GENETIC ANALYSIS OF EQUINE 2',5'-OLIGOADENYLATE SYNTHETASE (*OAS1*) AND RIBONUCLEASE L (*RNASEL*) POLYMORPHISMS AND ASSOCIATION TO SEVERE WEST NILE VIRUS DISEASE

A Dissertation

by

JONATHAN JOSEPH RIOS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2008

Major Subject: Genetics

GENETIC ANALYSIS OF EQUINE 2',5'-OLIGOADENYLATE SYNTHETASE (*OAS1*) AND RIBONUCLEASE L (*RNASEL*) POLYMORPHISMS AND ASSOCIATION TO SEVERE WEST NILE VIRUS DISEASE

A Dissertation

by

JONATHAN JOSEPH RIOS

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	David L. Adelson
	Thomas E. Spencer
Committee Members,	Bhanu P. Chowdhary
	John C. Huber, Jr.
	Loren C. Skow
Chair of Genetics Faculty,	James R. Wild

May 2008

Major Subject: Genetics

ABSTRACT

Genetic Analysis of Equine 2',5'-Oligoadenylate Synthetase (*OAS1)* and Ribonuclease L (*RNASEL*) Polymorphisms and Association to Severe West Nile Virus Disease.

(May 2008)

Jonathan Joseph Rios, B.S., Tarleton State University

Co-Chairs of Advisory Committee: Dr. David L. Adelson
Dr. Thomas E. Spencer

West Nile virus (WNV), a member of the *Flaviviridae* family of RNA viruses, was first introduced to the United States in 1999 with rapid transmission across a variety of hosts throughout the continental states. Genetic research to identify genes involved in resistance and susceptibility to WNV began in mice, where it was observed that natural populations were resistant or fatally susceptible. Further investigation led to the identification of the *Flavivirus* resistance gene as the oligoadenylate synthetase 1b gene in mice. A nonsense mutation was found within the coding region of this gene that associated absolutely with susceptibility to WNV.

A two-stage association study was conducted to identify similar genetic associations to West Nile encephalitis in naturally susceptible and resistant populations of horses in the United States.

Genomic sequence of a majority of the equine 2',5'-oligoadenylate synthetase 1 (*OAS1*) gene was assembled by shotgun-sequencing CHORI BAC 100:I10 (3.95X). A contig map spanning the entire gene was constructed, including 8 kilobases of promoter

sequence upstream of the first exon. Coding regions of equine *OAS1* and ribonuclease L (*RNASEL*) genes, as well as the *OAS1* promoter, were screened for mutations from a random sample of horses of multiple breeds. Numerous polymorphisms were identified for case-control analyses. Analysis using Fisher's Exact test identified allelic and genotypic associations. Odds ratios were also determined to measure strength of the associations. Case-control analysis of haplotype frequencies identified significant differences in haplotype frequencies between populations and association to West Nile encephalitis.

A conserved interferon-stimulated response element was mapped to within 518 basepairs upstream of the transcription start site of *OAS1*. Promoter polymorphisms were not found to affect induction by interferon-tau; however, additional analyses are necessary to further characterize the equine *OAS1* promoter and the host factors involved in regulating expression.

Statistical analyses of the genotype data from the case and control populations identified significant associations between polymorphisms of the *OAS1* and *RNASEL* genes with severe West Nile encephalitis. The similarity between human and horse *OAS* immunity genes suggests that the horse may provide a genetic model to further investigate mammalian SNP-associated viral susceptibility.

ACKNOWLEDGEMENTS

My estimation of a proper acknowledgement requires a sincere demonstration of my appreciation for all those who contributed in any way to my success in completing this work. At the forefront are my family and friends, in particular Renee and Don Lawhorne, Leo and Jaunie Rios, my brother Christopher and my fiancée Carrie Tibbals. Together, they supported me throughout my undergraduate and graduate studies and have facilitated the means by which I have been able to succeed in my accomplishments. I would like to thank each of them for their love and support and pray that I can continually express my gratitude for everything they have given me.

This work, in its entirety, would not have been possible without the leadership of Dr. David Adelson as well as the support and guidance of my entire Graduate Committee. Their ability to understand my goals and to challenge me as a student and scientist has driven me to complete this study and has given me the skill-set necessary to continue setting challenging personal and professional goals. For your support over the past four years, thank you.

The seemingly endless laboratory hours necessary to complete this study would not have been possible without the support of fellow lab mates. I would like to thank Mrs. Colette Abbey for her tireless support, encouragement and guidance. It has been my goal to constantly express my gratitude for her efforts and her ability to effectively manage the lab. I cannot thank her enough.

Thank you all.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

CHAPTER I

INTRODUCTION: UNDERSTANDING EQUINE WEST NILE VIRUS SUSCEPTIBILITY

Objective

The emergence of naturally variable levels of resistance to West Nile virus (WNV) among multiple species has prompted investigation of the equine innate immune system. The 2',5'-oligoadenylate synthetase (*OAS*)/ribonuclease L (*RNASEL*) system of viral immunity emerged as a leading candidate responsible for determining innate levels of resistance and susceptibility to WNV infection. This research investigates the potential roles *OAS1* and *RNASEL* play in the development of naturally susceptible and resistant populations of horses to severe WNV infection. Assembly of the genomic sequence of these genes allowed for the identification of single nucleotide polymorphisms (SNPs) from a random sample of horses. Associations of SNPs or SNP haplotypes between naturally susceptible and resistant populations of horses to WNV were measured using multiple statistical methods. Accompanying the sequence-dependent role of the equine *OAS/RNASEL* system in WNV susceptibility, this investigation examined the functional importance of a promoter microsatellite during interferon (IFN) stimulation of the equine *OAS1* proximal promoter. Together, this investigation proposes an equine model of SNP-dependent susceptibility to severe West Nile encephalitis.

This dissertation follows the style of *BMC Genomics*.

Present Status of the Question

The cellular response to viral infection represents an intricate convergence of many pathways involving hundreds of genes working to suppress and prevent viral propagation within and between neighboring cells. The genetic response is categorized by early (within cell) and delayed (between cells) cellular activity [1]. Two pathways, described as either toll-like receptor (*TLR*)-dependent or *TLR*-independent, mediate early viral recognition. Many TLR receptors recognize different viral products within the cell; however, TLR3 specifically recognizes viral double-stranded RNA (dsRNA) and is responsible, in part, for activating the *TLR*-dependent immune response (Figure 1.1). Both the *TLR*-dependent and independent modes of early response culminate in the activation of Type I interferon, *IFNA* and *IFNB*.

The secreted IFN binds the receptors of neighboring cells to initiate a complex cascade whose interferon-stimulated genes (ISGs) represent the anti-viral repertoire of the host immune response. Interferon acts as an inter-cellular signal whose role is to activate many downstream genetic cascades with the shared purpose of limiting viral replication at the levels of transcription and translation (Figure 1.2) [2].

However, with the wide array of host genetic factors limiting viral infection, viruses have evolved mechanisms to counteract the antiviral activity of specific host genes (Table 1.1).

With natural resistance to WNV infection, domestic mammals act as dead-end hosts during viral transmission [3]. Sheep infected with WNV developed neutralizing

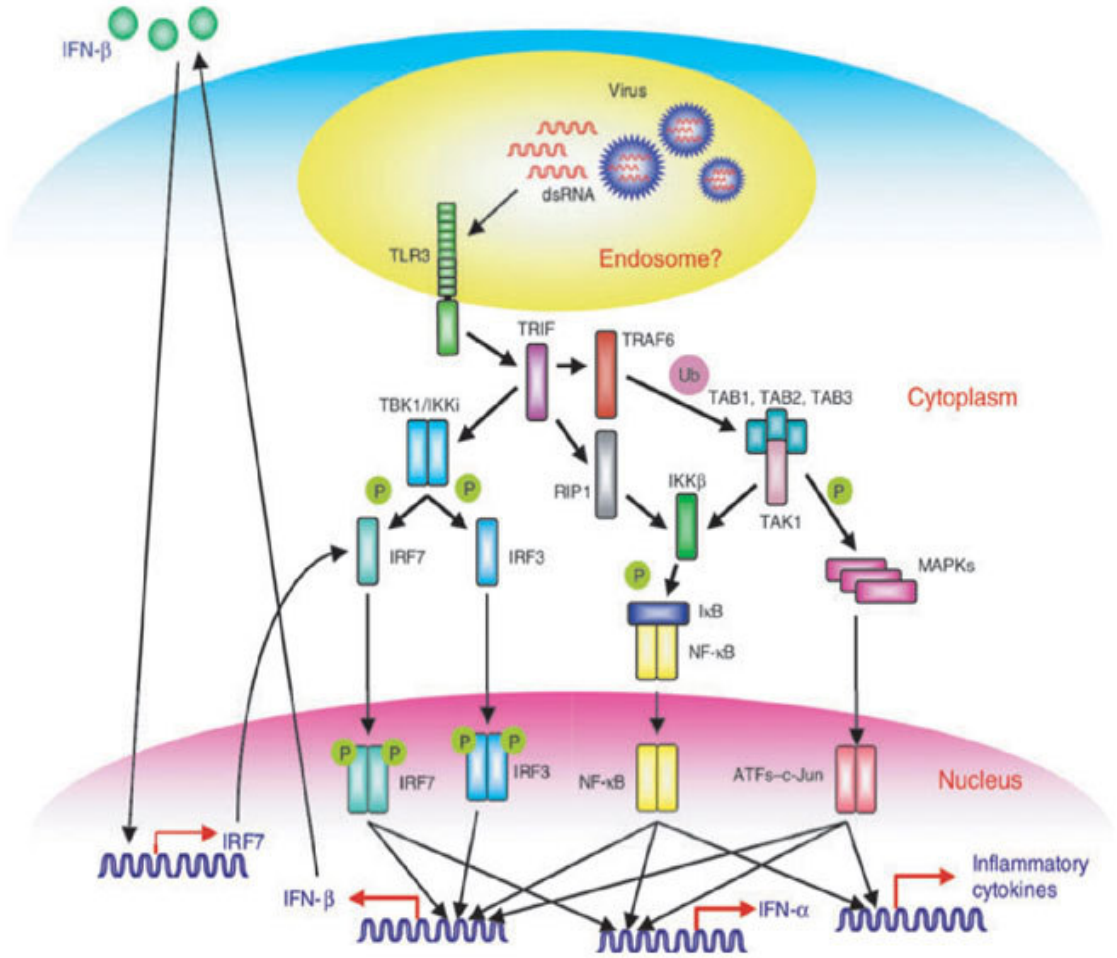


Figure 1.1 TLR3-dependent signaling pathway (*Source*: Kawai, 2006).
Viral dsRNA is recognized by TLR3 and activates the nuclear localization of IRF and NF-κB transcription factors by signaling through TICAM1 (TRIF). Nuclear transcription factors activate IFN expression with positive feedback signals through IRF7.

antibodies and suffered a moderate febrile reaction and low-grade viremia [4]. Calves experimentally infected with WNV did not produce a viremic response [5].

Cross-protection studies in pigs infected with WNV and Japanese Encephalitis virus (JEV) indicated that pigs were poor hosts for WNV but good hosts for JEV. Pigs infected subcutaneously first with WNV developed low viremia and haemagglutination-inhibition antibodies to both viruses, whereas pigs infected through mosquito bite

produced WNV antibodies but showed no detectable viremia [6]. Dogs infected with WNV showed similar results to the previously mentioned livestock species. Three dogs infected with WNV developed antibodies with only one dog displaying a low titer viremia [7].

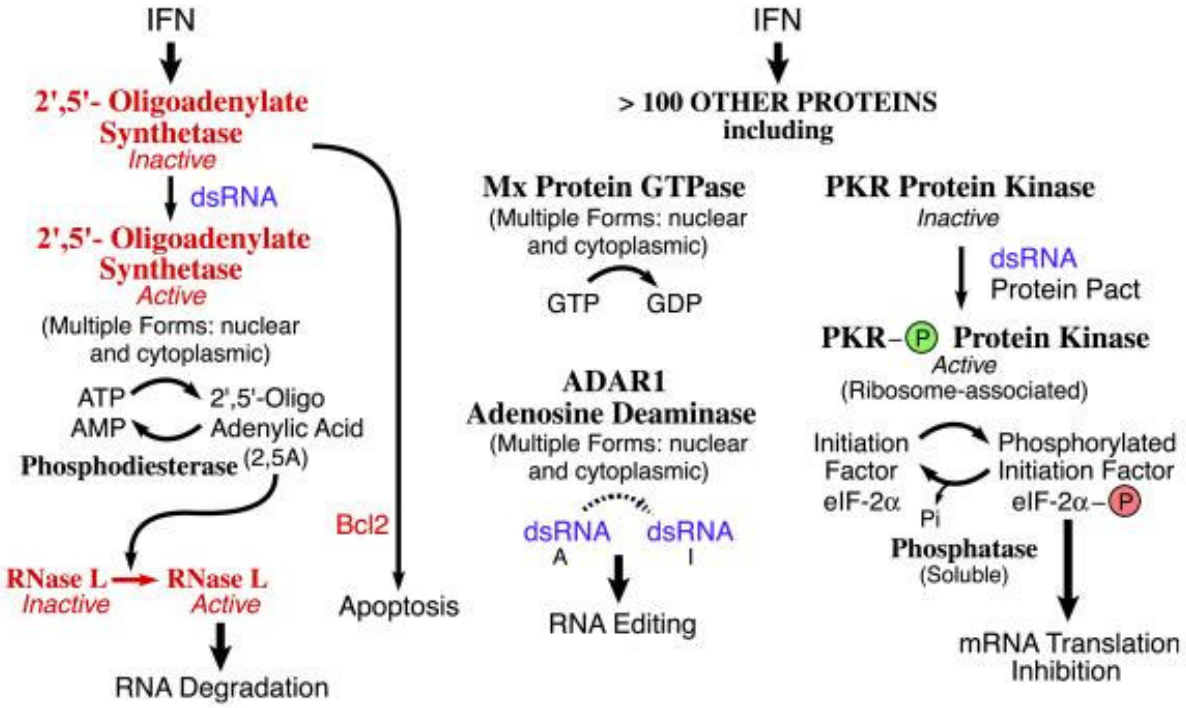


Figure 1.2 Biochemical functions of some IFN-inducible proteins (*Source:* Samuel, 2002).
Interferon activates expression of many antiviral pathways. The *OAS/RNASEL* and *EIF2AK2* (*PKR*) pathways are activated by interferon and act at the levels of transcription and translation, respectively.

Contrary to dogs and other livestock species, horses are particularly susceptible to WNV infection. Clinical symptoms of WNV infection in horses range from biphasic fever to weakness, muscle tremors, encephalomyelitis and paralysis, ultimately resulting in death [3]. Within a single survey of horses in southern France, 34% of the confirmed

Table 1.1 Recently identified mechanisms of IFN system evasion (*Source:* Grandvaux, 2002)

Virus	**Viral Protein**	**Mechanism of IFN system inhibition**
Molluscum contagiosum virus	MC159L	Blocks EIF2AK2-mediated apoptosis[8]
African swine fever virus	-	Inhibits NF-κB: encodes an NFKBIA inhibitor, and downregulates p65 subunit [9]
Influenza virus	NS1	Double-stranded RNA sequestering: inhibits NF-κB and IRF3 transduction pathways [10, 11]
		Affects ISG15 synthesis or activity [12]
Vaccinia virus	E3L	Double-stranded RNA binding protein: abrogate IRF3 and IRF7 transactivating potential [13]
Herpes simplex virus-1	Us11	Double-stranded RNA binding protein: inhibits EIF2AK2 activation [14]
Adenoviruses	E1A	Prevents IRF3 transcriptional activity through binding to CBP/p300 [15]
	-	Downregulates IFNGR2 chain expression [16]
Human herpesvirus-8	Virally encoded IRF1	Prevents IRF3 and IRF1 activities through binding to CBP/p300 [17, 18]
	Virally encoded IRF3	Inhibits IRF3 activity [19]
Vesicular stomatitis virus	Matrix protein	Blocks STAT3 phosphorylation [20]

infected horses died; a lower mortality rate than during the Italian outbreak (43%) [21]. Among 60 horses confirmed with West Nile encephalitis in the United States during the 2000 outbreak, 38% died or required euthanasia [22]. It seems apparent that horses possess a unique susceptibility to severe WNV infection.

Flavivirus susceptibility has been studied extensively in mice; such studies identified a single flavivirus resistance gene, *Flv* [23]. The *Flv* gene, predicted to map to mouse chromosome 5, was positionally cloned and subsequently identified as the murine gene *Oas1b* [24, 25]. A review of the cloning and characterization of the murine *Flv* is also reported [25]. Sequence comparison of *Oas1b* cDNA from susceptible and resistant mice identified 31SNPs. The nonsense transversion SNP C820T located in exon 4 resulted in a truncated transcript lacking 30% of the C-terminal sequence. Comparison between susceptible and resistant strains identified the truncated protein-encoding transcript in each susceptible strain; the truncated form was absent in all resistant strains analyzed [24].

The murine (*Mus musculus*) *Oas* gene cluster consists of eight small-form *Oas* genes (*Oas1a-Oas1h*) and psuedogene *Oas1i*, as well as *Oas2*, *Oas3* and two *Oas*-like genes *Oasl1* and *Oasl2* located on chromosome MMU5 [24, 26-33].

The human (*Homo sapien*) gene cluster is located on human chromosome HSA12q24.2 in the following orientation: *OAS1-OAS3-OAS2* [34, 35]. The small form synthetases p42 and p46 are translated from *OAS1* while the medium forms p69 and p71 and the large form p100 are encoded from *OAS2* and *OAS3*, respectively [36, 37]. Isoforms of both *OAS1* and *OAS2* are products of alternatively spliced transcripts.

Human *OAS1* isoform E16 corresponds to the p42 protein encoded from a 1.6 kilobase (kb) transcript while the E18 splice variant encodes the p46 protein from a 1.8 kb transcript [38]. Both transcripts contain five translated N-terminal exons (A-E) and are identical in their first 346 amino acids but differ at the C-terminus [39, 40]. Human *OAS2* contains 12 exons, including two groups of 5 exons (A1-E1, A2-E2) homologous to the exons of *OAS1* [40]. The strong conservation between the exons of *OAS1* and *OAS2* suggests that *OAS2* derived from an ancestral fusion of two *OAS1* genes [41]. The assembled mRNA sequence of *OAS3* identified three domains in the p100 protein homologous to the *OAS1* p42 protein. Furthermore, human *OAS3* contains three groups of 5 exons (A1-E1, A2-E2, A3-E3), each homologous to *OAS1* exons A-E, suggesting a possible second ancestral duplication event [40]. Phylogenetic analysis among mammalian *OAS* gene families provided a model for the ancestral evolution of the rodent and human *OAS* gene clusters [33]. Recent investigations of the human *OAS* gene cluster identified multiple SNPs for case-control analysis [42, 43].

Human *RNASEL* maps to chromosome HSA1q25, whose ~2.8 kb transcript encodes a 741 amino acid 83,539 Dalton protein [44, 45]. The RNASEL protein consists of three domains, an N-terminal domain of ankyrin repeats and P-loop motif between the seventh and eighth repeat, a domain of protein kinase homology, and a C-terminal ribonuclease domain [46]. Ribonuclease activation requires binding of a single [(pp)p(A2'p5')$_n$A] (2-5A) molecule to the N-terminal ankyrin repeats 2-4 [47, 48]. 2-5A binding causes a conformational change that releases the repression caused by the ankyrin repeats, ultimately concluding with a functional homodimer with ribonuclease

activity [46, 48-50]. In further understanding its functional domains, mutagenesis experiments have identified amino acids critical for 2-5A binding and ribonuclease activity [45, 46, 48, 50-52].

The *OAS/RNASEL* system is an interferon-inducible host cell defense pathway activated through binding dsRNA. dsRNA, present upon viral infection, activates OAS1, catalyzing the oligomerization of ATP to form 2',5' –linked oligoadenylate chains with general structure pppA(2'p5'A)$_n$ [53-55]. Originally discovered as a low molecular weight inhibitor of protein synthesis, pppA(2'p5'A)$_n$ induces the activation of the latent endoribonuclease L, responsible for the degradation of both cellular and viral RNA in a non-preferential manner [53, 56-58]. The *OAS/RNASEL* antiviral system has also been implicated in the induction of apoptosis [59-63].

The equine *OAS* gene cluster was recently characterized and mapped by fluorescent *in* situ hybridization (FISH) to ECA8p15-p14 [33, 64]. The organization of the equine gene cluster is most similar to the human cluster, with single copies of *OAS1*, *OAS3* and *OAS2* in the same organization [64].

Procedure

This investigation of the equine *OAS/RNASEL* antiviral system included identification and analysis of polymorphisms among naturally susceptible and resistant populations. The investigation is divided into three sequential phases:

Phase 1 – OAS1, RNASEL genomic assembly and SNP identification

The full-length cDNA sequence was assembled [GenBank: AY321355] and genomic Children's Hospital Oakland Research Institute (CHORI) Bacterial Artificial

Chromosome (BAC) clones identified as containing the full genomic sequence of equine *OAS1* [64]. CHORI BAC clone 100:I10 (~130 kb) was used for the construction of a shotgun library because of its smaller size relative to clone 77:F4 (~200 kb); both contain the full genomic sequence of equine *OAS1* and *OAS3*. Clone 77:F4 also contains nine 5' exons of the downstream *OAS2* [64].

Prior to the availability of the equine (*Equus caballus*) whole genome shotgun (WGS) sequence, BAC DNA from clone 100:I10 was randomly sheared into fragments with estimated fragment sizes of ~2.5 kb. These fragments were cloned into vector pCR®4Blunt-TOPO® and isolated into a library of 964 clones with verified inserts of the expected size. Additional clones were isolated with insert sizes smaller than expected; these clones were excluded during sequencing. A total of 900 clones were bi-directionally sequenced from the endogenous forward and reverse M13 vector regions providing 513,390 bases with quality scores >15 (3.95X coverage). The sequence data was analyzed using multiple computational programs. Sequence data was analyzed with Phred and Phrap software to assemble high quality overlapping sequences [65, 66]. Individual assembled contigs were visualized using the Consed software tool [67, 68]. Once the equine WGS sequence was released [69], traces identified using BLAST were added to the library clone sequence data to verify the contig assemblies. Individual contigs were aligned to human chromosome HSA12 using BLAST to identify the proper contig order and orientation. The following BLAST parameters were adjusted to allow for cross-species comparison: i) word size = 7; ii) reward for nucleotide match = 17; iii) penalty for nucleotide mismatch = -21; iv) threshold for extending hits = 280; v) cost to

open a gap = 29; vi) cost to extend a gap = 22; vii) dropoff value for gapped alignment = 240; viii) expectation value = 1; and ix) number of database sequences to show (-b and –V) = 5. The resulting genomic assembly of the equine *OAS1* gene [GenBank: DQ536887] included 4 contigs containing six exons (1-6) as well as 4.5 kb and 1.6 kb of sequence upstream of exon 1 and downstream of exon 6, respectively. While the complete sequences of introns 2, 3 and 5 were not assembled, additional contigs were identified within the downstream equine *OAS3* and upstream rabphilin 3A (*RPH3A*) genes.

The equine full-length *RNASEL* mRNA sequence was identified and used as a reference for primer design and subsequent polymerase chain reaction (PCR) amplification from CHORI BAC 159:N12. Amplification products were sequenced to extend the known mRNA sequence and verified using trace data from the equine WGS sequence [69]. The final genomic assembly of equine *RNASEL* [GenBank: EF070193] included 4 contigs containing 7 exons. The full-length sequences of introns 1, 2 and 5 were undetermined.

Using the genomic assemblies of equine *OAS1* and *RNASEL*, genomic primers were designed to amplify individual exons of each gene from the flanking intron sequence. In addition, primers were designed to amplify the proximal promoter of *OAS1* upstream of exon 1. Reaction conditions were optimized and amplification products from both the BAC and genomic DNA templates were verified by sequencing.

DNA from 13 horses was isolated from blood samples and used to identify SNPs. Horses were selected to minimize bloodline similarity and inbreeding, representing

American Quarter Horse (9), Arabian (1), American Paint Horse (1), Appaloosa (1) and Thoroughbred (1) breeds. Blood samples were collected at the Texas A&M University Horse Center with approval from the Animal Science Department and in accordance with ethical standards. Individual regions of *OAS1* and *RNASEL* were amplified and sequenced from each sample and aligned using Phred, Phrap and Consed. Computational analysis identified high quality sequence discrepancies as well as length polymorphisms of two microsatellite repeats within the proximal promoter and downstream of exon 6 of *OAS1*. Visual analysis of the individual chromatogram data identified polymorphic alleles within heterozygous individuals.

Analysis among the 13 individuals identified 33 and 31 SNPs within equine *OAS1* (Table 1.2) and *RNASEL* (Table 1.3), respectively, as well as 2 polymorphic microsatellites within *OAS1*. To avoid identifying artifacts of the PCR and sequencing processes, SNPs were verified by meeting the following criteria: 1) each homozygous allele found in at least two individuals; or 2) the heterozygous genotype found in at least two individuals. These criteria correspond to a minor allele frequency ≥ 0.08.

Phase 2 – OAS1, RNASEL polymorphism association to severe West Nile virus infection

The occurrence of naturally susceptible and resistant populations of horses within the United States suggests a possible role by which the equine innate immune response may affect the clinical severity of WNV infection. Comparing transcript sequences of susceptible and resistant mice, *Perelygin et al. (2002)* identified a SNP absolutely associated with WNV susceptibility within inbred strains of mice [24]. The specific aim of this phase was to identify potential polymorphisms or haplotypes significantly

Table 1.2 Equine *OAS1* single nucleotide polymorphisms and microsatellites (*Source*: Rios *et al.* 2007)

Region	Accession DQ536887	Alleles	Residue	Amino Acid Type	Frequency	Polymorphism
-	3640	C T	-	-	0.25 0.75	Transition
-	3687	G T	-	-	0.65 0.35	Transversion
-	3718	A G	-	-	0.65 0.35	Transition
-	3724	C T	-	-	0.65 0.35	Transition
-	3825	C T	-	-	0.35 0.65	Transition
-	3830	A T	-	-	0.85 0.15	Transversion
-	3973	C T	-	-	0.65 0.35	Transition
-	4032-4063	-	-	-	-	GT repeat
-	4234	C T	-	-	0.65 0.35	Transition
-	4333	C T	-	-	0.35 0.65	Transition
-	4455	C G	-	-	0.08 0.92	Transversion
-	4487	C T	-	-	0.88 0.12	Transition
-	4501	C T	-	-	0.65 0.35	Transition
-	4531	A G	-	-	0.08 0.92	Transition
5' UTR	4598	C G	-	-	0.65 0.35	Transversion
5' UTR	4625	A C	-	-	0.35 0.65	Transversion
Exon 1	4690	A G	18Tyr 18Cys	Uncharged Polar Uncharged Polar	0.85 0.15	Transition
Exon 1	4783	C T	49Ala 49Val	Nonpolar Nonpolar	0.35 0.65	Transition
Intron 1	5609	C T	-	-	0.67 0.33	Transition
Exon 2	5701	C T	77Leu 77Leu	Nonpolar	0.64 0.36	Transition
Exon 2	5743	C T	91Phe 91Phe	Nonpolar	0.33 0.67	Transition
Exon 2	5765	A G	99Lys 99Glu	Basic Polar Acidic Polar	0.65 0.35	Transition
Exon 2	5776	A G	102Arg 102Arg	Basic Polar	0.38 0.62	Transition
Exon 2	5786	A G	106Lys 106Glu	Basic Polar Acidic Polar	0.38 0.62	Transition
Exon 2	5920	G T	150Pro 150Pro	Nonpolar	0.08 0.92	Transversion
Exon 3	9374	C T	209Arg 209Cys	Basic Polar Uncharged Polar	0.85 0.15	Transition
Exon 4	12714	C G	264Asn 264Lys	Uncharged Polar Basic Polar	0.59 0.41	Transversion
Intron 4	12810	C T	-	-	0.64 0.36	Transition
Intron 4	12853	A G	-	-	0.55 0.45	Transition
Intron 5	13628	A T	-	-	0.55 0.45	Transversion
Intron 5	13649	C G	-	-	0.46 0.54	Transversion
Exon 6	15320	C T	370Arg 370Trp	Basic Polar Nonpolar	0.75 0.25	Transition
3' UTR	15410	G T	-	-	0.62 0.38	Transversion
3' UTR	15537	G T	-	-	0.19 0.81	Transversion
-	15798-15855	-	-	-	-	GT repeat

Table 1.3 Equine *RNASEL* single nucleotide polymorphisms (*Source:* Rios *et al.* 2007)

Region	Accession EF070193	Alleles	Residue	Amino Acid Type	Frequency	Polymorphism
-	143	C G	-	-	0.12 0.88	Transversion
5' UTR	1857	A C	-	-	0.55 0.45	Transversion
Exon 2	1991	C T	27His 27Tyr	Basic Polar Uncharged Polar	0.54 0.46	Transition
Exon 2	2020	C T	36Gly 36Gly	Uncharged Polar	0.92 0.08	Transition
Exon 2	2021	G T	37Asp 37Tyr	Acidic Polar Uncharged Polar	0.69 0.31	Transversion
Exon 2	2118	A C	69Asn 69Thr	Uncharged Polar Uncharged Polar	0.29 0.71	Transversion
Exon 2	2121	A G	70Tyr 70Cys	Uncharged Polar Uncharged Polar	0.92 0.08	Transition
Exon 2	2316	A C	135Lys 135Thr	Basic Polar Uncharged Polar	0.75 0.25	Transversion
Exon 2	2332	A G	140Ala 140Ala	Nonpolar	0.35 0.65	Transition
Exon 2	2374	G T	154Arg 154Ser	Basic Polar Uncharged Polar	0.83 0.17	Transversion
Exon 2	2635	A G	241Thr 241Thr	Uncharged Polar	0.21 0.79	Transition
Exon 2	2680	C G	256Ser 256Ser	Uncharged Polar	0.43 0.57	Transversion
Exon 2	2771	A G	287Lys 287Glu	Basic Polar Acidic Polar	0.57 0.43	Transition
Exon 2	3144	A G	411Asn 411Ser	Uncharged Polar Uncharged Polar	0.19 0.81	Transition
Exon 2	3152	C T	414Arg 414Cys	Basic Polar Uncharged Polar	0.81 0.19	Transition
Exon 2	3281	A G	457Lys 457Glu	Basic Polar Acidic Polar	0.19 0.81	Transition
Exon 2	3301	A C	463Lys 463Asn	Basic Polar Uncharged Polar	0.19 0.81	Transversion
Exon 2	3311	C T	467Pro 467Ser	Nonpolar Uncharged Polar	0.19 0.81	Transition
Exon 2	3372	A G	487Gln 487Arg	Uncharged Polar Basic Polar	0.58 0.42	Transition
Intron 2	3404	A G	-	-	0.19 0.81	Transition
Exon 3	5108	A G	513Lys 513Glu	Basic Polar Acidic Polar	0.11 0.89	Transition
Exon 3	5111	C T	514Pro 514Ser	Nonpolar Uncharged Polar	0.82 0.18	Transition
Exon 5	7314	A G	598Asn 598Asp	Uncharged Polar Acidic Polar	0.87 0.13	Transition
3' UTR	9994	C G	-	-	0.15 0.85	Transversion
3' UTR	9999	A T	-	-	0.88 0.12	Transversion
3' UTR	10247	C T	-	-	0.31 0.69	Transition
3' UTR	10914	C T	-	-	0.38 0.62	Transition
3' UTR	11105	C T	-	-	0.88 0.12	Transition
3' UTR	11146	C T	-	-	0.35 0.65	Transition
3' UTR	11184	C T	-	-	0.35 0.65	Transition
3' UTR	11228	C T	-	-	0.83 0.17	Transition

associated with WNV susceptibility or resistance in horses. Candidate polymorphisms for analysis include those identified in Phase 1 while additional polymorphic candidates may be identified during the analysis of specific susceptible and resistant animals. The SNPs verified in the previous section were identified from horses without specific exposure to WNV, where analysis may not identify highly associated polymorphisms. However, sequencing from susceptible and resistant horses may identify alleles highly associated to WNV phenotype not specifically identified in the random, neutral population.

Researchers have identified SNPs within the human *OAS* gene cluster for association studies to human WNV infection [42, 43]. In a similar manner, this analysis of SNP-associated WNV disease will include multiple computational analyses including statistical analysis using Fisher's Exact test to identify SNPs and/or haplotypes associated with WNV susceptibility or resistance.

A detailed phenotypic definition of case and control populations is critical to developing statistically relevant analyses. Statistical analyses were repeated using different case definitions to investigate the potential genetic effect on equine WNV susceptibility. For example, the first analysis grouped control horses as those that did not present with clinical symptoms or presented with clinical signs yet recovered (survivors). The case population consisted of horses that showed clinical signs of WNV disease and either died or required humane euthanasia (non-survivors). A second analysis grouped control horses that presented no clinical signs of WNV disease (subclinical) and a case population of horses presenting clinical signs (clinical),

including survivors and non-survivors. Data from both scenarios were analyzed in the same manner and compared to better understand the degree to which associated alleles, genotypes and/or haplotypes contribute to WNV disease in horses. During veterinary examination, case and control samples were confirmed as being infected with WNV by diagnostic PCR and histopathology examination of multiple tissues.

Phase 3 – Regulatory effect of polymorphic microsatellite within the equine OAS1 proximal promoter

Sequencing from a random population of 13 horses identified 2 polymorphic microsatellites within the proximal promoter and downstream of exon 6 of equine *OAS1*.

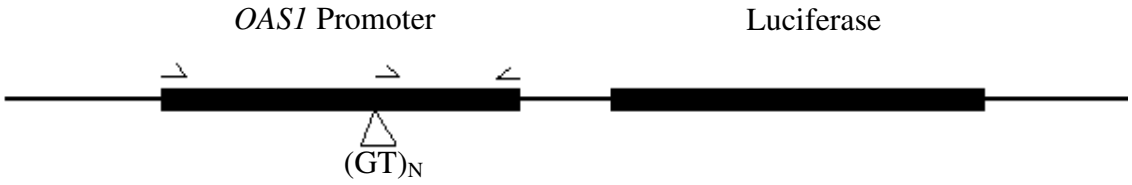
The microsatellite within the promoter region of *OAS1* is located ~575 bp upstream of the translation initiation ATG, between sequence regions conserved in the human *OAS1* promoter (Figure 1.3). This microsatellite's placement within the promoter suggests a possible functional role in *OAS1* gene expression. Dinucleotide repeats have been shown to modulate gene activity both positively and negatively [70-75]. Alternating dinucleotide repeats of purines and pyrimidines, such as those found within equine *OAS1*, alter DNA to form Z-DNA structures and are located near transcription start sites [76-79]. One Z-DNA-forming repeat was found to repress promoter activity, such that, when deleted, promoter activity increased 36-51% [80].

Microsatellite genotypes of the 13 random equine individuals found promoter repeat lengths of $(GT)_9$ and $(GT)_{18}$ were over-represented among this sample set. However, preliminary data from Phase 2 identified common alleles of $(GT)_9$ and $(GT)_{19}$. To identify a potential regulatory function of the equine *OAS1* promoter microsatellite,

Figure 1.3 Alignment of the proximal promoters of human *OAS1* and equine *OAS1* (*Source:* Rios *et al.* 2007).
BLAST2 alignment of the 1000 bp upstream of the transcription start for human *OAS1* and equine *OAS1* genes. The alignment shows that the sequence from ~800 bp to ~-350 bp in the horse promoter is similar to a region of the human promoter interrupted by a 200 bp *Alu* repeat (~-811 bp ~-590 bp). The horse microsatellite is shown in underlined bold. Numbering shown in the alignments is from the translation ATG start sites.

OAS1 promoter-reporter clone constructs containing repeat alleles of $(GT)_9$, $(GT)_{16}$ or $(GT)_{19}$ were transfected into mammalian cells to measure potential differences in IFN-induced activation (Figure 1.4). Additional constructs lacking the microsatellite and upstream sequence (5' deletion) were transfected and luciferase expression compared to the polymorphic, full-length promoter constructs. Constructs lacking the microsatellite

and downstream sequence (3' deletion constructs) were analyzed and compared to full-length constructs. Data analysis focused primarily on differences in IFN induction between the constructs as well as time-course and dose-response differences between clones. With a population over-representation of alleles corresponding to full DNA helical rotations, alleles altering this configuration may disrupt the alignment of yet unidentified flanking regulatory elements, including the as yet unidentified *OAS1* interferon-stimulated response element (ISRE).



Clone	Repeat Sequence
OAS1-9	...AGAGAGAGCT**GTGTGTGTGTGTGTGTGT**----------CTTAACCTAA...
OAS1-16	...AGAGAGAGCT**GTGTGTGTGTGTGTGTGTGTGTGTGTGTGT**------CTTAACCTAA...
OAS1-19	...AGAGAGAGCT**GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT**CTTAACCTAA...

