., M. Bauer, E. Naschberger, T. Mortinger, C. Hohenadl, E. Cornali, M. Thurau, and M. Sturzl. 2007. Unique features of different members of the human guanylate-binding protein family. *J Interferon Cytokine Res*. 27:44-52.

Trowsdale, J., R. Barten, A. Haude, C.A. Stewart, S. Beck, and M.J. Wilson. 2001. The genomic context of natural killer receptor extended gene families. *Immunol Rev*. 181:20-38.

Tucker, J., G. Sczakiel, J. Feuerstein, J. John, R.S. Goody, and A. Wittinghofer. 1986. Expression of p21 proteins in Escherichia coli and stereochemistry of the nucleotide-binding site. *Embo J*. 5:1351-8.

Tuma, P.L., and C.A. Collins. 1994. Activation of dynamin GTPase is a result of positive cooperativity. *J Biol Chem*. 269:30842-7.

Tuma, P.L., M.C. Stachniak, and C.A. Collins. 1993. Activation of dynamin GTPase by acidic phospholipids and endogenous rat brain vesicles. *J Biol Chem*. 268:17240-6.

Turan, K., M. Mibayashi, K. Sugiyama, S. Saito, A. Numajiri, and K. Nagata. 2004. Nuclear MxA proteins form a complex with influenza virus NP and inhibit the transcription of the engineered influenza virus genome. *Nucleic Acids Res*. 32:643-52.

Turco, J., and H.H. Winkler. 1986. Gamma-interferon-induced inhibition of the growth of Rickettsia prowazekii in fibroblasts cannot be explained by the degradation of tryptophan or other amino acids. *Infect Immun*. 53:38-46.

Turelli, P., B. Mangeat, S. Jost, S. Vianin, and D. Trono. 2004. Inhibition of hepatitis B virus replication by APOBEC3G. *Science*. 303:1829.

Urrutia, R., J.R. Henley, T. Cook, and M.A. McNiven. 1997. The dynamins: redundant or distinct functions for an expanding family of related GTPases? *Proc Natl Acad Sci U S A*. 94:377-84.

Uthaiah, R.C. 2002. Biochemical, structural and cellular studies on IIGP1, a member of the p47 family of GTPases. PhD thesis. Institute for Genetics, University of Cologne, Cologne.

Uthaiah, R.C., G.J. Praefcke, J.C. Howard, and C. Herrmann. 2003. IIGP1, an interferon-gamma-inducible 47-kDa GTPase of the mouse, showing cooperative enzymatic activity and GTP-dependent multimerization. *J Biol Chem*. 278:29336-43.

Uze, G., and D. Monneron. 2007. IL-28 and IL-29: newcomers to the interferon family. *Biochimie*. 89:729-34.

Vaishnava, S., and B. Striepen. 2006. The cell biology of secondary endosymbiosis--how parasites build, divide and segregate the apicoplast. *Mol Microbiol*. 61:1380-7.

Valente, G., L. Ozmen, F. Novelli, M. Geuna, G. Palestro, G. Forni, and G. Garotta. 1992. Distribution of interferon-gamma receptor in human tissues. *Eur J Immunol*. 22:2403-12.

Van Criekinge, W., and R. Beyaert. 1999. Yeast Two-Hybrid: State of the Art. *Biol Proced Online*. 2:1-38.

van der Bliek, A.M. 1999. Is dynamin a regular motor or a master regulator? *Trends Cell Biol*. 9:253-4.

van Pesch, V., H. Lanaya, J.C. Renauld, and T. Michiels. 2004. Characterization of the murine alpha interferon gene family. *J Virol*. 78:8219-28.

Vazquez-Torres, A., and F.C. Fang. 2001. Oxygen-dependent anti-Salmonella activity of macrophages. *Trends Microbiol*. 9:29-33.

Vegeto, E., G.F. Allan, W.T. Schrader, M.J. Tsai, D.P. McDonnell, and B.W. O'Malley. 1992. The mechanism of RU486 antagonism is dependent on the conformation of the carboxy-terminal tail of the human progesterone receptor. *Cell*. 69:703-13.

