Published in final edited form as:

Infect Genet Evol. 2016 October ; 44: 341–350. doi:10.1016/j.meegid.2016.07.005.

FUNCTIONAL EVOLUTION OF THE OAS1 VIRAL SENSOR: INSIGHTS FROM OLD WORLD PRIMATES

Ian Fish^{a,b} and Stéphane Boissinot^c

^aBiology Department, Queens College, the City University of New York, 65-30 Kissena Boulevard, Flushing, NY, 11367, USA

^bGraduate Center, Sub-program in Molecular, Cellular and Developmental Biology, the City University of New York, 365th avenue, New York, NY, 10016, USA

^cNew York University Abu Dhabi, P.O. 129188, Abu Dhabi, United Arab Emirates

Abstract

Infections with viral pathogens impose considerable selective pressure on host defensive genes. Those genes at the forefront, responsible for identifying and binding exogenous molecular viral components, will carry the hallmarks of this struggle. Oligoadenylate synthetase (OAS) enzymes play a major role in the innate defense against a large number of viruses by acting as sensors of viral infections. Following their up-regulation by the interferon pathway, OASs bind viral dsRNA and then signal ribonuclease L (RNase L) to degrade RNA, shutting down viral and host protein synthesis. We have investigated the evolution of OAS1 in twenty-two Old World monkey species. We identified a total of 35 codons with the earmarks of positive selection and we performed a comprehensive analysis of their functional significance using *in silico* modeling of the OAS1 protein. Subdividing OAS1 into functional domains revealed intense purifying selection in the active domain but significant positive directional selection in the RNA-binding domain (RBD), the region where OAS1 binds viral dsRNA. The modeling analysis revealed a concentration of rapidly evolving residues in one region of the RBD suggestive of the sub-functionalization of different regions of the RBD. This analysis also identified several positively selected residues circumscribing the entry to the active site suggesting adaptive evasion of viral antagonism and/or selection for production of oligoadenylate of different length.

Keywords

OAS1; positive selection; cercopithecidae; innate immunity; dsRNA sensor; immune system evolution

Corresponding author: Stéphane Boissinot, NYU Abu Dhabi, PO Box 129188, Abu Dhabi, UAE, sb5272@nyu.edu, Phone: (+971) 2 628 4790.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. INTRODUCTION

Infection by pathogens constitutes one of the most important selective pressures organisms are subjected to. The impact viral pathogens exert on their host is attested by the rapid evolution of immunity genes, particularly at sites directly interacting with pathogens (Daugherty and Malik 2012; Quintana-Murci, et al. 2007; Rausell and Telenti 2014). This pattern is best explained by a recurrent co-evolutionary process between rapidly evolving pathogens evading and antagonizing host defenses and the host immunity genes adapting in response. Since selection for adaptive changes is most likely to affect regions of immunoproteins that interact directly with pathogens, evolutionary studies can be of great help to better understand the function of immunoproteins and the mechanisms of resistance to infection.

In recent years, a number of functional and evolutionary studies have emphasized the significance of oligoadenylate synthetases (OASs) as key players in the innate defense of vertebrates against viral pathogens (Drappier and Michiels 2015; Silverman 2007). Shortly after viral infection, systemic interferon signaling upregulates genes of the OAS-RNase L pathway. Within the cytosol, an OAS enzyme will bind viral dsRNA by way of a large electropositive furrow, the RNA-binding domain (RBD). This leads to active reconfiguration of the catalytic site (Donovan, et al. 2013; Hartmann, et al. 2003). Oligoadenylate synthesis requires 2'-to-5' phosphodiester linkage of donor ATP to either an acceptor ATP or an acceptor oligoadenylate (Hartmann, et al. 2003). The oligoadenylate products are the ligands needed for RNase L to dimerize into its active form (Han, et al. 2012; Huang, et al. 2014). Active RNase L subsequently degrades viral and host RNA, leading to inhibition of protein synthesis and restriction of viral replication.

In primates, including human, the OAS family is comprised of OAS1, OAS2, OAS3, and OASL. OAS1 is by far the most studied member of this family. Association studies and experimental analyses in animal models have shown that OAS1 plays an important role in resistance to flaviviruses such as West Nile virus and dengue virus (Alagarasu, et al. 2013; Bonnevie-Nielsen, et al. 2005; Lim, et al. 2009; Mashimo, et al. 2002). Variation at the OAS1 locus has also been implicated in susceptibility to enteroviruses, respiratory syncytial virus and others (Behera, et al. 2002; Cai, et al. 2014).

Genes of the OAS family have been shown to harbor the signatures of positive selection, i.e. selection in favor of amino acid changes. Evolutionary analyses across Primates and Chiroptera have reported accelerated evolution of OAS genes, particularly for *OAS1* (Hancks, et al. 2015; Mozzi, et al. 2015). Moreover, house mouse, chimpanzee, and macaque populations have been shown to harbor an extremely high level of polymorphism at *OAS1* denoting the action of balancing selection at this locus (Ferguson, et al. 2012; Ferguson, et al. 2008; Fish and Boissinot 2015).

The relationship between OAS1's defensive role and its dynamic evolutionary history remains unknown. To bridge this gap in knowledge, we need to identify the functional advantages conferred by positively selected amino acids in OAS1. Here, we focus on the family *Cercopithecidae*, the Old World monkeys (OWM), to help uncover the functional

mechanisms underlying OAS1 evolution. Studying this group of primates is particularly relevant to human health since several species are used as animal models in infectious disease research and others are known reservoirs for viruses of high public health significance. We performed a detailed characterization of positively selected sites in primates using a combination of evolutionary tests of selection and *in silico* analysis of structural models.

2. MATERIALS AND METHODS

2.1. Genomic Samples and Sequencing

For fifteen primates species, full *OAS1* coding sequences were obtained by amplification of each of the 6 exons using intron-targeted primers. The DNAs were all from captive animals, either from zoos or from the Southwest National Primate Research Center (for the baboons and macaques). The *OAS1* exonic sequences of four additional species were obtained from GenBank (indicated with ^G on figure 1). We also included three sequences previously analyzed by Hancks et al. (2015) (indicated with ^H on figure 1). Amplicons were directly sequenced in both directions at the University of Washington Seattle High Throughput Genomics Unit. Read assembly and alignment as well as heterozygote calls were made using Geneious Pro and confirmed by chromatogram observation (www.geneious.com, Kearse, et al. (2012)). Alignments corresponding to the two most common isoforms inferred from known primate cDNA sequences were generated. Isoform 5i2 encompasses 364 amino acids and includes the longest readthrough of exons one through five. Isoform 6i1 contains a shorter fifth exon and extends through exon six to a total length of 400 amino acids. All evolutionary analyses were performed for both isoforms, yet modeling and structural analysis is limited to the shorter 5i2 protein because *OAS1* structure has only been solved to hOAS1 residue 346 (porcine *OAS1* residue 349). Sequences are deposited in Genbank under accession number KX001957-KX001966.