Figure 1.4 *OAS1* promoter-reporter constructs.
Schematic diagram of the *OAS1* promoter cloned upstream of the pGL3 luciferase reporter coding region. Half arrows indicate primer locations used to amplify and clone the *OAS1* promoter. Table represents the GT dinucleotide sequence of individual clone constructs. The GT microsatellite is bolded and underlined.

CHAPTER II

REVIEW OF THE CELLULAR ANTIVIRAL IMMUNE RESPONSE

The host cell recognition and response to viral challenge is propagated through the induction of the interferon response system. When cells are challenged by viral infection, they respond by activating a variety of genetic pathways involved in the production of interferon. The pathways to IFN production are redundant and ensure that the cell's response can limit viral replication while also preparing uninfected cells for viral infection. This chapter reviews the cellular response to viral infection and the genetic mechanisms of IFN production. This chapter concludes with a review of the viral mechanisms responsible for limiting the host immune response.

IFN Induction

Flaviviruses represent one class of positive-strand RNA virus whose genome consists of both structural and non-structural coding regions. After entry into the host cell, *Flavivirus* replication produces dsRNA, which is consequently recognized by the host cell to activate the complex immune response. Host cell response to viral infection is triggered by the recognition of pathogen-associated molecular patterns (PAMPs), such as dsRNAs produced by viruses of the *Flaviviridae* genus, including WNV. These viral recognition patterns interact with host pattern recognition receptors (PRRs) to stimulate downstream cascades, signaling the beginning of the early stages of the host immune response (ie. cytokine production). Cytokine production resulting from the presence of viral dsRNA is activated by both a *TLR* -dependent and -independent cascade.

The toll-like family of PRRs consists of membrane glycoproteins containing leucine-rich-repeat motifs and a cytoplasmic toll/ interleukin 1 receptor (TIR) homology domain, through which PAMP recognition is signaled [81]. The TLR family of receptors can be classified by their respective ligands (Table 2.1). Of particular interest to the viral immune response to WNV is TLR3, a PRR activated by binding dsRNA [82, 83]. Figure 1.1 shows the cascade of TLR3 activation regulating IFN production [84]. Activated TLR3 signals toll-like receptor adaptor molecule 1 (TICAM1) for the downstream activation of neural factor κB (NF-κB) through the inhibitor of kappa light polypeptide gene enhancer (IKBK) complex [85]. The interaction between TICAM1 and receptor interacting protein (RIP), responsible for NF-κB activation, is mediated by the RIP homotypic interaction motif (RHIM), a C-terminal motif required for proper TICAM1 activity [86]. Unstimulated NF-κB remains a cytoplasmic heterodimer bound by the nuclear factor of kappa light polypeptide gene enhancer inhibitor alpha (NFKBIA). Upon activation, the IKBK complex phosphorylates NFKBIA, releasing NF-κB into the nucleus [85]. Nuclear NF-κB stimulates cytokine production, including Type I IFN [87].

Contrary to the antiviral role *TLR3* plays in IFN production, research using *Tlr3*-deficient mice identified a pathogenic role of the gene during WNV infection. Severe central nervous system (CNS) disease and lethality from WNV infection results from viral progression through the blood brain barrier. Recombinant mice lacking *Tlr3* (*Tlr3*$^{-/-}$) and infected with lethal doses of WNV were more resistant than infected wild type mice, with survival rates of 40% and 0%, respectively [88]. Data indicated that WNV

Table 2.1 TLR recognition of microbial components (*Source:* Akira *et al.* 2006)

Microbial Components	Species	TLR Usage
Bacteria		
LPS	Gram-negative bacteria	TLR4
Diacyl lipopeptides	*Mycoplasma*	TLR6/TLR2
Triacyl lipopeptides	Bacteria and mycobacteria	TLR1/TLR2
LTA	Group B *Streptococcus*	TLR6/TLR2
PG	Gram-positive bacteria	TLR2
Porins	*Neisseria*	TLR2
Liparabinomannan	Mycobacteria	TLR2
Flagellin	Flagellated bacteria	TLR5
CpG-DNA	Bacteria and mycobacteria	TLR9
ND	Uropathogenic bacteria	TLR11
Fungus		
Zymosan	*Saccharomyces cerevisiae*	TLR6/TLR2
Phospholipomannan	*Candida albicans*	TLR2
Mannan	*Candida albicans*	TLR4
Glucoronoxylomannan	*Cryptococcus neoformans*	TLR2 and TLR4
Parasites		
tGPI-mutin	*Trpanosoma*	TLR2
Glycoinositolphopholipids	*Trpanosoma*	TLR4
Hemozoin	*Plasmodium*	TLR9
Profilin-like molecule	*Toxoplasma gondii*	TLR11
Viruses		
DNA	Viruses	TLR9
dsRNA	Viruses	TLR3
ssRNA	RNA viruses	TLR7 and TLR8
Envelope proteins	RSV, MMTV	TLR4
Hemagglutinin protein	Measles virus	TLR2
ND	HCMV, HSV1	TLR2
Host		
Heat-shock protein 60, 70		TLR4
Fibrinogen		TLR4
ND - not determined		

replication and inflammatory responses within the brains of infected mice were reduced in *Tlr3*$^{-/-}$ mice while the viral load within the blood was significantly higher, compared to wild-type mice. Indeed, the researchers showed that the *Tlr*-dependent response

mediated viral entry into the CNS through reversible damage of the blood brain barrier [88].

The observation of dendritic cell maturation, a response mediated by TLR activation, in *Tlr*-deficient mice suggested a *Tlr*-independent response to virus infection mediated by cytoplasmic retinoic acid-inducible gene 1 (RIG1) protein [89, 90]. Figure 2.1 shows a diagrammatic representation of NF-κB activation by the *RIG1* signaling pathway. RIG1 is a RNA helicase with two caspase-recruiting domain (CARD)-like domains. The helicase domain interacts with dsRNA while the CARD-like domains are responsible for the downstream activation of NF-κB [90]. The RIG1 helicase is linked to TANK binding kinase 1 (TBK1) and IKBKE by interferon β promoter stimulator 1 (IPS1), also known as MAVS, VISA and Cardif [91-94]. *RIG1* was shown to be essential for IFN induction, as induction from dendritic cells deficient for *RIG1* were greatly diminished [95]. With both TLR and RIG1 proteins localized within the cytoplasm and on endosomal membranes, it is unclear which pathway is activated for *IFN* induction. A possible manner by which cells differentially activate IFN production is by the route of viral infection. Other possible mechanisms are that actively replicating viruses signal the *RIG1* pathway, while plasmacytoid dendritic cells preferentially activate through a *RIG1*-independent manner [83, 95].

Recent publications have identified an additional level of complexity by which the cell mediates its response to dsRNA species [96, 97]. The presence of dsRNA species derives from both cellular activity (self) and viral infection (non-self), and the

cell must differentiate the two to properly activate downstream cascades. For example, the *OAS/RNASEL* innate immune response generates dsRNA species from the

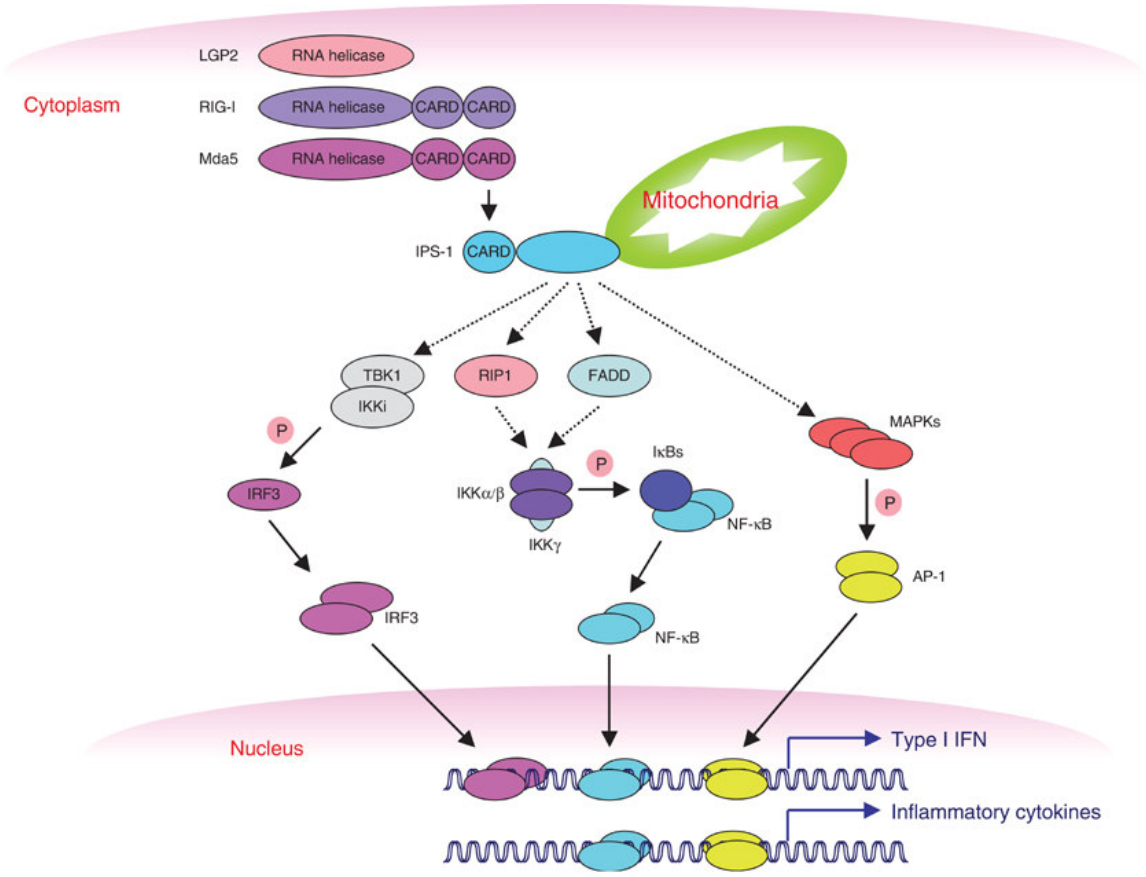


Figure 2.1 RIG1- and MDA5-mediated signaling pathway (*Source:* Kawai *et al.* 2006). RIG1 and MDA5 helicases recognize replicating viruses and interacts with mitochondrial IPS1. IPS1 signals activation of IFN genes by multiple transcription factors.

degradation of both cellular and viral dsRNA. This pathway contributes to the cellular population of dsRNA, which may then provide activating potential for pathways activated by dsRNA (i.e. OAS, RIG1). On the other hand, endogenous dsRNAs (self) are produced through the transcription of cellular miRNA genes. These dsRNA species

are utilized in the RNAi cascade. Thirdly, dsRNAs (non-self) are produced as replication intermediates of infecting viruses (WNV). Several lines of evidence have described mechanisms by which cells differentiate dsRNA-responsive mechanisms (RIG1 activation vs. RNAi). One mode by which dsRNAs can be differentially identified by host factors is size. The dsRNA binding protein of the RNAi system, dicer 1 (DICER1), utilizes small molecules 21-23 bp in length while longer molecules preferentially activate eukaryotic initiation factor 2 alpha kinase 2 (EIF2AK2) and OAS proteins [98, 99]. Additionally, DICER1 recognizes and is activated by the presence of a 2 nucleotide overhang at the 3' end of the short dsRNA molecule. This 3' overhang seems sufficient to distinguish dsRNA species between the RNAi and RIG1 pathways. RIG1 does bind the 3'-overhang dsRNAs but its helicase activity is not activated by these molecules [96]. Therefore, the long dsRNA species produced by RNASEL ribonuclease activity and viral infection are recognized as lacking the end-motifs and activate RIG1 while the endogenous short dsRNA species containing the 3' overhang are recognized and activate the DICER1 complex of the RNAi pathway [96, 97].

Both the *TLR*-dependent and *TLR*-independent mechanisms converge in their response to viral infection by stimulating the production of Type I IFN. Type I IFN include those previously designated as leukocyte (IFNA/IFNO) and fibroblast interferon (IFNB). Less extensively studied Type I interferon include IFNT, IFNK, IFNE, and IFNL [100-108]. Type II interferon represents those previously designated as immune interferon (IFNG) [109-111]. The remainder of this discussion will focus primarily on Type I IFNA and IFNB, encoded by genes clustered on human chromosome 9q21 [112].

The human *IFNA* genes, representing different subtypes of IFNA, and *IFNB* gene all lack introns and each encode preproteins with a secretory signal peptide sequence which is cleaved prior to protein secretion. The human IFNAs contain 23 amino acid signal peptides and are secreted as mature 166 amino acid proteins with molecular weights ranging between 16 and 27 kDa [113]. The *IFNA* genes are highly conserved, with 80-95% homology at the nucleotide level, suggesting this gene cluster originated from a common ancestor [114]. At the amino acid level, the IFNAs share >50% homology [115]. One hypothesis for the evolution of multiple *IFNA* genes suggests that each evolved under different selective pressures, resulting in genes with different functional roles [116]. This hypothesis is supported by the observation that *IFNA* genes are differentially expressed in response to similar stimuli [113].

Human IFNB preprotein contains a 21 amino acid signal peptide with a resulting 166 residue mature protein with molecular weight between 28 and 35 kDa depending on its degree of glycosylation [113]. Evidence suggests that in mammals, the *IFNA* genes evolved independently of *IFNB*; however, both human IFNAs and IFNB contain conserved regions of amino acid homology associated with receptor interactions [117, 118].

Comparison between human IFNAs and IFNB revealed only 20% homology while IFNO shares 36% identity, suggesting that these Type I IFN species evolved from a single ancestral gene [119]. Arguably, such evolutionary duplication occurred mainly from unequal crossing-over in conserved regions, including inter-genic repeats [112]. Phylogenetic analysis of all mammalian Type I IFNs identified three main subgroups of

IFNA, IFNB and IFNO; however, IFNB clustered as an out-group to the mammalian IFNAs and IFNOs [115].

As previously discussed, *IFN* gene regulation occurs primarily at the transcription level, along with regulating mRNA stability and the short half-life of *IFN* transcripts. The presence of activating factors produces a rapid increase in *IFN* expression through regulatory elements located within 200 base pairs (bp) of the transcription start site. The immediate promoters of *IFNA* and *IFNB* genes contain different binding motifs and, therefore, have different binding requirements for induction. Within the *IFNB* immediate promoter are four binding regions, positive regulatory domains (PRD) I-IV, responsible for binding activating transcription factors [120]. These regulatory domains were found to coordinate a mechanism by which binding factors cooperatively assemble into an enhanceosome complex with strictly organized protein-protein and protein-DNA interactions [121, 122] (Figure 2.2).

The organization of the protein complex is dependent on the proper spatial orientation of each factor, aligning each in its appropriate groove within the double helix. The helical relationship between the proteins of the enhanceosome complex enables the coordinated activation of the *IFNB* promoter [121]. Insertion of half-helical rotations between the PRD domains diminished the *in vitro* activation potential of the enhanceosome complex. However, enhanceosome activity was maintained when sequence representing an entire helical turn was inserted between domains. Additionally, dual insertions between domains that maintained the helical phase of PRDIV and PRDII were unable to activate the reporter construct. These data show that

the *IFNB* enhanceosome requires proper spacing of the PRD domains such that the proteins are in proper helical contact with each other as well as with the DNA molecule. The high mobility group protein HMG-I(Y) induces a conformational change within the protein complex that allows for the synergistic activity of the complex and, thus, is required for full transcription activation [121]. The coordinated assembly of the entire protein complex alters the inactive DNA conformation, specifically between interactions of the NF-κB, ATF-2/c-jun and HMG-I(Y) proteins [123].

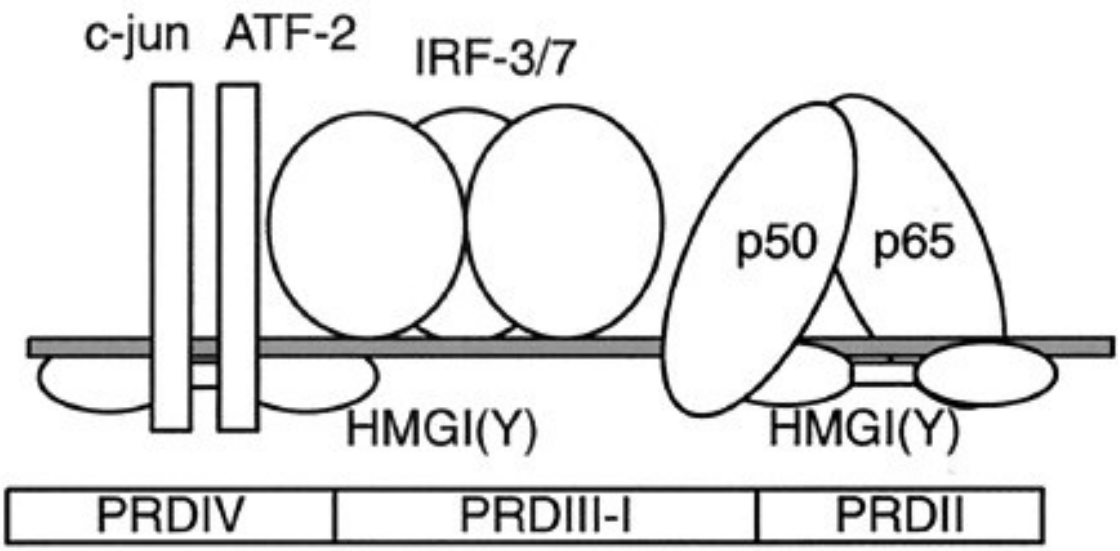


Figure 2.2 Diagram of the binding sites for the *IFNB* enhanceosome (*Source*: Falvo *et al.* 2000).
The IFNB gene promoter requires the cooperative binding and assembly of multiple transcription factors into an enhanceosome. Activation involves protein-DNA and protein-protein interactions and proper spacing of DNA binding regions.

Induction of these *IFN* genes is transient even under conditions of continued induction. The continued state of induction is suggested to be maintained through a system which prevents protein synthesis of repressors that inhibit *IFN* transcription [124]. Flanking the positive regulatory elements are two negative regulatory domains

(NRDI and NRDII). The NRD elements regulate the stable repression of the *IFNB* promoter except under inducing conditions. However, basal levels of IFN are detectable during the repressive state when the *IFN* promoter is not induced. While PRDI and PRDII bind transcriptional activators during promoter induction, these two elements also act to repress the promoter after induction in a trans-regulatory feedback mechanism [125-127].

Although induced under similar condition, the *IFNA* gene promoter differs from the *IFNB* promoter. *IFNA* induction, like *IFNB*, requires the activation of PRDI binding factors; however, the *IFNA* promoter contains a PRDI-like sequence element but does not contain NF-κB binding motifs. *IFNA* induction in response to various factors, including viral infection, is mediated through the sequence specificity of the PRDI-like sequence. Like the *IFNB* promoter, *IFNA* induction requires binding of activating factors to the PRDI domain. However, while the *IFNA* promoter does not bind NF-κB, it does require active TG-sequence binding proteins to the GAAATG binding motif [128].

IFN Signal Response

Mature Type I IFNs are secreted and act as inter-cellular signals to prepare uninfected cells for viral challenge. With a typical ligand-receptor relationship, extracellular IFN activates the interferon response of adjacent cells to stimulate transcription of host antiviral genes in anticipation of viral infection. The Type I IFN cellular receptor is composed of two subunits, interferon receptor 1 (IFNAR1) and IFNAR2, encoded by genes located on human chromosome 21 [129, 130]. The IFNAR1

subunit of the IFN receptor is a 110 kDa glycosylated protein while the IFNAR2 subunit exists in multiple forms resulting from differential splicing of the *IFNAR2* gene [129-132]. IFNAR2c is the longer form representing the major subunit with a molecular mass between 90-100 kDa. IFNAR2b is smaller with a molecular mass of 51 kDa [129, 131, 132]. IFNAR1 forms different active complexes with each of the subunits of IFNAR2; however, IFNAR2c binds the IFNA and IFNB ligands with significantly greater affinity [133-135].

The classical janus activated kinase (*JAK*)-signal transducer and activator of transcription (*STAT*) pathway is represented in Figure 2.3. IFNAR1 and IFNAR2 associate with two members of the JAK-family of kinases, tyrosine kinase 2 (TYK2) and JAK1, respectively, through interactions with proline-rich receptor sequences [115, 136-139]. Upon ligand binding, these JAK kinases autophosphorylate and activate in response to receptor rearrangement and dimerization [140]. The activated kinases phosphorylate multiple STAT proteins leading to the formation of multiple STAT homodimer and heterodimer complexes that subsequently translocate to the nucleus for gene promoter binding and transcription activation [115, 136, 141-143]. STAT phosphorylation, while mediated in part by tyrosine phosphorylation from the JAK kinases, also requires protein kinase C-delta (PRKCD)-mediated phosphorylation of serine 727 for full transcription activation [144-146]. Dimerization of the STAT proteins is mediated by the reciprocal recognition of the phosphorylated tyrosine by the required SH2 domains of STAT proteins [147]. The regulatory specificity of STAT activation is mediated by the SH2 domain during receptor activation. For example,

STAT1 protein containing a STAT2 SH2 domain was not activated by IFNG; however, a STAT2 protein with a STAT1 SH2 domain was activated by both IFNA and IFNG [148].

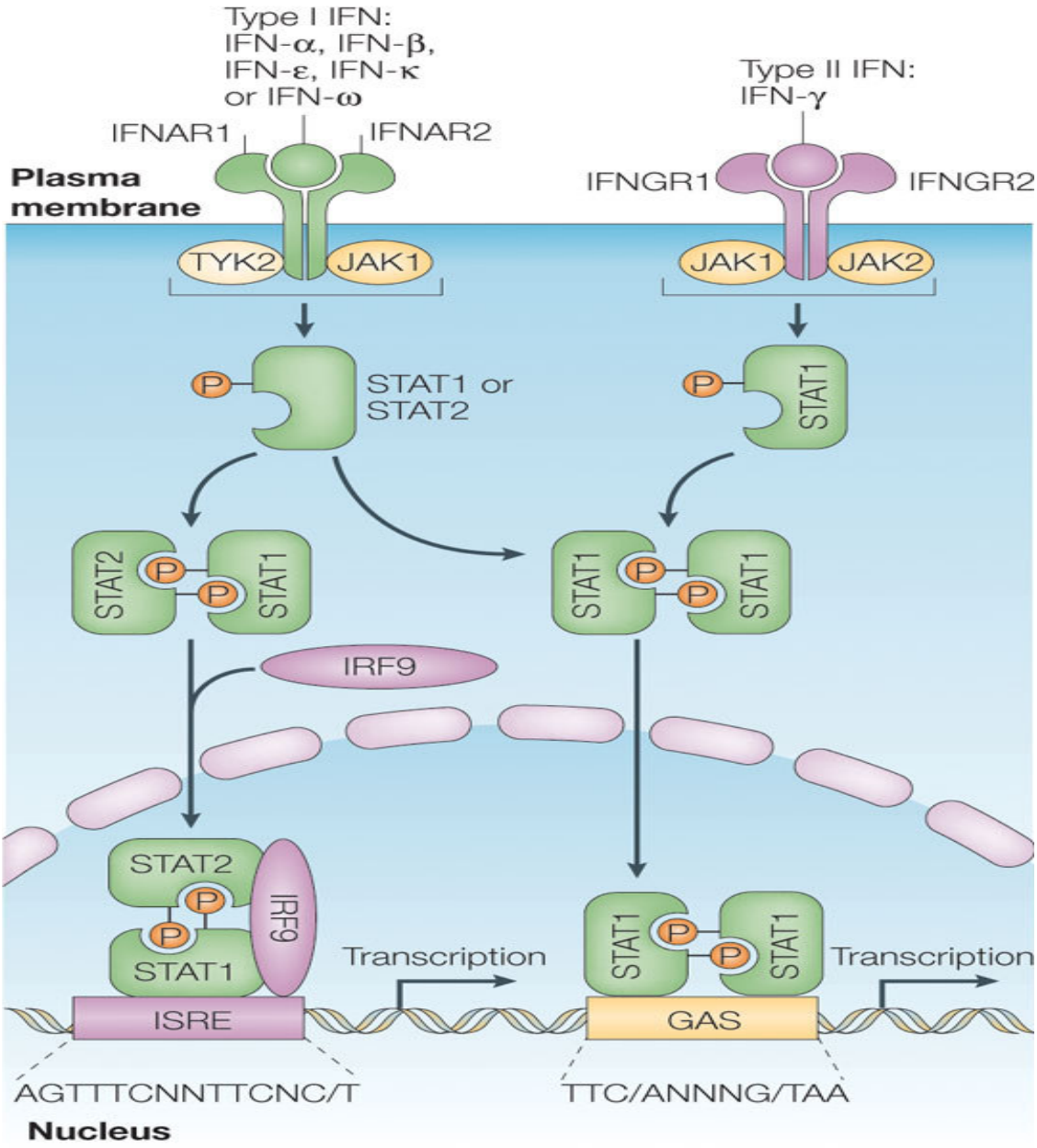


Figure 2.3 Interferon receptors and activation of classical *JAK–STAT* pathways by Type I and Type II IFN (*Source:* Platanias 2005).
Activated IFN receptors signal STAT dimerization through kinase phosphorylation. STAT dimmers bind gene promoter regulatory motifs for activation of transcription.

Specifically induced by Type I IFN, the phosphorylated STAT1-STAT2 complex translocates into the nucleus where it interacts with interferon regulatory factor 9 (IRF9) to form the mature heterotrimeric complex interferon-stimulated gene factor 3 (ISGF3) [141-143]. This transcription factor complex binds exclusively to the ISREs located within the proximal promoter region of many ISGs and activates transcription. Alternative STAT complexes bind other promoter elements to activate gene transcription, including the IFNG-activated sites (GAS) bound by STAT1 homodimers. The interactions of these STAT complexes on ISRE and GAS elements activate hundreds of genes whose promoters may contain one or both of these elements and may require specific complexes of STAT proteins [149].

The *JAK-STAT* pathway contains multiple sites of regulation potentially mediating the ligand-specific activation of immediate early ISGs. These include the 1) ligand-receptor-kinase interactions, 2) STAT protein activation and interaction and 3) the DNA binding sites located within the promoters of ISGs [150].

Activated STAT proteins may also interact with multiple co-activators including p300 and cAMP responsive element binding protein (CREB) -binding protein (CBP) [151, 152]. Both p300 and CBP are co-activators with histone-acetyltransferase activity [153]. Activated STAT proteins also interact with histone deacetylase 1 (HDAC1), recently shown to be required for IFN-dependent gene transcription [154-156]. Levels of IFN-activated gene transcription are also regulated by the degree of ligand-receptor occupancy [157].

Viral Antagonists of the IFN Response

Targeting IFN induction

To counteract the antiviral response of host cells, viruses have evolved mechanisms that interrupt multiple stages of the IFN response, including induction, signaling and altering the effects of specific ISGs. Many viruses block induction of Type I IFN through interactions with IRF3 and IPS1. Both Influenza and Poxviruses compete against activation of IRF3 by encoding dsRNA binding proteins NS1 and E3L, respectively [158, 159]. IRF3 phosphorylation by TBK1 is inhibited by the viral P protein of some negative-stranded RNA viruses [160]. As well, some viral proteins target multiple steps of the IFN induction cascade. Hepatitis C viral protease NS3/4A cleaves IPS1 as well as blocks TLR3- and RIG1 –mediated induction by cleaving TICAM1 [92, 161]. Other viruses target NF-κB to inhibit IFN production [9, 11].

Targeting IFN activation

Viruses have evolved mechanisms to alter the ISG activation response to Type I IFN. Poxviruses encode soluble receptors which compete with host receptors for ligand binding [162]. *JAK/STAT* signaling of the *IFN* response pathway is targeted by a number of viruses. Paramyxoviruses, Murine Polyoma virus, Human Papillomavirus and Herpesviruses all encode proteins targeting members of the JAK-family of signaling proteins [163-166]. The ISGF3 complex is targeted by Paramyxoviruses by inhibiting STAT synthesis [167]. IRF9 expression is inhibited after Herpesvirus infection, while the Papillomavirus E7 protein interacts directly with IRF9 to prevent proper ISGF3 activity [163, 168].

Targeting the activity of ISGs

Viruses encode genes producing small non-activating RNAs which compete with replication intermediate dsRNAs for binding dsRNA-activated proteins, particularly EIF2AK2 and OAS [14, 169, 170]. This competition prevents activation of these proteins and prevents them from binding activating dsRNAs.

CHAPTER III

DISCOVERY OF THE *FLAVIVIRUS* RESISTANCE GENE AND ANTIVIRAL CHARACTERIZATION OF THE *OAS/RNASEL* SYSTEM

Murine *Flv* Gene Discovery

Genetic resistance to *Flaviviruses* was originally identified in mice as early as the 1920s. Resistant mice were used in breeding studies to identify a dominant allele within a single autosomal gene as conferring the *Flavivirus* resistance phenotype [171]. A resistant mouse strain (PRI) was crossed to susceptible C3H/He mice and offspring backcrossed to C3H/He mice for eight generations [172]. The final strain of inbred mice contained an estimated donor linkage region of 31 cM containing the *Flavivirus* resistance (Flv^r) gene [173]. Phenotypic characterization of resistant and susceptible mice has shown resistance to be specific to *Flaviviruses*, including disease from mosquito-borne WNV [25]. Further characterization showed that while resistant mice were capable of infection, they produced lower viral yields than susceptible mice. Data measuring the amount of minus-strand and positive-strand viral RNA in resistant and susceptible strains showed genetic resistance to *Flaviviruses* was mediated at the level of viral replication and not attachment or entry [174, 175].

Coinheritance studies mapped the *Flv* locus to mouse chromosome 5 [176]. Three-point linkage analyses placed the *Flv* gene in a precise gene order relative to four other genes [177]. Further mapping of the *Flv* locus was completed using 20 microsatellite markers within the known locus region [178]. Microsatellites were

genotyped relative to multiple backcross mice and a linkage region of <0.15 cM was identified with three markers D5Mit408, D5Mit159 and D5Mit242 [179].

From the estimated linkage region, the *Flv* gene was positionally cloned [24]. A genomic physical map was constructed using BACs located between markers D5Mit408 and D5Mit242. Full-length cDNAs were compared between susceptible and resistant mouse strains and a premature stop codon was identified within the *Oas1b* gene in susceptible mice. This C820T transversion produces a gene product lacking 30% of its C-terminal sequence when compared to the full-length gene product in resistant mice. The presence of the full-length and truncated alleles correlated absolutely within resistant and susceptible mouse strains, respectively. The premature truncation was also identified by a second group and correlated to *Flavivirus* susceptibility [180]. The truncated form of the Oas1b protein lacks a CFK motif involved in forming a homotetramer complex required for synthetase activity. However, unpublished data suggests that the full-length Oas1b protein lacks functional synthetase activity. Additionally, a second gene on chromosome 5 has been hypothesized to provide a synergistic effect with *Oas1b* in resistance to *Flavivirus* disease [181].

Oligoadenylate Synthetase Gene Cluster

Human cluster

The human *OAS* gene cluster spans a 130 kb region on human chromosome 12q24.2 [35]. Evidence suggesting evolutionary gene duplication events explain the formation of the human cluster containing three genes, cen-*OAS1*-*OAS3*-*OAS2*-tel [41, 182]. Recent cross-species analysis of mammalian gene clusters has suggested an

evolutionary model of gene duplications and rearrangements of a common gene ancestor resulting in the formation of the three *OAS* genes [33] (Figure 3.1).

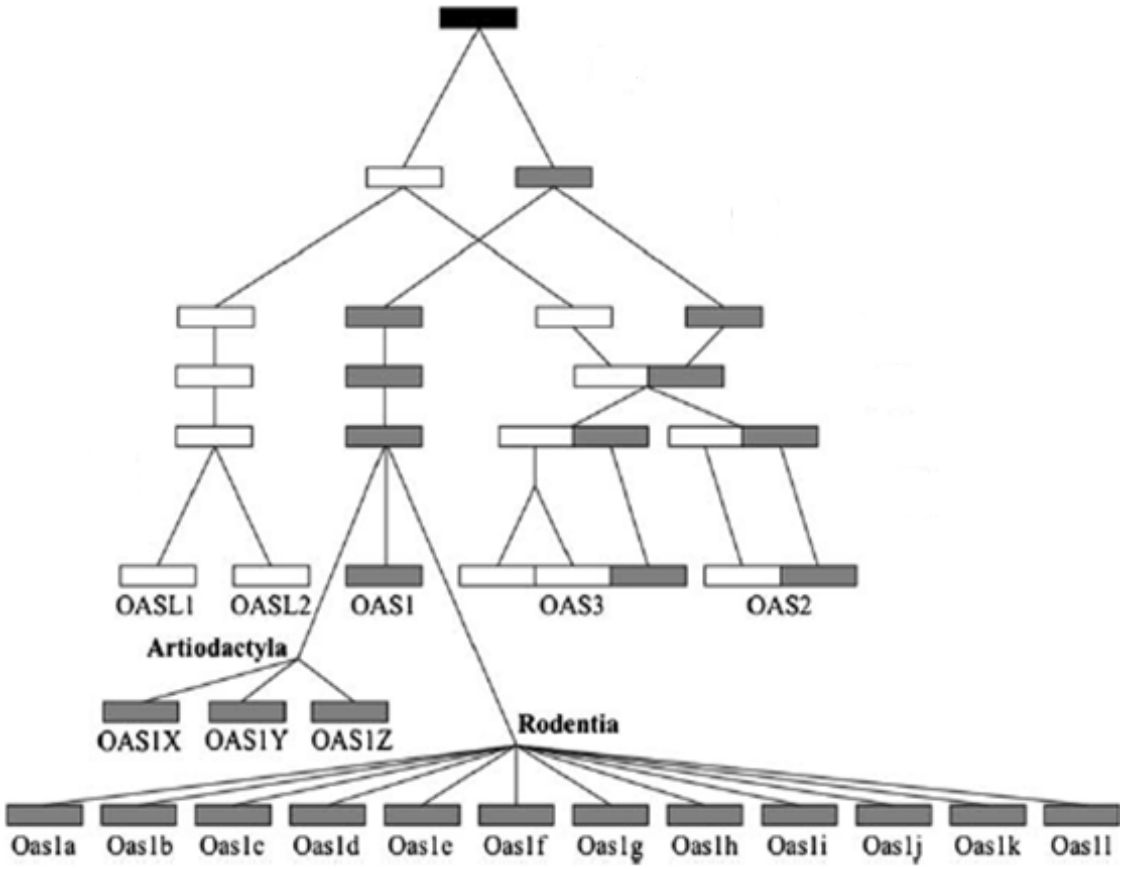


Figure 3.1 Ancestral evolution of the *OAS* gene domains (*Source:* Perelygin, Zharkikh *et al.* 2006).
Schematic representation of the hypothesized evolution of the *OAS* gene cluster. Sequence comparison suggests *OAS2* evolved from a fusion of two *OAS* domains and a subsequent duplication even leading to the formation of *OAS3*

Human *OAS1*, *OAS2* and *OAS3* genes represent the small, medium and large form synthetases, respectively, all with exon structures similar to *OAS1*. The exon structure of these genes provides evidence for evolution by gene duplication. Human *OAS1* contains 7 exons responsible for transcribing alternatively spliced variants which

share their first 5 exons (A-E). Exons A-E of *OAS1* are evident in two and three copies in *OAS2* (A2-E2) and *OAS3* (A2-E2, A3-E3), respectively [41, 182]. Table 3.1 shows the amino acid homology between the homologous exons of human *OAS* genes. As suggested from Figure 3.1 and Table 3.1, a single core exon clusters *OAS2,2* and *OAS3,3* show greater homology to the *OAS1* exon cluster.

The small form synthetases p42 and p46 are products of alternatively spliced 1.6 kb and 1.8 kb transcripts, respectively. The transcript encoding the smaller synthetase contains exons A-E and continues into exon 6; however, the larger product transcribes exons A-E but splices into exon 7. Thus, both OAS1 proteins share their N-terminal 346 amino acids and differ in their C-terminal ends [39, 40]. Figure 3.2 shows the alternatively spliced transcripts of the human *OAS1* gene.