Venezia, T.A., A.A. Merchant, C.A. Ramos, N.L. Whitehouse, A.S. Young, C.A. Shaw, and M.A. Goodell. 2004. Molecular signatures of proliferation and quiescence in hematopoietic stem cells. *PLoS Biol*. 2:e301.

Vestal, D.J., J.E. Buss, G.S. Kelner, D. Maciejewski, V.K. Asundi, and R.A. Maki. 1996. Rat p67 GBP is induced by interferon-gamma and isoprenoid-modified in macrophages. *Biochem Biophys Res Commun*. 224:528-34.

Vestal, D.J., J.E. Buss, S.R. McKercher, N.A. Jenkins, N.G. Copeland, G.S. Kelner, V.K. Asundi, and R.A. Maki. 1998. Murine GBP-2: a new IFN-gamma-induced member of the GBP family of GTPases isolated from macrophages. *J Interferon Cytokine Res*. 18:977-85.

Vestal, D.J., V.Y. Gorbacheva, and G.C. Sen. 2000. Different subcellular localizations for the related interferon-induced GTPases, MuGBP-1 and MuGBP-2: implications for different functions? *J Interferon Cytokine Res*. 20:991-1000.

Vetter, I.R., and A. Wittinghofer. 2001. The guanine nucleotide-binding switch in three dimensions. *Science*. 294:1299-304.

Vidal, S.M., D. Malo, K. Vogan, E. Skamene, and P. Gros. 1993. Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. *Cell*. 73:469-85.

Wang, Y., B.W. O'Malley, Jr., S.Y. Tsai, and B.W. O'Malley. 1994. A regulatory system for use in gene transfer. *Proc Natl Acad Sci U S A*. 91:8180-4.

Watanabe, T. 2007. Polymorphisms of the chicken antiviral MX gene. *Cytogenet Genome Res*. 117:370-5.