2.2. Molecular Evolutionary Analyses

Since *OAS1* belong to a multi-genic family, we first investigated the possibility of inter-locus recombination since this would affect phylogenetic reconstruction as well as detection of selection. To this end we searched for recombination using the dual multiple change-point model (Minin, et al. 2005), with the DualBrother program implemented in Geneious Pro. We used a window length of 100 and step size of 10. We ran the MCMC chain for 1,000,000 iterations including 10,000 as burn-in. Once inter-locus recombination was excluded, we calculated the ratios of non-synonymous to synonymous substitution dN/dS, which is an estimator of the type of selection acting on a gene. A value of dN/dS <1 indicates that mutations coding for amino acid replacements are removed by negative or purifying selection faster than synonymous, putatively neutral, mutations. A value >1 indicates that non-synonymous changes reach fixation at a faster rate than synonymous mutations and is indicative of positive selection. A value of 1 indicates that mutations at non-synonymous and synonymous sites are of the same fitness, a case of neutral evolution. dN/dS values were calculated using Selecton (<http://selecton.tau.ac.il/index.html>, Doron-Faigenboim, et al. (2005); Stern, et al. (2007)) or SLAC (Kosakovsky Pond and Frost 2005). For domain-based

dN/dS calculations, domains were defined based on the most recent OAS1 crystal structure analyses by Donovan, et al. (2013) and Lohofener, et al. (2015).

Because any region of a protein contains a mixture of neutral, positively and negatively selected sites, we used several methods that detect specific sites under positive selection. We used several programs included in the HyPhy suite (Pond, et al. 2005), and implemented in datamonkey.org (Delpont, et al. 2010a): random effects likelihood REL (Kosakovsky Pond and Frost 2005), mixed effects model evolution MEME (Murrell, et al. 2012), and TOGGLE (Delpont, et al. 2008). REL models variation of selective pressures at each individual codon and compares this to the observed variation. MEME differs from REL primarily in that it allows for different selective pressure among lineages at the same site, which can help identify a history of episodic or species-specific selection acting on a codon. MEME can be particularly conservative if synonymous divergences are low, which is the case here, but is valuable as it can highlight sites that may only be selected in few branches. A subset of selected mutations that can be difficult to detect are those that alternate between a single conserved residue and a distinct mutation. The program TOGGLE tests for this by modeling each sense codon followed by non-synonymous mutation and then reversion mutation. Posterior probabilities for REL and MEME are calculated by comparing Bayes factors between positively selected site models and negatively selected (null) site models. For all applicable analyses, we used the HKY85 model of mutation (Hasegawa, et al. 1985), found to best fit the observed codon evolution as determined by the codon model selection tool CMS (Delpont, et al. 2010b).

A complementary set of selection tests provided by the Selecton server incorporates a user-provided phylogeny (Stern, et al. 2007). We used the generally accepted primate phylogenetic tree based on Perelman, et al. (2011). We ran three models: the M8 positive selection model, which restricts dN/dS sample values to either >1 or <1 for each site; the M5 model, for which we used eight grades (bins) distributed across negative, neutral and positive selective forces; and the mechanistic empirical combination (MEC) model (Doron-Faigenboim and Pupko 2007), which takes into account the severity and probability of each residue change, encoded in the WAG matrix (Whelan and Goldman 2001), for which we also used an eight-bin distribution. M8 and MEC are statistically validated by comparison to the M8a null model (M8 via likelihood ratio test and MEC via Akaike information content scores). The dN/dS estimates are given confidence interval scores; for positively selected sites, we only included sites with $>95\%$ prior probability. In order to infer ancestral codons, we implemented the ancestral sequence reconstruction (ASR, Pupko, et al. (2000); Yang, et al. (1995)) program, which uses maximum likelihood evolution reconstructions. Gene trees were built using the maximum likelihood method with the HKY model of evolution using the MEGA 5.0 program (Tamura, et al. 2011). The robustness of the nodes was estimated using 1,000 bootstrap replicates.

2.3. Protein structure analysis

For enzyme structure modeling and analysis, the programs PolyPhen-2 (Adzhubei, et al. 2010), SWISS-MODEL (swissmodel.expasy.org, Arnold, et al. (2006); Bordoli, et al. (2009)), and PyMOL v. 1.7 (pymol.org) were used as published previously (Fish and

Boissinot 2015). Energy minimization for models was run using YASARA (Krieger, et al. 2009). SWISS-MODEL and YASARA-minimized structures are highly congruous, this lending to the overall accuracy of our predictions. Concordant with the very high (>90%) pairwise identity to hOAS1 across all OWM, model quality scores determined using SWISS-MODEL Global Model Quality Estimation (all scores as well as ProSA-web (Sippl 1993; Wiederstein and Sippl 2007) are of the same or very close to those of the human crystal model (4IG8).

In addition to hOAS1 (PDB ID:4IG8; Donovan, et al. (2013)), the most recent porcine crystal structures (PDB IDs: 4RWN–Q; Lohofener, et al. (2015)) were also used for active site visualization because this structure set includes the holo, pre-reactive, and pseudo-reactive states allowing for simulation of active enzyme reconfiguration and acceptor ligand visualization. Docked dsRNA appearing in figures is comprised of the 18-base 4IG8 chain B and chain C RNA molecules; RNA nucleoside enumeration is in concordance with 4IG8 as well.

3. RESULTS

3.1. Variation and phylogenetic analysis

We generated a dataset of 15 novel OAS1 sequences and combined them with 7 publicly available sequences, thus the dataset analyzed here is comprised of 22 sequences representing 12 OWM genera (figure 1). The number of heterozygote positions varies substantially among species, from 0 in patas monkey and sooty mangabey to 9 in proboscis monkey. The Golden snub-nosed monkey, patas monkey, Wolf's monkey, red-tailed guenon and stump-tail macaque OAS1 sequences contained in-frame insertions of 1, 2 or 3 codons. Wolf's monkey has a 25 codons insertion in exon 1 and the patas' contains a deletion consisting of an internal 35-codon in-frame loss within exon six, which maintains carboxy terminal –CTIL, a conserved motif implicated in posttranslational modification and localization (Maurer-Stroh and Eisenhaber 2005).

Incomplete lineage sorting, introgression, or recombination (within or between loci) can significantly affect the outcome of tests of selection. These concerns are particularly pertinent for OAS1 due to known examples of balancing selection (Ferguson, et al. 2012; Fish and Boissinot 2015), introgression (Mendez, et al. 2012, 2013) and the presence of adjacent OAS paralogs (Perelygin, et al. 2006). To rule out these contingencies, we first generated a gene tree using the phased alleles for the nineteen OAS1 coding sequences (figure 2). The phylogeny for OAS1 is generally consistent with the accepted primate phylogeny (Perelman, et al. 2011). The monophyly of the colobinae and cercopithecinae subfamilies was recovered with strong statistical confidence (100% bootstrap). Within the colobinae, the monophyly of the African and Asian species was also well supported. Within the cercopithecinae, the tribes papionini and cercopithecini are forming poorly supported clades and within papionini we found substantial mixing among genera, the stump-tailed macaque alleles branching with the magabeys and drill and the hamadryas baboon alleles branching with the macaques. However the number of informative sites within this group is very small and none of the nodes are strongly supported. Thus, this phylogenetic analysis does not support an impact of long-term balancing selection or introgression on the

evolution of OAS1 in OWM. Note that running selection tests with and without the stump-tailed macaque or hamadryas baboon yielded exactly the same set of positively selected sites. We also examined the possibility of recombination between OAS1 and its highly divergent paralogs (OASL, OAS2 and OAS3) using the DualBrother program and we did not find any evidence for inter-locus recombination. Finally, BLAST searches of available primate genomes elicited no evidence of additional duplications of OAS1, as reported in rodents and cetartiodactyls (Perelygin, et al. 2006).