Human *OAS2* contains 12 exons, including two groups of exons similar to exons A-E of *OAS1*. Like *OAS1*, *OAS2* produces differentially spliced transcripts encoding synthetases p69 and p71. The smaller variant consists of 687 amino acids while the larger is of 727 amino acids. Because these proteins result from alternative splicing, they share their first 683 N-terminal amino acids. The high level of homology between the two domains of *OAS2*, especially between *OAS1* and *OAS2* A2-E2, originally suggested that *OAS2* derived from a fusion event between two ancestral *OAS1* genes [41] (Table 3.1). The large form synthetase encoded from a 6,276 nucleotide transcript of *OAS3*, synthetase p100, is the only product of the gene [182]. *OAS3* contains three exon clusters homologous to *OAS1* exons A-E. The exon structures of the three clusters within *OAS3* are similar to *OAS1* and *OAS2*, with similar splice acceptor/donor sites and

Table 3.1 Exon size and identities between *OAS* genes (*Source:* Justesen 2000)

	Exon A		Exon B		Exon C		Exon D		Exon E	
	aa	Identity	aa	Identity	aa	Identity	aa	Identity	aa	Identity
OAS1	60	–	96	–	61	–	76	–	51	–
OAS2, 1	59	32	90	37	59	49	78	50	47	49
OAS2, 2	58	45	96	49	66	56	79	58	51	55
OAS3, 1	59	36	94	38	59	53	80	48	51	53
OAS3, 2	115	21	94	41	59	56	80	53	49	49
OAS3, 3	60	53	95	68	59	66	80	56	49	47
OASL	67	33	95	26	58	55	81	48	49	45

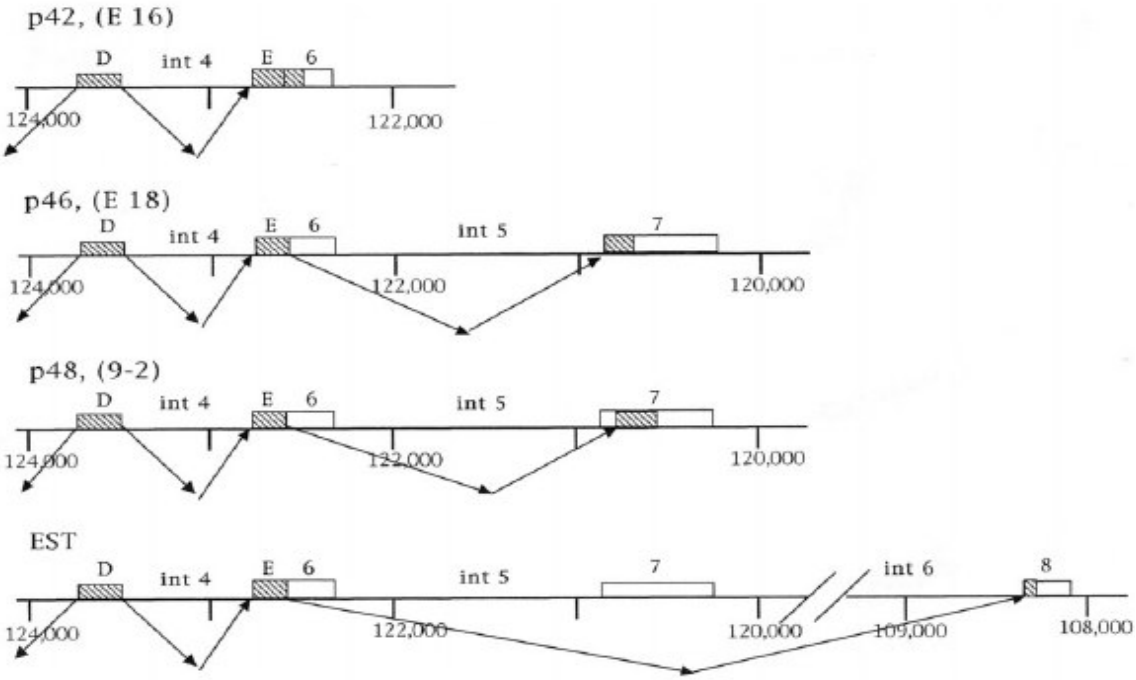


Figure 3.2 Alternatively spliced transcripts of human *OAS1* (*Source:* Justesen 2000). Known splice variants of human OAS1 are homologous in their N-terminal regions and are differentially spliced after exon 5.

reading frames across intron/exon boundaries. The enlarged *OAS3* exon A2 represents the only exon with altered splice sites, leading to an expanded 5' end of the exon [40, 41].

A fourth synthetase gene, *OASL*, contains an N-terminal cluster similar to *OAS1* exons A-E. Therefore, the N-terminal 349 amino acids are homologous to *OAS1*; however, unlike the other synthetases, the C-terminal 165 amino acids contain two ubiquitin-like repeats [40, 183]. Human *OASL* is alternatively spliced, producing a p56 variant as well as a variant lacking exon D, p30.

Investigation of the human *OAS1* promoter identified the regulatory regions of this ISG. An IFN-regulated sequence was originally identified using deletion constructs

driving expression of a chloramphenicol acetyltransferase (*CAT*) reporter gene [184]. They identified a 72 bp region from -159 to -87 responsible for IFN-stimulated induction. This region is conserved within the equine promoter with a nucleotide conservation rate of 55% upon alignment using ClustalX software (Figure 3.3). Shortly thereafter, a second group identified a sequence boundary (-113 to -74) conferring IFN-induced binding of nuclear factors [185]. This human promoter region was sufficient for the IFN-induced activation of a *CAT* reporter gene as well as for the IFN-responsive activation of a heterologous promoter [185, 186].

```
Horse   -175   GAGGCAAAGGAAACGAAACCAAACGGCAGCCCAGACTTG-   -137
                * *    * ** *   *    *   *   ***       *  **
Human   -159   GGGATCAGGGGAGTGT--CTGATTTGCAAAAGGAAAGTGC   -122

Horse   -136   GAAGACGACTTCCTGCTTCCAAGGAAACGAAACCA   -102
                *****  **  *   ****  **************
Human   -121   AAAGACAGCTCCTCCCTTCTGAGGAAACGAAACCA   -87
```

Figure 3.3 Local alignment of human *OAS1* IFN-regulated sequence and horse promoter. Local alignment of the human *OAS1* ISRE sequence and the homologous region of the equine *OAS1* promoter. The high degree of conservation suggests a role for this region as an ISRE in the horse *OAS1* promoter.

A typical ISRE contains repeats of the hexamer sequence AGTGA with a consensus sequence GGYAAAY[A/T]GAAACTY [187]. However, the ISRE located within the promoter of human *OAS1* (*OAS*-ISRE) is differentially activated in response to IFNB than the ISRE of the major histocompatibility complex (MHC) class I genes (MHC-ISRE) [188]. It was shown that in neurons, the ISGF3 complex preferentially bound the *OAS*-ISRE and not the MHC-ISRE. Flanking sequences may affect the binding affinity of the ISGF3 complex to ISRE elements [189]. The cell-type specific

activation of different ISGs may be mediated by the differential binding affinities of the ISRE-binding transcription factors expressed within those cells.

The promoter of the human *OAS1* gene also interacts with the core protein of the Hepatitis C virus (HCV) [190]. Both the HCV core(M) and core(P) proteins activated a luciferase reporter construct containing the human *OAS1* ISRE sequence region from -159 to +82. Activation by the core proteins was observed in a dose-dependent manner where increasing amounts of protein activated the reporter gene to greater levels. Further analysis using deletion constructs lacking regions of the ISRE showed that the activation by HCV core proteins occurs through this promoter element. The role of ISREs in HCV core protein-mediated gene activation was further investigated and it was found that a sub-class of ISREs was differentially activated [191]. Among IFN-induced genes, the sequences of the ISREs are variable [143]. Activation by HCV proteins of luciferase reporter constructs containing the ISREs from *OAS1*, *EIF2AK2* and guanylate binding protein 1 (*GBP1*) were compared and showed preferential activation of the ISRE of *OAS*. A synthetic ISRE converting an inactive ISRE sequence to the *OAS* ISRE sequence recovered the activating potential of the HCV proteins. The synthetic constructs were produced using a PCR-based method [192]. The preferential activation of the *OAS*-form ISRE was also observed using the adenosine deaminase (*ADAR1*) gene, whose ISRE is very similar to *OAS* [191].

Murine cluster

The genomic structure of the murine *Oas* gene cluster is different from that of the human *OAS* gene cluster, particularly with the murine *Oas1* gene(s). While the human

cluster contains single copies of all *OAS* genes, the murine cluster consists of multiple *Oas1* and *Oasl* genes (Figure 3.4). The murine cluster contains eight *Oas1* genes (*Oas1a-h*), an *Oas1i* psuedogene, single copies of *Oas2* and *Oas3* and two *Oas*-like genes (*Oasl1* and *Oasl2*) [33]. Within the identified *Oas1* genes, *Oas1a-g* contain a complete set of exons (A, B, C, D, E and T) [26, 193]. Murine genes *Oas1a* and *Oas1g* contain an additional C-terminal exon 7. All exons of the *Oas1* genes are spliced following typical GT/AG intron-exon boundaries. Alignment of protein residues between multiple *Oas1* genes identified a high degree of conservation within each exon between genes. Investigations using recombinant murine Oas1 proteins identified those proteins with synthetase activity (*Oas1a* and *Oas1g*) while *Oas1c-Oas1e* and *Oas1h* lacked this activity [28, 30]. Identification of inter-genic retrovirus-like elements and molecular evolutionary analysis suggest that the expansion of the mouse *Oas1* genes derived from multiple duplication events occurring as recently as ~11 MYA [28, 33, 194] (Figure 3.1).

The genomic exon structure of murine *Oas2* and *Oas3* genes, like human, contain duplications and triplications of the exon group (A-E), respectively [26]. The exon groups are followed by a single C-terminal T exon in both *Oas2* and *Oas3*. Murine *Oas2* is differentially spliced while *Oas3* is transcribed into a single transcript [193]. The mouse also has two *Oas*-like genes, *Oasl1* and *Oasl2*. Amino acid homology for exons A-E compared to human *OASL* were 74% and 49% for *Oasl1* and *Oasl2*, respectively, suggesting an ancestral role of *Oasl1* in the evolutionary duplication event that produced the *Oasl2* gene [26].

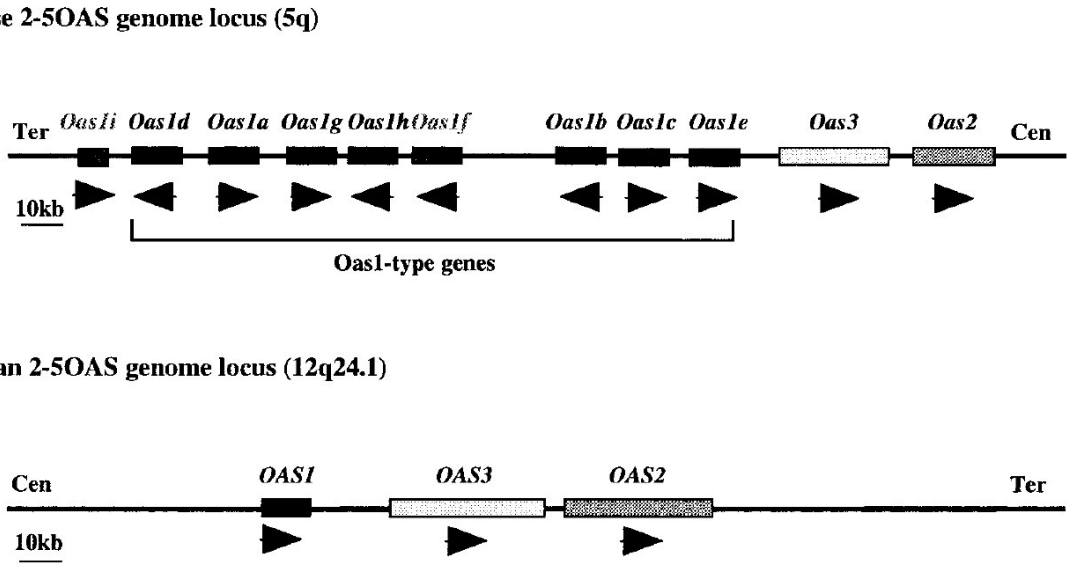


Figure 3.4 Schematic representation of the mouse and human *OAS* gene clusters (*Source:* Kakuta *et al.* 2002).
The multiple *Oas1* genes in the murine cluster may have evolved from duplication events occurring from the presence of retrovirus-like element in the cluster. The human and horse clusters contain single genes of *OAS1* , *OAS2* and *OAS3*.

Canine, bovine and equine clusters

The *OAS* gene clusters among these mammals are strikingly different from the human and murine clusters. The canine (*Canis familiaris*) cluster of *OAS* genes resembles attributes of both the human and murine *OAS* clusters [33]. Similar to the human *OAS* gene cluster, the canine cluster is located on chromosome CFA26 in the same orientation, *OAS1-OAS3-OAS2*. However, the canine *OAS* gene family also resembles the murine cluster, containing two *OAS*-like genes, *OASL1* and *OASL2*. Canine *OASL1* encodes a full-length protein containing two C-terminal ubiquitin domains while the *OASL2* is a pseudogene.

The bovine (*Bos taurus*) *OAS* gene cluster most resembles the human gene cluster, although not entirely. The bovine cluster was FISH mapped to BTA17q24

telomeric to *RPH3A* [33, 64]. Sequencing of bovine CHORI BAC clone 49:I16 identified three *OAS1* genes (*OAS1X-OAS1X-OAS1Y*) and a single *OAS2* gene. Sequencing the intergenic region between the *OAS1* genes and *OAS2* did not identify a bovine *OAS3* gene.

The equine (*Equus caballus*) *OAS* gene cluster was FISH mapped to equine chromosome ECA8p15 and found to contain single copies of each *OAS* gene in the same orientation as the human cluster (*RPH3A-OAS1-OAS3-OAS2*) and a single *OASL* gene [64, 195]. On average, exon lengths are more similar to the human genes than mouse. Furthermore, equine *OAS2* is differentially spliced into two transcripts [33].

OAS Activation and 2-5A Synthesis

The enzymatic products of the *OAS* family of proteins were originally discovered as interferon-induced, low molecular weight inhibitors of protein synthesis able to bind poly(I)·poly(C) columns during purification [53, 54, 56-58]. The mechanism through which this protein family was discovered involves a cascade of activation by dsRNA and downstream processes involving the activation of an endoribonuclease and subsequent RNA cleavage and inhibition of protein synthesis. Following is a review of the mechanisms by which the *OAS/RNASEL* system, in particular *OAS1*, acts within the innate immune system.

Activation by dsRNA

In response to Type I IFN, the *OAS* gene family is induced and synthetases poised for activation by dsRNA. dsRNAs play important roles in signaling many cellular events and thus must be able to differentiate the activation of specific cellular

pathways. Of late, investigation has identified means by which dsRNA species are differentiated for targeted downstream pathway recognition, i.e. RNAi vs. OAS activation. Both the DICER1 complex of the RNAi system and the OAS proteins bind and respond to dsRNA species; however, recent investigation has identified a mode by which the cell differentiates self and non-self (viral) dsRNA species to activate either the RNAi or *OAS* pathways [96].

Characteristics including size and termini of the dsRNA are used to differentiate dsRNA species. Activation of the OAS proteins was measured at maximal levels with dsRNA species of lengths between 65 and 80 nucleotides [99]. This agrees with previous identification of the minimal length threshold of dsRNAs to activate EIF2AK2 and OAS proteins and to inhibit protein synthesis [99, 196, 197]. The RNAi DICER1 complex uses dsRNA species with lengths between 21 and 23 nucleotides for targeted degradation [198]. Additionally, the terminal nature of the dsRNA species differentiates the activation of downstream targets. Upon processing self-dsRNA for targeted degradation, the DICER1 complex recognizes 2 nucleotide 3'-overhangs at the terminus of the dsRNA molecule [198]. These dsRNA species containing 3'-overhangs are endogenous and activate the RNAi pathway. Non-self dsRNA species do not contain the 2 nucleotide overhang and thus do not activate the RNAi pathway. These dsRNAs may be introduced during viral infection/replication or produced from non-specific ribonuclease activity, as seen following RNASEL activation. The specific cellular response is mediated, in part, from the recognition of the 3'-overhang. dsRNA species

containing the typical DICER1-recognized overhang bind RIG1 helicase but fail to activate its helicase activity [199].

dsRNA acts as a potent activator of the host antiviral system by inducing the IFN response as well as by activating antiviral genes such as those of the *OAS* cluster and *EIF2AK2*. As mentioned above, dsRNA species are used to differentially activate host pathways, and depending on the structural nature of the dsRNA molecules. The activating potential of dsRNA dependents on the degree of mismatches within the molecule [200]. Investigations have identified different dsRNA structural requirements for the activation of ISG proteins. The activating potential of partially methylated dsRNA differs between OAS and PKR, providing evidence of differential dsRNA structural requirements of these antiviral proteins [201]. Other structural requirements were identified by measuring OAS and PKR activation from multiple polynucleotides, including analogs of $(A)_n \cdot (U)_n$ and $(I)_n \cdot (C)_n$ [202]. However, the analog polynucleotide species failed to activate OAS to the levels of $(A)_n \cdot (U)_n$ and $(I)_n \cdot (C)_n$.

Synthetase activity of OAS proteins

Using molecular antibodies, the multiple synthetase proteins were identified with molecular weights 40, 46, 69 and 100 kDa [36, 37]. Each protein contains multiple functional domains, including distinct acceptor and donor ATP binding sites and a nucleotidyl transferase catalytic domain [40]. The functional synthetases of OAS1, OAS2 and OAS3 are found as tetramers, dimers and monomers, respectively [203, 204]. This suggests that four catalytic subunits are required for proper activity, provided by four molecules of OAS1 and two molecules of OAS2 [205]. The enzymatic properties

of the OAS2 synthetase were found to be mediated by two independent catalytic domains [206].

The synthetases are induced differentially by IFNA, IFNB AND IFNG in a tissue specific manner [36, 207]. The different synthetase forms also differ in their subcellular localization [36, 204]. The small form synthetases are present in the ribosomal fraction with differential cell type specific expression of the splice variants. The medium form synthetase is localized on the nuclear envelope and in a patterned manner throughout the cytoplasm [40]. The 100 kDa synthetase was localized in a diffuse pattern throughout the cytoplasm. Additionally, the medium form 69 kDa synthetase was found to be myristylated [204].

Upon activation, the synthetases use adenosine triphosphate (ATP) as a substrate for the oligomerization of 2-5A molecules represented with the general formula $pppA(2'p5'A)_n$, where $n \geq 2$ [53-55, 57] (Figure 3.5). These oligomers are phosphorylated at their 5' end and are uniquely structured 2'-5' compared to the normal 3'-5' bond linkage of the genomic DNA helix. Measuring the level of inhibition on type I DNA topoisomerase by the presence of 2-5A, researchers found that the level of inhibition was dependent upon the degree of 5' phosphorylation as well as oligomer length. Inhibition was most effective by a 5'-triphosphorylated hexamer [208]. Oligomerization by the OAS1 synthetase homodimer requires multiple active sites, including an acceptor binding site for ATP, a second for the donor ATP (distinct from the acceptor ATP binding site) and a third site with activity to catalyze the nucleotidyl transferase reaction [40].

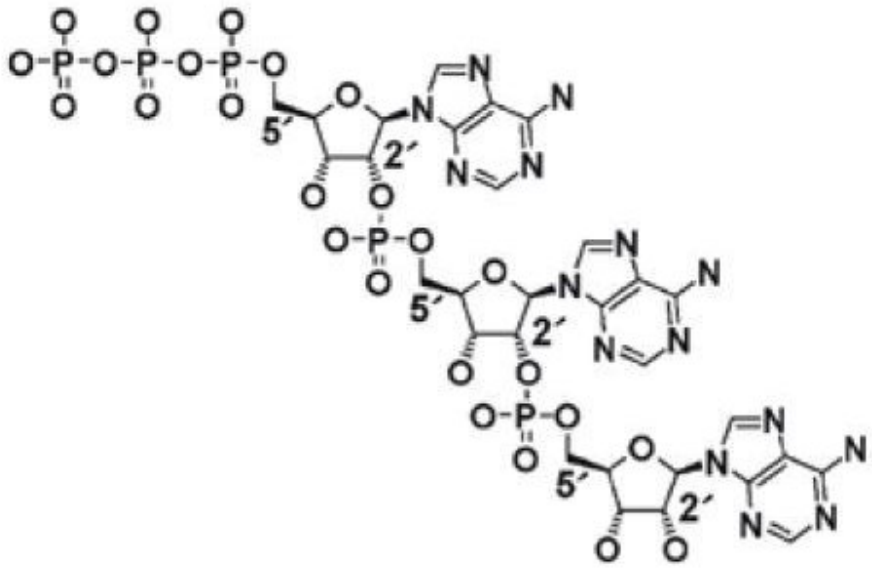


Figure 3.5 Structure of a [(pp)p(A2'p5')$_2$A] molecule (*Source:* Tanaka *et al.* 2004). Activated OAS1 synthetase homotetramer complexes catalyzes the oligomerization of 2-5A molecules from ATP. The preferential 2-5A for RNASEL activation is a tri-phosphorylated trimer.

RNASEL Activation and Activity

The only known role of 2-5A molecules in the immune response to viral infection is to activate the latent RNASEL [53]. The human *RNASEL* gene encodes a 741 amino acid protein from a ~2.8 kb transcript and maps to human chromosome 1q25 [45, 209]. The RNASEL protein contains multiple regions of functional importance, including an N-terminal domain of ankyrin repeats with P-loop motifs between the seventh and eighth repeat, a serine/threonine protein kinase domain and a C-terminal ribonuclease domain.

Deletion analysis identified the nine N-terminal ankyrin repeats as responsible for repressing the ribonuclease activity of unbound RNASEL as well as binding 2-5A for

proper activation [46, 48]. Detailed structural analysis of RNASEL identified a direct binding interaction between a single 2-5A molecule and akyrin repeats 2-4 [48]. Additionally, repeats 6-9 are required for full RNASEL activation from 2-5A binding [210]. The interaction between the 2-5A molecule and ankyrin repeats, and thus ribonuclease activation, is mediated by certain structural attributes of the 2-5A molecule [48, 211]. Structural considerations of the 2-5A molecule include the ribose-phosphate linkage, 5' phosphorylation, length and nature of the bases. Figure 3.6 summarizes the analogs tested and the interpretations regarding the structural requirements of 2-5A on ankyrin binding and RNASEL activation [48]. The activating potential of the oligonucleotide analogs showed that activation by 2-5A is dependent on the specific backbone linkage (2',5') and on containing at least a single phosphoryl group at the 5' end of a molecule of at least three adenylyl residues.

The N-terminal ankyrin repeats, while responsible for activating RNASEL through binding 2-5A molecules, are also responsible for the inactivated repression of ribonuclease activity. The repressive state of RNASEL was maintained when the first 237 N-terminal residues were deleted, corresponding to ankyrin repeats 1-6. These data provide evidence that the repressive function of the N-terminal ankyrin repeats involves at most domains 7-9 [46].

Other functional domains include a domain with protein kinase homology as well as a C-terminal ribonuclease domain [46]. Deletion analysis of the C-terminal region identified 10 residues required for ribonuclease activity. C-terminal deletions including

the most C-terminal 31 residues inactivated the protein. Ribonuclease activity requires amino acid residues between 710 and 720 for proper activity [46].

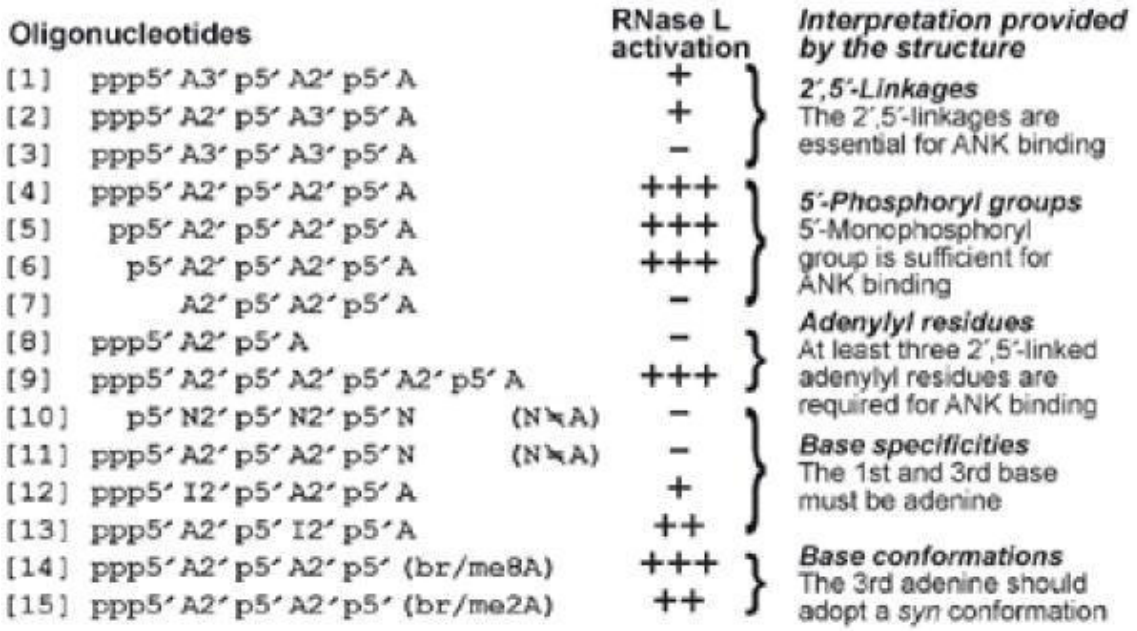


Figure 3.6 RNASEL activation by 2-5A analogs (*Source:* Tanaka *et al.* 2004). 2-5A molecules activate the dimerization of RNASEL proteins. Several characteristics of 2-5A molecules are necessary for proper RNASEL activation. Characteristics include phosphorylation and the unique 2',5' linkages of the 2-5A molecules.

Functional RNASEL ribonuclease activity cleaves both cellular and viral RNA without discrimination. However, the efficiency with which RNA is cleaved depends on the nucleotide preference between the protein and RNA molecule, with cleavage dependent on RNA sequence and not availability of ribonuclease binding. Specific sequence preference for RNASEL was identified by several groups. Cleavage of labeled RNA identified UpN sites as preferred for RNASEL ribonuclease activity. Levels of ribonuclease cleavage were variable among different UpN sites, with a strong preference for UA and UU sites. However, cleavage at all UU and UA sites was not achieved,

possibly due to the presence of secondary structures within the RNA molecule [212]. Cleavage specificity was expanded to investigate ribonuclease activity on different homopolymers. Comparison between reactions without RNASEL, with RNASEL and with both RNASEL and 2-5A showed appreciable levels of ribonuclease activity with poly(U) but not with poly(A), poly(G) or poly(C) [213]. These data show the specific nature of the ribonuclease activity of RNASEL, leading to the patterned cleavage of rRNA [214].

Contrary to the induction of *OAS1* by IFN, there is only a slight increase in expression of *RNASEL* in response to IFN treatment [215]. Basal transcription levels of *RNASEL* are sufficient for antiviral induction by 2-5A. Mapping the promoter region identified general transcription factor binding sites as well as several tissue-specific transcription factor binding sites that together may regulate the ubiquitous expression of human *RNASEL* [216]. A single GAS element was identified ~147 bp upstream of the transcription start site while no ISRE element was found within the promoter.

Additional regulatory motifs are located within the 3' untranslated region for post-transcriptional regulation of the *RNASEL* mRNA [217]. The human *RNASEL* mRNA contains multiple AU-rich elements (ARE) of sequence AUUUA. This regulatory sequence binds ARE-binding proteins that either positively or negatively regulate the transcript's accessibility to mRNA decay mechanisms, including deadenylation and decapping [218, 219]. Deletion analysis of the AREs identified the regulatory roles of these motifs within the 3'UTR [217]. Analysis of multiple single and double deletion constructs showed that AREs 7-8 act as positive regulators while AREs

2 and 3 independently decrease RNASEL levels. Between AREs 7 and 8 lies a binding site for ELAV-like 1 (ELAVL1), an ARE binding protein, which exhibits a regulatory effect independent of the flanking AREs. Finally, the 3' UTR of human *RNASEL* is predicted to contain regulatory binding sites for micro-RNAs.

CHAPTER IV

CHARACTERIZATION OF THE EQUINE 2'-5' OLIGOADENYLATE SYNTHETASE 1 (*OAS1*) AND RIBONUCLEASE L (*RNASEL*) INNATE IMMUNITY GENES*

Background

The innate immune responses are the first line of host defense against virus infection. An important component of the intracellular antiviral response is mediated by the *OAS/RNASEL* pathway. *OAS* genes are interferon-inducible and activated by binding dsRNA. dsRNA, present in virus infected cells, activates OAS proteins to catalyze the oligomerization of ATP to form 2',5' –linked oligoadenylate chains (pppA(2'p5'A)$_n$) [53-55]. Originally discovered as a low molecular weight inhibitor of protein synthesis, pppA(2'p5'A)$_n$ induces the activation of the latent endoribonuclease, RNASEL, which degrades both cellular and viral RNA in a non-preferential manner [53, 56-58]. The *OAS/RNASEL* pathway has also been implicated in the induction of apoptosis [59-63].

The murine *Flv* was positionally cloned and identified as *Oas1b* [24]. A cDNA sequence comparison among susceptible and resistant strains of mice identified a single nucleotide substitution that causes a premature stop codon in the *Oas1b* transcripts of susceptible mice [24].

---

*Reprinted from "Characterization of the equine 2-5 oligoadenylate synthetase 1 (OAS1) and ribonuclease L (RNASEL) innate immunity genes" by Rios JJ, Perelygin AA, Long MT, Lear TL *et al.* 2007. *BMC Genomics*, 8, 313, www. biomedcentral.com/bmcgenomics.

The human *OAS* gene cluster, consisting of genes *OAS1*, *OAS3* and *OAS2,* is located on chromosome 12q24.2 [35]. The small synthetases are transcribed from the *OAS1* gene while the medium and large synthetases are encoded by the *OAS2* and *OAS3* genes, respectively. Alternative splicing was previously reported in both *OAS1* and *OAS2* transcripts [36, 37]. For example, the human *OAS1* transcript E16 corresponds to the p42 protein, which is translated from a ~1.6 kb mRNA, while the alternatively spliced E18 transcript encoding the p46 protein is ~1.8 kb [38]. Both p42 and p46 proteins are identical in their first 346 N-terminal amino acids but differ at the C-terminus [39]. Variations in the human *OAS1* gene that may be relevant to the outcome of virus infections have been reported [42, 43, 220-222].

The human *RNASEL* gene maps to chromosome 1q25 [209]. The 741 amino acid, 83,539 Dalton protein is translated from a ~2.8 kb transcript [44, 45]. The RNASEL protein consists of three domains: 1) an N-terminal domain of ankyrin repeats with P-loop motifs between the seventh and eighth repeat, 2) a serine/threonine protein kinase domain, and 3) a C-terminal ribonuclease domain [46]. RNASEL activation requires binding of a single 2-5A molecule to the N-terminal ankyrin repeats 2-4 [47, 48]. 2-5A binding reverses the naturally repressive state of the RNASEL ankyrin repeats, ultimately producing a functional homodimer with ribonuclease activity [46, 48-50].

Previously, the equine *OAS* gene cluster was mapped to horse (*Equus caballus; ECA*) chromosome 8p15 and shown to have an organization similar to that in the human genome: *OAS1-OAS3-OAS2* [64]. Two clones were identified from segment 1 of the

CHORI-241 equine BAC library, 77:F4 (~200 kb) and 100:I10 (~130 kb), that contain complete *OAS1* and *OAS3* sequences. BAC clone 77:F4 also contains nine 5'-terminal exons of *OAS2* [64].

A subclone library generated from CHORI-241 BAC 100:I10 was sequenced and used to construct a contig assembly spanning the *OAS1* gene. The equine *RNASEL* gene was identified in multiple BAC clones of the CHORI-241 library and FISH mapped on metaphase spreads to ECA5p17-p16. Equine *RNASEL g*enomic sequence was obtained from BAC clone 159:N12 and an assembly similar to that for *OAS1* was constructed. Full-length *RNASEL* cDNAs from 8 species were determined and compared in a phylogenetic analysis. Re-sequencing of genomic DNA from multiple horses of different breeds identified a total of 64 SNPs and 2 microsatellites within the *OAS1* and *RNASEL* genes.

Results

BAC 100:I10 sequencing and OAS1 contig assembly

A shotgun subclone library was constructed from sheared fragments of CHORI-241 BAC 100:I10. Nine hundred sub-clones were bi-directionally sequenced, resulting in 513,390 bases with quality scores >15, providing 3.95X coverage. The individual chromatogram files were analyzed by Phred, Phrap and Consed [65-68, 223] and individual contigs scaffolded on the human genome sequence using BLAST. The scaffold was further validated by the addition of multiple sequences from TraceDB [224] retrieved via BLAST searches using full length equine *OAS1* mRNA [GenBank: AY321355]. The scaffold contained four genomic contigs spanning a substantial part of

the equine *OAS1* gene, including 4.5 kb of promoter sequence upstream of exon 1 and 1.6 kb of sequence downstream of exon 6, and was submitted to GenBank under accession number DQ536887. The genomic assembly also included sequence for the downstream equine *OAS3* gene as well as an upstream gene orthologous to human *RPH3A* (data not shown). This assembly completely overlaps two contigs of whole genome shotgun sequences, AAWR01028567 (55475 bp) and AAWR01028568 (31407 bp), that were recently submitted to GenBank from the Broad Institute.

Identification of OAS1 microsatellites

The genomic sequence assembly identified two microsatellites, one located within the promoter and the other downstream of exon 6. The promoter GT-microsatellite is located 575 bp upstream of the ATG translation initiation site. A shorter GT-microsatellite is in the same relative position in the human *OAS1* promoter (Figure 1.3). This microsatellite may affect the functions of flanking regulatory elements. Sequencing the *OAS1* promoter in 13 horses established this promoter microsatellite as polymorphic in length. The second polymorphic microsatellite was a GT-dinucleotide repeat located 43 bp downstream of exon 6.

OAS1 SNP identification

The assembled *OAS1* scaffold was aligned to the full length, 1.6 kb cDNA equine transcript [GenBank: AY321355] to delineate individual exons and flanking intron sequences from the genomic contigs. Genomic primers were designed within flanking intron sequences as well as for the proximal promoter (Table 4.1).

Table 4.1 Primers for amplifying genomic fragments for SNP detection

Gene	Region	Forward Primer	Reverse Primer
OAS1	Promoter	AACCCACAGAATAAACACCACA	GTCGATGGCTTCTCGGAC
	Exon 1	CCAGACTCAGGCAACGTAAG	GTTTTGCTCTCTCCCTTCCT
	Exon 2	GTGATTGTTGTCTGGTGATGG	AAACTGTGGGAGATTTCTGCT
	Exon 3	GTAACTTGGTTGTGTCCGTGG	AGACTGGATGGAGGGCCATA
	Exon 4	AGCGTGAAAACCACCACAGA	TCCCACATCCTCCATTTCC
	Exon 5	CACTGGGCTGGTCCTCCT	CCTCCAAACGGGGTCAAA
	Exon 6	GCAGGTGGCACGTCACAG	GGCACTGTGCCCTGAAGTTAA
RNASEL	Exon 1	CATCTCCCTTCTCCGTCCTCG	TGCAATGGATGAGTCCTGGT
	Exon 2	CAAAGTACTTCTCTCATCCCCAG	TCCGAAGAGCATGGAACAAA
	Exon 2	AAGATCCTCCTTGATGAGATGG	GGCCTTTCCTATCTGCAATG
	Exon 3	AAATGTAAGTCCTGCTCTTGGC	CAAGCAACTCCACACCAACC
	Exon 4	CTCGTAGCCTGCACCACAC	CACGGTAGATCGCGGAACTT
	Exon 5	CCATGTTAATTCTCTCATCTTCAG	TCTCTCACCTCTTGGTAGGGC
	Exon 6	GCTCCTACATTTTTGCGTAATG	GTTCTTCCCATCAAATAGCAGA
	Exon 7	ATCTCTGGAACCGGGTGCT	CACTACCAAATGGCCCTGAG
	Exon 7	CTCTGGGTGGCTCATTCATT	TCCCAGCTCTTCCCATTACA

Sequence data obtained from the screening population and from CHORI BAC 100:I10 were analyzed using Phred, Phrap and Consed programs [65-68, 223]. Both visual analysis of the chromatogram data to identify heterozygotes and computer analysis using the Consed visualization tool identified 33 single nucleotide substitutions within the proximal promoter and exons of equine *OAS1* (Table 1.2). Of these, 11 were within coding regions, 9 within non-coding regions and the remaining 13 within the proximal promoter upstream of exon 1. Four of the 9 non-coding polymorphisms were located within the 5' and 3' untranslated regions (UTR). Of the 11 coding polymorphisms, 4 were synonymous and 7 were non-synonymous. Five of the 7 non-synonymous SNPs resulted in substitutions of amino acids with different properties. Interestingly, the amino acids encoded by the major alleles of 4 of the 7 non-synonymous mutations were identical to the corresponding amino acids in the human

OAS1 protein [UniProtKB: P00973]. The genotypes of each individual were used to identify potential haplotypes within equine *OAS1* using PHASE v2.1 software [225, 226]. Only those SNPs verified within multiple individuals were used for the haplotype analysis (minor allele frequency ≥ 0.08). The best reconstruction produced 15 haplotypes from the 33 diallelic SNPs (Table 4.2). The polymorphic microsatellites were not included in the analysis.

