1 **Title**: Adaptively introgressed Neandertal haplotype at the OAS locus functionally
2 impacts innate immune responses in humans.

3 **Authors**

4 Aaron J. Sams[1,* †], Anne Dumaine[2†], Yohann Nédélec[2,3†], Vania Yotova[2], Caroline
5 Alfieri[2,4], Jerome E. Tanner[2], Philipp W. Messer[1*††], and Luis B. Barreiro[2,5,*††]

6 **Affiliations**

7 [1]Department of Biological Statistics and Computational Biology, Cornell
8 University, Ithaca, NY

9 [2]Department of Genetics, Sainte-Justine Hospital Research Centre, University of
10 Montreal, Montreal, Qc, Canada

11 [3]Department of Biochemistry, University of Montreal, Montreal, Qc, Canada

12 [4]Department of Microbiology and Immunology, University of Montreal, Montreal,
13 Qc, Canada

14 [5]Department of Pediatrics, University of Montreal, Montreal, Qc, Canada

15 [†]These authors contributed equally to this work

16 [††]These authors jointly supervised this work

17 [*]Correspondence to: (A.J.S.) as2847@cornell.edu, (P.W.M.) messer@cornell.edu,
18 (L.B.B.) luis.barreiro@umontreal.ca

Abstract

The 2'-5' oligoadenylate synthetase (OAS) locus encodes for three OAS enzymes (OAS1-3) involved in innate immune response. This region harbors high amounts of Neandertal ancestry in non-African populations; yet, strong evidence of positive selection in the OAS region is still lacking. Here we used a combination of neutral coalescent simulations and neutrality tests to firmly demonstrate a signal of adaptive introgression at the OAS locus. Furthermore, we characterized the functional consequences of the Neandertal haplotype in the transcriptional regulation of OAS genes at baseline and infected conditions. We found that cells from people with the Neandertal-like haplotype express lower levels of *OAS3* upon infection, as well as distinct isoforms of *OAS1* and *OAS2*. Notably, the Neandertal-introgressed haplotype reintroduced an ancestral splice variant of *OAS1* encoding a more active protein, suggesting that adaptive introgression occurred as a means to resurrect adaptive variation that had been lost outside Africa.

Introduction

Whole genome sequencing of several archaic human genomes(Green et al., 2010; Reich et al., 2010; Meyer et al., 2012; Prüfer et al., 2014; Sawyer et al., 2015) representing Neandertals and an as yet geographically and paleontologically unknown population referred to as Denisovans has revealed gene flow between these populations and the ancestors of present-day humans. Neandertal ancestry makes up approximately 0.5-2 percent of the ancestry of most living humans, with

higher amounts of Neandertal ancestry found outside of Africa(Sankararaman et al., 2014; Vernot and Akey, 2014; Llorente et al., 2015). While it seems that there may have been widespread purifying selection against Neandertal ancestry on the ancestral African genomic background(Sankararaman et al., 2014; Vernot and Akey, 2014), some positive selection on Neandertal genes (adaptive introgression) has also been observed(Racimo et al., 2015; 2016). Neandertals and other archaic populations inhabited Eurasia for several hundred thousand years(Hublin, 2009), therefore some adaptive introgression is expected, particularly across phenotypes that are strongly influenced by direct interactions with the surrounding environment (Racimo et al., 2015), such as our immune response to infectious agents (Ségurel and Quintana-Murci, 2014).

The OAS locus on chromosome 12, which harbors three genes (*OAS1, OAS2, OAS3*) encoding the 2'-5' oligoadenylate synthetase enzymes has received considerable attention due to its clear signs of multiple archaic haplotypes in populations outside of Africa(Mendez et al., 2012; 2013), and the critical role of OAS genes in the innate immune response to viruses (Player and Torrence, 1998). The major Neandertal haplotype at the OAS locus spans a ~190 kilobase region between two surrounding recombination hotspots. This haplotype is absent from sub-Saharan African samples in the 1000 Genomes Project data, yet found at relatively high frequencies outside of Africa, reaching highest frequencies in European population samples (up to 41%, Figure 1). Indeed, the Neandertal haplotypes in the OAS region are among the most common Neandertal haplotypes among Europeans(Sankararaman et al., 2014).