3.2. Tests of Positive Selection

We first calculated the ratio dN/dS, between each species and its 18 orthologous OAS1 coding sequences (figure 3). The values of dN/dS are elevated, with the vast majority being over 0.5 and 47% of the values being higher than 1, which is remarkable for a protein-coding gene. The average dN/dS across all species is 1.08, with species' averages ranging from 0.86 to 1.55. The OWM average dN/dS being 1.08 implies that overall, cercopithecoid OAS1 is either evolving neutrally or that it contains some sites evolving under purifying selection (dN/dS<1) while others are evolving under positive selection (dN/dS>1). Domain-level dN/dS calculation was next performed to resolve this question.

We divided each OWM OAS1 coding sequence into active domain, RNA-binding domain, and uncategorized regions and calculated dN/dS for each. The average dN/dS value of the active domain is 0.35, well under a neutral value of 1.0, which is consistent with a history of strong purifying selection. The uncategorized regions are comprised of a blend of purely structural sites and those yet uncharacterized (presumably a mixture of sites under varying degrees of selective pressure) with an average dN/dS of 0.68. In contrast, the average dN/dS in the RNA-binding domain is well over 1 (1.58) and supports a history of positive selection, focused on the RBD.

Lastly, we ran tests of selection at the codon level, the results of which are summarized in figure 1. An abundance of sites were identified as evolving under positive selection: 31 in isoform 5i2 (8.1%) and 32 in isoform 6i1 (7.9%). Half of the selected sites shared between isoforms (amino acids 1–346) fall within the RBD (figure 4), which is noteworthy, since this region comprises only ~21% of the total coding sequence. Figure 1 also shows the results of the MEME and MEC analyses, highlighting those sites with strong signature of episodic positive selection and those corresponding to drastic amino acid substitutions, respectively. These particular sets of positively selected sites are also overrepresented at the RBD. The subsequent analysis proceeds with those sites determined to have evolved under positive selection by Selecton M8 and REL (and others, as indicated in figure 1).

3.3. Protein Model Analysis: RNA-Binding Domain

In order to compare the structure of OAS1 orthologous proteins and to visualize the effects of amino acid changes on structure, we created *in silico* protein models for each species with SWISS-MODEL. With few exceptions (later discussed), orthologs did not show evidence of compromised secondary or higher order structure nor did they evoke significant electrostatic surface changes. Given the high sequence similarity to solved human OAS1 (~92%) and high model quality scores, we investigated each positively selected site of interest. Our

initial focus is on the RBD because of the overrepresentation of selected sites identified there and because it is the physical location of virus sensing.

Solved OAS1:RNA structures exhibit the docked A-form dsRNA cradled along the length of the long electropositive channel spanning two minor grooves (figure 5). The RNA sequence that optimally activates human OAS1 is the consensus sequence $N_3N_4W_5W_6N_{7-15}W_{16}G_{17}N_{18}$ (Kodym, et al. (2009); $W=A/U$). The permutation of this crystallized with hAOS1, GGCUUUUGACCUUUAUGC and its complement, is the nucleic acid shown in all figures. This sequence docks and activates porcine OAS1 in a homologous manner, though activation is several degrees of magnitude weaker (Kodym, et al. 2009; Lohofener, et al. 2015). Because of this homology seen between pig and human OAS1, we can be confident about dsRNA orientation in docking and the equivalence of the surface residues responsible for bonding. The global topology and predominant electropositive character described in the solved structure is conserved in all Old World monkey ortholog models.

Fourteen sites under positive selection in OWM lie on the surface RBD and seven of these contribute direct bonds to RNA (figures 4 and 5). The majority of these fall to one side of the RBD, the portion of the RBD associated with minor groove 1 (RBD_{mg1}). RBD_{mg1} includes seven positively selected sites in the region of minor groove 1 (figure 5, colored yellow). Five of these (24, 28, 200, 247 and 248) have the potential to directly bind RNA. In contrast, only selected residues 54 and 196 are similarly capable of binding RNA in the protein region associated with RNA minor groove 2 (RBD_{mg2}) (figure 5, in green). Residues 161 and 31 map to RBD_{mg2} but cannot plausibly form direct bonds with docked RNA. We next analyzed each of the positively selected residues within the RBD with a focus on their predicted and potential influence on RNA docking. For brevity, we present only the analysis of those sites most likely to have a substantial impact on function.

The residue that exhibits the most variation across primates is residue 54, which has been shown to be under positive selection in human, chimpanzee, and gorilla as well as under balancing selection in chimpanzee and macaque (Ferguson, et al. 2012; Fish and Boissinot 2015; Mozzi, et al. 2015). In OWM, ancestral Arg predominates followed in commonality by Cys and lastly His and Tyr. The latter two bulky side groups conflict with the backbone at terminal C_{18} . This suggests that site 54 could play a decisive role in determining activator length. Its importance should also be stressed, as its succeeding residues (55–68), are indelibly conserved throughout mammals, directly binding RNA and performing an integral role in conformational change.

An interesting group of positively selected sites are those located on helix N3. Upon RNA binding and reconfiguration, helix N3 residues 24, 25, 28 and 31 shift considerably, between 5.6 and 8.4 Å, including a ~20° rotation around the helix. Selected residue Arg/His25 lends its electrostatic profile to the apo state protein aiding in initial docking but becomes oriented ~90° away from the activator in the holoenzyme. Therefore, the derived His substitution may only have indirect effects on binding via its neighboring residues. Ancestral residues Thr24 and Asn31 form hydrogen bonds with the RNA spine in the majority of OWM models; common electropositive replacements Lys/Arg24 and Lys/His31 are predicted to do the

same, but also appreciably shift local electrostatics. Both replacements at residue 24 receive *probably damaging* (0.99) ratings from PolyPhen-2, yet we do not suspect this to be true as our models show no perturbation and, furthermore, Lys and Arg are common at this position in chimpanzee and other modern primate populations (Ferguson, et al. 2012; Fish and Boissinot 2015). These long side groups are equally capable of interaction with ribose or phosphate groups, though with a minimum distance of 5.5 Å from Lys/Arg24 to the nearest nitrogenous base (U₆), formation of a direct base-sensing bond is unlikely.

Another region of interest corresponds to the loop between helices C5 and C6, which has been evolving remarkably quickly in OWM (a.a. 242, 245, 247, 248 and 253). Among those sites, residues 247 and 248 are particularly interesting since they have the ability to bind RNA. Ancestral Thr247 donates a hydrogen bond to the phosphate spine at A₆. Derived Asn can follow suit or easily hydrogen bond ribose. The derived Arg247, present in seven species, overstretchs the spine, and is easily able to hydrogen bond with the sugar or nitrogenous base (at 2.6 Å from C8 of A₆). Arg here may favor a pyrimidine hydrogen bond acceptor over a purine. Neighboring residue Asp248, located at the end of minor groove 1, is interesting because a BLAST search indicates that, in mammals, it is most often encoded as a negatively charged R group yet it essentially abuts the phosphate spine. The one replacement in OWM is Tyr, which could in principle donate its hydroxy hydrogen to a backbone phosphate, but models predict more likely a steric conflict. As was the case for residue 54, this location at the margin of the RBD, may indicate a role for 248 in governing RNA size preference. The only residue in this region for which PolyPhen-2 predicts problematic mutations is 247. Here, Arg is deemed *possibly damaging* (0.66) and Gly (which is unique to Tanzanian colobus) *probably damaging* (0.96), since it entails complete loss of ancestral binding capability.