Table 4.2 Haplotypes of equine *OAS1* and *RNASEL*

Gene	Haplotype Sequence	Count Frequency
OAS1	CTGTCATTCGCTGGAACCCTAGGTCGTAACCTT	0.08
	CTGTCTTTCGCTGGAACCCTAGGTCCCATGCGT	0.04
	CTGTCTTTCGCTGGAACCCTAGGTCCCGTGCGT	0.08
	CTGTCTTTCGCTGGAACCCTAAGTCCCGTGCGT	0.04
	TGACTACCTGCCGCCATCCTAGGTCGTAACCTT	0.19
	TGACTACCTGCCGCCATCCTAGGTCGTAAGCTT	0.04
	TGACTACCTGCCGCCATCCTAGGTCGCAACCTT	0.04
	TGACTACCTGCCGCCATCCTAGGGCGTAACCTT	0.04
	TGACTACCTGCCGCCATTTCGAATCCCGTGTGT	0.12
	TGACTACCTGCCGCCGTTTCGAATCCCAACCGT	0.08
	TGACTACCTGTCGCCGTTTCGAATCCCGACCGG	0.08
	TGACTACCTCCCACCATTTCGAATCCCGTGTGT	0.04
	TGACTACCTCCCACCATTTCGAAGCCCGTGTGT	0.04
	TTGTCATTCGCTGGAACCCTAGGTTCCGTGTGG	0.08
	TTGTCATTCGTTGGCACCCTAGGTTCCGTGTGG	0.04
RNASEL	GACTGCAAAGGGAGCGCTGGGCAGTTTCTTT	0.08
	GATCGCAAGGACGGCGCTGGGCACACCCCCC	0.04
	GATCGCAAGGACGGCGCTGGGCACACCTCCC	0.12
	GATCTCACAGGGAGCGCTGGGCAGATTCTTC	0.12
	GATCTCACAGGGAATAACAAGTGGACCCCCC	0.12
	GATCTAGAGGACGGCGCTAGGCAGATTCTTC	0.08
	GCCCGCAAGGGCGGCGCTAGGCAGATTCTTC	0.23
	GCCCGAAAGGGGAATAACAAGTAGATTCTTC	0.08
	GCCCGAAAGTGGAGCGCTGGGCAGTTTCTTT	0.04
	CCCCGAAAGTGGAGCGCTAGACAGACCCCCC	0.12

Assembling full-length RNASEL mRNA sequences of cattle, dog, horse, cat, domestic pig, Guinea pig, elephant and opossum

A limited number of mammalian *RNASEL* mRNA sequences were previously deposited to GenBank and some of these sequences were predicted from whole genome annotations. However, this GenBank information was not sufficient to identify evolutionarily conserved regions in mammalian *RNASEL* sequences that could be used to design PCR primers to amplify equine *RNASEL* fragments. The predicted sequences of cattle [GenBank: XM_597290] and dog [GenBank: XM_547430] *RNASEL* open reading frames (ORFs) were amplified from commercial cDNA (BioChain, Hayward, CA), directly sequenced and extended to full-length cDNA sequences by DNA walking. The full-length cattle and dog *RNASEL* sequences were submitted to GenBank under accession numbers DQ497162 and DQ497163, respectively. These two sequences as well as the human full-length *RNASEL* sequence NM_021133 were aligned and degenerate primers were designed from conserved regions (Table 4.3) and used to amplify the middle portions of equine *RNASEL* cDNA. This partial sequence was extended to the full-length sequence by DNA walking and submitted to GenBank under accession number DQ497159.

Several additional mammalian *RNASEL* sequences were also determined and subsequently used to perform a phylogenetic analysis. The GenBank feline WGS database was searched with the canine *RNASEL* sequence [GenBank: DQ497163]. Four genomic contigs, AANG01026257, AANG01026302, AANG01630549 and AANG01026248, were detected. These contigs contain the first, second and third, as

Table 4.3 Primer for amplifying *RNASEL* sequence from multiple species

Species	Forward primer	Reverse primer
Cat	CAGGCATCCAGAAGGGAGAC	CAGAGGCAGCCAATCTCTCC
Cattle/Dog	GCTGGTCACCTTTGCATAATGC	CCCAACTCCAAAAGAAGGGATG
Domestic pig	ATGGAGACCAAGCGCCATAACA	TGTTCTCCCAAGTTCCGGATGA
Elephant	ATATCCCTACTAGCCTGACGAG	TTGCCTTGACACCCCCAATTCT
	AGCTGTAGGATGTAACTCTCACT	GATTAGAGGAACCACTGAGAGG
	GCGGTACCTCATTGTGGTTTTG	CCTCTGTATCTTCATGGTCTGG
	TGCCTTTGAATTGTGGTGTTGGT	CCATGTGGTGGATTCATTATAGG
	GTTGAGGTGTCAGGATCTGCAT	GGGGTAACACTGGAACTGTTTC
Guinea pig	TAATGGTCTGGACCATTCCTCC	GTTTGAGGAAAGTGCCTTGCGT
Horse	TTCACRGCYTTCATGGAAGC	CYTTKATCAAAATCTGCCAG

well as the fifth and sixth feline *RNASEL* coding exons, respectively. No contigs containing the fourth coding exon of the feline *RNASEL* gene were found in GenBank. Two primers were designed based on the 3'- end AANG01026302 sequence and the 5'-end AANG01630549 sequence (Table 4.3) and used to amplify and sequence this region from commercial cat genomic DNA (Novagen, Madison, Wisconsin). The sequence of this exon was submitted to GenBank under accession number EF062998. Using this sequence as well as the other exon sequences derived from GenBank (see above), the predicted full-length mRNA sequence of the feline *RNASEL* gene was assembled.

The TIGR porcine database [227] was searched using the cattle sequence [GenBank: DQ497162] and five partial *RNASEL* sequences were found. The TC212507 and TC212872 sequences correspond to the 5'-end of porcine *RNASEL* mRNA, while the TC218317, TC237301, and TC236970 sequences represent the 3'-end. An additional 5'-end cDNA sequence, 20060611S-038813, was detected in the Pig EST Data Explorer [228]. A pair of primers was designed based on the partial sequence

(Table 4.3) and used to amplify pooled cDNA (kindly provided by Dr. Jonathan E. Beever, University of Illinois at Urbana–Champaign). The middle portions of the porcine *RNASEL* cDNA were directly sequenced. The partial sequence was then extended to the full-length sequence by DNA walking and submitted to GenBank under accession number DQ497160.

The GenBank Guinea pig whole genome sequence database was searched using both mouse [GenBank: NM_011882] and rat [GenBank: NM_182673] full-length *RNASEL* sequences. Two Guinea pig sequences, AAKN01052053 and AAKN01424676, showed significant similarity to the 5' and 3' regions of the rodent *RNASEL* sequences, respectively. These two sequences were used to design primers (Figure 4.3) to amplify commercial cDNA (BioChain, Hayward, CA) and directly sequence the middle portions of Guinea pig *RNASEL* cDNA. This partial sequence was extended to the full-length sequence by DNA walking and submitted to GenBank under accession number DQ497161.

Cattle, dog, horse and pig *RNASEL* sequences were used to search the GenBank elephant genome trace archive using the discontiguous Mega BLAST program. The same sequences were also used to search the GenBank elephant WGS database using the BLASTN program. The sequences for all potential exons of the elephant *RNASEL* gene were identified. Based on these sequences, five primer pairs (Table 4.3) were designed to amplify genomic DNA (kindly provided by Drs. Alfred L. Roca and Stephen J. O'Brien, National Cancer Institute) and directly sequence each of the elephant *RNASEL*

exons. The resulting sequence was submitted to GenBank under accession number DQ497164.

The *RNASEL* ORF sequence of the laboratory opossum (*Monodelphis domestica*) was predicted by searching the UCSC genome browser [229] using the BLAT program. No sequence traces similar to *RNASEL* were detected in frog (*Xenopus tropicalis*) or several fish species (*Danio rerio*, *Takifugu rubripes* and *Tetraodon nigroviridis*).

Phylogenetic analysis of vertebrate RNASEL gene sequences.

Human, chimpanzee, orangutan, rhesus macaque, mouse, rat and chicken ORF sequences of *RNASEL* genes were downloaded from GenBank and aligned to orthologous sequences described above to build a phylogenetic tree (Figure 4.1). Rodents showed the highest rate of nucleotide substitutions, while primates showed the lowest rate of evolution. Evolution rates were found to be fairly uniform in the three different RNASEL domains: ankyrin repeats, serine/threonine protein kinase domain, and ribonuclease domain.

Assignment of the RNASEL gene to horse chromosome ECA5p17-p16

The horse CHORI-241 BAC library was searched with a probe derived from the partial equine *RNASEL* cDNA fragment. Twelve positive clones were identified and two of them, 108:P15 and 189:I19, were FISH mapped to assign the *RNASEL* gene to the horse chromosomal location ECA5p17-p16 (Figure 4.2).

Exon/intron structures of vertebrate RNASEL genes

A partial sequence of the equine *RNASEL* gene was obtained by sequencing PCR fragments of BAC 159:N12. The mRNA sequence [GenBank: DQ497159] was used as

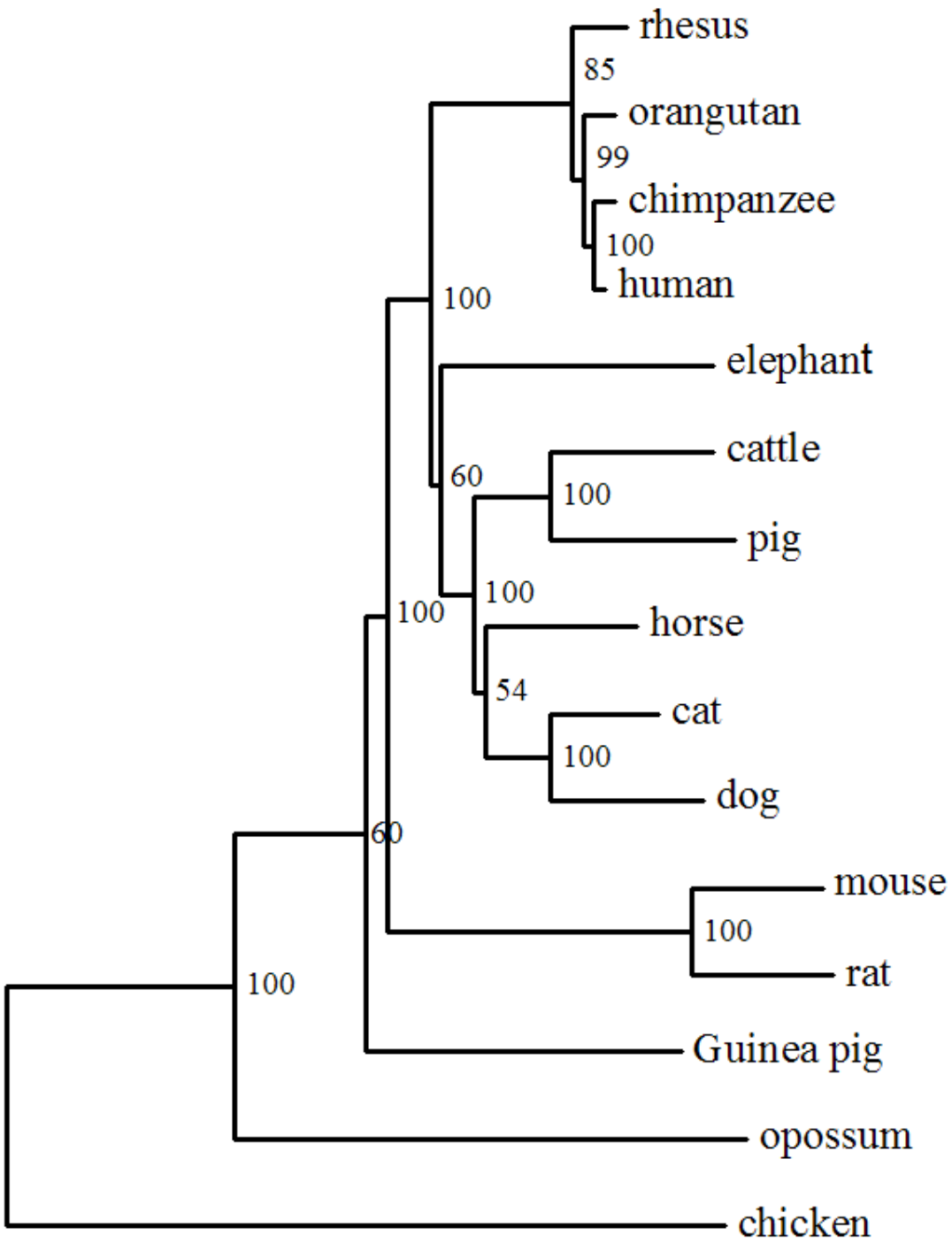


Figure 4.1 Phylogenetic tree of vertebrate *RNASEL* genes.
RNASEL ORF sequences from 15 vertebrate species were aligned and the njtree program was used for tree construction.

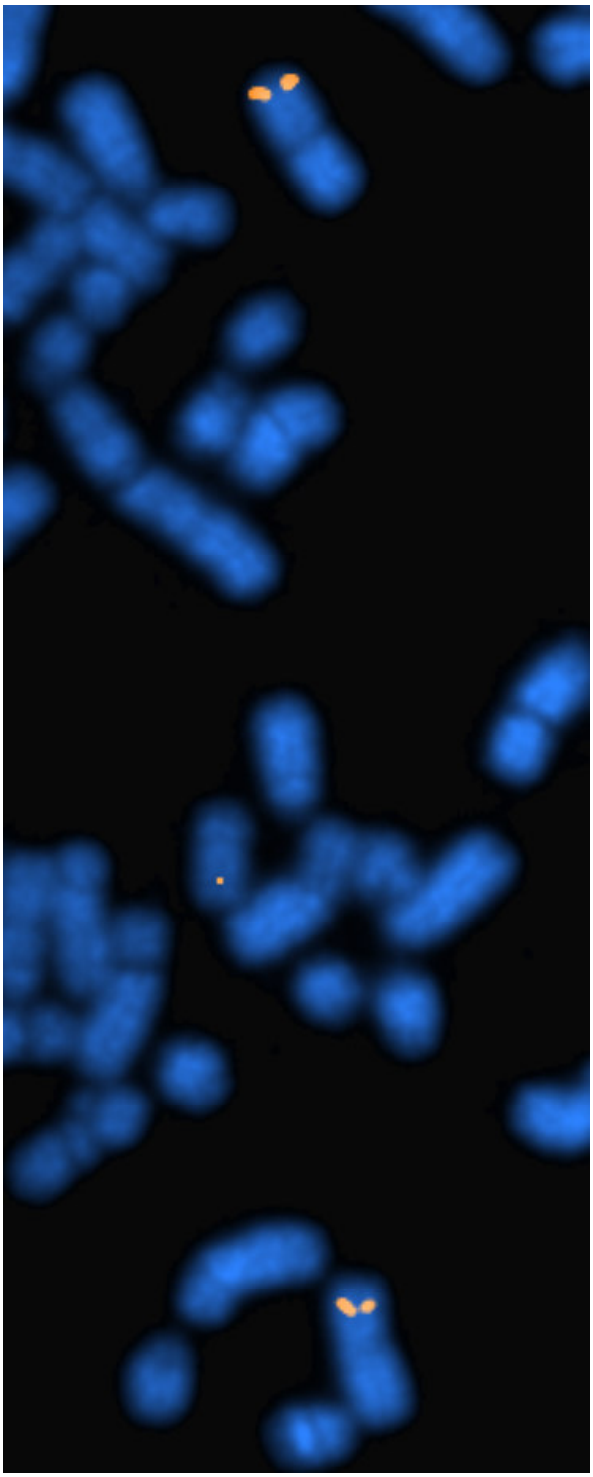


Figure 4.2 FISH mapping equine *RNASEL.*
FISH map position ECA5p17-p16 of horse *RNASEL* gene (orange) on DAPI counterstained metaphase chromosomes (blue).

a reference for determining intron/exon junctions. Sufficient genomic sequence was obtained to build a scaffold as described for the equine *OAS1* gene. The scaffold was verified using sequences from TraceDB [224] and submitted to GenBank under accession number EF070193. This scaffold completely overlaps the WGS sequence AAWR01030439 (193510 bp) that was recently submitted to GenBank from the Broad Institute. Comparison of genomic and mRNA sequences revealed six coding and one 5'-terminal non-coding exon in the equine *RNASEL* gene. This exonic composition is similar to that of a number of other mammalian *RNASEL* genes. However, two and three 5'-terminal non-coding exons were found in the chicken and mouse *RNASEL* genes, respectively. The coding vertebrate *RNASEL* exons were designated A through F. Comparison of the genomic and mRNA sequences of vertebrate *RNASEL* genes revealed significant length variation in both the 5'- (1402-1510 bp) and 3'-terminal (130-187 bp) coding exons (Table 4.4).

SNP identification in the equine RNASEL gene

After identification of the equine *RNASEL* introns, exon-specific genomic primers were designed (Table 4.1). Exon-specific sequencing of DNA from the screening population identified 31 SNPs within the *RNASEL* gene (Table 1.3). Of the 10 non-coding polymorphisms, one was within the second intron and the others were located in the 5' and 3' UTRs. Seventeen of the 31 SNPs were located within the ankyrin repeat-encoding exon 2, 13 of which are non-synonymous, with 10 resulting in substitutions of amino acids with different properties. Three non-synonymous polymorphisms were identified within exons 3 and 5. The remaining exons, including

Table 4.4 Lengths of coding exons (bp) within ORFs of vertebrate *RNASEL* genes

Species	Exon A	Exon B	Exon C	Exon D	Exon E	Exon F
Horse	1480	86	206	133	134	130
Cat	1477	86	206	133	134	139
Dog	1477	86	206	133	134	139
Cattle	1474	86	206	130	131	145
Elephant	1510	86	206	133	137	187
Human	1480	86	206	133	134	187
Chimpanzee	1480	86	206	133	134	187
Orangutan	1480	86	206	133	134	187
Rhesus	1480	86	206	133	134	187
Mouse	1474	86	206	133	137	172
Rat	1489	86	206	133	131	172
Guinea pig	1462	86	206	133	134	187
Opossum	1453	86	206	129	131	139
Chicken	1402	89	191	124	122	136

the non-coding exon 1 were invariant among these horses. The amino acids encoded by the major allele of 11 of the 16 non-synonymous mutations were identical to the corresponding human RNASEL amino acid [UniProtKB: Q05823]. Using MOTIF Search [230] to identify putative transcription factor binding motifs in the TRANSFAC database, the promoter SNP was identified within a potential (Score: 90) cAMP-response element binding site upstream of the first exon. Haplotypes were assembled in the same manner as for the equine *OAS1* gene. The best reconstruction from Phase analysis produced 10 haplotypes among the 31 verified diallelic SNPs with minor allele frequencies ≥ 0.08 (Table 4.2). As with *OAS1*, only good quality, unambiguous resequencing data were used for the haplotype analysis.

Identifying SNPs by sequencing DNA from multiple individuals enhances the possibility of artifacts either from PCR or sequencing error. The 64 SNPs identified from the equine *OAS1* and *RNASEL* genes were considered valid if each allele was identified in at least two individuals. Eight additional SNPs were identified but could not be verified in more than one individual (minor allele frequency < 0.08). Within the 3,864 bp and 5,406 bp re-sequenced during the SNP identification for *OAS1* and *RNASEL*, respectively, equine *OAS1* contained an average of one polymorphism per 117 bp, while equine *RNASEL* averaged one polymorphism per 174 bp.

Discussion

Sequence characterization of equine *OAS1* in CHORI-241 BAC 100:I10 enabled a partial genomic sequence assembly [GenBank: DQ536887] and comparison among multiple equine individuals. Re-sequencing identified 2 polymorphic microsatellites and 33 SNPs from a group of 13 individuals and BAC 100:I10 (Table 1.2). Although the effects of the majority of mutations detected are unknown, a single mutation that results in a Arg209Cys substitution may significantly change OAS1 enzymatic activity. Arg209 in the equine OAS1 protein corresponds to Arg544 in the human OAS2 protein, which is located in the donor binding domain. Substitution of Arg544 with either Ala or Tyr significantly decreased enzymatic activity of the OAS2 protein [231]. In addition, the equine *OAS1* promoter SNP at position 4531 is located in a potential ISRE [230]. Inactivation of this regulatory element by a single nucleotide substitution may alter expression of the equine *OAS1* gene.

RNASEL enzymatic activity was previously reported in reptiles, birds, and mammals [232]. However, no *RNASEL* genes have been found for amphibians or fishes to date. This observation is in good agreement with the absence of *OAS* genes in these same classes of vertebrates [33]. These data suggest that the *OAS* and *RNASEL* genes, which are functionally connected, co-evolved in birds and mammals.

The equine *RNASEL* gene was FISH mapped to chromosomal location ECA5p17-p16. Orthologous genes are located on primate chromosome 1 (human, chimpanzee and rhesus macaque), cattle chromosome 16, dog chromosome 7, mouse chromosome 1, rat chromosome 13 and chicken chromosome 8 [233]. Using comparative chromosome painting (Zoo-FISH), similarities between human chromosome 1 and horse chromosome 5 [234], mouse chromosome 1, rat chromosome 13 [235], dog chromosome 7 [236, 237] and cattle chromosome 16 [238] were previously established. Our results further confirm the conservation of *RNASEL*-containing syntenic chromosomal segments in horses.

Thirty one SNPs were identified in equine *RNASEL* (Table 1.3). Interestingly, all but three of the 17 coding SNPs identified are located within exon 2. The RNASEL protein contains 9 N-terminal ankyrin repeats responsible for binding 2-5A molecules that are essential for activation [46]. Exon 2 of the human *RNASEL* gene encodes the entire ankyrin repeat region (amino acid 24 to 329). The high frequency of non-synonymous polymorphisms within exon 2 suggests that a single SNP or haplotype could ablate 2-5A binding and/or other RNASEL interactions. As well, the SNP identified within the promoter upstream of the first exon is located within a potential

cAMP-response element binding site. Mutations within this promoter element have been shown to affect gene expression [239-241].

A number of SNPs were detected within the 3' UTR region of the equine *RNASEL* gene. Of the eight SNPs found within this region, six result in transitions. The 3' UTR regions of mRNAs contain regulatory regions capable of protein and microRNA binding that control mRNA stability, translation and localization. A simple analysis of octamer motifs containing equine 3' UTR SNPs identified SNP 10247 as being within a human miRNA target site [242]. If this target site is functionally conserved in horses, this SNP could significantly affect RNASEL synthesis.

Genotype analysis using PHASE v2.1 [225, 226] identified 15 and 10 haplotypes among equine *OAS1* and *RNASEL* genes, respectively, and suggested the existence of haplotype blocks spanning most of each gene (Table 4.2). Even if efforts to show an association between viral-induced disease susceptibility and *OAS1* and/or *RNASEL* SNPs are successful, it may prove difficult to unambiguously identify a single causal SNP because of potential linkage disequilibrium at these loci. As determined from our screening population, a single haplotype occurred more frequently than any other, with a frequency of 0.19 and 0.23 in *OAS1* and *RNASEL*, respectively (Table 4.2).

The frequency of SNP identification in this study in two equine genes was high considering the previously estimated equine SNP frequency of 1 per 1500 bp [243]. In dogs, the estimated SNP frequency is ~1 per 1600 bp (based on entire genome re-sequencing), but a higher frequency of ~1 per 900 bp was estimated between breeds [244]. Re-sequencing of specific genes in several breeds of the domestic dog identified

polymorphisms at frequencies comparable to our estimates, with 1 SNP per ~250-330 bp [S. Canterbury, personal communication]. Furthermore, re-sequencing within an Elk (*Cervus elaphus nelsoni*) putative promoter region, which is highly conserved between mule deer, cow and sheep, detected an average one SNP per 69 bp [unpublished data].

The microsatellite identified within the promoter region in this study may also alter expression of the equine *OAS1* gene. The alleles observed to date indicate that dinucleotide repeat lengths of 9 and 18 may represent the major alleles at this locus. The over-representation of these alleles may be due to the fact that they correspond to complete rotations of the DNA helix. If this microsatellite separates cis-regulatory elements, alterations in its length could affect binding of transcriptional regulators to these elements and significantly alter gene expression. In support of this hypothesis, there is a high degree of conservation between human and horse *OAS1* promoters in the regions flanking the microsatellite (Figure 1.3).

Conclusion

This research reports the genomic sequences of the equine *OAS1* and *RNASEL* genes and identifies 64 single nucleotide polymorphisms and 2 polymorphic microsatellites in these genes. On the basis of the allelic variants characterized, a number of these are plausible candidates for regulatory or structural mutations, which may influence *OAS1* transcription or enzymatic activity of OAS1 and RNASEL proteins. Also, *RNASEL* cDNA sequences were determined for 8 mammals and utilized in a phylogenetic analysis. The chromosomal location of the *RNASEL* gene was assigned by FISH to ECA5p17-p16.

Methods

RNASEL cDNA and FISH

Preparation of horse cDNA was described previously [64]. Partial *RNASEL* sequences were extended using a DNA Walking SpeedUp Kit (Seegene USA, Del Mar, CA) according to the manufacturer's protocol. Four high-density filters for segment 1 of the CHORI-241 equine genomic BAC library were purchased from the Children's Hospital Oakland Research Institute, Oakland, CA. These filters were screened using a P^{32}-labeled equine *RNASEL* cDNA probe according to the supplier's protocol. Two positive equine BAC clones were purchased from CHORI. Each of these BAC clones was grown individually in 500 mL of LB media. BAC DNA was isolated using the NucleoBond BAC Maxi Kit (BD Biosciences Clontech, Palo Alto, CA) and used as the template for direct partial sequencing with a BigDye terminator v1.1 Cycle Sequencing Kit on an ABI 3100 Genetic Analyzer according to the manufacturer's recommendations. DNA from equine BAC clones 108:P15 and 189:I19 was FISH mapped as described previously [245]. International cytogenetic nomenclature of the domestic horse [246] was used to identify individual horse chromosomes.

The njtree program was used to construct a phylogenetic tree as described previously [33]. This program is available upon request.

Construction of subclone library

BAC clone 100:I10 was isolated from segment 1 of the CHORI-241 equine BAC library at Texas A&M University and confirmed by PCR as containing *OAS1*. The colony-isolated clone was cultured and BAC DNA isolated by standard alkaline/lysis

miniprep using Millipore Solutions and treated with Plasmid-Safe ATP-dependent DNAse (Epicentre, Madison, WI). BAC DNA was fragmented using a HydroShear® DNA Shearing Device (GeneMachine, San Carlos, CA) at Speed Code 8 for an estimated fragment size of 2.5 kb. The fragmented product was analyzed by agarose gel electrophoresis stained with ethidium bromide and gel extracted using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA). Extractions were eluted in water according to the manufacturer's protocol. Purified fragments were cloned into vector pCR®4Blunt-TOPO® using the TOPO® Shotgun Subcloning Kit (Invitrogen, Carlsbad, CA) following the manufacturer's protocol. Ligation reactions were incubated 30 minutes at room temperature and electroporated into *E. coli*. Colonies were screened for lack of β-galactosidase activity and selected for ampicillin resistance on LB-agarose plates containing 50 µg/mL ampicillin. White colonies were cultured and screened for appropriate insert size by PCR using vector-sequence M13 primer sites flanking the cloned insert, prior to sequencing.

Sequencing of clones

Individual *OAS1* inserts were amplified directly from individual colonies by PCR using vector-sequence M13 primer sites flanking the cloned insert. Amplification products were purified by centrifugation with the Qiaquick PCR Purification Kit (Qiagen, Valencia, CA) in 96-well plate format according to manufacturer's protocol. Purified products were sequenced in separate reactions with each M13 primer using a cycle sequence of 96C, 10 sec; 50C, 5 sec; 60C, 4 min with BigDye® Terminator Mix

v1.1 (Applied Biosystems, Foster City, CA). Sequencing reactions were analyzed using an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA).

Primers were designed to amplify the immediate promoter and exons of *OAS1* and *RNASEL* genes from 13 individual horses by PCR (Table 4.1). Sequencing was carried out in the same manner as used for the library subclones. Sequences obtained were compared between individuals to identify SNPs within the amplified regions.

Sequence analysis and contig assembly

Sequences were assembled and analyzed using Phrap assembly software [65, 66, 223] and viewed with the Consed visualization tool [67, 68, 223]. Contig and singleton reads were assembled by scaffolding onto the human genome using BLASTN [247-249].

Additional sequences were added to the assembly data and re-analyzed with Phrap and BLAST until the consensus sequence spanned the genes from the promoter to the 3' UTR. The genomic equine consensus sequence was confirmed using data from the Equine Genome Sequencing Project (2x) [69] and intron/exon boundaries were assigned by local alignment to the full-length equine *OAS1* [GenBank: AY321355] and *RNASEL* [GenBank: DQ497159] cDNAs. The equine genomic sequences of *OAS1* and *RNASEL* were submitted to GenBank and assigned accession numbers DQ536887 and EF070193, respectively.

Genotyping population

Blood samples were collected at the Texas A&M University Equestrian Center in accordance with ethical standards. The sampled set used for screening consisted of 13 horses, including 10 geldings/stallions and 3 mares, ranging in age from 21 months to 20

years. Breeds represented include American Quarter Horse (9), Arabian (1), American Paint Horse (1), Appaloosa (1) and Thoroughbred (1).

White blood cells (WBC) were digested with Proteinase K (Promega, Madison, Wisconsin) and washed twice with phenol/chloroform and ethanol precipitated. Purity and concentration were analyzed by Nanodrop©.

CHAPTER V

OAS1 AND *RNASEL* POLYMORPHISMS ARE ASSOCIATED WITH SUSCEPTIBILITY TO WEST NILE ENCEPHALITIS IN HORSES

Background

Naturally susceptible and resistant mouse populations led to the identification of the *Flv* gene as *Oas1b* [24, 180]. A nonsense mutation identified within the *Oas1b* gene of susceptible mice results in the translation of a truncated protein lacking C-terminal functional domains**.** The full-length *Oas1b*^{r} genotype was recently shown to confer resistance to WNV infection in susceptible mice [250]. The interferon-induced *OAS* genes encode dsRNA-activated proteins which catalyze the synthesis of 2'-5'-linked oligoadenylates from ATP [53-55]. The only known function of the 2-5A molecules are to activate the dimerization of RNASEL proteins for the degradation of cellular and viral RNA [53, 56-58]. Interestingly, the full-length murine Oas1b protein lacks synthetase activity, suggesting an antiviral function of Oas1b independent of RNASEL activation [251].

Susceptibility to severe West Nile encephalitis among mammalian species is naturally variable [3]. Experimental infections in sheep [4], calves [5], pigs [6] and dogs [7] showed these domestic species are poor hosts for, or develop only mild clinical symptoms to, infection from WNV. Conversely, horses are particularly susceptible to infection from WNV and typically show clinical symptoms including fever, ataxia, paralysis and death. Treatments for human and equine patients are similar, providing only supportive care targeted to minimize symptom severity.

To investigate a potential role of the mammalian innate immune response to WNV infection, a two-stage association study was conducted using SNPs within the equine *OAS1* and *RNASEL* genes. The known structure of the *OAS* gene clusters of several domesticated mammals is variable. Both the canine and bovine clusters contain gene duplications in *OASL* and *OAS1* genes, respectively [33, 64]. The equine cluster is more similar to the human *OAS* gene cluster than to any other domesticated mammal, with single copies of each gene, *OAS1-OAS3-OAS2*, and a single *OASL* gene [35, 64].

Numerous single nucleotide polymorphisms were identified in both equine *OAS1* and *RNASEL* genes and two polymorphic microsatellites within the *OAS1* gene [252]. Each SNP was genotyped among case and control populations infected with WNV. Control individuals consisted of unvaccinated horses infected with WNV through natural mosquito transmission yet failed to exhibit clinical symptoms. Horses genotyped and included in the case population were unvaccinated, naturally infected and subsequently developed clinical symptoms requiring treatment from veterinary services. Veterinary examination of clinical horses noted a variety of symptoms, the most common including forelimb and/or hindlimb ataxia. Diagnostic tests confirmed WNV infection in both case and control individuals. The relevance of case-control analyses is strongly affected by the comparability in infection between case and control populations [253]. Previous case-control studies finding SNP association to West Nile susceptibility in human patients lacked such highly comparable case and control populations, likely a result of the few numbers of known infected patients failing to exhibit clinical signs [43]. In this study, recently hospitalized human West Nile patients were genotyped and compared to

HAPMAP samples collected prior to the first reported infection of WNV in the United States. This report attempts to accommodate the need for highly comparable case and control samples by investigating equine susceptibility to WNV.

Results and Discussion

This study involved genotyping 66 equine *OAS1* and *RNASEL* SNPs among 20 control and 58 case samples. Genotype data was analyzed using STATA 9 [254] software to identify statistically significant allelic (2x2) and genotypic (2x3) associations to WN encephalitis using Fisher's Exact tests. Odds ratios were also calculated for homozygous and heterozygous genotypes. Analyses using Fisher's Exact Test identified 13 SNPs in *OAS1* (Table 5.1) and a single Glutamine to Arginine mutation in exon 2 of *RNASEL* (Table 5.2) significantly associated with WNV susceptibility. Eleven of the significant polymorphisms are located in the promoter and 5'UTR, flanking the polymorphic microsatellite of *OAS1*. Only the *OAS1* promoter polymorphisms and the *RNASEL* polymorphism had statistically significant odds ratios. Using the data collected from our case and control populations, horses genotyped with susceptibility-associated alleles are up to 11.77 times more likely to suffer severe West Nile encephalitis upon infection (Table 5.3). Discrepancies between allelic and genotypic significance values as well as the homozygous and heterozygous odds ratios likely resulted from sampling error. The case-control association reported here is exploratory in nature and confirms the contribution of the *OAS/RNASEL* antiviral system in equine resistance to West Nile virus.

Table 5.1 Fisher's Exact test for *OAS1* allelic (2x2) and genotypic (2x3) associations

				Allele Frequency		Fisher p-value	
SNP ID	Gene	Region	Allele	Clinical	Subclinical	2x2	2x3
snp6567078	**OAS1**	**Promoter**	**C**	**0.51**	**0.21**	**0.002**	**0.001**
snp6567031	**OAS1**	**Promoter**	**T**	**0.48**	**0.21**	**0.004**	**0.001**
snp6567000	**OAS1**	**Promoter**	**G**	**0.37**	**0.20**	**0.074**	**0.000**
snp6566994	**OAS1**	**Promoter**	**T**	**0.45**	**0.20**	**0.007**	**0.000**
snp6566893	**OAS1**	**Promoter**	**C**	**0.44**	**0.20**	**0.007**	**0.000**
snp6566888	**OAS1**	**Promoter**	**T**	**0.29**	**0.15**	**0.133**	**0.003**
snp6566745	**OAS1**	**Promoter**	**T**	**0.46**	**0.20**	**0.004**	**0.000**
snp6566713	OAS1	Promoter	G	0.10	0.03	0.298	0.169
snp6566498	**OAS1**	**Promoter**	**T**	**0.46**	**0.20**	**0.004**	**0.001**
snp6566399	OAS1	Promoter	C	0.24	0.20	0.792	0.091
snp6566277	OAS1	Promoter	G	0.95	0.88	0.134	0.122
snp6566231	**OAS1**	**Promoter**	**T**	**0.55**	**0.20**	**0.000**	**0.000**
snp6566201	OAS1	Promoter	G	0.94	0.85	0.093	0.205
snp6566134	**OAS1**	**5'UTR**	**G**	**0.42**	**0.18**	**0.013**	**0.092**
snp6566107	**OAS1**	**5'UTR**	**A**	**0.39**	**0.20**	**0.039**	**0.197**
snp6566042	OAS1	Exon 1	G	0.12	0.08	0.540	1.000
snp6565949	OAS1	Exon 1	C	0.37	0.20	0.095	0.278
snp6565123	OAS1	Intron 1	T	0.23	0.25	0.828	0.593
snp6565031	OAS1	Exon 2	T	0.29	0.25	0.681	0.395
snp6564989	OAS1	Exon 2	T	0.81	0.75	0.490	0.168
snp6564967	OAS1	Exon 2	A	0.83	0.75	0.348	0.203
snp6564956	OAS1	Exon 2	G	0.81	0.75	0.490	0.168
snp6564946	OAS1	Exon 2	G	0.81	0.75	0.498	0.238
snp6550610	OAS1	Exon 4	C	0.74	0.55	0.045	0.149
snp6550514	OAS1	Intron 4	C	0.61	0.50	0.248	0.524
snp6550471	OAS1	Intron 4	G	0.68	0.47	0.044	0.135
snp6549905	OAS1	Intron 4	G	0.52	0.39	0.251	0.555
snp6549803	OAS1	Exon 5	G	0.32	0.16	0.100	0.339
snp6549696	OAS1	Intron 5	T	0.50	0.45	0.702	0.360
snp6549675	OAS1	Intron 5	G	0.48	0.42	0.570	0.346
snp6548520	OAS1	Exon 6	T	0.37	0.24	0.211	0.198
snp6548430	OAS1	3' UTR	T	0.41	0.50	0.424	0.189

Equine *OAS1* polymorphisms genotyped among case and control populations for statistical analysis by Fisher's Exact test. Allele frequencies and allelic and genotypic p-values are shown. Bolded SNPs were significantly associated with WNV phenotype.