The elevated frequency of the introgressed haplotype along with the key role OAS genes play in protective immunity against viral infections raises the possibility that introgressed Neandertal haplotypes at OAS may have been adaptive in modern humans. While some studies provide suggestive evidence of adaptive introgression at the OAS locus (Sankararaman et al., 2014; Racimo et al., 2016), strong evidence of positive selection in the OAS region is still lacking. Indeed, most studies failed to reject a model of neutral evolution for the Neandertal haplotype when using standard neutrality tests(Mendez et al., 2013; Deschamps et al., 2016). We hypothesize that the overall lack of signals of selection in the OAS region stems from the low power of standard neutrality tests to detect adaptive introgression(Sankararaman et al., 2016), and the particular genomic architecture of the region, which is characterized by the presence of two strong recombination hotspots (Figure 2A). The presence of these recombination hotspots is likely to significantly reduce power to detect signals of selection, particularly for tests based on haplotype-lengths or levels of linkage disequilibrium associated with the selected allele. Here we circumvent these issues by testing the hypothesis of adaptive introgression using extensive neutral coalescent simulations specifically tailored to match the genomic features of the OAS region in combination with ancient DNA data from Eurasia. We firmly demonstrate a population genetic signal of adaptive introgression at the OAS locus and characterize the functional consequences of the Neandertal haplotype in the transcriptional regulation of OAS genes in macrophages and peripheral blood mononuclear cells (PBMCs) at baseline and infected conditions.

87 **Results**

88 To investigate the hypothesis of non-neutral evolution at the OAS locus we
89 simulated neutral sequence data using a demographic model based on several
90 previously inferred parameters of human demographic history(Gravel et al., 2011;
91 Tennessen et al., 2012; Vernot and Akey, 2015)(Figure S1) and explicitly
92 incorporated estimated recombination rates in the OAS region (Methods). Our
93 model included a single pulse of Neandertal introgression occurring over a span of
94 500-1000 years (sampled from a uniform distribution) into the ancestral Eurasian
95 population after their population split from Africa. We investigated patterns of
96 haplotype homozygosity surrounding Neandertal-like sites (NLS), derived alleles
97 that are shared between Neandertals and a non-African population sample (CEU),
98 but absent in a sub-Saharan African sample (YRI). This class of sites is frequently
99 used as a conservative indicator of Neandertal introgression(Yang et al., 2012; Fu
100 et al., 2014; Sankararaman et al., 2014). In each simulation we chose a NLS at
101 random to measure two summary statistics, NLS frequency, and $H_{D/A}$, a statistic
102 which measures the ratio of average pairwise homozygosity tract lengths among
103 haplotypes carrying a derived versus ancestral allele at each NLS (Methods). This
104 statistic is based on the *H* statistic that provides a general measure for haplotype
105 lengths and has been shown to be robust in detecting both hard and soft selective
106 sweeps(Schlamp et al., 2016). If Neandertal haplotypes have experienced a greater
107 rate of positive selection than ancestral haplotypes, this should lead to a relative
108 increase of *H* among the subgroup of haplotypes that carry Neandertal alleles
109 versus those carrying only ancestral alleles, leading to higher values of $H_{D/A}$

110 compared to a neutrally evolving locus with similar recombination rates (Figure
111 S2). Additionally, we expected the frequency of NLS at the OAS locus to be
112 elevated compared to the NLS frequency distribution observed across neutral
113 simulations.

114 In practice, we observed both signals. We first examined the probability in our
115 model that neutrally evolving Neandertal like sites would be segregating at their
116 current frequencies. We found that the highest frequency NLS (DAF = 38.4%) fall
117 in the extreme 1% of all simulations (lowest $P = 0.0055$) (Figure 2A). Next, we
118 compared $H_{D/A}$ at each NLS in the OAS region to distributions of simulated NLS
119 drawn to match local recombination rate and frequency (Methods), thereby
120 controlling for associations between frequency and recombination, and haplotype
121 lengths. We found several NLS in the OAS region that have $H_{D/A}$ values greater
122 than 99% of comparable simulated NLS (lowest $P = 0.0023$) (Figure 2B). Further,
123 we measured mean $H_{D/A}$ across all NLS in the OAS region and compared this to
124 randomly sampled distributions of simulated NLS (again matching frequency and
125 recombination rate). We found that the true mean $H_{D/A}$ in OAS is greater than four
126 standard errors from the neutral expectation ($P < 10^{-5}$) (Figure 2D). Lastly, we
127 considered the joint probability of observing a neutral NLS in our simulations with
128 frequencies and $H_{D/A}$ values as high as those observed in the OAS region. Only
129 413 simulated NLS (out of 1,000,079) had both a higher frequency and $H_{D/A}$ value
130 than the highest frequency/ $H_{D/A}$ pair in our NLS sample ($P < 5 \times 10^{-4}$) (Figure 2E).
131 Combined, we identified the strongest evidence to date in support of adaptive
132 introgression in the OAS region.