Wegner, M. 1999. From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res*. 27:1409-20.
Weiss, L.M., and K. Kim. 2000. The development and biology of bradyzoites of Toxoplasma gondii. *Front Biosci*. 5:D391-405.
Welchman, R.L., C. Gordon, and R.J. Mayer. 2005. Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat Rev Mol Cell Biol*. 6:599-609.
Whelan, S., and N. Goldman. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*. 18:691-9.
Wickerham, L. 1951. Taxonomy of yeasts. 1. Techniques of classification. 2. A classification of the genus Han.senula. *US Dept Agric Tech Bull*. 1029:1-56.
Wildfuhr, W. 1966. [Electron microscopic studies on the morphology and reproduction of Toxoplasma gondii. II. Observations on the reproduction of Toxoplasma gondii (endodyogeny)]. *Zentralbl Bakteriol [Orig]*. 201:110-30.
Wittinghofer, A., and E.F. Pai. 1991. The structure of Ras protein: a model for a universal molecular switch. *Trends Biochem Sci*. 16:382-7.
Wong, R.T., C.C. Hon, F. Zeng, and F.C. Leung. 2007. Screening of differentially expressed transcripts in infectious bursal disease virus-induced apoptotic chicken embryonic fibroblasts by using cDNA microarrays. *J Gen Virol*. 88:1785-96.
Wong, S.K., and D.W. Lazinski. 2002. Replicating hepatitis delta virus RNA is edited in the nucleus by the small form of ADAR1. *Proc Natl Acad Sci U S A*. 99:15118-23.
WTCCC, and T.W.T.C.C. Consortium. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 447:661-78.
Wynn, T.A., C.M. Nicolet, and D.M. Paulnock. 1991. Identification and characterization of a new gene family induced during macrophage activation. *J Immunol*. 147:4384-92.
Yamaguchi, T., N. Omatsu, A. Omukae, and T. Osumi. 2006. Analysis of interaction partners for perilipin and ADRP on lipid droplets(*). *Mol Cell Biochem*. 284:167-73.
Yang, Y.L., L.F. Reis, J. Pavlovic, A. Aguzzi, R. Schafer, A. Kumar, B.R. Williams, M. Aguet, and C. Weissmann. 1995. Deficient signaling in mice devoid of double-stranded RNA-dependent protein kinase. *Embo J*. 14:6095-106.
Yap, G.S., and A. Sher. 1999a. Cell-mediated immunity to Toxoplasma gondii: initiation, regulation and effector function. *Immunobiology*. 201:240-7.
Yap, G.S., and A. Sher. 1999b. Effector cells of both nonhemopoietic and hemopoietic origin are required for interferon (IFN)-gamma- and tumor necrosis factor (TNF)-alpha-dependent host resistance to the intracellular pathogen, Toxoplasma gondii. *J Exp Med*. 189:1083-92.
Young, H.A., and J.H. Bream. 2007. IFN-gamma: recent advances in understanding regulation of expression, biological functions, and clinical applications. *Curr Top Microbiol Immunol*. 316:97-117.
Zerrahn, J., U.E. Schaible, V. Brinkmann, U. Guhlich, and S.H. Kaufmann. 2002. The IFN-inducible Golgi- and endoplasmic reticulum- associated 47-kDa GTPase IIGP is transiently expressed during listeriosis. *J Immunol*. 168:3428-36.
Zhang, B., Y. Zhang, Z. Wang, and Y. Zheng. 2000. The role of Mg2+ cofactor in the guanine nucleotide exchange and GTP hydrolysis reactions of Rho family GTP-binding proteins. *J Biol Chem*. 275:25299-307.
Zhang, H.M., J. Yuan, P. Cheung, H. Luo, B. Yanagawa, D. Chau, N. Stephan-Tozy, B.W. Wong, J. Zhang, J.E. Wilson, B.M. McManus, and D. Yang. 2003. Overexpression of interferon-gamma-inducible GTPase inhibits coxsackievirus B3-induced apoptosis through the activation of the phosphatidylinositol 3-kinase/Akt pathway and inhibition of viral replication. *J Biol Chem*. 278:33011-9.
Zeng, J. 2007. Interferon-induced and constitutive expression of immunity-related GTPases (IRG) in mouse tissues. PhD thesis. Institute for Genetics, University of Cologne, Cologne.
Zheng, J., D.R. Knighton, N.H. Xuong, S.S. Taylor, J.M. Sowadski, and L.F. Ten Eyck. 1993. Crystal structures of the myristylated catalytic subunit of cAMP-dependent protein kinase reveal open and closed conformations. *Protein Sci*. 2:1559-73.
Zhou, Y., Y. Wang, M. Kovacs, J. Jin, and J. Zhang. 2005. Microglial activation induced by neurodegeneration: a proteomic analysis. *Mol Cell Proteomics*. 4:1471-9.
Zurcher, T., J. Pavlovic, and P. Staeheli. 1992a. Mechanism of human MxA protein action: variants with changed antiviral properties. *Embo J*. 11:1657-61.
Zurcher, T., J. Pavlovic, and P. Staeheli. 1992b. Mouse Mx2 protein inhibits vesicular stomatitis virus but not influenza virus. *Virology*. 187:796-800.
Zurcher, T., J. Pavlovic, and P. Staeheli. 1992c. Nuclear localization of mouse Mx1 protein is necessary for inhibition of influenza virus. *J Virol*. 66:5059-66.

VII. Summary

Members of the immunity-related (p47) GTPases (IRGs) are essential, interferon-inducible resistance factors active against a broad spectrum of important intracellular bacterial and protozoal pathogens including *Toxoplasma gondii*. Despite the strong, partially overlapping but non-redundant susceptibility phenotypes of mice genetically deficient for individual members of the family, little is known about molecular mechanisms of function and regulation of the IRG resistance system.

This study demonstrates that IRG GTPases function in a system of direct, nucleotide-dependent regulatory interactions between family members. Spontaneous GTP-dependent homo-oligomerisation of the GKS subfamily member Irga6 leads to the activation of GTPase function *in vitro. In vivo,* accumulation of activated, GTP-bound GKS proteins at the parasitophorous vacuolar membrane is associated with the IFN-induced, cell autonomous destruction of avirulent *Toxoplasma gondii*. In contrast, direct GDP-dependent interactions of the unusual GMS subfamily members with the GKS proteins Irga6 and Irgb6 via the G-domains prevent premature activation of the resistance system in absence of infection. Lack of GMS GTPases results in spontaneous activation and aggregation of Irga6 and Irgb6 on endomembranes and hinders association with the parasitophorous vacuolar membrane upon infection with *T. gondii*. The three GMS GTPases are both necessary and sufficient to regulate Irga6 and Irgb6. No other interferon-inducible proteins are required for the regulation of the resting localisation of Irga6 and Irgb6 in the IFN-induced cell or for their infection-induced assembly on the membrane of the parasitophorous vacuole. Nucleotide binding is essential for the translocation from the resting localisation to the vacuole. Virulent *Toxoplasma* efficiently inhibited the recruitment of IRG proteins to the parasitophorous vacuole. These findings provide the first link between the enzymatic properties of IRG proteins as GTPases and their function in pathogen resistance.