Table 5.2 Fisher's Exact test for *RNASEL* allelic (2x2) and genotypic (2x3) associations

				Allele Frequency		Fisher p-value	
SNP ID	Gene	Region	Allele	Clinical	Subclinical	2x2	2x3
snp2758810	RNASEL	Promoter	G	0.83	0.65	0.444	0.736
snp2756586	RNASEL	Exon 2	C	0.50	0.32	0.127	0.252
snp2756461	RNASEL	Exon 2	A	0.84	0.71	0.139	0.271
snp2756452	RNASEL	Exon 2	T	0.36	0.39	0.835	0.706
snp2756423	RNASEL	Exon 2	C	0.85	0.71	0.128	0.126
snp2756422	RNASEL	Exon 2	G	0.31	0.26	0.663	0.852
snp2756421	RNASEL	Exon 2	C	0.28	0.34	0.517	0.562
snp2756325	RNASEL	Exon 2	C	0.76	0.71	0.643	0.545
snp2756127	RNASEL	Exon 2	C	0.26	0.21	0.653	0.843
snp2756111	RNASEL	Exon 2	G	0.59	0.47	0.324	0.688
snp2756069	RNASEL	Exon 2	T	0.25	0.24	1.000	1.000
snp2756056	RNASEL	Exon 2	G	0.75	0.79	0.762	0.724
snp2756043	RNASEL	Exon 2	G	0.75	0.76	1.000	1.000
snp2756001	RNASEL	Exon 2	C	0.46	0.42	0.798	0.652
snp2755808	RNASEL	Exon 2	G	0.77	0.83	0.475	0.907
snp2755763	RNASEL	Exon 2	G	0.66	0.81	0.132	0.421
snp2755672	RNASEL	Exon 2	G	0.37	0.25	0.285	0.597
snp2755299	RNASEL	Exon 2	G	1.00	0.97	0.340	0.340
snp2755162	RNASEL	Exon 2	G	1.00	0.97	0.360	0.360
snp2755142	RNASEL	Exon 2	C	0.98	0.97	1.000	1.000
snp2755132	RNASEL	Exon 2	T	0.97	0.97	1.000	0.605
snp2755071	**RNASEL**	**Exon 2**	**A**	**0.50**	**0.22**	**0.005**	**0.049**
snp2755039	RNASEL	Exon 2	G	0.99	0.97	0.473	0.474
snp2750857	RNASEL	Exon 3	T	0.87	0.80	0.402	0.681
snp2750736	RNASEL	Exon 3	G	0.86	0.85	1.000	0.723
snp2750733	RNASEL	Exon 3	T	0.00	0.03	0.476	0.476
snp2743998	RNASEL	3' UTR	G	0.87	0.82	0.564	0.754
snp2743993	RNASEL	3' UTR	A	0.80	0.79	1.000	1.000
snp2743789	RNASEL	3' UTR	T	0.84	0.75	0.402	0.681
snp2743745	RNASEL	3' UTR	C	0.41	0.29	0.272	0.078
snp2743078	RNASEL	3' UTR	T	0.67	0.69	1.000	0.892
snp2742898	RNASEL	3' UTR	T	0.17	0.15	1.000	0.554
snp2742846	RNASEL	3' UTR	T	0.64	0.71	0.661	0.924
snp2742808	RNASEL	3' UTR	T	0.65	0.71	0.781	0.789
snp2742764	RNASEL	3' UTR	T	0.33	0.31	1.000	1.000

Equine *RNASEL* polymorphisms genotyped among case and control populations for statistical analysis by Fisher's Exact test. Allele frequencies and allelic and genotypic p-values are shown. Bolded SNPs were significantly associated with WNV phenotype.

Table 5.3 Odds ratio analysis of significantly associated polymorphisms

				Homozygous			Heterozygous		
SNP ID	Gene	Region	Allele	Odds Ratio	95% CI	p-value	Odds Ratio	95% CI	p-value
snp6567078	OAS1	Promoter	C	7.09	1.30:38.76	0.024	8.86	2.38:33.06	0.001
snp6567031	OAS1	Promoter	T	5.32	0.94:29.99	0.058	10.05	2.71:37.26	0.001
snp6567000	OAS1	Promoter	G	0.47	0.04:5.73	0.552	8.87	2.51:31.31	0.001
snp6566994	OAS1	Promoter	T	3.82	0.64:22.74	0.141	11.77	3.21:43.16	0.000
snp6566893	OAS1	Promoter	C	3.50	0.59:20.68	0.167	10.50	2.89:38.11	0.000
snp6566888	OAS1	Promoter	T				5.05	1.48:17.25	0.010
snp6566745	OAS1	Promoter	T	4.45	0.77:25.86	0.096	11.77	3.21:43.16	0.000
snp6566498	OAS1	Promoter	T	5.00	0.92:27.01	0.062	8.25	2.31:29.52	0.001
snp6566231	OAS1	Promoter	T	10.18	1.92:54.02	0.006	8.91	2.40:33.08	0.001
snp2755071	RNASEL	Exon 2	A	6.00	1.15:31.23	0.033	3.00	0.80:11.31	0.105

Odds ratios were determined for all SNPs genotyped in the case and control populations. Shown are the statistically significant ORs including 95% confidence intervals (95% CI) and statistical p-values.

To investigate potential haplotype effects between case and control populations, several analyses were conducted using multiple haplotype analysis software. Equine *OAS1* haplotypes were inferred with Phase v2.1 [225, 226] using (1) SNPs with genotype success rates ≥80% and (2) population samples genotyped at ≥80% of SNPs to minimize the occurrence of unknown genotypes. From the assembled best reconstruction, tagSNPs were identified by *htsubsets* using STATA 9 [254, 255]. Six tagSNPs (snp6549803, snp6549905, snp6550610, snp6564946, snp6566498 and snp6567078) were identified which provide a mean percentage of diversity explained (PDE) of 99.440%. These tagSNPs were used to re-construct haplotypes among all samples and to conduct a haplotype frequency comparison between case and control populations using Phase v2. The tagSNPs reduced the occurrence of minor haplotypes; however, case-control haplotype frequencies were not significantly variable (p=0.11). Alternatively, sixteen SNPs genotyped within the promoter, 5'UTR and exon 1 of *OAS1* were assembled into haplotypes using all case and control samples. Six tagSNPs (snp6566042, snp6566107, snp6566201, snp6566888, snp6567000 and snp6567078) with a total mean PDE of 99.729% were used to re-construct haplotypes and case-control variation analyzed by Phase v2.1. As expected, reducing minor haplotypes through the exclusion of *OAS1* SNPs with lower linkage disequilibrium values to the significant promoter and 5'UTR polymorphisms (data not shown) resulted in a significant case-control haplotype association (p=0.02). A single haplotype (ACGAAT, Haplotype H) accounted for 57.5% and 30.17% of chromosomes genotyped within

control and case populations, respectively (Table 5.4). Fisher's Exact test showed deviations from this haplotype are statistically associated (p=0.004) with susceptibility to severe West Nile disease, with an odds ratio of 3.13 (p=0.003). The significant haplotype association supports the SNP associations identified by Fisher Exact test.

In a similar manner, *RNASEL* haplotypes were assembled from 65 individuals using 19 SNPs genotyped at ≥65% of all samples. Six tagSNPs (snp2742846, snp2755071, snp2755672, snp2756325, snp2756421 and snp2756422) were identified with total mean PDE of 99.134%. Haplotypes were re-constructed from the same 65 samples and were not significantly variable between case and control populations (p=0.55).

Human *OAS1* is interferon inducible from a ISRE proximal to the transcription start site in the minimal promoter (Figure 5.1). Since the SNPs associated with WNV susceptibility were present in the *OAS1* promoter, their potential effect on interferon induction was investigated. *OAS1* promoter-luciferase reporter constructs were transfected into 2fTGH cells and treated with IFNT as previously described [256]. To investigate promoter polymorphism effects on interferon inducibility, the full-length clones lack 5'UTR and exon 1 regions and their tagSNPs (snp6566107 and snp6566134). Three full length promoters were cloned, each representing multiple tagSNP haplotypes.

Full-length clone EcOAS1_A-Luc represents multiple tagSNP haplotypes, including the common haplotype identified above (ACGAAT) and was genotyped at a frequency of 41.4% and 65% among case and control populations, respectively. Clone

Table 5.4 *OAS1* promoter haplotype distribution among case and control populations

	OAS1 Promoter Haplotypes													
	A	B	C	D	E	F	G	H	I	J	K	L	M	Total
Case	1	6	13	32	5	1	6	35	5	1	2	7	2	116
	0.86%	5.17%	11.21%	27.59%	4.31%	0.86%	5.17%	30.17%	4.31%	0.86%	1.72%	6.03%	1.72%	100.00%
Control	0	0	2	5	6	0	0	23	0	0	1	3	0	40
	0.00%	0.00%	5.00%	12.50%	15.00%	0.00%	0.00%	57.50%	0.00%	0.00%	2.50%	7.50%	0.00%	100.00%
Total	1	6	15	37	11	1	6	58	5	1	3	10	2	156
	0.64%	3.85%	9.62%	23.72%	7.05%	0.64%	3.85%	37.18%	3.21%	0.64%	1.92%	6.41%	1.28%	100.00%

EcOAS1_B-Luc was genotyped with frequencies of 17.2% and 5% while full-length clone EcOAS1_C-Luc was identified at frequencies of 29.3% and 15% among case and control populations, respectively. Upon treatment with 10^4 U/mL IFNT, no difference in reporter activity was identified between full length clones (Figure 5.2). The full length *OAS1* promoter clones are induced strongly by IFNT treatment (Figures 5.2). The full-length clones were similarly induced 7-8 fold relative to basal levels (Figure 5.3).
To further localize the interferon regulatory region of the equine *OAS1* promoter, deletion constructs lacking the microsatellite and promoter sequence upstream or downstream were transfected and treated with IFNT. Constructs lacking promoter sequence downstream of the microsatellite, proximal to the transcription start site, were entirely inactive prior to and after IFNT treatment (data not shown). Constructs lacking the polymorphic microsatellite and upstream sequence (EcOAS1_AΔ5'-Luc and EcOAS1_BΔ5'-Luc) were induced to levels comparable to the full length clones (Figures 5.2 and 5.3). This proximal promoter region from the start site of transcription to the microsatellite is necessary and sufficient for strong induction by interferon.

To further characterize potential polymoprhism effects on IFN stimulation of the *OAS1* promoter, full-length constructs were transfected into 2fTGH cells and treated with different levels of IFNT. Reporter activation was similar between the clones and increased with increasing IFNT treatments (Figure 5.4). However, greater differences in activation were observed between treatments of 10^2 and 10^3 U/mL. To further characterize the relationship between treatment dose and reporter activation, cells were

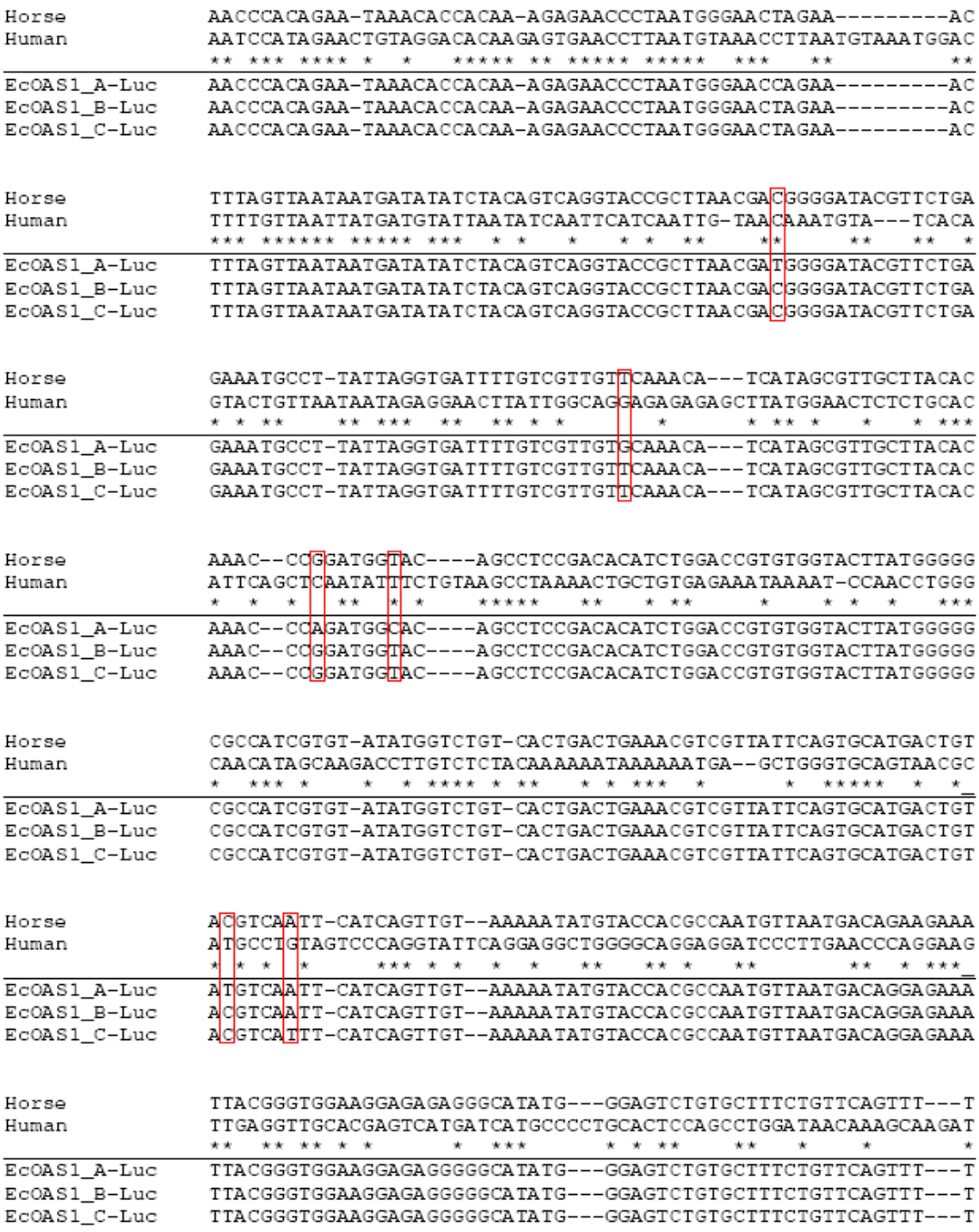


Figure 5.1 Local alignment of human and horse *OAS1* promoters.
ClustalX alignment of equine and human *OAS1* promoters. Also aligned are the sequences of the full-length clones transfected in 2fTGH fibroblast cells. Statistically significant polymorphisms are outlined in red. The known human ISRE is double underlined in the human sequence.

Figure 5.1 continued

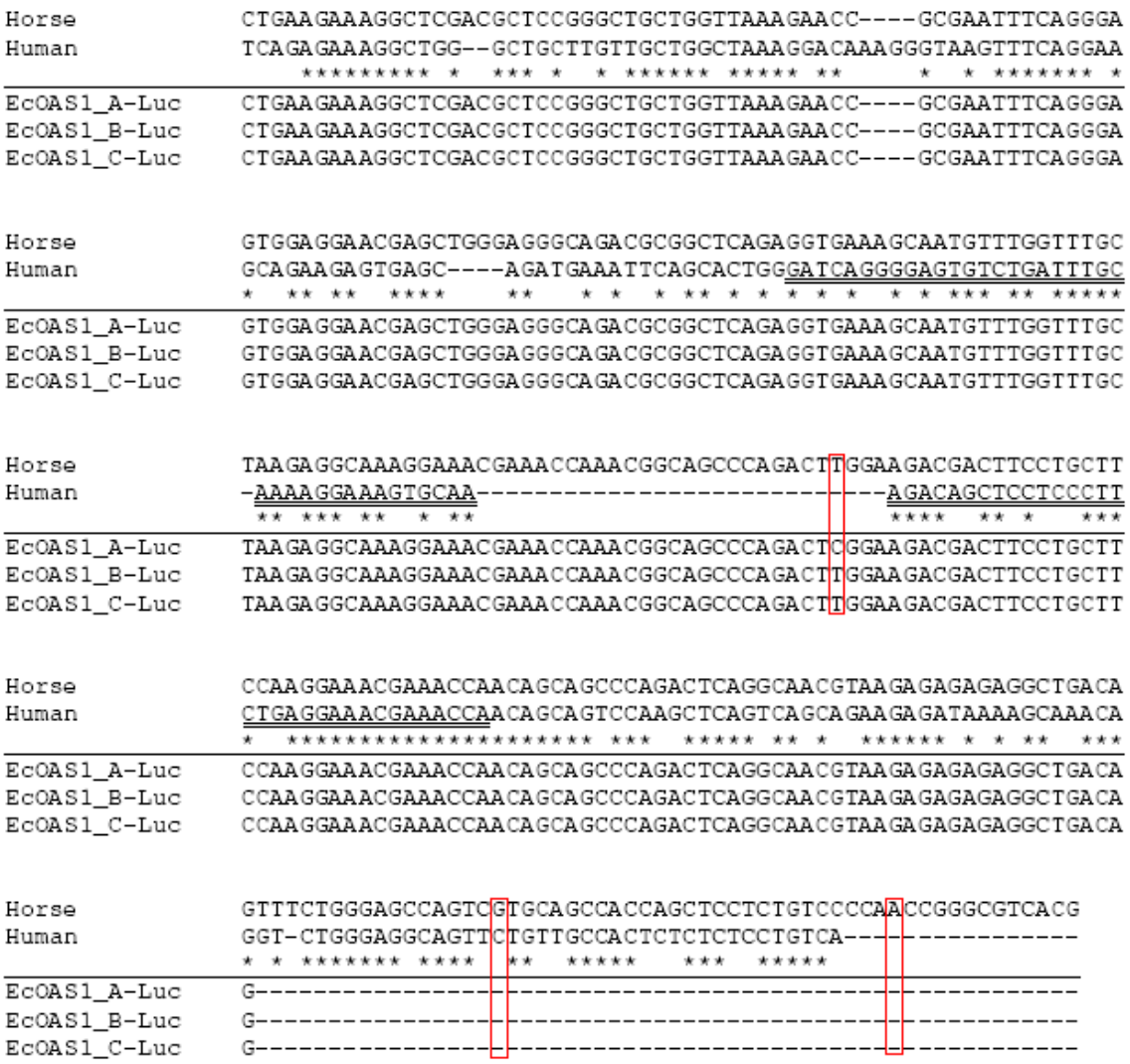


Figure 5.1 continued

treated with low dose levels of IFNT. As observed in previous experiments, levels of activation to low dose treatments of IFNT were similar between clones and increased with IFNT dose. Relative differences in reporter activation were greater between cells treated with 500 U/mL and 250 U/mL (Figure 5.5).

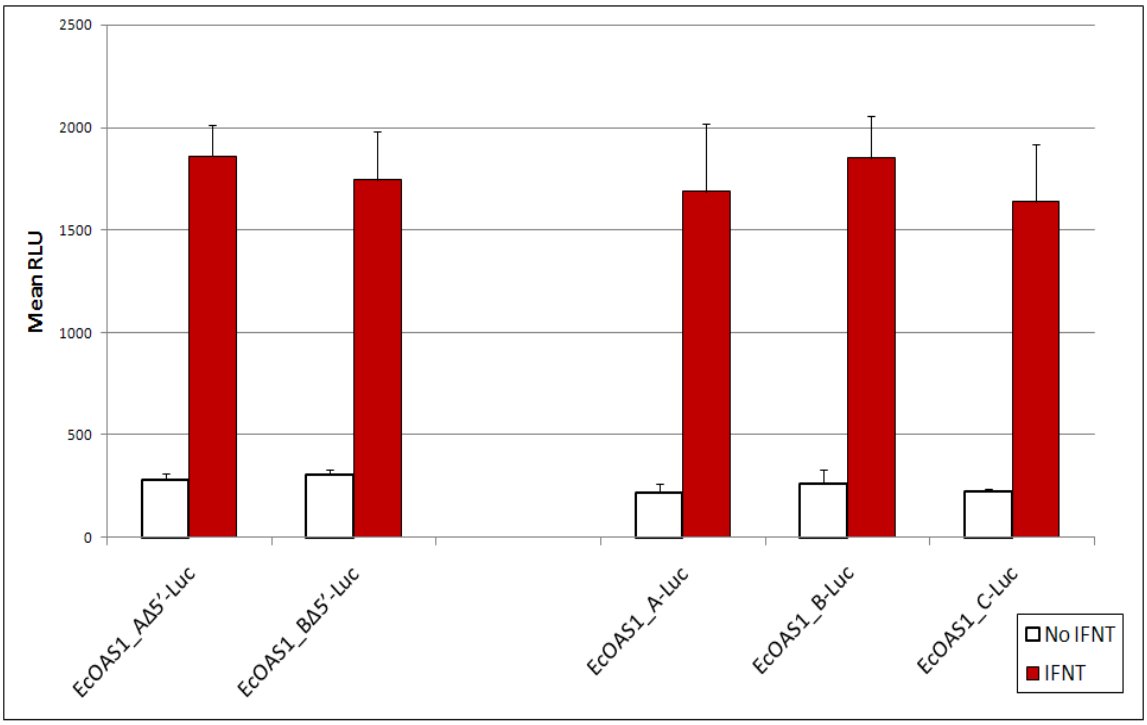


Figure 5.2 Effect of IFNT on *OAS1*-luciferase activity in 2fTGH fibroblast cells.

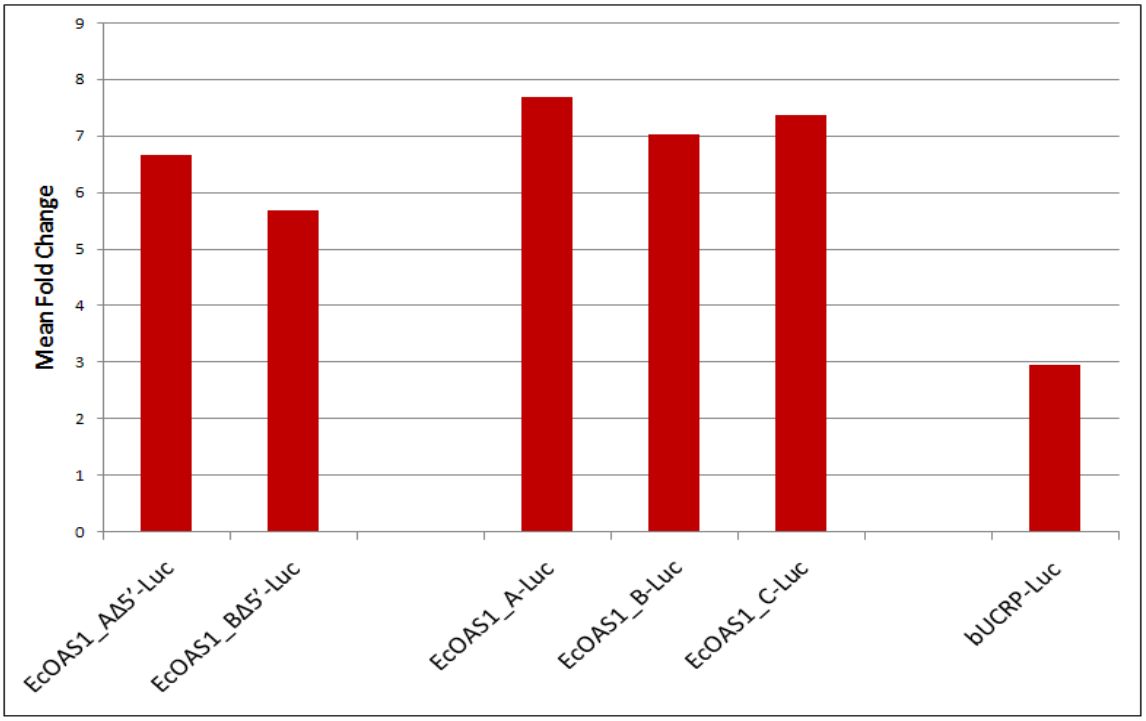


Figure 5.3 Fold IFNT-induced stimulation of *OAS1*-luciferase activity in 2fTGH cells.

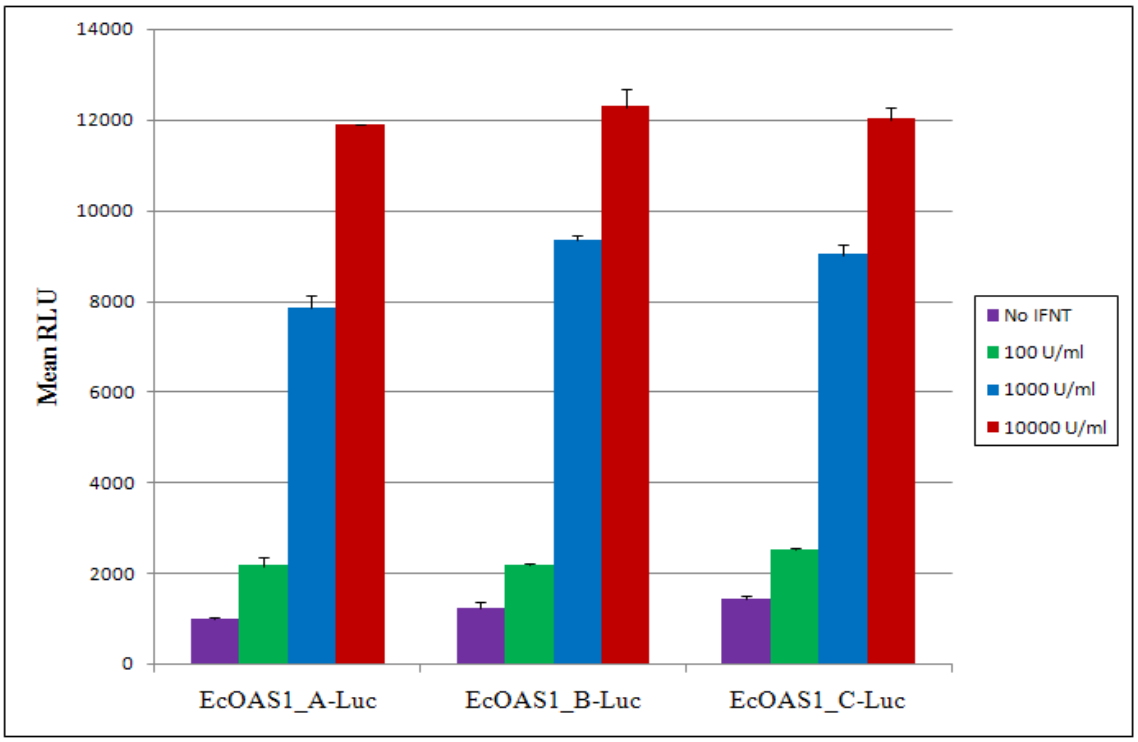


Figure 5.4 Effect of IFNT dose on *OAS1*-luciferase activity in 2fTGH fibroblast cells.

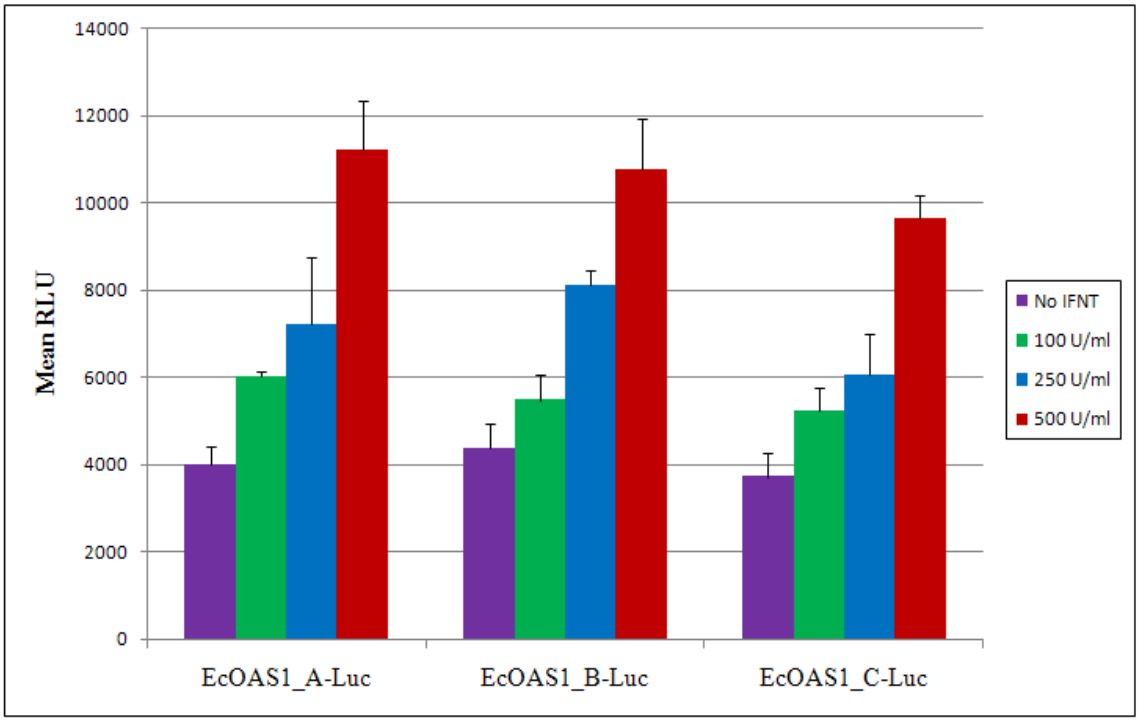


Figure 5.5 Effect of low IFNT dose on OAS1-luciferase activity in 2fTGH cells.

These data suggest that West Nile susceptibility may involve other regulatory factors independent of the interferon-induced expression of equine *OAS1*. However, additional experiments are needed to further characterize the effect of the polymorphisms on IFN induction in other cell types and in response to WNV infection.

This investigation reports a strong association of *OAS1* and *RNASEL* single nucleotide polymorphisms with severe West Nile encephalitis. With comparable immune gene structure to humans, the horse may provide a model for which to study SNP-associated disease susceptibility. The strongest association occurred among SNPs of the *OAS1* promoter and 5' gene sequence; however, while we show the proximal region of the *OAS1* promoter is induced strongly after IFN treatment, preliminary *in vitro* analysis suggests the functional association of these polymorphisms with severe West Nile disease is independent of the gene's response to interferon.

Methods

DNA extraction and SNP genotyping of equine samples

White blood cells were isolated from whole blood and DNA extracted (see below). Control DNA samples were genotyped at each SNP as previously described [252]. Case samples consisted of frozen or archived formalin-fixed paraffin-embedded (FFPE) liver, kidney or nervous (spinal cord or brain) tissues. DNA was extracted from frozen tissue samples after Proteinase K (Promega, Madison, Wisconsin) digestion, washed twice with phenol/chloroform and ethanol precipitated. FFPE liver and kidney samples were deparrafinized with xylene and DNA extracted using the RecoverAll Nucleic Acid Extraction Kit (Ambion, Austin, Texas). Additional FFPE brain and spinal

cord samples were deparaffinized with xylene and DNA extracted in a manner similar to frozen samples after treatment with 6 mg Proteinase K for 3 days at 55°C. All FFPE DNA samples were amplified using the Whole Genome Amplification Kit (Sigma, St. Louis, Missouri) using ~100ng input DNA without further digestion and amplified for 25 cycles. Amplification products were processed using either the GeneElute Purification System (Sigma, St. Louis, Missouri) or the Qiaquick PCR Purification Kit (Qiagen, Valencia, California). Amplification products from FFPE DNA resulted in fragmented template < 500 bp in length (data not shown). FFPE samples were genotyped by sequencing short PCR products <200bp. PCR primer sequences are available upon request.

Transfection experiment

Genotyped samples were amplified with Easy-A high fidelity taq (Stratagene, La Jolla, California) and TA-cloned into pCRII (Invitrogen, Carlsbad, California). Full-length promoters were amplified using PCR primers F:CGACGGCCAGCTCGAGAACCCACAGAATAAACACCACA and R:CAGCTATGACAAGCTTCTGTCAGCCTCTCTCTCTTACG. PCR primers F:CGACGGCCAGCTCGAGAACCCACAGAATAAACACCACA and R:CAGCTATGACAAGCTTAGCTCTCTCTTCTGTTTTATA were used to amplify the 3' deletion promoters. Primers F: CGACGGCCAGCTCGAGCTTAACCTAGAAACGCGTCTGA and R: CAGCTATGACAAGCTTCTGTCAGCCTCTCTCTCTTACG were used to amplify the 5' deletion constructs. Individual clones were cultured and verified by sequencing.

Each primer pair contains XhoI and HindIII sites used to directionally clone the promoter regions into pGL3-Basic (Promega, Madison, Wisconsin). Final constructs were verified by sequencing (Figure 5.1).

The 2fTGH immortalized cells [257] were maintained in DMEM-F12 medium (Sigma-Aldrich Corp., St. Louis, MO) supplemented with penicillin/ streptomycin/ amphotericin B (Invitrogen, Carlesbad,CA) and 5% FBS (Hyclone, Logan,UT). Cells were seeded into 12-well plates, allowed to grow until monolayers were 67-75% confluent and transiently transfected as described previously [256]. Briefly, luciferase constructs (500 ng/well) were co-transfected with an equivalent amount of pEF1-Myc-His LacZ (500 ng/well; Invitrogen) and GenePorter Transfection Reagent (Gene Therapy Systems, San Diegeo, CA) according to the manufacturer's instructions. Transfected cells were grown overnight (14-16 h) in medium containing 10% FBS before treatment. Recombinant ovine interferon tau (IFN τ , 10^8 antiviral units/ml) was produced and assayed as described previously [258]. Transfected cells were treated with 10^4 AVU IFN τ/ml or left untreated in serum-free medium for 24 h. Cells were lysed in Cell Culture Lysis Reagent (Promega, Madison, WI), and luciferase activity was assayed according to the manufacturer's instructions (Promega). Each construct-treatment combination was tested in four wells/combination, and transfection assays were repeated a minimum of three times. Lower IFN concentrations (10^2-10^4 AVU/ml) were tested in two experiments.

Statistical analysis

Statistical association analyses were conducted using STATA v9 software. Allelic association was conducted using Fisher's Exact tests on 2x2 tables. Fisher's Exact tests were conducted on 2x3 tables to identify genotypic association. Genotypes were coded such that alleles with greater case population frequencies were coded alike. Allelic odds ratios were determined for each SNP as well as for the heterozygous and homozygous genotypes of the associated alleles. Significance is reported with $\alpha=0.05$. Haplotype associations were computed using a 2x2 design by comparing single haplotypes to all others. Case-control haplotype analysis was also conducted with Phase v2.

CHAPTER VI

SUMMARY AND CONCLUSION

Prior to the release of the equine genome sequence and the advancements made as a result, this work successfully identified polymorphisms implicating the involvement of the *OAS/RNASEL* innate immune system in the mammalian response to WNV infection. Using data provided from years of study with susceptible and resistant strains of mice, which identified a role of the murine *Oas1b* gene in resistance to severe WNV infection, this research conducted a two-stage association study to identify a similar role of the *OAS/RNASEL* system in the horse. In doing so, multiple polymorphisms were utilized within a case-control study to identify significantly associated mutations with susceptibility to equine West Nile encephalitis.

Genomic sequence was assembled by shotgun sequencing CHORI BAC 100:I10 (*OAS1*) or from extending full-length transcript sequence (*RNASEL*) and polymorphisms were identified. Screening for polymorphisms within these genes from a random population of horses successfully identified a high frequency of SNPs in both genes. These polymorphisms were then genotyped in case and control samples collected across the United States.

The ability to detect associations in a case-control study design requires a high degree of comparability between case and control samples. Through collaborations cultivated across the United States, this research succeeded in assembling a highly comparable collection of equine samples for case-control study of WNV susceptibility.

Control animals were maintained under controlled conditions where naïve horses were infected with WNV through natural mosquito transmission and their response closely monitored. These horses failed to respond clinically to viral infection. Accordingly, these control horses are considered to be naturally resistant. Samples included in the case population were greater in number and collected during the initial United States epidemic. All case horses tested positive, as did control horses, for the presence of West Nile viral infection. Case horses suffered debilitating symptoms, albeit to varying degrees, with the most apparent being incoordination and ataxia of the forelimb and/or hindlimb. Other symptoms suffered by clinical case horses included paralysis, seizures, fever, recumbancy and eventual death. These case horses, like the control population, were infected with WNV through natural mosquito transmission and all horses in both populations were unvaccinated at the time of infection and/or death.