133 We further supported and contextualized our simulation results by examining the
134 OAS region in a dataset of 230 ancient Eurasians(Mathieson et al., 2015).
135 Assuming neutrality, the expected frequency of an allele in contemporary
136 European populations can be predicted as a linear combination of allele
137 frequencies sampled from representative ancient populations that have contributed
138 ancestry to present-day European populations in different proportions(Lazaridis et
139 al., 2014; Mathieson et al., 2015) (Methods). Therefore, using ancient allele
140 frequencies estimated by Mathieson and colleagues, we calculated the expected
141 allele frequency in four present day European samples from the 1,000 Genomes
142 Project(Consortium, 2015) at 11 NLS falling within the bounds of the three OAS
143 genes. To set up our null expectations we performed a similar analysis on a dataset
144 of approximately one million SNPs scattered around the genome, generated by
145 Mathieson and colleagues (by merging 213 ancient samples dated between 6500
146 and 300 BCE with sequencing data from four European samples from the 1,000
147 Genomes Project). We found that OAS NLS SNPs are outliers in the genome with
148 respect to deviations from ancient frequencies. More specifically, we found that
149 the allele frequency of 6 out of the 11 OAS SNPs tested in the *OAS1-OAS3* region
150 have increased above the frequency predicted by ancient Eurasian samples by
151 more than 20%, significantly more than what we observed for other SNPs
152 genome-wide with comparable present-day frequencies (lowest $P = 0.00476$,
153 Figure 2C, Table S1). Our findings at this single locus are consistent with results
154 from the genome-wide selection scan performed by Mathieson and
155 colleagues(2015) where the OAS region was also identified as an outlier.

156 Our population genetic results provide evidence that Neandertal alleles at the OAS
157 locus have likely experienced positive selection at some point after their
158 introduction into the human population, suggesting a possible functional role of
159 these alleles in human innate immune responses. To study this possibility, we
160 analyzed RNA-sequencing data collected on primary macrophages from 99
161 European-descent individuals, before and after *in-vitro* infection with *Salmonella*
162 *typhimurium.* After 2 hours of infection, we found that all OAS genes were
163 strongly up-regulated (up to 19-fold, $P < 1 \times 10^{-10}$, Figure S3), confirming the
164 ability of *Salmonella* to activate the interferon (IFN) production pathway(Nauciel
165 and Espinasse-Maes, 1992; Gulig et al., 1997; LaRock et al., 2015). Using
166 genotype data available for the same individuals (673 SNPs spanning the OAS
167 region, see methods) we tested if NLS were associated with variation in the
168 expression levels of *OAS1*, *OAS2* or *OAS3,* in either infected or non-infected
169 macrophages. We found that SNPs linked with the Neandertal haplotype reduced
170 the expression levels of *OAS3* (i.e., they were expression quantitative trait loci, or
171 *cis* eQTL for *OAS3*) (Figure 3A, false discovery rate (FDR) < 5%). Interestingly,
172 these *cis* eQTL showed a much stronger effect in infected macrophages (best P
173 $_{salmonella} = 3.5\times10^{-3}$ *vs* best $P_{\text{non-infected}} = 0.027$), supporting an interaction between
174 the Neandertal haplotype and the *OAS3* response to *Salmonella* infection.