The IRG resistance GTPases are an ancient family that underwent extensive expansion and diversification as well as contraction in the euchordates – a feature characteristic for multigene families associated with pathogen resistance due to host-pathogen coevolution. Despite its essential role in resistance to vacuolar pathogens in mice, the IRG resistance system as such is not conserved in higher primates. While all other groups of mammals possess multiple divergent IRG genes, humans and higher primates contain a largely reduced set of typically three genes. All of these genes lack IFN-inducible elements in their putative promoters and are either significantly truncated, drastically damaged in the nucleotide binding domain or show testis-specific expression paralleling sexual maturity. Consequently, mice and humans must deploy their immune resources against vacuolar pathogens in radically different ways.

VIII. Zusammenfassung

Die Mitglieder der (p47) immun-verwandten GTPasen (‘immune-related GTPases’, IRGs) sind essentielle, Interferon-induzierbare Resistenzfaktoren und aktiv gegen ein breites Spektrum bedeutender, intrazellulärer Pathogene einschließlich *Toxoplasma gondii*. Obwohl Mäuse, die genetisch defizient für einzelne Mitglieder des IRG Resistenzsystems sind, ausgeprägte, teilweise überlappende aber nicht redundante Anfälligkeiten gegenüber Pathogenen aufweisen, ist wenig über die molekularen Mechanismen der Funktion und der Regulation des IRG Resistenzsystems bekannt.

Diese Studie zeigt, dass die IRG GTPasen in einem System von direkten, Nukleotid-abhängigen Interaktionen zwischen Familienmitgliedern funktionieren. Spontane GTP-abhängige Oligomerisierung von Irga6, einem Mitglied der GKS Unterfamilie, führt zur Aktivierung der GTPase Funktion *in vitro*. Die Akkumulation von aktivierten, GTP-gebundenen GKS Proteinen an der Membran der parasitophoren Vakuole *in vivo* ist mit der IFN-induzierten, zellautonomen Zerstörung von avirulenten *Toxoplasma gondii* assoziiert. Direkte GDP-abhängige Interaktionen von Mitgliedern der ungewöhnlichen GMS Unterfamilie mit den GKS Proteinen mittels der G Domänen verhindert dagegen die verfrühte Aktivierung des Resistenzsystems in Abwesenheit von Infektion. Die Abwesenheit von GMS GTPasen resultiert in der spontanen Aktivierung und Aggregation von Irga6 und Irgb6 an Endomembranen und behindert die Assoziation mit der Membran der parasitophoren Vakuole im Zuge der Infektion mit *T. gondii*. Die drei GMS GTPasen sind sowohl notwendig als auch hinreichend für die Regulation von Irga6 und Irgb6. Keine anderen Interferon-induzierbaren Proteine sind für die Infektions-induzierte Assemblierung an der Membran der parasitophoren Vakuole notwendig. Die Nukleotid-Bindung ist essentiell für den Transfer von der Ruhelokalisation zur Vakuole. Virulente Toxoplasmen inhibieren die Rekrutierung der IRG Proteine an die parasitophore Vakuole. Diese Ergebnisse stellen die erste Verbindung zwischen den enzymatischen Eigenschaften der IRG Proteine als GTPasen und ihrer Funktion in der Resistenz gegen Pathogene her.