A total of 66 SNPs were genotyped in the case and control samples. Statistical analyses identified highly significant associations between West Nile susceptibility and polymorphisms in both genes. A majority of the *OAS1* polymorphisms associated with susceptibility were identified within the promoter and 5' UTR region, while only a single polymorphism associated significantly in *RNASEL*. Haplotype analyses of the *OAS1* promoter polymorphisms identified a number of haplotypes whose frequencies were significantly variable between case and control populations. A single haplotype was found at greater frequency in both populations. While this haplotype was not significantly identified as conveying a resistant phenotype to West Nile infection, differences from this common haplotype were significantly associated with WNV

susceptibility. The direct consequences of these deviations or of the common haplotype on *OAS1* expression are still to be determined; however, *OAS1* is highly inducible by interferon. The most direct hypothesis is that these polymorphisms affect the interferon inducibility of the gene and thus its ability to provide adequate immediate earlyhost resistance.

To further investigate the effect of the single nucleotide and microsatellite polymorphisms on *OAS1* expression, luciferase clones were constructed to measure *in vitro* IFN inducibility of multiple promoter haplotypes. Full-length and deletion promoter-reporter constructs were transfected into 2fTGH cells and treated with IFNT to induce promoter activity. The equine *OAS1* ISRE was localized to within 518 bp of the transcription start site, which is in agreement with many ISRE-stimulated genes throughout the human genome, including human *OAS1*. This proximal region was identified as necessary and sufficient for the activation of equine *OAS1* transcription in response to interferon.

Further investigation will identify potential roles of the polymorphic microsatellite on the immediate early response of the *OAS1* gene to IFN.

Although multiple polymorphisms were statistically associated with WNV susceptibility, this work has only begun to identify the functional consequences of these polymorphisms on *OAS1* and *RNASEL* that may contribute to the susceptible phenotype. The contribution of the *OAS1/RNASEL* immune system in equine resistance to WNV infection is still unknown; however, the strong statistical associations and evidence presented herein are suggestive.

The strict pathogenesis of West Nile infection introduces the potential for a cellular context by which the immune response acts to provide host resistance. Initial infection progresses to particular tissues within the animal and may require a specific cellular environment for which the *OAS/RNASEL* pathway responds to infection and limits viral replication. To understand further the functional involvement of *OAS1* and the consequences of the detected polymorphisms, additional work will need to focus on characterizing the *in vivo* environment of the promoter and its specific response to cellular signals. An absolute role for the *OAS/RNASEL* system in resistance to WNV infection is unlikely, although the odds ratios are quite definitive. With advancements in technology and the release of the equine genome sequence, further detection of genome-wide immune response loci involved in host resistance and susceptibility will provide a greater understanding of the host innate immune response in horses and other mammals.

REFERENCES

1. Grandvaux N, tenOever BR, Servant MJ, Hiscott J: **The interferon antiviral response: from viral invasion to evasion**. *Current Opinion in Infectious Diseases* 2002, **15(3):**259-267.

2. Samuel CE: **Host genetic variability and West Nile virus susceptibility**. *Proc Natl Acad Sci U S A* 2002, **99(18)**:11555-11557.

3. McLean RG, Ubico SR, Bourne D, Komar N: **West Nile virus in livestock and wildlife**. *Curr Top Microbiol Immunol* 2002, **267**:271-308.

4. Barnard BJ, Voges SF: **Flaviviruses in South Africa: pathogenicity for sheep**. *The Onderstepoort Journal of Veterinary Research* 1986, **53(4)**:235-238.

5. McIntosh B: **Arboviral Zoonoses in Africa, West Nile Fever.** In *Zoonoses Section B: Viral Zoonoses* Edited by: Steele K, vol. 1. Boca Raton, Florida: CRC Press; 1982.

6. Ilkal MA, Prasanna Y, Jacob PG, Geevarghese G, Banerjee K: **Experimental studies on the susceptibility of domestic pigs to West Nile virus followed by Japanese encephalitis virus infection and vice versa**. *Acta Virol* 1994, **38(3)**:157-161.

7. Blackburn NK, Reyers F, Berry WL, Shepherd AJ: **Susceptibility of dogs to West Nile virus: a survey and pathogenicity trial**. *J Comp Pathol* 1989, **100(1)**:59-66.

8. Gil J, Rullas J, Alcami J, Esteban M: **MC159L protein from the poxvirus molluscum contagiosum virus inhibits NF-kappaB activation and apoptosis induced by PKR**. *The Journal of General Virology* 2001, **82(Pt 12)**:3027-3034.

9. Vallee I, Tait SW, Powell PP: **African swine fever virus infection of porcine aortic endothelial cells leads to inhibition of inflammatory responses, activation of the thrombotic state, and apoptosis**. *J Virol* 2001, **75(21)**:10372-10382.

10. Talon J, Horvath CM, Polley R, Basler CF, Muster T, Palese P, Garcia-Sastre A: **Activation of interferon regulatory factor 3 is inhibited by the influenza A virus NS1 protein**. *J Virol* 2000, **74(17)**:7989-7996.

11. Wang X, Li M, Zheng H, Muster T, Palese P, Beg AA, Garcia-Sastre A: **Influenza A virus NS1 protein prevents activation of NF-kappaB and**

induction of alpha/beta interferon. *Journal of Virology* 2000, **74(24)**:11566-11573.

12. Yuan W, Krug RM: **Influenza B virus NS1 protein inhibits conjugation of the interferon (IFN)-induced ubiquitin-like ISG15 protein**. *Embo J* 2001, **20(3)**:362-371.

13. Smith EJ, Marie I, Prakash A, Garcia-Sastre A, Levy DE: **IRF3 and IRF7 phosphorylation in virus-infected cells does not require double-stranded RNA-dependent protein kinase R or Ikappa B kinase but is blocked by Vaccinia virus E3L protein**. *J Biol Chem* 2001, **276(12)**:8951-8957.

14. Poppers J, Mulvey M, Khoo D, Mohr I: **Inhibition of PKR activation by the proline-rich RNA binding domain of the herpes simplex virus type 1 Us11 protein**. *J Virol* 2000, **74(23)**:11215-11221.

15. Juang YT, Lowther W, Kellum M, Au WC, Lin R, Hiscott J, Pitha PM: **Primary activation of interferon A and interferon B gene transcription by interferon regulatory factor 3**. *Proc Natl Acad Sci U S A* 1998, **95(17)**:9837-9842.

16. Joseph TD, Look DC: **Specific inhibition of interferon signal transduction pathways by adenoviral infection**. *J Biol Chem* 2001, **276(50)**:47136-47142.

17. Li M, Damania B, Alvarez X, Ogryzko V, Ozato K, Jung JU: **Inhibition of p300 histone acetyltransferase by viral interferon regulatory factor**. *Mol Cell Biol* 2000, **20(21)**:8254-8263.

18. Lin R, Genin P, Mamane Y, Sgarbanti M, Battistini A, Harrington WJ, Jr., Barber GN, Hiscott J: **HHV-8 encoded vIRF-1 represses the interferon antiviral response by blocking IRF-3 recruitment of the CBP/p300 coactivators**. *Oncogene* 2001, **20(7)**:800-811.

19. Lubyova B, Pitha PM: **Characterization of a novel human herpesvirus 8-encoded protein, vIRF-3, that shows homology to viral and cellular interferon regulatory factors**. *J Virol* 2000, **74(17)**:8194-8201.

20. Terstegen L, Gatsios P, Ludwig S, Pleschka S, Jahnen-Dechent W, Heinrich PC, Graeve L: **The vesicular stomatitis virus matrix protein inhibits glycoprotein 130-dependent STAT activation**. *J Immunol* 2001, **167(9)**:5209-5216.

21. Murgue B, Murri S, Zientara S, Durand B, Durand JP, Zeller H: **West Nile outbreak in horses in southern France, 2000: the return after 35 years**. *Emerging Infectious Diseases* 2001, **7(4)**:692-696.

22. Ostlund EN, Crom RL, Pedersen DD, Johnson DJ, Williams WO, Schmitt BJ: **Equine West Nile encephalitis, United States**. *Emerging Infectious Diseases* 2001, **7(4)**:665-669.

23. Green M (ed.): Genetic Variants and Strains of Laboratory Mice. Stuttgart: Gustav Fischer; 1989.

24. Perelygin AA, Scherbik SV, Zhulin IB, Stockman BM, Li Y, Brinton MA: **Positional cloning of the murine flavivirus resistance gene**. *Proc Natl Acad Sci U S A* 2002, **99(14)**:9322-9327.

25. Brinton MA, Perelygin AA: **Genetic resistance to flaviviruses**. *Advances in Virus Research* 2003, **60**:43-85.

26. Eskildsen S, Hartmann R, Kjeldgaard NO, Justesen J: **Gene structure of the murine 2'-5'-oligoadenylate synthetase family**. *Cell Mol Life Sci* 2002, **59(7)**:1212-1222.

27. Ichii Y, Fukunaga R, Shiojiri S, Sokawa Y: **Mouse 2-5A synthetase cDNA: nucleotide sequence and comparison to human 2-5A synthetase**. *Nucleic Acids Res* 1986, **14(24)**:10117.

28. Kakuta S, Shibata S, Iwakura Y: **Genomic structure of the mouse 2',5'-oligoadenylate synthetase gene family**. *J Interferon Cytokine Res* 2002, **22(9)**:981-993.

29. Rutherford MN, Kumar A, Nissim A, Chebath J, Williams BR: **The murine 2-5A synthetase locus: three distinct transcripts from two linked genes**. *Nucleic Acids Res* 1991, **19(8)**:1917-1924.

30. Shibata S, Kakuta S, Hamada K, Sokawa Y, Iwakura Y: **Cloning of a novel 2',5'-oligoadenylate synthetase-like molecule, Oasl5 in mice**. *Gene* 2001, **271(2)**:261-271.

31. Smith JB, Nguyen TT, Hughes HJ, Herschman HR, Widney DP, Bui KC, Rovai LE: **Glucocorticoid-attenuated response genes induced in the lung during endotoxemia**. *American Journal of Physiology* 2002, **283(3)**:L636-647.

32. Tiefenthaler M, Marksteiner R, Neyer S, Koch F, Hofer S, Schuler G, Nussenzweig M, Schneider R, Heufler C: **M1204, a novel 2',5' oligoadenylate synthetase with a ubiquitin-like extension, is induced during maturation of murine dendritic cells**. *J Immunol* 1999, **163(2)**:760-765.

33. Perelygin AA, Zharkikh AA, Scherbik SV, Brinton MA: **The mammalian 2'-5' oligoadenylate synthetase gene family: evidence for concerted evolution of paralogous Oas1 genes in Rodentia and Artiodactyla**. *Journal of Molecular Evolution* 2006, **63(4)**:562-576.

34. Hovnanian A, Rebouillat D, Levy ER, Mattei MG, Hovanessian AG: **The human 2',5'-oligoadenylate synthetase-like gene (OASL) encoding the interferon-induced 56-kDa protein maps to chromosome 12q24.2 in the proximity of the 2',5'-OAS locus**. *Genomics* 1999, **56(3)**:362-363.

35. Hovnanian A, Rebouillat D, Mattei MG, Levy ER, Marie I, Monaco AP, Hovanessian AG: **The human 2',5'-oligoadenylate synthetase locus is composed of three distinct genes clustered on chromosome 12q24.2 encoding the 100-, 69-, and 40-kDa forms**. *Genomics* 1998, **52(3)**:267-277.

36. Chebath J, Benech P, Hovanessian A, Galabru J, Revel M: **Four different forms of interferon-induced 2',5'-oligo(A) synthetase identified by immunoblotting in human cells**. *J Biol Chem* 1987, **262(8)**:3852-3857.

37. Hovanessian AG, Laurent AG, Chebath J, Galabru J, Robert N, Svab J: **Identification of 69-kd and 100-kd forms of 2-5A synthetase in interferon-treated human cells by specific monoclonal antibodies**. *Embo J* 1987, **6(5)**:1273-1280.

38. Merlin G, Chebath J, Benech P, Metz R, Revel M: **Molecular cloning and sequence of partial cDNA for interferon-induced (2'-5')oligo(A) synthetase mRNA from human cells**. *Proc Natl Acad Sci U S A* 1983, **80(16)**:4904-4908.

39. Benech P, Merlin G, Revel M, Chebath J: **3' end structure of the human (2'-5') oligo A synthetase gene: prediction of two distinct proteins with cell type-specific expression**. *Nucleic Acids Research* 1985, **13(4)**:1267-1281.

40. Justesen J, Hartmann R, Kjeldgaard NO: **Gene structure and function of the 2'-5'-oligoadenylate synthetase family**. *Cell Mol Life Sci* 2000, **57(11)**:1593-1612.

41. Marie I, Hovanessian AG: **The 69-kDa 2-5A synthetase is composed of two homologous and adjacent functional domains**. *The Journal of Biological Chemistry* 1992, **267(14)**:9933-9939.

42. Bonnevie-Nielsen V, Field LL, Lu S, Zheng DJ, Li M, Martensen PM, Nielsen TB, Beck-Nielsen H, Lau YL, Pociot F: **Variation in antiviral 2',5'-oligoadenylate synthetase (2'5'AS) enzyme activity is controlled by a single-nucleotide polymorphism at a splice-acceptor site in the OAS1 gene**. *Am J Hum Genet* 2005, **76(4)**:623-633.

43. Yakub I, Lillibridge KM, Moran A, Gonzalez OY, Belmont J, Gibbs RA, Tweardy DJ: **Single nucleotide polymorphisms in genes for 2'-5'-oligoadenylate synthetase and RNase L inpatients hospitalized with West Nile virus infection**. *J Infect Dis* 2005, **192(10)**:1741-1748.

44. Squire J, Meurs EF, Chong KL, McMillan NA, Hovanessian AG, Williams BR: **Localization of the human interferon-induced, ds-RNA activated p68 kinase gene (PRKR) to chromosome 2p21-p22**. *Genomics* 1993, **16(3)**:768-770.

45. Zhou A, Hassel BA, Silverman RH: **Expression cloning of 2-5A-dependent RNAase: a uniquely regulated mediator of interferon action**. *Cell* 1993, **72(5)**:753-765.

46. Dong B, Silverman RH: **A bipartite model of 2-5A-dependent RNase L**. *J Biol Chem* 1997, **272(35)**:22236-22242.

47. Cole JL, Carroll SS, Kuo LC: **Stoichiometry of 2',5'-oligoadenylate-induced dimerization of ribonuclease L. A sedimentation equilibrium study**. *J Biol Chem* 1996, **271(8)**:3979-3981.

48. Tanaka N, Nakanishi M, Kusakabe Y, Goto Y, Kitade Y, Nakamura KT: **Structural basis for recognition of 2',5'-linked oligoadenylates by human ribonuclease L**. *Embo J* 2004, **23(20)**:3929-3938.

49. Dong B, Silverman RH: **2-5A-dependent RNase molecules dimerize during activation by 2-5A**. *J Biol Chem* 1995, **270(8)**:4133-4137.

50. Nakanishi M, Goto Y, Kitade Y: **2-5A induces a conformational change in the ankyrin-repeat domain of RNase L**. *Proteins* 2005, **60(1)**:131-138.

51. Dong B, Niwa M, Walter P, Silverman RH: **Basis for regulated RNA cleavage by functional analysis of RNase L and Ire1p**. *Rna* 2001, **7(3)**:361-373.

52. Nakanishi M, Yoshimura A, Ishida N, Ueno Y, Kitade Y: **Contribution of Tyr712 and Phe716 to the activity of human RNase L**. *Eur J Biochem* 2004, **271(13)**:2737-2744.

53. Baglioni C, Minks MA, Maroney PA: **Interferon action may be mediated by activation of a nuclease by pppA2'p5'A2'p5'A**. *Nature* 1978, **273(5664)**:684-687.

54. Clemens MJ, Williams BR: **Inhibition of cell-free protein synthesis by pppA2'p5'A2'p5'A: a novel oligonucleotide synthesized by interferon-treated L cell extracts**. *Cell* 1978, **13(3)**:565-572.

55. Kerr IM, Brown RE: **pppA2'p5'A2'p5'A: an inhibitor of protein synthesis synthesized with an enzyme fraction from interferon-treated cells**. *Proc Natl Acad Sci U S A* 1978, **75(1)**:256-260.

56. Hovanessian AG, Brown RE, Kerr IM: **Synthesis of low molecular weight inhibitor of protein synthesis with enzyme from interferon-treated cells**. *Nature* 1977, **268(5620)**:537-540.

57. Kerr IM, Brown RE, Hovanessian AG: **Nature of inhibitor of cell-free protein synthesis formed in response to interferon and double-stranded RNA**. *Nature* 1977, **268(5620)**:540-542.

58. Roberts WK, Hovanessian A, Brown RE, Clemens MJ, Kerr IM: **Interferon-mediated protein kinase and low-molecular-weight inhibitor of protein synthesis**. *Nature* 1976, **264(5585)**:477-480.

59. Castelli JC, Hassel BA, Maran A, Paranjape J, Hewitt JA, Li XL, Hsu YT, Silverman RH, Youle RJ: **The role of 2'-5' oligoadenylate-activated ribonuclease L in apoptosis**. *Cell Death Differ* 1998, **5(4)**:313-320.

60. Diaz-Guerra M, Rivas C, Esteban M: **Activation of the IFN-inducible enzyme RNase L causes apoptosis of animal cells**. *Virology* 1997, **236(2)**:354-363.

61. Mullan PB, Hosey AM, Buckley NE, Quinn JE, Kennedy RD, Johnston PG, Harkin DP: **The 2,5 oligoadenylate synthetase/RNaseL pathway is a novel effector of BRCA1- and interferon-gamma-mediated apoptosis**. *Oncogene* 2005, **24(35)**:5492-5501.

62. Zhou A, Paranjape J, Brown TL, Nie H, Naik S, Dong B, Chang A, Trapp B, Fairchild R, Colmenares C *et al*: **Interferon action and apoptosis are defective in mice devoid of 2',5'-oligoadenylate-dependent RNase L**. *Embo J* 1997, **16(21)**:6355-6363.

63. Domingo-Gil E, Esteban M: **Role of mitochondria in apoptosis induced by the 2-5A system and mechanisms involved**. *Apoptosis* 2006, **11(5)**:725-738.

64. Perelygin AA, Lear TL, Zharkikh AA, Brinton MA: **Structure of equine 2'-5'oligoadenylate synthetase (OAS) gene family and FISH mapping of OAS genes to ECA8p15-->p14 and BTA17q24-->q25**. *Cytogenet Genome Res* 2005, **111(1)**:51-56.

65. Ewing B, Green P: **Base-calling of automated sequencer traces using phred. II. Error probabilities**. *Genome Res* 1998, **8(3)**:186-194.

66. Ewing B, Hillier L, Wendl MC, Green P: **Base-calling of automated sequencer traces using phred. I. Accuracy assessment**. *Genome Res* 1998, **8(3)**:175-185.

67. Gordon D: **Viewing and Editing Assembled Sequences Using Consed**. In *Current Protocols in Bioinformatics.* Edited by Baxevanis AD, Davison DB. New York: John Wiley & Co.; 2004: 11.12.11-11.12.43.

68. Gordon D, Abajian C, Green P: **Consed: a graphical tool for sequence finishing**. *Genome Res* 1998, **8(3)**:195-202.

69. **NCBI Trace Archive** [http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?] 2007

70. Borrmann L, Seebeck B, Rogalla P, Bullerdiek J: **Human HMGA2 promoter is coregulated by a polymorphic dinucleotide (TC)-repeat**. *Oncogene* 2003, **22(5)**:756-760.

71. Gebhardt F, Zanker KS, Brandt B: **Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1**. *J Biol Chem* 1999, **274(19)**:13176-13180.

72. Hamada H, Seidman M, Howard BH, Gorman CM: **Enhanced gene expression by the poly(dT-dG).poly(dC-dA) sequence**. *Mol Cell Biol* 1984, **4(12)**:2622-2630.

73. Huang TS, Lee CC, Chang AC, Lin S, Chao CC, Jou YS, Chu YW, Wu CW, Whang-Peng J: **Shortening of microsatellite deoxy(CA) repeats involved in GL331-induced down-regulation of matrix metalloproteinase-9 gene expression**. *Biochemical and Biophysical Research Communications* 2003, **300(4)**:901-907.

74. Naylor LH, Clark EM: **d(TG)n.d(CA)n sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription**. *Nucleic Acids Res* 1990, **18(6)**:1595-1601.

75. Wang B, Ren J, Ooi LL, Chong SS, Lee CG: **Dinucleotide repeats negatively modulate the promoter activity of Cyr61 and is unstable in hepatocellular carcinoma patients**. *Oncogene* 2005, **24(24)**:3999-4008.

76. Rich A: **DNA comes in many forms**. *Gene* 1993, **135**(1-2):99-109.

77. Rich A, Nordheim A, Wang AH: **The chemistry and biology of left-handed Z-DNA**. *Annual Review of Biochemistry* 1984, **53**:791-846.

78. Schroth GP, Chou PJ, Ho PS: **Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes**. *J Biol Chem* 1992, **267(17)**:11846-11855.

79. Wells RD: **Unusual DNA structures**. *J Biol Chem* 1988, **263(3)**:1095-1098.

80. Rothenburg S, Koch-Nolte F, Rich A, Haag F: **A polymorphic dinucleotide repeat in the rat nucleolin gene forms Z-DNA and inhibits promoter activity**. *Proc Natl Acad Sci U S A* 2001, **98(16)**:8985-8990.

81. Bowie A, O'Neill LA: **The interleukin-1 receptor/Toll-like receptor superfamily: signal generators for pro-inflammatory interleukins and microbial products**. *Journal of Leukocyte Biology* 2000, **67(4)**:508-514.

82. Alexopoulou L, Holt AC, Medzhitov R, Flavell RA: **Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3**. *Nature* 2001, **413(6857)**:732-738.

83. Akira S, Uematsu S, Takeuchi O: **Pathogen recognition and innate immunity**. *Cell* 2006, **124(4)**:783-801.

84. Kawai T, Akira S: **Innate immune recognition of viral infection**. *Nature Immunology* 2006, **7(2)**:131-137.

85. Kawai T, Akira S: **TLR signaling**. *Cell Death and Differentiation* 2006, **13(5)**:816-825.

86. Meylan E, Burns K, Hofmann K, Blancheteau V, Martinon F, Kelliher M, Tschopp J: **RIP1 is an essential mediator of Toll-like receptor 3-induced NF-kappa B activation**. *Nature Immunology* 2004, **5(5)**:503-507.

87. Wathelet MG, Lin CH, Parekh BS, Ronco LV, Howley PM, Maniatis T: **Virus infection induces the assembly of coordinately activated transcription factors on the IFN-beta enhancer in vivo**. *Molecular Cell* 1998, **1(4)**:507-518.

88. Wang T, Town T, Alexopoulou L, Anderson JF, Fikrig E, Flavell RA: **Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis**. *Nature Medicine* 2004, **10(12)**:1366-1373.

89. Lopez CB, Moltedo B, Alexopoulou L, Bonifaz L, Flavell RA, Moran TM: **TLR-independent induction of dendritic cell maturation and adaptive immunity by negative-strand RNA viruses**. *J Immunol* 2004, **173(11)**:6882-6889.

90. Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, Taira K, Akira S, Fujita T: **The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses**. *Nature Immunology* 2004, **5(7)**:730-737.

91. Kawai T, Takahashi K, Sato S, Coban C, Kumar H, Kato H, Ishii KJ, Takeuchi O, Akira S: **IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction**. *Nature Immunology* 2005, **6(10)**:981-988.

92. Meylan E, Curran J, Hofmann K, Moradpour D, Binder M, Bartenschlager R, Tschopp J: **Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus**. *Nature* 2005, **437(7062)**:1167-1172.

93. Seth RB, Sun L, Ea CK, Chen ZJ: **Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3**. *Cell* 2005, **122(5)**:669-682.

94. Xu LG, Wang YY, Han KJ, Li LY, Zhai Z, Shu HB: **VISA is an adapter protein required for virus-triggered IFN-beta signaling**. *Molecular Cell* 2005, **19(6)**:727-740.

95. Kato H, Sato S, Yoneyama M, Yamamoto M, Uematsu S, Matsui K, Tsujimura T, Takeda K, Fujita T, Takeuchi O *et al*: **Cell type-specific involvement of RIG-I in antiviral response**. *Immunity* 2005, **23(1)**:19-28.

96. Gantier MP, Williams BR: **The response of mammalian cells to double-stranded RNA**. *Cytokine Growth Factor Rev* 2007.

97. Malathi K, Dong B, Gale M, Jr., Silverman RH: **Small self-RNA generated by RNase L amplifies antiviral innate immunity**. *Nature* 2007, **448(7155)**:816-819.

98. Manche L, Green SR, Schmedt C, Mathews MB: **Interactions between double-stranded RNA regulators and the protein kinase DAI**. *Mol Cell Biol* 1992, **12(11)**:5238-5248.

99. Minks MA, West DK, Benvin S, Baglioni C: **Structural requirements of double-stranded RNA for the activation of 2',5'-oligo(A) polymerase and**

protein kinase of interferon-treated HeLa cells. *The Journal of Biological Chemistry* 1979, **254(20)**:10180-10183.

100. Adolf GR: **Human interferon omega--a review**. *Multiple sclerosis (Houndmills, Basingstoke, England)* 1995, **1 Suppl 1**:S44-47.

101. Conklin DC, Grant FJ, Rixon MW, Kindsvogel W: **Interferon-ε**. In.; 2002.

102. Foster GR, Finter NB: **Are all type I human interferons equivalent?** *Journal of Viral Hepatitis* 1998, **5(3)**:143-152.

103. Kawamoto S, Oritani K, Asada H, Takahashi I, Ishikawa J, Yoshida H, Yamada M, Ishida N, Ujiie H, Masaie H *et al*: **Antiviral activity of limitin against encephalomyocarditis virus, herpes simplex virus, and mouse hepatitis virus: diverse requirements by limitin and alpha interferon for interferon regulatory factor 1**. *Journal of Virology* 2003, **77(17)**:9622-9631.

104. Kotenko SV, Gallagher G, Baurin VV, Lewis-Antes A, Shen M, Shah NK, Langer JA, Sheikh F, Dickensheets H, Donnelly RP: **IFN-lambdas mediate antiviral protection through a distinct class II cytokine receptor complex**. *Nature Immunology* 2003, **4(1)**:69-77.

105. LaFleur DW, Nardelli B, Tsareva T, Mather D, Feng P, Semenuk M, Taylor K, Buergin M, Chinchilla D, Roshke V *et al*: **Interferon-kappa, a novel type I interferon expressed in human keratinocytes**. *The Journal of Biological Chemistry* 2001, **276(43)**:39765-39771.

106. Martal JL, Chene NM, Huynh LP, L'Haridon RM, Reinaud PB, Guillomot MW, Charlier MA, Charpigny SY: **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* 1998, **80(8-9)**:755-777.

107. Oritani K, Medina KL, Tomiyama Y, Ishikawa J, Okajima Y, Ogawa M, Yokota T, Aoyama K, Takahashi I, Kincade PW *et al*: **Limitin: An interferon-like cytokine that preferentially influences B-lymphocyte precursors**. *Nature Medicine* 2000, **6(6)**:659-666.

108. Pestka S, Langer JA, Zoon KC, Samuel CE: **Interferons and their actions**. *Annu Rev Biochem* 1987, **56**:727-777.

109. Bach EA, Aguet M, Schreiber RD: **The IFN gamma receptor: a paradigm for cytokine receptor signaling**. *Annual Review of Immunology* 1997, **15**:563-591.

110. Farrar MA, Schreiber RD: **The molecular cell biology of interferon-gamma and its receptor**. *Annual Review of Immunology* 1993, **11**:571-611.

111. Gray PW, Goeddel DV: **Structure of the human immune interferon gene**. *Nature* 1982, **298(5877)**:859-863.

112. Diaz MO, Pomykala HM, Bohlander SK, Maltepe E, Malik K, Brownstein B, Olopade OI: **Structure of the human type-I interferon gene cluster determined from a YAC clone contig**. *Genomics* 1994, **22(3)**:540-552.

113. Le Page C, Genin P, Baines MG, Hiscott J: **Interferon activation and innate immunity**. *Reviews in Immunogenetics* 2000, **2(3)**:374-386.

114. Diaz MO: **The Human Type I Interferon Gene Cluster**. *Seminars in Virology* 1995, **6**:143-149.

115. Chen J, Baig E, Fish EN: **Diversity and relatedness among the type I interferons**. *J Interferon Cytokine Res* 2004, **24(12)**:687-698.

116. Finter NB: **Why Are There So Many Subtypes of Alpha-interferons?** *Journal of Interferon Research* 1991(January):185-194.

117. Hughes AL: **The evolution of the type I interferon gene family in mammals**. *Journal of Molecular Evolution* 1995, **41(5)**:539-548.

118. Kontsekova E, Liptakova H, Mucha V, Kontsek P: **Structural and functional heterogeneity of the amino-terminal receptor-binding domain of human interferon-alpha 2**. *International Journal of Biological Macromolecules* 1999, **24(1)**:11-14.

119. Weissmann C, Weber H: **The interferon genes**. *Progress in Nucleic Acid Research and Molecular Biology* 1986, **33**:251-300.

120. Braganca J, Civas A: **Type I interferon gene expression: differential expression of IFN-A genes induced by viruses and double-stranded RNA**. *Biochimie* 1998, **80(8-9)**:673-687.

121. Thanos D, Maniatis T: **Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome**. *Cell* 1995, **83(7)**:1091-1100.

122. Falvo JV, Parekh BS, Lin CH, Fraenkel E, Maniatis T: **Assembly of a functional beta interferon enhanceosome is dependent on ATF-2-c-jun heterodimer orientation**. *Molecular and Cellular Biology* 2000, **20(13)**:4814-4825.

123. Falvo JV, Thanos D, Maniatis T: **Reversal of intrinsic DNA bends in the IFN beta gene enhancer by transcription factors and the architectural protein HMG I(Y)**. *Cell* 1995, **83(7)**:1101-1111.

124. Sen GC, Lengyel P: **The interferon system. A bird's eye view of its biochemistry**. *The Journal of Biological Chemistry* 1992, **267(8)**:5017-5020.

125. Lenardo MJ, Fan CM, Maniatis T, Baltimore D: **The involvement of NF-kappa B in beta-interferon gene regulation reveals its role as widely inducible mediator of signal transduction**. *Cell* 1989, **57(2)**:287-294.

126. Keller AD, Maniatis T: **Identification and characterization of a novel repressor of beta-interferon gene expression**. *Genes & Development* 1991, **5(5)**:868-879.

127. Harada H, Fujita T, Miyamoto M, Kimura Y, Maruyama M, Furia A, Miyata T, Taniguchi T: **Structurally similar but functionally distinct factors, IRF-1 and IRF-2, bind to the same regulatory elements of IFN and IFN-inducible genes**. *Cell* 1989, **58(4)**:729-739.

128. MacDonald NJ, Kuhl D, Maguire D, Naf D, Gallant P, Goswamy A, Hug H, Bueler H, Chaturvedi M, de la Fuente J *et al*: **Different pathways mediate virus inducibility of the human IFN-alpha 1 and IFN-beta genes**. *Cell* 1990, **60(5)**:767-779.

129. Novick D, Cohen B, Rubinstein M: **The human interferon alpha/beta receptor: characterization and molecular cloning**. *Cell* 1994, **77(3)**:391-400.

130. Uze G, Lutfalla G, Gresser I: **Genetic transfer of a functional human interferon alpha receptor into mouse cells: cloning and expression of its cDNA**. *Cell* 1990, **60(2)**:225-234.

131. Domanski P, Witte M, Kellum M, Rubinstein M, Hackett R, Pitha P, Colamonici OR: **Cloning and expression of a long form of the beta subunit of the interferon alpha beta receptor that is required for signaling**. *The Journal of Biological Chemistry* 1995, **270(37)**:21606-21611.

132. Lutfalla G, Holland SJ, Cinato E, Monneron D, Reboul J, Rogers NC, Smith JM, Stark GR, Gardiner K, Mogensen KE *et al*: **Mutant U5A cells are complemented by an interferon-alpha beta receptor subunit generated by alternative processing of a new member of a cytokine receptor gene cluster**. *The EMBO Journal* 1995, **14(20)**:5100-5108.

133. Cohen B, Novick D, Barak S, Rubinstein M: **Ligand-induced association of the type I interferon receptor components**. *Molecular and Cellular Biology* 1995, **15(8)**:4208-4214.

134. Colamonici OR, Domanski P, Krolewski JJ, Fu XY, Reich NC, Pfeffer LM, Sweet ME, Platanias LC: **Interferon alpha (IFN alpha) signaling in cells expressing the variant form of the type I IFN receptor**. *The Journal of Biological Chemistry* 1994, **269(8)**:5660-5665.

135. Colamonici OR, Pfeffer LM, D'Alessandro F, Platanias LC, Gregory SA, Rosolen A, Nordan R, Cruciani RA, Diaz MO: **Multichain structure of the IFN-alpha receptor on hematopoietic cells**. *J Immunol* 1992, **148(7)**:2126-2132.

136. Darnell JE, Jr., Kerr IM, Stark GR: **Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins**. *Science* 1994, **264(5164)**:1415-1421.

137. Ihle JN: **The Janus protein tyrosine kinase family and its role in cytokine signaling**. *Advances in Immunology* 1995, **60**:1-35.

138. Platanias LC: **The p38 mitogen-activated protein kinase pathway and its role in interferon signaling**. *Pharmacology & Therapeutics* 2003, **98(2)**:129-142.

139. Darnell JE, Jr.: **The JAK-STAT pathway: summary of initial studies and recent advances**. *Recent Progress in Hormone Research* 1996, **51**:391-403; discussion 403-394.

140. Silvennoinen O, Ihle JN, Schlessinger J, Levy DE: **Interferon-induced nuclear signalling by Jak protein tyrosine kinases**. *Nature* 1993, **366(6455)**:583-585.

141. Aaronson DS, Horvath CM: **A road map for those who don't know JAK-STAT**. *Science* 2002, **296(5573)**:1653-1655.

142. Darnell JE, Jr.: **STATs and gene regulation**. *Science* 1997, **277(5332)**:1630-1635.

143. Stark GR, Kerr IM, Williams BR, Silverman RH, Schreiber RD: **How cells respond to interferons**. *Annu Rev Biochem* 1998, **67**:227-264.

144. Wen Z, Darnell JE, Jr.: **Mapping of Stat3 serine phosphorylation to a single residue (727) and evidence that serine phosphorylation has no influence on DNA binding of Stat1 and Stat3**. *Nucleic Acids Research* 1997, **25(11)**:2062-2067.

145. Wen Z, Zhong Z, Darnell JE, Jr.: **Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation**. *Cell* 1995, **82(2)**:241-250.

146. Uddin S, Sassano A, Deb DK, Verma A, Majchrzak B, Rahman A, Malik AB, Fish EN, Platanias LC: **Protein kinase C-delta (PKC-delta) is activated by type I interferons and mediates phosphorylation of Stat1 on serine 727**. *The Journal of Biological Chemistry* 2002, **277(17)**:14408-14416.

147. Waksman G, Kominos D, Robertson SC, Pant N, Baltimore D, Birge RB, Cowburn D, Hanafusa H, Mayer BJ, Overduin M *et al*: **Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides**. *Nature* 1992, **358(6388)**:646-653.

148. Heim MH, Kerr IM, Stark GR, Darnell JE, Jr.: **Contribution of STAT SH2 groups to specific interferon signaling by the Jak-STAT pathway**. *Science* 1995, **267(5202)**:1347-1349.

149. Platanias LC: **Mechanisms of type-I- and type-II-interferon-mediated signalling**. *Nature Reviews* 2005, **5(5)**:375-386.

150. Schindler C, Darnell JE, Jr.: **Transcriptional responses to polypeptide ligands: the JAK-STAT pathway**. *Annu Rev Biochem* 1995, **64**:621-651.

151. Bhattacharya S, Eckner R, Grossman S, Oldread E, Arany Z, D'Andrea A, Livingston DM: **Cooperation of Stat2 and p300/CBP in signalling induced by interferon-alpha**. *Nature* 1996, **383(6598)**:344-347.

152. Zhang JJ, Vinkemeier U, Gu W, Chakravarti D, Horvath CM, Darnell JE, Jr.: **Two contact regions between Stat1 and CBP/p300 in interferon gamma signaling**. *Proc Natl Acad Sci U S A* 1996, **93(26)**:15092-15096.

153. Hebbes TR, Thorne AW, Crane-Robinson C: **A direct link between core histone acetylation and transcriptionally active chromatin**. *The EMBO Journal* 1988, **7(5)**:1395-1402.

154. Chang HM, Paulson M, Holko M, Rice CM, Williams BR, Marie I, Levy DE: **Induction of interferon-stimulated gene expression and antiviral responses require protein deacetylase activity**. *Proc Natl Acad Sci U S A* 2004, **101(26)**:9578-9583.