3.4. Active Site Entry Evolution

Five positively selected amino acids circumscribe the entrance to the active site: 126, 131, 181, 313, and 314 (figure 6). Residue 131 sits quite near (4.9 Å) to Mg²⁺-coordinating Asp148 while the other four are relatively distant from the catalytic core. Although the orientation and binding side groups of oligoadenylate acceptors is unknown (only ‘monoadenylate’ ATP is), a longer acceptor substrate seated in its catalytic binding site would almost certainly come into contact with at least one of these sites under positive selection. These selected sites may also represent key sites of exogenous antagonism.

Between strands β3 and β4, the loop containing residues 126 and 131 exhibits a highly dynamic structure, changing conformation significantly between each enzyme complex stage (apo > RNA-docking > donor binding > acceptor binding). Amino acid 126 is predicted to relocate 12.6 Å (19.3 Å as measured from carbon 6), positioned above the modeled acceptor substrate (ATP) within 2.2 Å of its beta phosphate group (figure 7). Here, ancestral Trp appears to interfere with reactant entrance and product exit. Interestingly though, the residue is predicted to withdraw upon product disengagement. In this way, OAS1 alleles featuring Trp126 show a preference for ATP as acceptor because longer acceptors would easily conflict with the bulky side group in the reactive state. Arg126 replacement found in three OWM species relocates essentially the same way, but the electrostatics and

pliancy of the Arg group along with its increased distance from acceptor (4.5 Å) would provide for an increased affinity for longer acceptors (oligos, instead of solely ATP).

Positively selected site 181 occurs predominantly as Gln in OWM, but is also represented with Arg or Lys. Residue 181 caps helix C1, positioned above the acceptor seat. Arginine replacement would unlikely interact with a monomeric or dimeric acceptor upon docking, but could influence binding of longer substrates. More intriguing at this helical corner is the density of mutated sites in line with 181: 166, 172 (variable, but not necessarily under positive selection), 175, and 179. These variable residues run along the length of helix C1 from one side of the enzyme to the opposite side (figure 8). This helix has not been characterized to have particular non-structural function, so this array of selected sites is more suggestive of escape mutations that maintain structure while averting allosteric antagonism.

Sites 313 and 314, located in helix C8, are among the most variable in OWM, each being occupied by four different amino acids (Pro, Thr, Ser, Arg at position 313 and Lys, Met, Thr, Asn at position 314). Modeled replacements of the ancestral Pro313 and Lys314 do not indicate structural change in any models (except for golden snubnose, discussed below). Mutations at these sites would not interfere with ATP or a diadenylate acceptor, but may influence longer substrate use. The most remarkable observation here might be the diversity represented by the side groups. They each represent unique mutations from the ancestral state, yet without evidence of replacement constraint or preference. This suggests that these mutations could be escape mutations in which the strongest fitness advantage of the new amino acid is to be different from the ancestral state in order to best counteract antagonist binding.

4. DISCUSSION

4.1 Positive selection at OAS1

Evolutionary analysis of the OAS1 gene in 22 OWMs reveals the signature of positive selection at 35 amino acid sites (~8% of total) with an overabundance of these sites falling within the region of the virus-sensing interface, the RBD. These data corroborate recent findings of elevated positive selection in OAS1 across Primates (Hancks, et al. 2015; Mozzi, et al. 2015). There is considerable overlap between the positively selected sites we detected and those identified by other authors. However, we identified 15 additional sites, mostly because we included 22 OWM species in our study *versus* five in Mozzi et al. (2015) and ten in Hancks et al. (2015). However, we failed to recover a substantial number of positively selected sites identified in these two studies because their phylogenetic scope was broader than ours, including hominidae, New World monkeys and some prosimians.

Comparative protein modeling reveals new patterns in OAS1 evolution including a concentration of rapid evolution in one subdomain of the RBD, minor groove 1, as well as a cluster of selected sites around the active site entrance. The rapid evolution observed in these regions is highly suggestive of a history of virus-mediated pressures on the gene. Mutations granting an advantage in efficient virus sensing and downstream signaling could have led to rapid amino acid replacement throughout Old World monkeys at the RBD, whereas changes at the opening to the active site could have been in response to antagonism by viral products

for optimal regulation of oligomer synthesis. Overall, this analysis strongly indicates that the primary driving force behind the evolution of OAS1 is viral in nature and that this force is relatively constant over time and across diverse primate lineages. There are many shared sites under selective pressure, most likely due to shared pathogens and mechanisms of evasion, yet there is also heterogeneity representing the differences between each species' multi-viral landscape.

It should be noted that our interpretation of OAS1 evolution relies on a partial understanding of OAS1 structure and function. It has been known for a long time that OAS1 forms homotetramers (Ghosh, et al. 1997) but only the structure of the monomer has been resolved so far. It has been proposed that the *PCFK* motif (a.a. 330–333), in particular, and the planar surface encoded by the C-terminus, in general, likely play important roles in OAS tetramerization (Ghosh, et al. 1997; Torshin 2005). Since OAS1 isoforms differ at the C-terminus, it was predicted that different isoforms could form different isomeric structures with distinct enzymatic activity (Torshin 2005). It is noteworthy that several positively selected sites are indeed located at the C terminus of OAS1, including one within the PCFK motif (a.a. 333) and seven in alternatively spliced regions of *OAS1* (figure 1). The functional significance of positive selection on the C-terminus is unclear at this point and will require a better understanding of the tetramerization of OAS1 isoforms.

We noticed that many of the positively selected changes were identical to polymorphisms previously identified in macaque and Central African chimpanzee (figure 1). Within the first five exons, 14 of the 28 positively selected loci are polymorphic in either macaque or chimpanzee (Ferguson, et al. 2012; Fish and Boissinot 2015). Interestingly, five of these loci are polymorphic in both of these groups *and* the variations observed are identical. For example, OAS1 residue 126 is found with allelic variants Trp/Arg in macaque and chimp while the same residue occurs exclusively as Trp/Arg across OWM. Additionally, when looking at the RBD in particular, 10 of the 14 positively selected sites (in OWM) are balanced polymorphisms in macaque and chimpanzee. This high degree of overlapping sites under both balancing and directional selective pressures is surprising but coherent. The adaptive advantage that a heterozygote has in producing two different viral sensors from one gene given a vast viral landscape is easily understood. Over millions of years, sister lineages will sort alleles (completely and incompletely), whilst gaining, losing and maintaining novel polymorphisms that provide selective advantage to numerous pathogens. Given this narrative, a contemporary snapshot of the OAS1 gene should reveal balanced polymorphism at the population level while a broad survey of species' coding sequences should exhibit variation like we see here.

4.2. Concentration of mutation in minor groove 1 of the RBD

We describe fourteen residues with strong signatures of positive selection within the RNA binding domain. Models reveal that the majority of these are distributed in one region, RBD_{mg1}, which corresponds with minor groove I of docked dsRNA. Five of these selected sites participate in direct bonding with a nucleic acid activator. Ortholog models demonstrate that RBD_{mg1} nucleobases A₆, G₈, and A₉ are each potential targets for sequence-specific recognition. In RBD_{mg1}, we see Arg or Lys replacing Thr24, Arg replacing Gln200, and Arg

or Asn replacing Thr247. Each of these sites represents an ancestral polar side group taking part in hydrogen bonding with the RNA spine replaced by longer residues capable of more extensive bond formation with a sugar or base group. It should be noted that the ability of a particular amino acid to interact with a nucleobase can be just as significant as the inability of a different amino acid at that locus. Minor groove 2 contains at least three nucleobases known to interact with side groups (U₁₆, G₁₇, and G₁₈), yet those residues responsible for these bonds in RBD_{mg2} are not under positive selection. The distribution of amino acids found to be under balancing selection also appear to fit into this scheme, especially for macaques (figure 1). These observations support the hypothesis that RBD_{mg1} has evolved a distinct faculty from RBD_{mg2}.