Die IRG Resistenz-GTPasen sind eine alte Familie die in den Euchordaten eine extensive Expansion, Diversifikation und Kontraktion durchlaufen hat – eine Eigenschaft, die aufgrund von Wirts-Pathogen-Koevolution charakteristisch für Pathogenresistenz-assozierte Multigenfamilien ist. Trotz ihrer essentiellen Rolle in der Resistenz gegen vakuoläre Pathogene ist das IRG Resistenzsystem als solches in höheren Primaten nicht konserviert. Während alle anderen Gruppen von Säugetieren mehrere, divergente IRG Gene besitzen, haben Menschen und höhere Primaten einen stark reduzierten Satz von typischerweise drei Genen. Alle diese Gene besitzen keine Interferon-induzierbaren Elemente in ihren mutmaßlichen Promotoren und sind entweder verkürzt, stark beschädigt in ihrer Nukleotidbindungsdomäne oder zeigen eine Testis-spezifische Expression parallel zu sexuellen Reife. Deshalb müssen Mensch und Maus ihre Immunressourcen gegen vakuoläre Pathogene radikal unterschiedlich nutzen.

IX. Acknowledgements

First of all I would like to thank Prof. Dr. Jonathan C. Howard for his support, advice and criticism and the fruitful discussions about my project that started with obscure aggregates and evolved to a really fascinating story.

Prof. Dr. Thomas Langer danke ich für die Übernahme des Zweitgutachtens dieser Arbeit und Prof. Dr. Siegfried Roth für die Übernahme des Vorsitzes der Prüfungskommission.

Prof. Dr. Mats Paulsson und Prof. Dr. Dagmar Knebel-Mörsdorf möchte ich für ihre Mitwirkung in meinem Thesis Committee und für ihre Anregungen danken.

Prof. Dr. Otto Haller, PD Dr. Georg Kochs und ihrer Arbeitsgruppe möchte ich dafür danken, dass sie mich so freundlich aufgenommen und mich in die Welt der Viren eingeführt haben.

My thanks to the International Graduate School in Genetics and Functional Genomics for funding.

PD Dr. Gaby Reichmann danke ich für Ihre Kooperation beim Aufbau des *Toxoplasma*-Labors.

Dr. Matthias Kramer danke ich für seine unschätzbare Hilfe bei bürokratischen Anbelangen.

Stephanie Könen-Waisman danke ich für die produktive Zusammenarbeit am induzierbaren Expressionssystem und die vielen kurzweiligen Stunden, in denen wir fast dieselbe Pipette gehalten haben.

Rita Lange möchte ich für Ihre Ratschläge, Ihr „goldenes Händchen“ beim Klonieren und Ihr stets offenes Ohr danken sowie dafür, dass Sie mich adoptiert hat.

Claudia Poschner danke ich für ihre Hilfe und die langen gemeinsamen Stunden in der Zellkultur.

I want to thank Nataša Papić for working, eating and suffering together and for making sense of more and more protein bands.

Niko Pawlowski danke ich für die Einführung in die Welt der Proteinbiochemie und der Kristallstrukturen.

Christoph Rohde für seinen unerschütterlichen Optimismus, den gemeinsamen Ausflug in die Welt der Hefen und den fruchtbaren Austausch von neu entdeckten IRGC Sequenzen.

Cemalettin Bekpen for working hand in hand on various genome databases and for long, lively and encouraging discussions with or without “late workers motivator”.

All the former and current members of Jonathan’s lab for numerous reasons.

Gerrit Praefcke für den Schreibtisch im „Pumakäfig“, für sein unerschöpfliches Wissen und für das Korrekturlesen dieser Arbeit.

Julia Fres für das Korrekturlesen dieser Arbeit und für Ihre Unterstützung.

Kirstin Keusekotten für die „last minute“ Hilfe.

Stefan Weißhaar möchte ich ganz besonders für seine unschätzbare Hilfe bei Computerproblemen und beim Formatieren dieser Arbeit danken.

Allen Mitarbeitern der zweiten Etage danke ich für die wunderbare Arbeitsatmosphäre.

Viktor Steimle und Felix Schnappauf dafür, dass sie mich in die Laborpraxis eingeführt und die Faszination und die Liebe für die Wissenschaft in mir geweckt haben.

Maurijn van der Zee wil ik voor zijn hulp en geduld en voor al die andere dingen danken.