155. Nusinzon I, Horvath CM: **Interferon-stimulated transcription and innate antiviral immunity require deacetylase activity and histone deacetylase 1**. *Proc Natl Acad Sci U S A* 2003, **100(25)**:14742-14747.

156. Sakamoto S, Potla R, Larner AC: **Histone deacetylase activity is required to recruit RNA polymerase II to the promoters of selected interferon-stimulated early response genes**. *The Journal of Biological Chemistry* 2004, **279(39)**:40362-40367.

157. Hannigan G, Williams BR: **Transcriptional regulation of interferon-responsive genes is closely linked to interferon receptor occupancy**. *The EMBO Journal* 1986, **5(7)**:1607-1613.

158. Talon J, Salvatore M, O'Neill RE, Nakaya Y, Zheng H, Muster T, Garcia-Sastre A, Palese P: **Influenza A and B viruses expressing altered NS1 proteins: A vaccine approach**. *Proc Natl Acad Sci U S A* 2000, **97(8)**:4309-4314.

159. Xiang Y, Condit RC, Vijaysri S, Jacobs B, Williams BR, Silverman RH: **Blockade of interferon induction and action by the E3L double-stranded RNA binding proteins of vaccinia virus**. *Journal of Virology* 2002, **76(10)**:5251-5259.

160. Unterstab G, Ludwig S, Anton A, Planz O, Dauber B, Krappmann D, Heins G, Ehrhardt C, Wolff T: **Viral targeting of the interferon-{beta}-inducing Traf family member-associated NF-{kappa}B activator (TANK)-binding kinase-1**. *Proc Natl Acad Sci U S A* 2005, **102(38)**:13640-13645.

161. Li K, Foy E, Ferreon JC, Nakamura M, Ferreon AC, Ikeda M, Ray SC, Gale M, Jr., Lemon SM: **Immune evasion by hepatitis C virus NS3/4A protease-mediated cleavage of the Toll-like receptor 3 adaptor protein TRIF**. *Proc Natl Acad Sci U S A* 2005, **102(8)**:2992-2997.

162. Symons JA, Alcami A, Smith GL: **Vaccinia virus encodes a soluble type I interferon receptor of novel structure and broad species specificity**. *Cell* 1995, **81(4)**:551-560.

163. Koromilas AE, Li S, Matlashewski G: **Control of interferon signaling in human papillomavirus infection**. *Cytokine & Growth Factor Reviews* 2001, **12(2-3)**:157-170.

164. Levy DE, Garcia-Sastre A: **The virus battles: IFN induction of the antiviral state and mechanisms of viral evasion**. *Cytokine & Growth Factor Reviews* 2001, **12(2-3)**:143-156.

165. Miller DM, Rahill BM, Boss JM, Lairmore MD, Durbin JE, Waldman JW, Sedmak DD: **Human cytomegalovirus inhibits major histocompatibility complex class II expression by disruption of the Jak/Stat pathway**. *The Journal of Experimental Medicine* 1998, **187(5)**:675-683.

166. Samuel CE: **Antiviral actions of interferons**. *Clinical Microbiology Reviews* 2001, **14(4)**:778-809, table of contents.

167. Garcin D, Curran J, Itoh M, Kolakofsky D: **Longer and shorter forms of Sendai virus C proteins play different roles in modulating the cellular antiviral response**. *Journal of Virology* 2001, **75(15)**:6800-6807.

168. Miller DM, Zhang Y, Rahill BM, Waldman WJ, Sedmak DD: **Human cytomegalovirus inhibits IFN-alpha-stimulated antiviral and immunoregulatory responses by blocking multiple levels of IFN-alpha signal transduction**. *J Immunol* 1999, **162(10)**:6107-6113.

169. Katze MG, He Y, Gale M, Jr.: **Viruses and interferon: a fight for supremacy**. *Nature Reviews* 2002, **2(9)**:675-687.

170. Sen GC: **Viruses and interferons**. *Annual Review of Microbiology* 2001, **55**:255-281.

171. Darnell MB, Koprowski H, Lagerspetz K: **Genetically determined resistance to infection with group B arboviruses. I. Distribution of the resistance gene among various mouse populations and characteristics of gene expression in vivo**. *J Infect Dis* 1974, **129(3)**:240-247.

172. Groschel D, Koprowski H: **Development of a virus-resistant inbred mouse strain for the study of innate resistance to Arbo B viruses**. *Archiv fur die gesamte Virusforschung* 1965, **17(3)**:379-391.

173. Urosevic N, Silvia OJ, Sangster MY, Mansfield JP, Hodgetts SI, Shellam GR: **Development and characterization of new flavivirus-resistant mouse strains bearing Flv(r)-like and Flv(mr) alleles from wild or wild-derived mice**. *The Journal of General Virology* 1999, **80 (Pt 4)**:897-906.

174. Brinton MA: Host susceptibility to viral disease. Philadelphia: Lippincott-Raven; 1997.

175. Urosevic N, van Maanen M, Mansfield JP, Mackenzie JS, Shellam GR: **Molecular characterization of virus-specific RNA produced in the brains of flavivirus-susceptible and -resistant mice after challenge with Murray**

Valley encephalitis virus. *The Journal of General Virology* 1997, **78 (Pt 1)**:23-29.

176. Jerrells TR, Osterman JV: **Host defenses in experimental scrub typhus: inflammatory response of congenic C3H mice differing at the Ric gene**. *Infection and Immunity* 1981, **31(3)**:1014-1022.

177. Sangster MY, Urosevic N, Mansfield JP, Mackenzie JS, Shellam GR: **Mapping the Flv locus controlling resistance to flaviviruses on mouse chromosome 5**. *Journal of Virology* 1994, **68(1)**:448-452.

178. Urosevic N, Mansfield JP, Mackenzie JS, Shellam GR: **Low resolution mapping around the flavivirus resistance locus (Flv) on mouse chromosome 5**. *Mamm Genome* 1995, **6**(7):454-458.

179. Urosevic N: **The use of microsatellites in high-resolution genetic mapping around the mouse flavivirus resistance locus (Flv)**. *Arbovirus Res Aust* 1997**(7)**:296-299.

180. Mashimo T, Lucas M, Simon-Chazottes D, Frenkiel MP, Montagutelli X, Ceccaldi PE, Deubel V, Guenet JL, Despres P: **A nonsense mutation in the gene encoding 2'-5'-oligoadenylate synthetase/L1 isoform is associated with West Nile virus susceptibility in laboratory mice**. *Proc Natl Acad Sci U S A* 2002, **99(17)**:11311-11316.

181. Urosevic N: **Is flavivirus resistance interferon type I-independent?** *Immunology and Cell Biology* 2003, **81(3)**:224-229.

182. Rebouillat D, Hovnanian A, David G, Hovanessian AG, Williams BR: **Characterization of the gene encoding the 100-kDa form of human 2',5' oligoadenylate synthetase**. *Genomics* 2000, **70(2)**:232-240.

183. Hartmann R, Olsen HS, Widder S, Jorgensen R, Justesen J: **p59OASL, a 2'-5' oligoadenylate synthetase like protein: a novel human gene related to the 2'-5' oligoadenylate synthetase family**. *Nucleic Acids Research* 1998, **26(18)**:4121-4128.

184. Benech P, Vigneron M, Peretz D, Revel M, Chebath J: **Interferon-responsive regulatory elements in the promoter of the human 2',5'-oligo(A) synthetase gene**. *Mol Cell Biol* 1987, **7(12)**:4498-4504.

185. Rutherford MN, Hannigan GE, Williams BR: **Interferon-induced binding of nuclear factors to promoter elements of the 2-5A synthetase gene**. *Embo J* 1988, **7(3)**:751-759.

186. Williams BR, Rutherford MN, Hannigan GE: **Interferon and growth factor modulation of nuclear factors binding to 5' upstream elements of the 2-5A synthetase gene**. *J Cell Biochem* 1988, **38(4)**:261-267.

187. Levy DE, Kessler DS, Pine R, Reich N, Darnell JE, Jr.: **Interferon-induced nuclear factors that bind a shared promoter element correlate with positive and negative transcriptional control**. *Genes Dev* 1988, **2(4)**:383-393.

188. Massa PT, Whitney LW, Wu C, Ropka SL, Jarosinski KW: **A mechanism for selective induction of 2'-5' oligoadenylate synthetase, anti-viral state, but not MHC class I genes by interferon-beta in neurons**. *J Neurovirol* 1999, **5(2)**:161-171.

189. Veals SA, Santa Maria T, Levy DE: **Two domains of ISGF3 gamma that mediate protein-DNA and protein-protein interactions during transcription factor assembly contribute to DNA-binding specificity**. *Mol Cell Biol* 1993, **13(1)**:196-206.

190. Naganuma A, Nozaki A, Tanaka T, Sugiyama K, Takagi H, Mori M, Shimotohno K, Kato N: **Activation of the interferon-inducible 2'-5'-oligoadenylate synthetase gene by hepatitis C virus core protein**. *J Virol* 2000, **74(18)**:8744-8750.

191. Dansako H, Naganuma A, Nakamura T, Ikeda F, Nozaki A, Kato N: **Differential activation of interferon-inducible genes by hepatitis C virus core protein mediated by the interferon stimulated response element**. *Virus Res* 2003, **97(1)**:17-30.

192. Tanaka T, Sugiyama K, Ikeda M, Naganuma A, Nozaki A, Saito M, Shimotohno K, Kato N: **Hepatitis C virus NS5B RNA replicase specifically binds ribosomes**. *Microbiol Immunol* 2000, **44(6)**:543-550.

193. Mashimo T, Glaser P, Lucas M, Simon-Chazottes D, Ceccaldi PE, Montagutelli X, Despres P, Guenet JL: **Structural and functional genomics and evolutionary relationships in the cluster of genes encoding murine 2',5'-oligoadenylate synthetases**. *Genomics* 2003, **82(5)**:537-552.

194. Kumar S, Mitnik C, Valente G, Floyd-Smith G: **Expansion and molecular evolution of the interferon-induced 2'-5' oligoadenylate synthetase gene family**. *Mol Biol Evol* 2000, **17(5)**:738-750.

195. Rios JJ, Perelygin AA, Long MT, Lear TL, Zharkikh AA, Brinton MA, Adelson DL: **Characterization of the equine 2-5 oligoadenylate synthetase 1 (OAS1)**

and ribonuclease L (RNASEL) innate immunity genes. *BMC Genomics* 2007, **8(1)**:313.

196. De Clercq E: **Synthetic interferon inducers**. *Top Curr Chem* 1974, **52**:173-208.

197. Hunter T, Hunt T, Jackson RJ, Robertson HD: **The characteristics of inhibition of protein synthesis by double-stranded ribonucleic acid in reticulocyte lysates**. *The Journal of Biological Chemistry* 1975, **250(2)**:409-417.

198. Elbashir SM, Lendeckel W, Tuschl T: **RNA interference is mediated by 21- and 22-nucleotide RNAs**. *Genes Dev* 2001, **15(2)**:188-200.

199. Marques JT, Devosse T, Wang D, Zamanian-Daryoush M, Serbinowski P, Hartmann R, Fujita T, Behlke MA, Williams BR: **A structural basis for discriminating between self and nonself double-stranded RNAs in mammalian cells**. *Nat Biotechnol* 2006, **24(5)**:559-565.

200. Carter WA, Pitha PM, Marshall LW, Tazawa I, Tazawa S, Ts'o PO: **Structural requirements of the rI n -rC n complex for induction of human interferon**. *J Mol Biol* 1972, **70(3)**:567-587.

201. Minks MA, West DK, Benvin S, Greene JJ, Ts'o PO, Baglioni C: **Activation of 2',5'-oligo(A) polymerase and protein kinase of interferon-treated HeLa cells by 2'-O-methylated poly (inosinic acid) . poly(cytidylic acid), Correlations with interferon-inducing activity**. *The Journal of Biological Chemistry* 1980, **255(13)**:6403-6407.

202. Baglioni C, Minks MA, De Clercq E: **Structural requirements of polynucleotides for the activation of (2' - 5')An polymerase and protein kinase**. *Nucleic Acids Research* 1981, **9(19)**:4939-4950.

203. Ghosh A, Sarkar SN, Guo W, Bandyopadhyay S, Sen GC: **Enzymatic activity of 2'-5'-oligoadenylate synthetase is impaired by specific mutations that affect oligomerization of the protein**. *J Biol Chem* 1997, **272(52)**:33220-33226.

204. Marie I, Svab J, Robert N, Galabru J, Hovanessian AG: **Differential expression and distinct structure of 69- and 100-kDa forms of 2-5A synthetase in human cells treated with interferon**. *The Journal of Biological Chemistry* 1990, **265(30)**:18601-18607.

205. Rebouillat D, Hovanessian AG: **The human 2',5'-oligoadenylate synthetase family: interferon-induced proteins with unique enzymatic properties**. *J Interferon Cytokine Res* 1999, **19(4)**:295-308.

206. Sarkar SN, Ghosh A, Wang HW, Sung SS, Sen GC: **The nature of the catalytic domain of 2'-5'-oligoadenylate synthetases**. *The Journal of Biological Chemistry* 1999, **274(36)**:25535-25542.

207. Witt PL, Marie I, Robert N, Irizarry A, Borden EC, Hovanessian AG: **Isoforms p69 and p100 of 2',5'-oligoadenylate synthetase induced differentially by interferons in vivo and in vitro**. *Journal of Interferon Research* 1993, **13(1)**:17-23.

208. Castora FJ, Erickson CE, Kovacs T, Lesiak K, Torrence PF: **2',5'-oligoadenylates inhibit relaxation of supercoiled DNA by calf thymus DNA topoisomerase I**. *Journal of Interferon Research* 1991, **11(3)**:143-149.

209. Squire J, Zhou A, Hassel BA, Nie H, Silverman RH: **Localization of the interferon-induced, 2-5A-dependent RNase gene (RNS4) to human chromosome 1q25**. *Genomics* 1994, **19(1)**:174-175.

210. Diaz-Guerra M, Rivas C, Esteban M: **Full activation of RNaseL in animal cells requires binding of 2-5A within ankyrin repeats 6 to 9 of this interferon-inducible enzyme**. *J Interferon Cytokine Res* 1999, **19(2)**:113-119.

211. Player MR, Torrence PF: **The 2-5A system: modulation of viral and cellular processes through acceleration of RNA degradation**. *Pharmacology & Therapeutics* 1998, **78(2)**:55-113.

212. Wreschner DH, McCauley JW, Skehel JJ, Kerr IM: **Interferon action--sequence specificity of the ppp(A2'p)nA-dependent ribonuclease**. *Nature* 1981, **289(5796)**:414-417.

213. Floyd-Smith G, Slattery E, Lengyel P: **Interferon action: RNA cleavage pattern of a (2'-5')oligoadenylate--dependent endonuclease**. *Science* 1981, **212(4498)**:1030-1032.

214. Wreschner DH, James TC, Silverman RH, Kerr IM: **Ribosomal RNA cleavage, nuclease activation and 2-5A(ppp(A2'p)nA) in interferon-treated cells**. *Nucleic Acids Research* 1981, **9(7)**:1571-1581.

215. Rusch L, Zhou A, Silverman RH: **Caspase-dependent apoptosis by 2',5'-oligoadenylate activation of RNase L is enhanced by IFN-beta**. *J Interferon Cytokine Res* 2000, **20(12)**:1091-1100.

216. Zhou A, Molinaro RJ, Malathi K, Silverman RH: **Mapping of the human RNASEL promoter and expression in cancer and normal cells**. *J Interferon Cytokine Res* 2005, **25(10)**:595-603.

217. Li XL, Andersen JB, Ezelle HJ, Wilson GM, Hassel BA: **Post-transcriptional regulation of RNase-L expression is mediated by the 3'-untranslated region of its mRNA**. *The Journal of Biological Chemistry* 2007, **282(11)**:7950-7960.

218. Ford LP, Watson J, Keene JD, Wilusz J: **ELAV proteins stabilize deadenylated intermediates in a novel in vitro mRNA deadenylation/degradation system**. *Genes Dev* 1999, **13(2)**:188-201.

219. Gao M, Wilusz CJ, Peltz SW, Wilusz J: **A novel mRNA-decapping activity in HeLa cytoplasmic extracts is regulated by AU-rich elements**. *EMBO J* 2001, **20(5)**:1134-1143.

220. Hamano E, Hijikata M, Itoyama S, Quy T, Phi NC, Long HT, Ha le D, Ban VV, Matsushita I, Yanai H *et al*: **Polymorphisms of interferon-inducible genes OAS-1 and MxA associated with SARS in the Vietnamese population**. *Biochemical and Biophysical Research Communications* 2005, **329(4)**:1234-1239.

221. He J, Feng D, de Vlas SJ, Wang H, Fontanet A, Zhang P, Plancoulaine S, Tang F, Zhan L, Yang H *et al*: **Association of SARS susceptibility with single nucleic acid polymorphisms of OAS1 and MxA genes: a case-control study**. *BMC Infectious Diseases* 2006, **6**:106.

222. Knapp S, Yee LJ, Frodsham AJ, Hennig BJ, Hellier S, Zhang L, Wright M, Chiaramonte M, Graves M, Thomas HC *et al*: **Polymorphisms in interferon-induced genes and the outcome of hepatitis C virus infection: roles of MxA, OAS-1 and PKR**. *Genes and Immunity* 2003, **4(6)**:411-419.

223. **Laboratory of Phil Green** [http://www.phrap.org] 2007

224. NCBI: **NCBI Trace Archive**. [http://wwwncbinihgov/Traces/tracecgi?] 2006

225. Stephens M, Scheet P: **Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation**. *American Journal of Human Genetics* 2005, **76(3)**:449-462.

226. Stephens M, Smith NJ, Donnelly P: **A new statistical method for haplotype reconstruction from population data**. *American Journal of Human Genetics* 2001, **68(4)**:978-989.

227. **TIGR Institute for Genomic Research** [www.tigrblast.tigr.org] 2007

228. **Pig EST Data Explorer** [pede.dna.affrc.go.jp] 2007

229. **UCSC Genome Bioinformatics** [www.genome.ucsc.edu] 2007

230. **MOTIF Search** [http://motif.genome.jp/] 2007

231. Sarkar SN, Miyagi M, Crabb JW, Sen GC: **Identification of the substrate-binding sites of 2'-5'-oligoadenylate synthetase**. *J Biol Chem* 2002, **277(27)**:24321-24330.

232. Cayley PJ, White RF, Antoniw JF, Walesby NJ, Kerr IM: **Distribution of the ppp(A2'p)nA-binding protein and interferon-related enzymes in animals, plants, and lower organisms**. *Biochemical and Biophysical Research Communications* 1982, **108(3)**:1243-1250.

233. **Ensembl** [www.ensembl.org] 2007

234. Raudsepp T, Fronicke L, Scherthan H, Gustavsson I, Chowdhary BP: **Zoo-FISH delineates conserved chromosomal segments in horse and man**. *Chromosome Res* 1996, **4(3)**:218-225.

235. Nilsson S, Helou K, Walentinsson A, Szpirer C, Nerman O, Stahl F: **Rat-mouse and rat-human comparative maps based on gene homology and high-resolution zoo-FISH**. *Genomics* 2001, **74(3)**:287-298.

236. Breen M, Thomas R, Binns MM, Carter NP, Langford CF: **Reciprocal chromosome painting reveals detailed regions of conserved synteny between the karyotypes of the domestic dog (Canis familiaris) and human**. *Genomics* 1999, **61(2)**:145-155.

237. Yang F, O'Brien PC, Milne BS, Graphodatsky AS, Solanky N, Trifonov V, Rens W, Sargan D, Ferguson-Smith MA: **A complete comparative chromosome map for the dog, red fox, and human and its integration with canine genetic maps**. *Genomics* 1999, **62(2)**:189-202.

238. Solinas-Toldo S, Lengauer C, Fries R: **Comparative genome map of human and cattle**. *Genomics* 1995, **27(3)**:489-496.

239. Feng Y, Goulet AC, Nelson MA: **Identification and characterization of the human Cdc2l2 gene promoter**. *Gene* 2004, **330**:75-84.

240. Ogasawara K, Terada T, Asaka J, Katsura T, Inui K: **Human organic anion transporter 3 gene is regulated constitutively and inducibly via a cAMP-response element**. *The Journal of Pharmacology and Experimental Therapeutics* 2006, **319(1)**:317-322.

241. Rani CS, Qiang M, Ticku MK: **Potential role of cAMP response element-binding protein in ethanol-induced N-methyl-D-aspartate receptor 2B subunit gene transcription in fetal mouse cortical cells**. *Molecular Pharmacology* 2005, **67(6)**:2126-2136.

242. **Patrocles Targets Database** [http://www.patrocles.org/] 2007

243. Lindblad-Toh K: In: *Plant & Animal Genome XV: January 13-17, 2007; San Diego, California, USA*; 2007.

244. Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, Kamal M, Clamp M, Chang JL, Kulbokas EJ, 3rd, Zody MC *et al*: **Genome sequence, comparative analysis and haplotype structure of the domestic dog**. *Nature* 2005, **438(7069)**:803-819.

245. Perelygin AA, Lear TL, Zharkikh AA, Brinton MA: **Comparative analysis of vertebrate EIF2AK2 (PKR) genes and assignment of the equine gene to ECA15q24-q25 and the bovine gene to BTA11q12-q15**. *Genet Sel Evol* 2006, **38(5)**:551-563.

246. Bowling AT, Breen M, Chowdhary BP, Hirota K, Lear T, Millon LV, Ponce de Leon FA, Raudsepp T, Stranzinger G: **International system for cytogenetic nomenclature of the domestic horse. Report of the Third International Committee for the Standardization of the domestic horse karyotype, Davis, CA, USA, 1996**. In: *Chromosome Res.* **vol. 5**; 1997: 433-443.

247. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *J Mol Biol* 1990, **215(3)**:403-410.

248. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucleic Acids Res* 1997, **25(17)**:3389-3402.

249. **National Center for Biotechnology Information** [http://www.ncbi.nlm.nih.gov/] 2007

250. Scherbik SV, Kluetzman K, Perelygin AA, Brinton MA: **Knock-in of the Oas1b(r) allele into a flavivirus-induced disease susceptible mouse generates the resistant phenotype**. *Virology* 2007, **368(2)**:232-237.

251. Scherbik SV, Paranjape JM, Stockman BM, Silverman RH, Brinton MA: **RNase L plays a role in the antiviral response to West Nile virus**. *J Virol* 2006, **80(6)**:2987-2999.

252. Rios JJ, Perelygin AA, Long MT, Lear TL, Zharkikh AA, Brinton MA, Adelson DL: **Characterization of the equine 2'-5' oligoadenylate synthetase 1 (OAS1) and ribonuclease L (RNASEL) innate immunity genes**. *BMC Genomics* 2007, **8**:313.

253. Ian Dohoo WM, Henrik Stryhn: Veterinary Epidemiologic Research. Prince Edward Island: AVC Inc.; 2003.

254. StataCorp: Stata Statistical Software: Release 9. In. College Station, Texas: StataCorp LP; 2005.

255. Chapman JM, Cooper JD, Todd JA, Clayton DG: **Detecting disease associations due to linkage disequilibrium using haplotype tags: a class of tests and the determinants of statistical power**. *Hum Hered* 2003, **56(1-3)**:18-31.

256. Fleming JG, Spencer TE, Safe SH, Bazer FW: **Estrogen regulates transcription of the ovine oxytocin receptor gene through GC-rich SP1 promoter elements**. *Endocrinology* 2006, **147(2)**:899-911.

257. Pellegrini S, John J, Shearer M, Kerr IM, Stark GR: **Use of a selectable marker regulated by alpha interferon to obtain mutations in the signaling pathway**. *Mol Cell Biol* 1989, **9(11)**:4605-4612.

258. Van Heeke G, Ott TL, Strauss A, Ammaturo D, Bazer FW: **High yield expression and secretion of the ovine pregnancy recognition hormone interferon-tau by Pichia pastoris**. *J Interferon Cytokine Res* 1996, **16(2)**:119-126.

APPENDIX A

CASE-CONTROL STATISTICAL ANALYSES OF EQUINE *OAS1* POLYMORPHISMS

Allelic Fisher's Exact Tests

snp6567078

	Allele T	C	Total	
Control	30	8	38	Fisher's exact = 0.002
Case	52	54	106	1-sided Fisher's exact = 0.001
Total	82	62	144	

snp6567031

	Allele G	T	Total	
Control	30	8	38	Fisher's exact = 0.004
Case	56	52	108	1-sided Fisher's exact = 0.003
Total	86	60	146	

snp6567000

	Allele A	G	Total	
Control	32	8	40	Fisher's exact = 0.074
Case	68	40	108	1-sided Fisher's exact = 0.036
Total	100	48	148	

snp6566994

	Allele C	T	Total	
Control	32	8	40	Fisher's exact = 0.007
Case	59	49	108	1-sided Fisher's exact = 0.004
Total	91	57	148	

snp6566893

	Allele T	C	Total	
Control	32	8	40	Fisher's exact = 0.007
Case	60	48	108	1-sided Fisher's exact = 0.005
Total	92	56	148	

snp6566888

	Allele A	Allele T	Total	
Control	34	6	40	Fisher's exact = 0.133
Case	77	31	108	1-sided Fisher's exact = 0.064
Total	111	37	148	

snp6566745

	Allele C	Allele T	Total	
Control	32	8	40	Fisher's exact = 0.004
Case	59	51	110	1-sided Fisher's exact = 0.003
Total	91	59	150	

snp6566713

	Allele C	Allele G	Total	
Control	37	1	38	Fisher's exact = 0.298
Case	99	11	110	1-sided Fisher's exact = 0.135
Total	136	12	148	

snp6566498

	Allele C	Allele T	Total	
Control	32	8	40	Fisher's exact = 0.004
Case	61	53	114	1-sided Fisher's exact = 0.002
Total	93	61	154	

snp6566399

	Allele T	Allele C	Total	
Control	32	8	40	Fisher's exact = 0.792
Case	32	10	42	1-sided Fisher's exact = 0.441
Total	64	18	82	

snp6566277

	Allele C	Allele G	Total	
Control	5	35	40	Fisher's exact = 0.134
Case	5	103	108	1-sided Fisher's exact = 0.096
Total	10	138	148	

snp6566231

	Allele		
	C	T	Total
Control	32	8	40
Case	50	60	110
Total	82	68	150

Fisher's exact = 0.000
1-sided Fisher's exact = 0.000

snp6566201

	Allele		
	A	G	Total
Control	6	34	40
Case	6	98	104
Total	12	132	144

Fisher's exact = 0.093
1-sided Fisher's exact = 0.077

snp6566134

	Allele		
	C	G	Total
Control	31	7	38
Case	44	32	76
Total	75	39	114

Fisher's exact = 0.013
1-sided Fisher's exact = 0.009

snp6566107

	Allele		
	C	A	Total
Control	32	8	40
Case	45	29	74
Total	77	37	114

Fisher's exact = 0.039
1-sided Fisher's exact = 0.028

snp6566042

	Allele		
	A	G	Total
Control	37	3	40
Case	67	9	76
Total	104	12	116

Fisher's exact = 0.540
1-sided Fisher's exact = 0.351

snp6565949

	Allele		
	T	C	Total
Control	32	8	40
Case	52	30	82
Total	84	38	122

Fisher's exact = 0.095
1-sided Fisher's exact = 0.047

snp6565123

	Allele C	Allele T	Total	
Control	30	10	40	Fisher's exact = 0.828
Case	85	25	110	1-sided Fisher's exact = 0.464
Total	115	35	150	

snp6565031

	Allele C	Allele T	Total	
Control	30	10	40	Fisher's exact = 0.681
Case	72	30	102	1-sided Fisher's exact = 0.380
Total	102	40	142	

snp6564989

	Allele C	Allele T	Total	
Control	10	30	40	Fisher's exact = 0.490
Case	20	88	108	1-sided Fisher's exact = 0.257
Total	30	118	148	

snp6564967

	Allele G	Allele A	Total	
Control	10	30	40	Fisher's exact = 0.348
Case	18	86	104	1-sided Fisher's exact = 0.207
Total	28	116	144	

snp6564956

	Allele A	Allele G	Total	
Control	10	30	40	Fisher's exact = 0.490
Case	20	88	108	1-sided Fisher's exact = 0.257
Total	30	118	148	

snp6564946

	Allele A	Allele G	Total	
Control	10	30	40	Fisher's exact = 0.498
Case	21	87	108	1-sided Fisher's exact = 0.300
Total	31	117	148	

snp6550610

	Allele G	Allele C	Total	
Control	18	22	40	Fisher's exact = 0.045
Case	28	78	106	1-sided Fisher's exact = 0.027
Total	46	100	146	

snp6550514

	Allele T	Allele C	Total	
Control	20	20	40	Fisher's exact = 0.248
Case	31	49	80	1-sided Fisher's exact = 0.164
Total	51	69	120	

snp6550471

	Allele A	Allele G	Total	
Control	20	18	38	Fisher's exact = 0.044
Case	26	54	80	1-sided Fisher's exact = 0.030
Total	46	72	118	

snp6549905

	Allele A	Allele G	Total	
Control	23	15	38	Fisher's exact = 0.251
Case	47	51	98	1-sided Fisher's exact = 0.130
Total	70	66	136	

snp6549803

	Allele A	Allele G	Total	
Control	32	6	38	Fisher's exact = 0.100
Case	42	20	62	1-sided Fisher's exact = 0.054
Total	74	26	100	

snp6549696

	Allele A	Allele T	Total	
Control	21	17	38	Fisher's exact = 0.702
Case	48	48	96	1-sided Fisher's exact = 0.361
Total	69	65	134	

snp6549675

	Allele		
	C	G	Total
Control	22	16	38
Case	50	46	96
Total	72	62	134

Fisher's exact = 0.570
1-sided Fisher's exact = 0.340

snp6548520

	Allele		
	C	T	Total
0	29	9	38
1	53	31	84
Total	82	40	122

Fisher's exact = 0.211
1-sided Fisher's exact = 0.108

snp6548430

	Allele		
	G	T	Total
Control	19	19	38
Case	45	31	76
Total	64	50	114

Fisher's exact = 0.424
1-sided Fisher's exact = 0.231

Genotypic Fisher's Exact Tests

snp6567078

	Genotype			
	CC	CT	TT	Total
Control	2	4	13	19
Case	12	30	11	53
Total	14	34	24	72

Fisher's exact = 0.001

snp6567031

	Genotype			
	GG	GT	TT	Total
Control	13	4	2	19
Case	11	34	9	54
Total	24	38	11	73

Fisher's exact = 0.001

snp6567000

	Genotype			
	AA	AG	GG	Total
Control	14	4	2	20
Case	15	38	1	54
Total	29	42	3	74

Fisher's exact = 0.000

snp6566994

	Genotype			
	CC	CT	TT	Total
Control	14	4	2	20
Case	11	37	6	54
Total	25	41	8	74

Fisher's exact = 0.000

snp6566893

	Genotype			
	CC	CT	TT	Total
Control	2	4	14	20
Case	6	36	12	54
Total	8	40	26	74

Fisher's exact = 0.000

snp6566888

	Genotype			
	AA	AT	TT	Total
Control	15	4	1	20
Case	23	31	0	54
Total	38	35	1	74

Fisher's exact = 0.003

snp6566745

	Genotype			
	CC	CT	TT	Total
Control	14	4	2	20
Case	11	37	7	55
Total	25	41	9	75

Fisher's exact = 0.000

snp6566713

	Genotype		
	CC	CG	Total
Control	18	1	19
Case	44	11	55
Total	62	12	74

Fisher's exact = 0.169
1-sided Fisher's exact = 0.124

snp6566498

	Genotype			
	CC	CT	TT	Total
Control	14	4	2	20
Case	14	33	10	57
Total	28	37	12	77

Fisher's exact = 0.001

snp6566399

	Genotype			
	CC	CT	TT	Total
Control	2	4	14	20
Case	0	10	11	21
Total	2	14	25	41

Fisher's exact = 0.091

snp6566277

	Genotype		
	CG	GG	Total
Control	5	15	20
Case	5	49	54
Total	10	64	74

Fisher's exact = 0.122
1-sided Fisher's exact = 0.088

snp6566231

	Genotype			
	CC	CT	TT	Total
Control	14	4	2	20
Case	11	28	16	55
Total	25	32	18	75

Fisher's exact = 0.000

snp6566201

	Genotype			
	AA	AG	GG	Total
Control	1	4	15	20
Case	1	4	47	52
Total	2	8	62	72

Fisher's exact = 0.205

snp6566134

	Genotype			
	CC	CG	GG	Total
Control	14	3	2	19
Case	16	12	10	38
Total	30	15	12	57

Fisher's exact = 0.092

snp6566107

	Genotype			
	AA	AC	CC	Total
Control	2	4	14	20
Case	8	13	16	37
Total	10	17	30	57

Fisher's exact = 0.197

snp6566042

	Genotype			
	AA	AG	GG	Total
Control	17	3	0	20
Case	30	7	1	38
Total	47	10	1	58

Fisher's exact = 1.000

snp6565949

	Genotype			
	CC	CT	TT	Total
Control	2	4	14	20
Case	8	14	19	41
Total	10	18	33	61

Fisher's exact = 0.278

snp6565123

	Genotype			
	CC	CT	TT	Total
Control	11	8	1	20
Case	35	15	5	55
Total	46	23	6	75

Fisher's exact = 0.593

snp6565031

	Genotype			
	CC	CT	TT	Total
Control	11	8	1	20
Case	29	14	8	51
Total	40	22	9	71

Fisher's exact = 0.395

snp6564989

	Genotype			
	CC	CT	TT	Total
Control	1	8	11	20
Case	5	10	39	54
Total	6	18	50	74

Fisher's exact = 0.168

snp6564967

	Genotype			
	AA	AG	GG	Total
Control	11	8	1	20
Case	38	10	4	52
Total	49	18	5	72

Fisher's exact = 0.203

snp6564956

	Genotype			
	AA	AG	GG	Total
Control	1	8	11	20
Case	5	10	39	54
Total	6	18	50	74

Fisher's exact = 0.168

snp6564946

	Genotype			
	AA	AG	GG	Total
Control	1	8	11	20
Case	5	11	38	54
Total	6	19	49	74

Fisher's exact = 0.238

snp6550610

	Genotype			
	CC	CG	GG	Total
Control	7	8	5	20
Case	31	16	6	53
Total	38	24	11	73

Fisher's exact = 0.149

snp6550514

	Genotype			
	CC	CT	TT	Total
Control	5	10	5	20
Case	16	17	7	40
Total	21	27	12	60

Fisher's exact = 0.524

snp6550471

	Genotype			
	AA	AG	GG	Total
Control	6	8	5	19
Case	5	16	19	40
Total	11	24	24	59

Fisher's exact = 0.135

snp6549905

	Genotype			
	AA	AG	GG	Total
Control	9	5	5	19
Case	16	15	18	49
Total	25	20	23	68

Fisher's exact = 0.555

snp6549803

	Genotype			
	AA	AG	GG	Total
Control	14	4	1	19
Case	17	8	6	31
Total	31	12	7	50

Fisher's exact = 0.339

snp6549696

	Genotype			
	AA	AT	TT	Total
Control	7	7	5	19
Case	11	26	11	48
Total	18	33	16	67

Fisher's exact = 0.360

snp6549675

	Genotype			
	CC	CG	GG	Total
Control	8	6	5	19
Case	13	24	11	48
Total	21	30	16	67

Fisher's exact = 0.346

snp6548520

	Genotype			
	CC	CT	TT	Total
Control	11	7	1	19
Case	21	11	10	42
Total	32	18	11	61

Fisher's exact = 0.198

snp6548430

	Genotype			
	GG	GT	TT	Total
Control	5	9	5	19
Case	10	25	3	38
Total	15	34	8	57

Fisher's exact = 0.189

Genotypic Odds Ratios

snp6567078

```
                                                  Number of obs   =         72
                                                  LR chi2(2)      =      13.88
                                                  Prob > chi2     =     0.0010
Log likelihood =  -34.60889                       Pseudo R2       =     0.1671
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   7.090909   6.145649     2.26   0.024     1.297091    38.76442
Heterozygous |   8.863636   5.953402     3.25   0.001     2.376232    33.06245
------------------------------------------------------------------------------
```

snp6567031

```
                                                  Number of obs   =         73
                                                  LR chi2(2)      =      14.60
                                                  Prob > chi2     =     0.0007
Log likelihood = -34.554474                       Pseudo R2       =     0.1744
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   10.04545   6.718028     3.45   0.001     2.708429    37.25818
Homozygous   |   5.318182   4.693708     1.89   0.058     .9430001    29.99264
------------------------------------------------------------------------------
```

snp6567000

```
                                                  Number of obs   =         74
                                                  LR chi2(2)      =      15.96
                                                  Prob > chi2     =     0.0003
Log likelihood = -35.202238                       Pseudo R2       =     0.1848
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   8.866667   5.707895     3.39   0.001     2.510738    31.31261
Homozygous   |   .4666667   .5972778    -0.60   0.552     .0379813    5.733812
------------------------------------------------------------------------------
```