Mutations in the RBD that bestow improved activation and therefore fitness in response to viruses will be positively selected for. Naturally, this process is complex, as the protein needs to maintain sufficient response to numerous pathogens while avoiding both endogenous activators and exogenous antagonists. Basically, because OAS1 is a pathogen sensor, the genetic conflict that has directed much of its evolution has chiefly put pressure on the interface between host and virus, namely the RBD. It appears that the RBD has responded in an innovative manner by adapting RBD_{mg1} and RBD_{mg2}. These subdomains can be seen as a reflection of the characteristics of activating ligands, where RBD_{mg1} is responsible for proficiently docking and responding to quickly mutating and divergent pathogen patterns, while RBD_{mg2} is charged with docking conserved exogenous motifs. Full support of this hypothesis will require experimental validation.

Our analysis of OWM OAS1 structure suggests that RNA length could play a role in driving some of the adaptive changes in the RBD. The most distal amino acids that docked RNA can interact with are residues 54 and 248. A subset of encoded residues at each of these sites clashes with the spine. We thus propose that dsRNA length preference could be regulated by these amino acid substitutions. It has been demonstrated experimentally that OAS1 is activated by dsRNA longer than 18bp but also that longer dsRNA generally activate more strongly (Ghosh, et al. 1997). A relationship between amino acid changes at OAS1 and the length of viral RNA is an intriguing possibility that needs to be tested experimentally.

4.3. Selection for changes affecting the active site

The group of amino acid changes under positive selection around the entrance to the active site indicates a history of adaptive evolution that relates directly to OAS1's catalytic activity. There are two putative reasons for this: response to antagonism by viral products or regulation of oligomer synthesis. While the simplest way a virus might prevent effective OAS1-mediated restriction would be by producing poorly activating dsRNA, this would be a competitive form of inhibition because other viral products could still activate. Far more potent would be an inhibitor that arrests OAS1 activity in a non-competitive fashion. Though such viral products have not been identified for OAS1, many key virulence factors fall in this category such as henipavirus V protein and influenza A NS1 protein, which inhibit members of the viral RNA sensor RIG-I-like receptor family, and vaccinia pE3 which antagonizes PKR's substrate-binding region (Sharp, et al. 1998; Zinzula and Tramontano 2013). We have identified a set of positively selected amino acids surrounding the active site passageway as

likely targets for viral antagonism whose rapid replacement indicates escape mutations (figure 6). Though the path to the active site is not very narrow, measuring from 17.2 Å to 20.5 Å between Trp126, Gln181, and Lys314, a relatively small peptide or other viral product capable of binding at or near these side groups could easily interfere with oligo synthesis. Four of these sites (and indeed most of the positively selected sites throughout the protein) are found in superficial loops which exposes them as potential targets for viral proteases in a similar fashion to hepatitis C virus N3/4A, which cleaves the antiviral MAVS protein (Gokhale, et al. 2014).

Comparative modeling of the orthologs encoding these amino acid substitutions also suggests they could affect the regulation of acceptor substrate size and that this could an adaptive advantage. Within the deep catalytic pocket, ATP binds at the donor-binding site followed by entry of acceptor substrate and lastly synthesis of the oligomer. The longer an acceptor is, the more opportunity it will have to interact, either positively or negatively, with shallower residues of the pocket such as those selected sites. Selection for amino acid replacements around this entryway is predicted to occur if fitness is related to the length of oligoadenylates synthesized. The human OAS1 produces short oligoadenylates, far less capable of activating RNase L compared with OAS2 and OAS3 products (Ibsen, et al. 2014; Marie, et al. 1997; Morin, et al. 2010; Rebouillat, et al. 1999). OAS1 gaining or losing the ability to directly signal this pathway would be significant. Furthermore, oligoadenylate length is important because viruses encode proteins known to sequester or degrade OAS products (Sanchez and Mohr 2007). Trp126's replacement arginine is a departure from the considerable steric effects of ancestral tryptophan in the β 3- β 4 linker (figures 6 and 7). With this change, the acceptor-binding site becomes open to larger substrates. Contrarily, in golden snub-nosed monkey, the acquisition of a longer linker may provide preference for production of short oligos. Some evidence for correlation of linker length to product size is that in both OAS2 and OAS3, which produce longer products, a two-residue deletion has taken place in their analogous loops. OAS1 may have evolved enhanced synthesis of longer oligos (perhaps even transiently) in species with these divergent forms that could translate to stronger RNase L response (Lim, et al. 2009). Structural study of hOAS3 has just revealed specialization in binding long dsRNA, perhaps the reason for its very high degree of conservation in primates (Donovan, et al. 2015). Future research will elucidate the distinct roles and coordinate relationships within the OAS protein family, a necessary step towards practical understanding and application of antiviral interventions.

Acknowledgments

The work was conducted in part with equipment from the Core Facilities for Imaging, Cellular and Molecular Biology at Queens College. This research was supported by the Professional Staff Congress-City University of New York grant 66642-00 44 to S.B. This investigation used resources that were supported by the Southwest National Primate Research Center grant P51 RR013986 from the National Center for Research Resources, National Institutes of Health, and that are currently supported by the Office of Research Infrastructure Programs through P51 OD011133. We thank two anonymous reviewers for their comments on an earlier version of the manuscript.

BIBLIOGRAPHY

The PyMOL Molecular Graphics System, Version 1.5.0.4. Schrödinger, LLC;

- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010; 7:248–249. DOI: 10.1038/nmeth0410-248 [PubMed: 20354512]
- Alagarasu K, Honap T, Damle IM, Mulay AP, Shah PS, Cecilia D. Polymorphisms in the oligoadenylate synthetase gene cluster and its association with clinical outcomes of dengue virus infection. *Infect Genet Evol*. 2013; 14:390–395. DOI: 10.1016/j.meegid.2012.12.021 [PubMed: 23337612]
- Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*. 2006; 22:195–201. DOI: 10.1093/bioinformatics/bti770 [PubMed: 16301204]
- Behera AK, Kumar M, Lockey RF, Mohapatra SS. 2'-5' Oligoadenylate synthetase plays a critical role in interferon-gamma inhibition of respiratory syncytial virus infection of human epithelial cells. *J Biol Chem*. 2002; 277:25601–25608. DOI: 10.1074/jbc.M200211200 [PubMed: 11980899]
- 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; 76:623–633. DOI: 10.1086/429391 [PubMed: 15732009]
- Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. *Nat Protoc*. 2009; 4:1–13. DOI: 10.1038/nprot.2008.197 [PubMed: 19131951]
- Cai Y, Chen Q, Zhou W, Chu C, Ji W, Ding Y, Xu J, Ji Z, You H, Wang J. Association analysis of polymorphisms in OAS1 with susceptibility and severity of hand, foot and mouth disease. *Int J Immunogenet*. 2014; 41:384–392. DOI: 10.1111/iji.12134 [PubMed: 25059424]
- Daugherty MD, Malik HS. Rules of engagement: molecular insights from host-virus arms races. *Annu Rev Genet*. 2012; 46:677–700. DOI: 10.1146/annurev-genet-110711-155522 [PubMed: 23145935]
- Delpont W, Poon AF, Frost SD, Kosakovsky Pond SL. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*. 2010a; 26:2455–2457. DOI: 10.1093/bioinformatics/btq429 [PubMed: 20671151]
- Delpont W, Scheffler K, Botha G, Gravenor MB, Muse SV, Kosakovsky PSL. CodonTest: modeling amino acid substitution preferences in coding sequences. *PLoS Comput Biol*. 2010b; 6doi: 10.1371/journal.pcbi.1000885
- Delpont W, Scheffler K, Seoighe C. Frequent toggling between alternative amino acids is driven by selection in HIV-1. *PLoS Pathog*. 2008; 4:e1000242.doi: 10.1371/journal.ppat.1000242 [PubMed: 19096508]
- Donovan J, Dufner M, Korennykh A. Structural basis for cytosolic double-stranded RNA surveillance by human oligoadenylate synthetase 1. *Proc Natl Acad Sci USA*. 2013; 110:1652–1657. DOI: 10.1073/pnas.1218528110 [PubMed: 23319625]
- Donovan J, Whitney G, Rath S, Korennykh A. Structural mechanism of sensing long dsRNA via a noncatalytic domain in human oligoadenylate synthetase 3. *Proc Natl Acad Sci USA*. 2015; 112:3949–3954. DOI: 10.1073/pnas.1419409112 [PubMed: 25775560]
- Doron-Faigenboim A, Pupko T. A combined empirical and mechanistic codon model. *Mol Biol Evol*. 2007; 24:388–397. DOI: 10.1093/molbev/msl175 [PubMed: 17110464]
- Doron-Faigenboim A, Stern A, Mayrose I, Bacharach E, Pupko T. Selecton: a server for detecting evolutionary forces at a single amino-acid site. *Bioinformatics*. 2005; 21:2101–2103. DOI: 10.1093/bioinformatics/bti259 [PubMed: 15647294]
- Drappier M, Michiels T. Inhibition of the OAS/RNase L pathway by viruses. *Curr Opin Virol*. 2015; 15:19–26. DOI: 10.1016/j.coviro.2015.07.002 [PubMed: 26231767]
- Ferguson W, Dvora S, Fikes RW, Stone AC, Boissinot S. Long-term balancing selection at the antiviral gene OAS1 in Central African chimpanzees. *Mol Biol Evol*. 2012; 29:1093–1103. DOI: 10.1093/molbev/msr247 [PubMed: 22104212]
- Ferguson W, Dvora S, Gallo J, Orth A, Boissinot S. Long-term balancing selection at the west Nile virus resistance gene, Oas1b, maintains transspecific polymorphisms in the house mouse. *Mol Biol Evol*. 2008; 25:1609–1618. DOI: 10.1093/molbev/msn106 [PubMed: 18460447]

- Fish I, Boissinot S. Contrasted patterns of variation and evolutionary convergence at the antiviral OAS1 gene in old world primates. *Immunogenetics*. 2015; 67:487–499. DOI: 10.1007/s00251-015-0855-0 [PubMed: 26156123]
- 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; 272:33220–33226. [PubMed: 9407111]
- Gokhale NS, Vazquez C, Horner SM. Hepatitis C Virus. Strategies to Evade Antiviral Responses. *Future Virol*. 2014; 9:1061–1075. DOI: 10.2217/fvl.14.89 [PubMed: 25983854]
- Han Y, Whitney G, Donovan J, Korennykh A. Innate immune messenger 2-5A tethers human RNase L into active high-order complexes. *Cell Rep*. 2012; 2:902–913. DOI: 10.1016/j.celrep.2012.09.004 [PubMed: 23084743]
- Hancks DC, Hartley MK, Hagan C, Clark NL, Elde NC. Overlapping Patterns of Rapid Evolution in the Nucleic Acid Sensors cGAS and OAS1 Suggest a Common Mechanism of Pathogen Antagonism and Escape. *PLoS Genet*. 2015; 11:e1005203.doi: 10.1371/journal.pgen.1005203 [PubMed: 25942676]
- Hartmann R, Justesen J, Sarkar SN, Sen GC, Yee VC. Crystal structure of the 2'-specific and double-stranded RNA-activated interferon-induced antiviral protein 2'-5'-oligoadenylate synthetase. *Mol Cell*. 2003; 12:1173–1185. [PubMed: 14636576]
- Hasegawa M, Kishino H, Yano T. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*. 1985; 22:160–174. [PubMed: 3934395]
- Huang H, Zeqiraj E, Dong B, Jha BK, Duffy NM, Orlicky S, Thevakumaran N, Talukdar M, Pillon MC, Ceccarelli DF, Wan LC, Juang YC, Mao DY, Gaughan C, Brinton MA, Pereygin AA, Kourinov I, Guarne A, Silverman RH, Sicheri F. Dimeric structure of pseudokinase RNase L bound to 2-5A reveals a basis for interferon-induced antiviral activity. *Mol Cell*. 2014; 53:221–234. DOI: 10.1016/j.molcel.2013.12.025 [PubMed: 24462203]
- Ibsen MS, Gad HH, Thavachelvam K, Boesen T, Despres P, Hartmann R. The 2'-5'-oligoadenylate synthetase 3 enzyme potentially synthesizes the 2'-5'-oligoadenylates required for RNase L activation. *J Virol*. 2014; 88:14222–14231. DOI: 10.1128/JVI.01763-14 [PubMed: 25275129]
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012; 28:1647–1649. DOI: 10.1093/bioinformatics/bts199 [PubMed: 22543367]
- Kodym R, Kodym E, Story MD. 2'-5'-Oligoadenylate synthetase is activated by a specific RNA sequence motif. *Biochem Biophys Res Commun*. 2009; 388:317–322. DOI: 10.1016/j.bbrc.2009.07.167 [PubMed: 19665006]
- Kosakovsky P, Frost SD. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol*. 2005; 22:1208–1222. DOI: 10.1093/molbev/msi105 [PubMed: 15703242]
- Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, Tyka M, Baker D, Karplus K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins*. 2009; 77(Suppl 9):114–122. DOI: 10.1002/prot.22570 [PubMed: 19768677]
- Lim JK, Lisco A, McDermott DH, Huynh L, Ward JM, Johnson B, Johnson H, Pape J, Foster GA, Kryzstof D, Follmann D, Stramer SL, Margolis LB, Murphy PM. Genetic variation in OAS1 is a risk factor for initial infection with West Nile virus in man. *PLoS Pathog*. 2009; 5:e1000321.doi: 10.1371/journal.ppat.1000321 [PubMed: 19247438]
- Lohofener J, Steinke N, Kay-Fedorov P, Baruch P, Nikulin A, Tishchenko S, Manstein DJ, Fedorov R. The Activation Mechanism of 2'-5'-Oligoadenylate Synthetase Gives New Insights Into OAS/ cGAS Triggers of Innate Immunity. *Structure*. 2015; 23:851–862. DOI: 10.1016/j.str.2015.03.012 [PubMed: 25892109]
- Marie I, Blanco J, Rebouillat D, Hovanessian AG. 69-kDa and 100-kDa isoforms of interferon-induced (2'-5')oligoadenylate synthetase exhibit differential catalytic parameters. *Eur J Biochem*. 1997; 248:558–566. [PubMed: 9346316]