Ich danke jenen Menschen außerhalb des Labors, die mich unterstützen und immer für mich da sind dafür, dass sie eine so unerschöpfliche Quelle der Freude sind.

Ich danke meiner Familie für Ihre Unterstützung in so vielerlei Hinsicht.

X. Kollaborationen

Die Mifepristone-induzierbaren gs3T3-IRG Zelllinien (außer der zweiten Transfektion von Irgb6 und den gs3T3-Irgm2 Zellen) wurden in Zusammenarbeit mit Stephanie Könen-Waisman generiert.

Die Dokumentation der direkte Nukleotid-abhängige Interaktion von zellulärem Irgm3 mit Irga6 wurde in Kollaboration mit Natasa Papic durchgeführt (Kapitel III.2.10, siehe auch (Papic 2007)).

Nina Schöder war im Zuge Ihrer Diplomarbeit unter meiner Betreunung an der Generierung der Daten zur Relokalisation von Irga6 und Irgb6 in uninfizierten Zellen durch Kotransfektion von Expressionkonstrukter der anderen IRG Proteine beteiligt (entsprechende Teile des Kapitels III.2.8).

XI. Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation mit Ausnahme der im Kapitel X. (Kollaborationen) aufgeführten Teile selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt.

Die von mir vorgelegte Dissertation ist von Prof. Dr. Jonathan C. Howard betreut worden.

Köln, Dezember 2007

Teilpublikationen:

Bekpen, C., **J.P. Hunn**, C. Rohde, I. Parvanova, L. Guethlein, D.M. Dunn, E. Glowalla, M. Leptin, and J.C. Howard. 2005. The interferon-inducible p47 (IRG) GTPases in vertebrates: loss of the cell autonomous resistance mechanism in the human lineage. *Genome Biol*. 6:R92.

J.P. Hunn, S. Könen-Waisman, N. Papic, N. Schröder, N. Pawlowski, R. Lange, F. Kaiser, J. Zerrahn, S. Martens, J.C. Howard. Regulatory interactions between IRG resistance GTPases in the cellular response to *Toxoplasma gondii.* Manuscript in preparation.

XII. Lebenslauf

Persönliche Daten

Name:	Julia Hunn
Geburtstag:	12.05.1975
Geburtsort:	Singen am Hohentwiel
Anschrift:	Nauheimer Str. 5, 50969 Köln
Staatsangehörigkeit:	Deutsch

Schulbildung

1981 - 1985	Beethoven-Grundschule in Singen
1985 - 1994 Singen	Naturwissenschaftliches Friedrich-Wöhler-Gymnasium
Juni 1994	Schulabschluss mit allgemeiner Hochschulreife

Hochschulbildung

Okt. 1994 – April 2001	Studium der Biologie an der Albert-Ludwigs-Universität Freiburg
Okt. 1996	Abschluss des Grundstudiums mit dem Vordiplom
Juli – Sept. 1998	Dreimonatiges Volontariat bei der „Desert Research Foundation of Namibia“ in Gobabeb, Namibia
Dez. 1999	Diplomprüfung im Hauptfach molekulare Immunologie und den Nebenfächern Neurobiologie, Verhaltensbiologie und klinische Neuropsychologie
Jan. 2000 – Jan. 2001	Diplomarbeit am Max-Planck-Institut für Immunbiologie in Freiburg bei Prof. Dr. Michael Reth in der Abteilung Dr. Viktor Steimle zum Thema „Untersuchungen zur differentiellen Genexpression von CD4 und CD8 T-Zellen und zur Repression des CD4 Gens“
April 2001	Studienabschluss mit dem Diplom in Biologie
Okt. 2001	Beginn des Promotionsstudiums in Biologie (Fachrichtung Genetik) als Mitglied der „International Graduate School in Genetics and Functional Genomics“ der Universität zu Köln
Okt. 2001 – März 2002	Rotationsperioden in den Laboren von Prof. Dr. Martin Scheffner, Prof. Dr. Mats Paulsson und Prof. Dr. Jonathan Howard
April 2002	Beginn der Arbeit am Promotionsthema am Institut für Genetik der Universität zu Köln in der Abteilung von Prof. Dr. Jonathan Howard zum Thema „Evolution and cellular resistance mechanisms of the immunity-related GTPases“