snp6566994

```
                                                  Number of obs   =         74
                                                  LR chi2(2)      =      16.85
                                                  Prob > chi2     =     0.0002
Log likelihood = -34.754241                       Pseudo R2       =     0.1952
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   11.77273   7.803504     3.72   0.000     3.211158     43.1611
Homozygous   |   3.818182   3.476442     1.47   0.141     .6409739    22.74431
------------------------------------------------------------------------------
```

snp6566893

```
                                                  Number of obs   =         74
                                                  LR chi2(2)      =      15.47
                                                  Prob > chi2     =     0.0004
Log likelihood = -35.446828                       Pseudo R2       =     0.1791
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |        3.5   3.172144     1.38   0.167      .592381    20.67926
Heterozygous |       10.5   6.905613     3.58   0.000     2.893165    38.10705
------------------------------------------------------------------------------
```

snp6566888

```
                                                  Number of obs   =         73
                                                  LR chi2(1)      =       7.85
                                                  Prob > chi2     =     0.0051
Log likelihood = -37.929555                       Pseudo R2       =     0.0938
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   5.054348   3.166148     2.59   0.010     1.480658    17.25343
------------------------------------------------------------------------------
```

snp6566745

```
                                                  Number of obs   =         75
                                                  LR chi2(2)      =      16.94
                                                  Prob > chi2     =     0.0002
Log likelihood = -35.022915                       Pseudo R2       =     0.1948
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   11.77273   7.803504     3.72   0.000     3.211158     43.1611
Homozygous   |   4.454545   3.997182     1.66   0.096     .7673574    25.85884
------------------------------------------------------------------------------
```

snp6566713

```
                                                  Number of obs   =         74
                                                  LR chi2(1)      =       2.72
                                                  Prob > chi2     =     0.0991
Log likelihood = -40.793328                       Pseudo R2       =     0.0323
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |        4.5   4.865811     1.39   0.164     .5405249    37.46359
------------------------------------------------------------------------------
```

snp6566498

```
                                                  Number of obs   =         77
                                                  LR chi2(2)      =      13.23
                                                  Prob > chi2     =     0.0013
Log likelihood = -37.488891                       Pseudo R2       =     0.1500
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |       8.25   5.366688     3.24   0.001     2.305371    29.52345
Homozygous   |          5   4.309458     1.87   0.062     .9232654    27.07781
------------------------------------------------------------------------------
```

snp6566399

```
                                                  Number of obs   =         39
                                                  LR chi2(1)      =       2.79
                                                  Prob > chi2     =     0.0951
Log likelihood = -25.524019                       Pseudo R2       =     0.0518
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   3.181818   2.277474     1.62   0.106     .7823558    12.94036
------------------------------------------------------------------------------
```

snp6566277

```
                                                  Number of obs   =         74
                                                  LR chi2(1)      =       2.80
                                                  Prob > chi2     =     0.0941
Log likelihood = -41.780042                       Pseudo R2       =     0.0324
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   3.266667   2.279831     1.70   0.090     .8318498    12.82817
------------------------------------------------------------------------------
```

snp6566231

```
                                                  Number of obs   =         75
                                                  LR chi2(2)      =      16.02
                                                  Prob > chi2     =     0.0003
Log likelihood = -35.483868                       Pseudo R2       =     0.1842
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   8.909091    5.96345     3.27   0.001     2.399205    33.08259
Homozygous   |   10.18182   8.668535     2.73   0.006     1.919263    54.01522
------------------------------------------------------------------------------
```

snp6566201

```
                                                  Number of obs   =         72
                                                  LR chi2(2)      =       2.61
                                                  Prob > chi2     =     0.2713
Log likelihood = -41.236113                       Pseudo R2       =     0.0307
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |          1   1.581139    -0.00   1.000     .0450954    22.17521
Homozygous   |   3.133333   4.527578     0.79   0.429     .1845204      53.207
------------------------------------------------------------------------------
```

snp6566134

```
                                                  Number of obs   =         57
                                                  LR chi2(2)      =       5.28
                                                  Prob > chi2     =     0.0713
Log likelihood =  -33.64047                       Pseudo R2       =     0.0728
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |        3.5   2.597073     1.69   0.091     .8174456    14.98571
```

```
Homozygous   |      4.375   3.748022     1.72   0.085     .8161342    23.45279
------------------------------------------------------------------------------

snp6566107
                                                  Number of obs   =         57
                                                  LR chi2(2)      =       3.86
                                                  Prob > chi2     =     0.1454
Log likelihood = -35.006831                       Pseudo R2       =     0.0522
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |        3.5   3.049075     1.44   0.150     .6346461     19.302
Heterozygous |    2.84375    1.93051     1.54   0.124     .7516995    10.75817
------------------------------------------------------------------------------

snp6566042
                                                  Number of obs   =         57
                                                  LR chi2(1)      =       0.14
                                                  Prob > chi2     =     0.7077
Log likelihood = -36.865032                       Pseudo R2       =     0.0019
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   1.322222   .9968051     0.37   0.711     .3017139    5.794468
------------------------------------------------------------------------------

snp6565949
                                                  Number of obs   =         61
                                                  LR chi2(2)      =       3.12
                                                  Prob > chi2     =     0.2102
Log likelihood = -37.032342                       Pseudo R2       =     0.0404
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   2.947368   2.550895     1.25   0.212     .5404239    16.07438
Heterozygous |   2.578947   1.721317     1.42   0.156     .6971216    9.540616
------------------------------------------------------------------------------

snp6565123
                                                  Number of obs   =         75
                                                  LR chi2(2)      =       1.25
                                                  Prob > chi2     =     0.5343
Log likelihood = -42.866923                       Pseudo R2       =     0.0144
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   .5892857   .3287069    -0.95   0.343      .197479    1.758454
Homozygous   |   1.571429   1.805078     0.39   0.694     .1653985    14.92993
------------------------------------------------------------------------------


snp6565031
                                                  Number of obs   =         71
                                                  LR chi2(2)      =       2.25
                                                  Prob > chi2     =     0.3244
Log likelihood = -41.086839                       Pseudo R2       =     0.0267
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   .6637931   .3765646    -0.72   0.470     .2183498    2.017961
```

```
Homozygous   |   3.034483   3.393187     0.99   0.321     .3390496    27.15852
------------------------------------------------------------------------------

snp6564989
                                                Number of obs   =         74
                                                LR chi2(2)      =       3.53
                                                Prob > chi2     =     0.1709
Log likelihood = -41.414074                     Pseudo R2       =     0.0409
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |        .25   .2984332    -1.16   0.246     .0240898    2.594462
Homozygous   |   .7090909   .8136176    -0.30   0.764     .0748213    6.720144
------------------------------------------------------------------------------

snp6564967
                                                Number of obs   =         72
                                                LR chi2(2)      =       3.16
                                                Prob > chi2     =     0.2061
Log likelihood = -40.961393                     Pseudo R2       =     0.0371
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   .8636364   1.009836    -0.13   0.900     .0873041     8.54333
Heterozygous |      .3125   .3795299    -0.96   0.338     .0289114    3.377775
------------------------------------------------------------------------------

snp6564956
                                                Number of obs   =         74
                                                LR chi2(2)      =       3.53
                                                Prob > chi2     =     0.1709
Log likelihood = -41.414074                     Pseudo R2       =     0.0409
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |        .25   .2984332    -1.16   0.246     .0240898    2.594462
Homozygous   |   .7090909   .8136176    -0.30   0.764     .0748213    6.720144
------------------------------------------------------------------------------

snp6564946
                                                Number of obs   =         74
                                                LR chi2(2)      =       2.90
                                                Prob > chi2     =     0.2342
Log likelihood =   -41.7294                     Pseudo R2       =     0.0336
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |       .275   .3272277    -1.08   0.278     .0266976    2.832654
Homozygous   |   .6909091   .7929587    -0.32   0.747     .0728608    6.551606
------------------------------------------------------------------------------

snp6550610
                                                Number of obs   =         73
                                                LR chi2(2)      =       3.71
                                                Prob > chi2     =     0.1565
Log likelihood = -41.008741                     Pseudo R2       =     0.0433
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   3.690476   2.716406     1.77   0.076     .8720642    15.61767
```

```
Heterozygous |   1.666667   1.240706     0.69   0.493     .3874308    7.169739
------------------------------------------------------------------------------

snp6550514
                                                Number of obs   =         60
                                                LR chi2(2)      =       1.43
                                                Prob > chi2     =     0.4882
Log likelihood = -37.473799                     Pseudo R2       =     0.0188
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   2.285714   1.778393     1.06   0.288     .4974474     10.5026
Heterozygous |   1.214286   .8600715     0.27   0.784     .3029841    4.866559
------------------------------------------------------------------------------

snp6550471
                                                Number of obs   =         59
                                                LR chi2(2)      =       3.88
                                                Prob > chi2     =     0.1440
Log likelihood = -35.137203                     Pseudo R2       =     0.0523
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |        2.4   1.786617     1.18   0.240     .5579003    10.32442
Homozygous   |       4.56   3.588515     1.93   0.054     .9752311    21.32172
------------------------------------------------------------------------------

snp6549905
                                                Number of obs   =         68
                                                LR chi2(2)      =       1.32
                                                Prob > chi2     =     0.5177
Log likelihood = -39.624644                     Pseudo R2       =     0.0163
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |     1.6875   1.119714     0.79   0.430      .459666    6.195056
Homozygous   |      2.025   1.326594     1.08   0.281     .5607845    7.312301
------------------------------------------------------------------------------

snp6549803
                                                Number of obs   =         50
                                                LR chi2(2)      =       2.70
                                                Prob > chi2     =     0.2587
Log likelihood = -31.851158                     Pseudo R2       =     0.0407
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   1.647059   1.170747     0.70   0.483     .4089458    6.633648
Homozygous   |   4.941176   5.627127     1.40   0.161     .5302273    46.04671
------------------------------------------------------------------------------

snp6549696
                                                Number of obs   =         67
                                                LR chi2(2)      =       1.87
                                                Prob > chi2     =     0.3931
Log likelihood = -39.018723                     Pseudo R2       =     0.0234
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   2.363636   1.522822     1.34   0.182     .6686153    8.355741
```

```
Homozygous   |        1.4   1.014083     0.46   0.642     .3385047    5.790171
------------------------------------------------------------------------------

snp6549675
                                                Number of obs   =         67
                                                LR chi2(2)      =       2.10
                                                Prob > chi2     =     0.3507
Log likelihood = -38.904552                     Pseudo R2       =     0.0262
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   2.461538   1.576646     1.41   0.160     .7014574    8.637975
Homozygous   |   1.353846   .9504272     0.43   0.666     .3419863    5.359569
------------------------------------------------------------------------------

snp6548520
                                                Number of obs   =         61
                                                LR chi2(2)      =       3.73
                                                Prob > chi2     =     0.1548
Log likelihood = -35.971199                     Pseudo R2       =     0.0493
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   .8231293     .50224    -0.32   0.750     .2489428    2.721677
Homozygous   |   5.238095   5.829428     1.49   0.137     .5913968    46.39464
------------------------------------------------------------------------------

snp6548430
                                                Number of obs   =         57
                                                LR chi2(2)      =       3.58
                                                Prob > chi2     =     0.1667
Log likelihood = -34.489559                     Pseudo R2       =     0.0494
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   1.388889   .9328445     0.49   0.625     .3723571    5.180544
Homozygous   |         .3   .2738613    -1.32   0.187     .0501284    1.795388
------------------------------------------------------------------------------
```

APPENDIX B

CASE-CONTROL STATISTICAL ANALYSES OF EQUINE *RNASEL* POLYMORPHISMS

Allelic Fisher's Exact Tests

snp2758810

	Allele C	G	Total	
Control	9	17	26	Fisher's exact = 0.444
Caes	2	10	12	1-sided Fisher's exact = 0.231
Total	11	27	38	

snp2756586

	Allele A	C	Total	
Control	26	12	38	Fisher's exact = 0.127
Case	25	25	50	1-sided Fisher's exact = 0.064
Total	51	37	88	

snp2756461

	Allele A	G	Total	
Control	11	27	38	Fisher's exact = 0.139
Case	11	57	68	1-sided Fisher's exact = 0.097
Total	22	84	106	

snp2756452

	Allele C	T	Total	
Control	23	15	38	Fisher's exact = 0.835
Case	45	25	70	1-sided Fisher's exact = 0.427
Total	68	40	108	

snp2756423

	Allele T	C	Total	
Control	11	27	38	Fisher's exact = 0.128
Case	10	56	66	1-sided Fisher's exact = 0.077
Total	21	83	104	

snp2756422

	Allele		
	T	G	Total
Control	28	10	38
Case	47	21	68
Total	75	31	106

Fisher's exact = 0.663
1-sided Fisher's exact = 0.396

snp2756421

	Allele		
	A	C	Total
Control	25	13	38
Case	52	20	72
Total	77	33	110

Fisher's exact = 0.517
1-sided Fisher's exact = 0.313

snp2756325

	Allele		
	A	C	Total
Control	11	27	38
Case	16	52	68
Total	27	79	106

Fisher's exact = 0.643
1-sided Fisher's exact = 0.348

snp2756127

	Allele		
	A	C	Total
Control	30	8	38
Case	61	21	82
Total	91	29	120

Fisher's exact = 0.653
1-sided Fisher's exact = 0.382

snp2756111

	Allele		
	A	G	Total
Control	20	18	38
Case	34	48	82
Total	54	66	120

Fisher's exact = 0.324
1-sided Fisher's exact = 0.172

snp2756069

	Allele		
	G	T	Total
Control	29	9	38
Case	18	6	24
Total	47	15	62

Fisher's exact = 1.000
1-sided Fisher's exact = 0.569

snp2756056

	Allele A	G	Total	
Control	8	30	38	Fisher's exact = 0.762
Case	6	18	24	1-sided Fisher's exact = 0.475
Total	14	48	62	

snp2756043

	Allele A	G	Total	
Control	9	29	38	Fisher's exact = 1.000
Case	6	18	24	1-sided Fisher's exact = 0.569
Total	15	47	62	

snp2756001

	Allele A	C	Total	
Control	22	16	38	Fisher's exact = 0.798
Case	13	11	24	1-sided Fisher's exact = 0.489
Total	35	27	62	

snp2755808

	Allele A	G	Total	
Control	6	30	36	Fisher's exact = 0.475
Case	19	63	82	1-sided Fisher's exact = 0.296
Total	25	93	118	

snp2755763

	Allele C	G	Total	
Control	7	29	36	Fisher's exact = 0.132
Case	29	57	86	1-sided Fisher's exact = 0.085
Total	36	86	122	

snp2755672

	Allele A	G	Total	
Control	27	9	36	Fisher's exact = 0.285
Case	49	29	78	1-sided Fisher's exact = 0.142
Total	76	38	114	

snp2755299

	Allele		
	A	G	Total
Control	1	35	36
Case	0	70	70
Total	1	105	106

Fisher's exact = 0.340
1-sided Fisher's exact = 0.340

snp2755162

	Allele		
	A	G	Total
Control	1	35	36
Case	0	64	64
Total	1	99	100

Fisher's exact = 0.360
1-sided Fisher's exact = 0.360

snp2755142

	Allele		
	A	C	Total
Control	1	35	36
Case	1	63	64
Total	2	98	100

Fisher's exact = 1.000
1-sided Fisher's exact = 0.593

snp2755132

	Allele		
	C	T	Total
Control	1	35	36
Case	2	60	62
Total	3	95	98

Fisher's exact = 1.000
1-sided Fisher's exact = 0.696

snp2755071

	snp63_		
	G	A	Total
Control	28	8	36
Case	48	48	96
Total	76	56	132

Fisher's exact = 0.005
1-sided Fisher's exact = 0.003

snp2755039

	Allele		
	A	G	Total
Control	1	35	36
Case	1	95	96
Total	2	130	132

Fisher's exact = 0.473
1-sided Fisher's exact = 0.473

snp2750857

	Allele		
	C	T	Total
Control	8	32	40
Case	7	47	54
Total	15	79	94

Fisher's exact = 0.402
1-sided Fisher's exact = 0.261

snp2750736

	Allele		
	A	G	Total
Control	6	34	40
Case	6	38	44
Total	12	72	84

Fisher's exact = 1.000
1-sided Fisher's exact = 0.551

snp2750733

	Allele		
	C	G	Total
Control	39	1	40
Case	44	0	44
Total	83	1	84

Fisher's exact = 0.476
1-sided Fisher's exact = 0.476

snp2743998

	Allele		
	C	G	Total
Control	6	28	34
Case	9	59	68
Total	15	87	102

Fisher's exact = 0.564
1-sided Fisher's exact = 0.376

snp2743993

	Allele		
	T	A	Total
Control	6	22	28
Case	13	51	64
Total	19	73	92

Fisher's exact = 1.000
1-sided Fisher's exact = 0.554

snp2743789

	Allele		
	C	T	Total
Control	7	21	28
Case	15	77	92
Total	22	98	120

Fisher's exact = 0.402
1-sided Fisher's exact = 0.219

snp2743745

	Allele		
	T	C	Total
Control	20	8	28
Case	51	35	86
Total	71	43	114

Fisher's exact = 0.272
1-sided Fisher's exact = 0.178

snp2743078

	Allele		
	C	T	Total
0	11	25	36
1	10	20	30
Total	21	45	66

Fisher's exact = 1.000
1-sided Fisher's exact = 0.508

snp2742898

	Allele		
	C	T	Total
Control	29	5	34
Case	58	12	70
Total	87	17	104

Fisher's exact = 1.000
1-sided Fisher's exact = 0.496

snp2742846

	Allele		
	C	T	Total
Control	10	24	34
Case	26	46	72
Total	36	70	106

Fisher's exact = 0.661
1-sided Fisher's exact = 0.325

snp2742808

	Allele		
	C	T	Total
Control	10	24	34
Case	9	17	26
Total	19	41	60

Fisher's exact = 0.781
1-sided Fisher's exact = 0.439

snp2742764

	Allele		
	C	T	Total
Control	22	10	32
Case	16	8	24
Total	38	18	56

Fisher's exact = 1.000
1-sided Fisher's exact = 0.547

Genotypic Fisher's Exact Tests

snp2758810

	Genotype			
	CC	CG	GG	Total
Control	4	1	8	13
Case	1	0	5	6
Total	5	1	13	19

Fisher's exact = 0.736

snp2756586

	Genotype			
	AA	AC	CC	Total
Control	11	4	4	19
Case	8	9	8	25
Total	19	13	12	44

Fisher's exact = 0.252

snp2756461

	Genotype			
	AA	AG	GG	Total
Control	10	7	2	19
Case	24	9	1	34
Total	34	16	3	53

Fisher's exact = 0.271

snp2756452

	Genotype			
	CC	CT	TT	Total
Control	8	7	4	19
Case	18	9	8	35
Total	26	16	12	54

Fisher's exact = 0.706

snp2756423

	Genotype			
	CC	CT	TT	Total
Control	10	7	2	19
Case	23	10	0	33
Total	33	17	2	52

Fisher's exact = 0.126

snp2756422

	Genotype			
	TT	GT	GG	Total
Control	11	6	2	19
Case	19	9	6	34
Total	30	15	8	53

Fisher's exact = 0.852

snp2756421

	Genotype			
	AA	AC	CC	Total
Control	8	9	2	19
Case	20	12	4	36
Total	28	21	6	55

Fisher's exact = 0.562

snp2756325

	Genotype			
	AA	AC	CC	Total
Control	2	7	10	19
Case	4	8	22	34
Total	6	15	32	53

Fisher's exact = 0.545

snp2756127

	Genotype			
	AA	AC	CC	Total
Control	13	4	2	19
Case	27	7	7	41
Total	40	11	9	60

Fisher's exact = 0.843

snp2756111

	Genotype			
	AA	AG	GG	Total
Control	7	6	6	19
Case	11	12	18	41
Total	18	18	24	60

Fisher's exact = 0.688

snp2756069

	Genotype			
	GG	GT	TT	Total
Control	11	7	1	19
Case	7	4	1	12
Total	18	11	2	31

Fisher's exact = 1.000

snp2756056

	Genotype		
	AG	GG	Total
Control	8	11	19
Case	6	6	12
Total	14	17	31

Fisher's exact = 0.724
1-sided Fisher's exact = 0.475

snp2756043

	Genotype			
	AA	AG	GG	Total
Control	1	7	11	19
Case	1	4	7	12
Total	2	11	18	31

Fisher's exact = 1.000

snp2756001

	Genotype			
	AA	AC	CC	Total
Control	7	8	4	19
Case	5	3	4	12
Total	12	11	8	31

Fisher's exact = 0.652

snp2755808

	Genotype			
	AA	AG	GG	Total
Control	1	4	13	18
Case	5	9	27	41
Total	6	13	40	59

Fisher's exact = 0.907

snp2755763

	Genotype			
	CC	CG	GG	Total
Control	1	5	12	18
Case	7	15	21	43
Total	8	20	33	61

Fisher's exact = 0.421

snp2755672

	Genotype			
	AA	AG	GG	Total
Control	11	5	2	18
Case	18	13	8	39
Total	29	18	10	57

Fisher's exact = 0.597

snp2755299

	Genotype		
	AG	GG	Total
Control	1	17	18
Case	0	35	35
Total	1	52	53

Fisher's exact = 0.340
1-sided Fisher's exact = 0.340

snp2755162

	Genotype		
	AG	GG	Total
Control	1	17	18
Case	0	32	32
Total	1	49	50

Fisher's exact = 0.360
1-sided Fisher's exact = 0.360

snp2755142

	Genotype		
	AC	CC	Total
Control	1	17	18
Case	1	31	32
Total	2	48	50

Fisher's exact = 1.000
1-sided Fisher's exact = 0.595

snp2755132

	Genotype			
	CC	CT	TT	Total
Control	0	1	17	18
Case	1	0	30	31
Total	1	1	47	49

Fisher's exact = 0.605

snp2755071

	Genotype			
	AA	AG	GG	Total
Control	2	4	12	18
Case	16	16	16	48
Total	18	20	28	66

Fisher's exact = 0.049

snp2755039

	Genotype		
	AG	GG	Total
Control	1	17	18
Case	1	47	48
Total	2	64	66

Fisher's exact = 0.474
1-sided Fisher's exact = 0.474

snp2750857

	Genotype			
	CC	CT	TT	Total
Control	2	4	14	20
Case	2	3	22	27
Total	4	7	36	47

Fisher's exact = 0.681

snp2750736

	Genotype			
	AA	AG	GG	Total
Control	1	4	15	20
Case	0	6	16	22
Total	1	10	31	42

Fisher's exact = 0.723

snp2750733

	Genotype		
	CC	CT	Total
Control	19	1	20
Case	22	0	22
Total	41	1	42

Fisher's exact = 0.476
1-sided Fisher's exact = 0.476

snp2743998

	Genotype			
	CC	CG	GG	Total
Control	2	2	13	17
Case	2	5	27	34
Total	4	7	40	51

Fisher's exact = 0.754

snp2743993

	Genotype			
	AA	AT	TT	Total
Control	9	4	1	14
Case	21	9	2	32
Total	30	13	3	46

Fisher's exact = 1.000

snp2743789

	Genotype			
	CC	CT	TT	Total
Control	2	3	9	14
Case	4	7	35	46
Total	6	10	44	60

Fisher's exact = 0.681

snp2743745

	Genotype			
	CC	CT	TT	Total
Control	4	0	10	14
Case	12	11	20	43
Total	16	11	30	57

Fisher's exact = 0.078

snp2743078

	Genotype			
	CC	CT	TT	Total
Control	3	5	10	18
Case	2	6	7	15
Total	5	11	17	33

Fisher's exact = 0.892

snp2742898

	Genotype			
	CC	CT	TT	Total
Control	13	3	1	17
Case	24	10	1	35
Total	37	13	2	52

Fisher's exact = 0.554

snp2742846

	Genotype			
	CC	CT	TT	Total
Control	3	4	10	17
Case	9	8	19	36
Total	12	12	29	53

Fisher's exact = 0.924

snp2742808

	Genotype			
	CC	CT	TT	Total
Control	3	4	10	17
Case	2	5	6	13
Total	5	9	16	30

Fisher's exact = 0.789

snp2742764

	Genotype			
	CC	CT	TT	Total
Control	8	6	2	16
Case	6	4	2	12
Total	14	10	4	28

Fisher's exact = 1.000

Genotypic Odds Ratios

snp2758810

```
Logistic regression                               Number of obs   =         18
                                                  LR chi2(1)      =       0.59
                                                  Prob > chi2     =     0.4435
Log likelihood = -11.163632                       Pseudo R2       =     0.0256
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |        2.5   3.137474     0.73   0.465     .2136442    29.25425
------------------------------------------------------------------------------
```

snp2756586

```
Logistic regression                               Number of obs   =         44
                                                  LR chi2(2)      =       2.99
                                                  Prob > chi2     =     0.2245
Log likelihood = -28.594273                       Pseudo R2       =     0.0496
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |    3.09375    2.35007     1.49   0.137     .6980665    13.71114
Homozygous   |       2.75   2.113942     1.32   0.188     .6095493    12.40671
------------------------------------------------------------------------------
```

snp2756461

```
Logistic regression                               Number of obs   =         53
                                                  LR chi2(2)      =       2.23
                                                  Prob > chi2     =     0.3285
Log likelihood = -33.471685                       Pseudo R2       =     0.0322
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |        4.8   6.150122     1.22   0.221     .3895988    59.13775
Heterozygous |   2.571429   3.405535     0.71   0.476     .1918071    34.47341
------------------------------------------------------------------------------
```

snp2756452

```
Logistic regression                               Number of obs   =         54
                                                  LR chi2(2)      =       0.74
                                                  Prob > chi2     =     0.6893
Log likelihood = -34.651483                       Pseudo R2       =     0.0106
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   .5714286   .3766762    -0.85   0.396     .1569871    2.079984
Homozygous   |   .8888889   .6625383    -0.16   0.874     .2062527    3.830852
------------------------------------------------------------------------------
```

snp2756423

Logistic regression — Number of obs = 52; LR chi2(2) = 4.75; Prob > chi2 = 0.0930; Log likelihood = -31.759937; Pseudo R2 = 0.0696

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Homozygous	2.15e+08	.	.	.	.	.
Heterozygous	1.34e+08	8.31e+07	30.10	0.000	3.96e+07	4.52e+08

snp2756422

Logistic regression — Number of obs = 53; LR chi2(2) = 0.55; Prob > chi2 = 0.7586; Log likelihood = -34.308589; Pseudo R2 = 0.0080

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Heterozygous	.8684211	.5636838	-0.22	0.828	.2433456	3.099111
Homozygous	1.736842	1.563358	0.61	0.540	.2975644	10.13771

snp2756421

Logistic regression — Number of obs = 55; LR chi2(2) = 1.08; Prob > chi2 = 0.5824; Log likelihood = -34.911704; Pseudo R2 = 0.0152

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Heterozygous	.5333333	.3241704	-1.03	0.301	.1620399	1.755398
Homozygous	.8	.7694154	-0.23	0.817	.1214592	5.269257

snp2756325

Logistic regression — Number of obs = 53; LR chi2(2) = 1.05; Prob > chi2 = 0.5903; Log likelihood = -34.057699; Pseudo R2 = 0.0152

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Heterozygous	.5714286	.5765079	-0.55	0.579	.0791032	4.127905
Homozygous	1.1	1.040913	0.10	0.920	.1721528	7.028641

snp2756127

Logistic regression — Number of obs = 60; LR chi2(2) = 0.52; Prob > chi2 = 0.7718; Log likelihood = -37.200896; Pseudo R2 = 0.0069

	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Heterozygous	.8425926	.5998509	-0.24	0.810	.2087555	3.400927
Homozygous	1.685185	1.466031	0.60	0.549	.306297	9.271553

snp2756111

```
Logistic regression                               Number of obs   =         60
                                                  LR chi2(2)      =       0.96
                                                  Prob > chi2     =     0.6200
Log likelihood = -36.981771                       Pseudo R2       =     0.0128
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   1.272727   .8852239     0.35   0.729     .3256091    4.974783
Homozygous   |   1.909091    1.28915     0.96   0.338     .5082021    7.171612
------------------------------------------------------------------------------
```

snp2756069

```
Logistic regression                               Number of obs   =         31
                                                  LR chi2(2)      =       0.13
                                                  Prob > chi2     =     0.9368
Log likelihood = -20.625066                       Pseudo R2       =     0.0032
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   .8979592   .7108204    -0.14   0.892     .1903046    4.237053
Homozygous   |   1.571429   2.348624     0.30   0.762     .0839666    29.40916
------------------------------------------------------------------------------
```

snp2756056

```
Logistic regression                               Number of obs   =         31
                                                  LR chi2(1)      =       0.18
                                                  Prob > chi2     =     0.6672
Log likelihood = -20.597935                       Pseudo R2       =     0.0045
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   .7272727   .5389883    -0.43   0.667     .1701627    3.108353
------------------------------------------------------------------------------
```

snp2756043

```
Logistic regression                               Number of obs   =         31
                                                  LR chi2(2)      =       0.13
                                                  Prob > chi2     =     0.9368
Log likelihood = -20.625066                       Pseudo R2       =     0.0032
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   .5714286    .883935    -0.36   0.718     .0275581    11.84879
Homozygous   |   .6363636   .9510957    -0.30   0.762      .034003    11.90949
------------------------------------------------------------------------------
```

snp2756001

```
Logistic regression                               Number of obs   =         31
                                                  LR chi2(2)      =       1.10
                                                  Prob > chi2     =     0.5773
Log likelihood = -20.140975                       Pseudo R2       =     0.0266
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |       .525   .4699235    -0.72   0.472     .0908354    3.034334
Homozygous   |        1.4   1.285302     0.37   0.714     .2315599    8.464334
------------------------------------------------------------------------------
```

snp2755808

```
Logistic regression                               Number of obs   =         59
                                                  LR chi2(2)      =       0.68
                                                  Prob > chi2     =     0.7112
Log likelihood = -35.950751                       Pseudo R2       =     0.0094
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |        .45   .5622485    -0.64   0.523     .0388764    5.208812
Homozygous   |   .4153846    .476146    -0.77   0.443     .0439276    3.927928
------------------------------------------------------------------------------
```

snp2755763

```
Logistic regression                               Number of obs   =         61
                                                  LR chi2(2)      =       2.23
                                                  Prob > chi2     =     0.3285
Log likelihood = -35.891763                       Pseudo R2       =     0.0301
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   .4285714   .5088144    -0.71   0.475     .0418262    4.391353
Homozygous   |        .25   .2821579    -1.23   0.219     .0273682    2.283668
------------------------------------------------------------------------------
```

snp2755672

```
Logistic regression                               Number of obs   =         57
                                                  LR chi2(2)      =       1.32
                                                  Prob > chi2     =     0.5163
Log likelihood = -34.887224                       Pseudo R2       =     0.0186
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   1.588889   1.033862     0.71   0.477     .4438452    5.687946
Homozygous   |   2.444444    2.14703     1.02   0.309     .4370629    13.67151
------------------------------------------------------------------------------
```

snp2755142

```
Logistic regression                               Number of obs   =         50
                                                  LR chi2(1)      =       0.17
                                                  Prob > chi2     =     0.6798
Log likelihood = -32.585713                       Pseudo R2       =     0.0026
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   1.823529   2.636928     0.42   0.678     .1071585    31.03123
------------------------------------------------------------------------------
```

snp2755071

```
Logistic regression                               Number of obs   =         66
                                                  LR chi2(2)      =       6.53
                                                  Prob > chi2     =     0.0382
Log likelihood = -35.408453                       Pseudo R2       =     0.0844
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |          6   5.049752     2.13   0.033     1.152812    31.22799
Heterozygous |          3    2.03101     1.62   0.105     .7958937    11.30804
------------------------------------------------------------------------------
```

snp2755039

```
Logistic regression                               Number of obs   =         66
                                                  LR chi2(1)      =       0.48
                                                  Prob > chi2     =     0.4888
Log likelihood = -38.433248                       Pseudo R2       =     0.0062
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   2.764706   3.987411     0.71   0.481     .1636813    46.69806
------------------------------------------------------------------------------
```

snp2750857

```
Logistic regression                               Number of obs   =         47
                                                  LR chi2(2)      =       0.89
                                                  Prob > chi2     =     0.6409
Log likelihood = -31.609891                       Pseudo R2       =     0.0139
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |        .75   .9437292    -0.23   0.819      .063678    8.833501
Homozygous   |   1.571429   1.660728     0.43   0.669     .1980259    12.47003
------------------------------------------------------------------------------
```

snp2750736

```
Logistic regression                               Number of obs   =         42
                                                  LR chi2(2)      =       1.73
                                                  Prob > chi2     =     0.4219
Log likelihood = -28.201547                       Pseudo R2       =     0.0297
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   1.21e+08          .        .       .            .           .
Homozygous   |   8.61e+07   6.36e+07    24.73   0.000     2.02e+07    3.66e+08
------------------------------------------------------------------------------
```

snp2743998

```
Logistic regression                               Number of obs   =         51
                                                  LR chi2(2)      =       0.56
                                                  Prob > chi2     =     0.7569
Log likelihood = -32.183717                       Pseudo R2       =     0.0086
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |        2.5   3.259601     0.70   0.482     .1941373    32.19371
Homozygous   |   2.076923   2.192074     0.69   0.489     .2624373    16.43672
------------------------------------------------------------------------------
```

snp2743993

```
Logistic regression                               Number of obs   =         46
                                                  LR chi2(2)      =       0.02
                                                  Prob > chi2     =     0.9925
Log likelihood = -28.259615                       Pseudo R2       =     0.0003
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |   1.166667   1.502569     0.12   0.905     .0934696    14.56208
Heterozygous |      1.125   1.534753     0.09   0.931     .0776108    16.30734
------------------------------------------------------------------------------
```

snp2743789

```
Logistic regression                               Number of obs   =         60
                                                  LR chi2(2)      =       0.75
                                                  Prob > chi2     =     0.6863
Log likelihood = -32.219869                       Pseudo R2       =     0.0116
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   1.166667   1.291891     0.14   0.889     .1331624    10.22144
Homozygous   |   1.944444   1.834058     0.70   0.481     .3061377    12.35021
------------------------------------------------------------------------------
```

snp2743745

```
Logistic regression                               Number of obs   =         46
                                                  LR chi2(1)      =       0.35
                                                  Prob > chi2     =     0.5548
Log likelihood = -28.092787                       Pseudo R2       =     0.0062
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Homozygous   |        1.5   1.042833     0.58   0.560     .3839878    5.859562
------------------------------------------------------------------------------
```

snp2743078

```
Logistic regression                               Number of obs   =         33
                                                  LR chi2(2)      =       0.55
                                                  Prob > chi2     =     0.7590
Log likelihood = -22.461565                       Pseudo R2       =     0.0121
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |        1.8   1.971801     0.54   0.592     .2102945    15.40696
Homozygous   |       1.05   1.089266     0.05   0.962     .1374537    8.020885
------------------------------------------------------------------------------
```

snp2742898

```
Logistic regression                               Number of obs   =         52
                                                  LR chi2(2)      =       0.94
                                                  Prob > chi2     =     0.6265
Log likelihood = -32.395277                       Pseudo R2       =     0.0142
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   1.805556   1.341375     0.80   0.426     .4209594    7.744287
Homozygous   |   .5416667   .7884162    -0.42   0.674     .0312447    9.390484
------------------------------------------------------------------------------
```

snp2742846

```
Logistic regression                               Number of obs   =         53
                                                  LR chi2(2)      =       0.37
                                                  Prob > chi2     =     0.8298
Log likelihood = -33.067579                       Pseudo R2       =     0.0056
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   .6666667   .6034878    -0.45   0.654      .113077     3.93046
Homozygous   |   .6333333   .4893811    -0.59   0.554     .1392846    2.879796
------------------------------------------------------------------------------
```

snp2742808

```
Logistic regression                               Number of obs   =         30
                                                  LR chi2(2)      =       0.79
                                                  Prob > chi2     =     0.6742
Log likelihood = -20.132724                       Pseudo R2       =     0.0192
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |      1.875   2.124081     0.55   0.579     .2035746    17.26947
Homozygous   |         .9   .9439279    -0.10   0.920     .1152113    7.030563
------------------------------------------------------------------------------
```

snp2742764

```
Logistic regression                               Number of obs   =         28
                                                  LR chi2(2)      =       0.12
                                                  Prob > chi2     =     0.9436
Log likelihood = -19.063419                       Pseudo R2       =     0.0030
------------------------------------------------------------------------------
             | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Heterozygous |   .8888889   .7481115    -0.14   0.889     .1707869     4.62637
Homozygous   |   1.333333   1.515354     0.25   0.800      .143726    12.36922
------------------------------------------------------------------------------
```

VITA

Name: Jonathan Joseph Rios

Address: c/o Mrs. Julia Williams
Texas A&M University
2128 TAMU
College Station, Texas 77843-2128

Email Address: julia-williams@tamu.edu

Education: B.S., Animal Science, Tarleton State University, 2003
Ph.D., Genetics, Texas A&M University, 2008