- 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 USA*. 2002; 99:11311–11316. DOI: 10.1073/pnas.172195399 [PubMed: 12186974]
- Maurer-Stroh S, Eisenhaber F. Refinement and prediction of protein prenylation motifs. *Genome Biol*. 2005; 6:R55.doi: 10.1186/gb-2005-6-6-r55 [PubMed: 15960807]
- Mendez FL, Watkins JC, Hammer MF. Global genetic variation at OAS1 provides evidence of archaic admixture in Melanesian populations. *Mol Biol Evol*. 2012; 29:1513–1520. DOI: 10.1093/molbev/msr301 [PubMed: 22319157]
- Mendez FL, Watkins JC, Hammer MF. Neandertal origin of genetic variation at the cluster of OAS immunity genes. *Mol Biol Evol*. 2013; 30:798–801. DOI: 10.1093/molbev/mst004 [PubMed: 23315957]
- Minin VN, Dorman KS, Fang F, Suchard MA. Dual multiple change-point model leads to more accurate recombination detection. *Bioinformatics*. 2005; 21:3034–3042. DOI: 10.1093/bioinformatics/bti459 [PubMed: 15914546]
- Morin B, Rabah N, Boretto-Soler J, Tolou H, Alvarez K, Canard B. High yield synthesis, purification and characterisation of the RNase L activators 5'-triphosphate 2'-5'-oligoadenylates. *Antiviral Res*. 2010; 87:345–352. DOI: 10.1016/j.antiviral.2010.06.003 [PubMed: 20547186]
- Mozzi A, Pontremoli C, Forni D, Clerici M, Pozzoli U, Bresolin N, Cagliani R, Sironi M. OASes and STING: adaptive evolution in concert. *Genome Biol Evol*. 2015; doi: 10.1093/gbe/evv046
- Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet*. 2012; 8:e1002764.doi: 10.1371/journal.pgen.1002764 [PubMed: 22807683]
- Perelman P, Johnson WE, Roos C, Seuanez HN, Horvath JE, Moreira MA, Kessing B, Pontius J, Roelke M, Rumpfer Y, Schneider MP, Silva A, O'Brien SJ, Pecon-Slattery J. A molecular phylogeny of living primates. *PLoS Genet*. 2011; 7:e1001342.doi: 10.1371/journal.pgen.1001342 [PubMed: 21436896]
- 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. *J Mol Evol*. 2006; 63:562–576. DOI: 10.1007/s00239-006-0073-3 [PubMed: 17024523]
- Pond SL, Frost SD, Muse SV. HyPhy: hypothesis testing using phylogenies. *Bioinformatics*. 2005; 21:676–679. DOI: 10.1093/bioinformatics/bti079 [PubMed: 15509596]
- Pupko T, Pe'er I, Shamir R, Graur D. A fast algorithm for joint reconstruction of ancestral amino acid sequences. *Mol Biol Evol*. 2000; 17:890–896. [PubMed: 10833195]
- Quintana-Murci L, Alcais A, Abel L, Casanova JL. Immunology in natura: clinical, epidemiological and evolutionary genetics of infectious diseases. *Nat Immunol*. 2007; 8:1165–1171. DOI: 10.1038/ni1535 [PubMed: 17952041]
- Rausell A, Telenti A. Genomics of host-pathogen interactions. *Curr Opin Immunol*. 2014; 30:32–38. DOI: 10.1016/j.coi.2014.06.001 [PubMed: 24972384]
- Rebouillat D, Hovnanian A, Marie I, Hovanessian AG. The 100-kDa 2',5'-oligoadenylate synthetase catalyzing preferentially the synthesis of dimeric pppA2'p5'A molecules is composed of three homologous domains. *J Biol Chem*. 1999; 274:1557–1565. [PubMed: 9880533]
- Sanchez R, Mohr I. Inhibition of cellular 2'-5' oligoadenylate synthetase by the herpes simplex virus type 1 Us11 protein. *J Virol*. 2007; 81:3455–3464. DOI: 10.1128/JVI.02520-06 [PubMed: 17229694]
- Sharp TV, Moonan F, Romashko A, Joshi B, Barber GN, Jagus R. The vaccinia virus E3L gene product interacts with both the regulatory and the substrate binding regions of PKR: implications for PKR autoregulation. *Virology*. 1998; 250:302–315. DOI: 10.1006/viro.1998.9365 [PubMed: 9792841]
- Silverman RH. Viral encounters with 2',5'-oligoadenylate synthetase and RNase L during the interferon antiviral response. *J Virol*. 2007; 81:12720–12729. DOI: 10.1128/JVI.01471-07 [PubMed: 17804500]

- Sippl MJ. Recognition of errors in three-dimensional structures of proteins. *Proteins*. 1993; 17:355–362. DOI: 10.1002/prot.340170404 [PubMed: 8108378]
- Stern A, Doron-Faigenboim A, Erez E, Martz E, Bacharach E, Pupko T. Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. *Nucleic Acids Res*. 2007; 35:W506–511. DOI: 10.1093/nar/gkm382 [PubMed: 17586822]
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011; 28:2731–2739. DOI: 10.1093/molbev/msr121 [PubMed: 21546353]
- Torshin IY. Three-dimensional models of human 2′–5′ oligoadenylate synthetases: a new computational method for reconstructing an enzyme assembly. *Med Sci Monit*. 2005; 11:BR235–247. [PubMed: 15990685]
- Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*. 2001; 18:691–699. [PubMed: 11319253]
- Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007; 35:W407–410. DOI: 10.1093/nar/gkm290 [PubMed: 17517781]
- Yang Z, Kumar S, Nei M. A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics*. 1995; 141:1641–1650. [PubMed: 8601501]
- Zinzula L, Tramontano E. Strategies of highly pathogenic RNA viruses to block dsRNA detection by RIG-I-like receptors: hide, mask, hit. *Antiviral Res*. 2013; 100:615–635. DOI: 10.1016/j.antiviral.2013.10.002 [PubMed: 24129118]

Highlights

OAS1 participates in the first line of defense against many viral pathogens.

OAS1 is evolving under positive selection in Old World monkeys.

Most positively selected sites are located in the RNA-binding domain.

Positively selected sites are concentrated in a region of the RNA binding domain.

Residues circumscribing the entry to the active site are also positively selected.

Codon #	Minor groove										Major groove										Isoform 2					Isoform 1											
	mg1	mg1	mg1	mg2	mg2	mg2					mg2	mg2	mg1	mg1	mg1	mg1	mg2					mg2	mg1	mg1	mg1												
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1												
	2	2	2	3	5	0	2	3	6	6	7	7	8	9	0	4	4	4	5	8	8	1	1	3	3	4	5	5	6	6	4	4	7	8			
	4	5	8	1	4	0	6	1	1	2	6	5	9	1	6	0	2	5	7	8	3	4	5	1	3	4	5	5	6	7	7	9	8	9			
RNA-binding ability	d/I		I	d/I	d/I											d/I		d	d/I																		
Ancestral amino acid	T	R	N	R	R	W	A	G	G											D	Q		D	G	P	K	K	W	S							R	
Black and white colobus (<i>Colobus angolensis</i>) ^a	K	R	T	H	R	R	W	A	G	G	D	K	S	Q	D	Q	Q	M	R	D	Q	G	E	P	K	E	W	P	F	M	L	T	N	Q	R		
King colobus (<i>Colobus polykomos</i>)	K	H	R	H	R	R	W	A	G	G	D	K	S	Q	D	Q	Q	M	G	D	Q	G	E	R	K	K	W	S	S	I	P	T	N	Q	R		
Golden snub-nosed monkey (<i>Rhinopithecus roosei</i>) ^a	T	R	T	H	C	Q	R	A	G	G	D	K	S	K	D	Q	Q	M	T	D	Q	G	E	S	N	K	L	S	F	M	L	T	S	Q	R		
Proboscis monkey (<i>Nasalis larvatus</i>)	T	R	M	N	H	Q	W	A	G	G	H	K	S	Q	D	Q	Q	M/R	N/T	D	Q	G	Q	S	K	K	L	S	F	M	L	T	S	Q	R		
Douc (<i>Pygathrix nemus</i>)	T	R	T	N	R	Q	W	A	G	G	H	K	S	Q	D	P	Q	Q	M	D	Q	G	Q	G	K	R	L	S	F	M	L	T	S	Q	R		
Woolf's monkey (<i>Cercopithecus woolfi</i>) ^a	T	R	M	K	H	Q	W	A	G	G	N	E	F	Q	N	Q	Q	M	E	D	Q	G	K	P	E	K	W	S	7	7	7	A	S	H	R		
Red-tailed guenon (<i>Cercopithecus escuinii</i>) ^a	T	R	M	N	R	Q	W	A	G	G	N	E	F	Q	N	Q	Q	M	T	D	R	E	K	P	E	K	W	7	7	7	7	7	7	A	S	H	R
Hamlyn's guenon (<i>Cercopithecus hamlyni</i>)	T	R	M	K	Y	Q	W	A	G	G	D	E	F	Q	N	Q	Q	M	T	D	R	E	Q	P	K	E	W	P	S	I	P	A	S	H	R		
Alleni's Swamp monkey (<i>Allenopithecus nigroviridis</i>) ^a	T	H	M	N	R	Q	R	A	G	G	D	E	S	Q	N	Q	Q	R	M	T	D	R	G	K	P	K	W	P	7	7	7	V	S	H	R		
Green monkey (<i>Chlorocebus sabaeus</i>) ^a	T	H	T	N	R	Q	W	A	G	G	D	E	S	Q	N	Q	Q	M	T	Y	R	G	K	P	T	E	W	P	S	I	P	A	S	H	R		
Patas (<i>Erythrocebus patas</i>)	T	R	T	N	R	Q	W	V	S	G	N	K	S	Q	D	Q	Q	M	T	Y	R	G	K	P	M	E	W	P	S	I	P	T	S	–	–	–	
Sooty mangabey (<i>Cercopithecus atys</i>) ^a	T	R	M	N	R	Q	W	A	G	S	N	K	D	Q	D	Q	Q	M	R	D	R	G	K	P	K	K	W	S	S	I	P	A	S	Q	R		
Collared mangabey (<i>Cercopithecus torquatus</i>)	R	M	N	R	R	Q	W	A	G	G	N	K	D	Q	D	Q	Q	M	R	D	R	G	K	P	K	K	W	S	S	I	P	A	S	Q	R		
Drill (<i>Mandrillus leucophaea</i>)	T	R	M	N	R	Q	W	A	G	G	D/N	E	F	Q	D	Q	Q	M	R	D	R	G	K	P	K	K	W	S	S	I	P	A	S	Q	R		
Olive baboon (<i>Papio anubis</i>)	R	M	N	R	Q	W	V	G	G	D	E	F	Q	D	Q	Q	M	T	D	R	E	E	K	P	K	K	W	S	S	I	P	A	G	Q	R		
Hamadryas baboon (<i>Papio hamadryas</i>)	R	M	N	R	R	R	A	G	G	D	E	F	Q	D	Q	Q	M	M/T	T	D	R	E/G	K	T	K	E	R	S	S	I	P	A	S	H/Q	P/R		
Guinea baboon (<i>Papio papio</i>)	R	M	N	R	Q	W	A	G	G	D	E	F	Q	D	Q	Q	M	T	D	R	E	K	P	K	K	W	S	–	–	–	–	–	–	–	–		
Yellow baboon (<i>Papio cyncephalus</i>)	R/T	R	M	N	R	Q	W	A	G	G	D	E	F	Q	D	Q	Q	M	T	D	R	E	K	P	K	K	W	S	S	I	P	A	G	Q	R		
Rhesus macaque (<i>Macaca mulatta</i>)	T	H	M	N	R	R	A	G	G	D	E	S	Q	N	Q	Q	R	M	T	D	R	G	K	P	K	K	W	S	S	I	P	A	S	H	P		
Stump-tailed macaque (<i>Macaca arctoides</i>)	T	H	M	N	R	R/Q	R/W	A	G	G	N	K	S	Q	D/N	Q	Q	R	T	R	Y	R	G	K	P	K	W	S	S	I	A	S	Q	R			
Pig-tailed macaque (<i>Macaca nemestrina</i>)	T	H	M	N	R	R	W	A	G	G	D	E	S	Q	N	Q	Q	R	T	R	D	R	G	K	T	K	W	S	S	I	P	A	S	H	R		
Cynomolgus macaque (<i>Macaca fascicularis</i>)	T	H	T	N	C	R	W	A	G	G	D	K	S	Q	N	Q	Q	M	T	D	Q	G	K	T	K	K	C	S	S	I	P	A	S	H	P		
SELECTOR M8 REL	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
MIM	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
SELECTOR M5	2	3	4	2	4	3	3	2	2	2	4	4	4	3	2	2	2	4	2	3	3	3	3	3	3	3	3	3	2	2	3	2	4	3	X		
MEME	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
SELECTOR MEC	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
TOGGLE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		
Balancing selection	C	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	C/M	

Figure 1.

Positively selected sites in OWM OAS1. All the sequences were generated for this study except four collected from Genbank (^G) and three previously published by Hancks et al (2015; ^H). Only amino acids determined to have evolved under positive selection by at least two tests (Selecton M8 and HyPhy REL) are shown. Residues located in the RBD are indicated in grey and residues located near the active site in gold. Sites that are directly or indirectly binding with RNA (Donovan, et al. 2013; Lohofener, et al. 2015) are respectively indicated with a d or I. The sites previously shown to evolve under balancing selection in chimpanzee and macaque are annotated with C or M, respectively. MIM = minimum number of independent mutations necessary to produce the observed OWM genotypes.

Figure 2. OAS1 gene tree. The tree was built using the maximum likelihood method using the HKY98 model and human as an outgroup. Numbers at the nodes indicates the bootstrap support.

Figure 3.

Pairwise values of dN/dS . Each column displays the distribution of dN/dS values between the indicated species and all other species.

Figure 4.

Structural model of Old World monkey OAS1. Blue sites indicate those amino acids in the RNA binding domain identified as evolving under positive selection. Red ribbon backbone is the minor groove 1 half of docked dsRNA and blue is minor groove 2 half.

Figure 5.

Overview of nucleotides and positively selected amino acids at the RNA binding domain. dsRNA docks OAS1 with two minor grooves contacting the RBD. Yellow residues interact with base pairs 1 through 9 (RBD_{mg1}) and green residues with nucleotides 10 through 18 (RBD_{mg2}). RNA enumeration is in concordance with 4IG8.

Figure 6. Positive selection at OAS1 active site entrance. Yellow indicates selected amino acids located between the site of catalysis and the entry/exit to the reactant binding sites.

Figure 7.

OWM variation affecting reactant use. Divergent replacement arginine at 126 is predicted to decrease OAS1 preference for ATP at acceptor substrate binding site. Red residues are catalytic metal-binding aspartates (indicating the approximate site of catalysis).

Figure 8.

Cluster of sites under positive selection at helix C1. The high density of positively selected sites (purple) at this surface structure is suggestive of a history of escape mutation from viral antagonism.