175 In addition to overall changes in expression, we took advantage of the power of
176 RNA-sequencing data to test if NLS in the OAS regions influenced the ratio of
177 alternative isoforms used for each of the OAS genes (i.e., alternative splicing
178 QTL: asQTL). We found that SNPs associated with the Neandertal haplotype are

179 significant asQTL for *OAS1* and *OAS2* in both infected and non-infected
180 macrophages (FDR<< 1%; Figure 3B-C). The effect of the splice site variant
181 rs10774671 at determining what isoform is primarily encoded by *OAS1* was
182 particularly strong ($P \leq 2x10^{-32}$). The ancestral G allele at this SNP (AG at
183 acceptor site) retains the splice site whereas the derived allele, A, (AA at acceptor
184 site) disrupts the splice site leading to the usage of a distinct isoform (Figure S4).
185 The Neandertal haplotype harbors the ancestral allele (encoding the p46 isoform),
186 which is associated with high enzyme activity(Bonnevie-Nielsen et al., 2005),
187 whereas the derived allele – present in virtually all non-Neandertal haplotypes
188 found outside Africa (98.4% in CEU) – is associated with reduced enzymatic
189 activity.

190 Because OAS genes are primarily involved in the control of viral infections we
191 decided to validate our functional findings on peripheral blood mononuclear cells
192 (PBMCs) from 30 individuals stimulated/infected with viral-ligands (polyI:C and
193 gardiquimod), and live viruses (Influenza, Herpes simplex virus (HSV) 1 and
194 HSV2). The individuals were chosen based on their genotype for the NLS
195 rs1557866, a SNP that is a strong proxy for the presence or absence of the
196 Neandertal haplotype in the OAS region (9 were homozygous for the Neandertal
197 haplotype, 9 were heterozygous, and 12 homozygous for the modern human
198 sequence).

199 As expected, we found that all viral-associated immune triggers led to a marked
200 increase in *OAS1-3* gene expression levels, as measured by real-time PCR (up to

201 30-fold, $P \leq 5.4x10^{-6}$, Figure 4A), concomitantly with the up-regulation of type-I
202 and type-II interferon genes (Figure S5). Confirming the QTL results obtained in
203 macrophages, we found that rs10774671 was a strong asQTL for *OAS1* in both
204 non-infected and infected PBMCs ($P \leq 3.6x10^{-5}$, Figure 4B). Likewise, we found
205 that the presence of the Neandertal haplotype was associated with reduced
206 expression levels of *OAS3,* particularly in PBMCs infected with influenza (P =
207 $3.9x10^{-4}$) and the synthetic ligand gardiquimod (P $=1.0x10^{-3}$), which mimics a
208 single strand RNA infection (Figure 4B). Interestingly, the Neandertal haplotype
209 harbors additional regulatory variants that only impact expression levels in a cell-
210 type and immune stimuli specific fashion. For example, we found that the
211 Neandertal haplotype is associated with increased expression levels of *OAS2* in
212 non-infected (*P*=0.009), HSV1-infected (*P*=0.04), and gardiquimod-stimulated
213 PBMCs (*P*=0.0055), but not in macrophages nor in PBMCs treated with other
214 viral agents. Collectively, our functional data shows a pervasive impact of the
215 Neandertal haplotype on the regulation of *OAS* genes. Moreover, they show that
216 the regulatory impact of the Neandertal haplotype varies depending on the cell
217 type and the immune stimuli the cells are responding to.

218 **Discussion**

219 The Neandertal lineage was present in Eurasia for at least 400,000 years(Meyer et
220 al., 2016), providing ample time for Neandertals to adapt to local disease
221 environments. The admixture process, which likely fostered the transmission of
222 pathogens between Neandertals and humans migrating out of Africa, could have

223 also led to the exchange of genes useful in responding to local pathogens. Here,
224 we have demonstrated that a previously reported case of Neandertal introgression
225 at the OAS locus(Mendez et al., 2013) displays signatures of positive selection in
226 the European population. Additionally, we have strengthened the case for adaptive
227 introgression by providing direct functional evidence of a role for the Neandertal
228 OAS haplotype to the regulatory responses in innate immune cells to infectious
229 agents.

230 Under a model of constant directional selection with codominant fitness effects for
231 Neandertal alleles we can use present-day frequencies and a plausible
232 establishment frequency at the time of introgression to estimate the selection
233 coefficient associated with the Neandertal haplotype at the OAS locus (Methods).
234 Assuming an establishment frequency of 0.02, and introgression approximately
235 2000-2400 generations ago the selection coefficient in heterozygotes would be $s \sim$
236 0.0014 - 0.0017. However, the observed allele frequency shift of 0.26 in only 200-
237 340 generations (maximum shift in CEU from ancient samples; see Figure 2C**)**
238 predicts that the selection coefficient associated with the Neandertal haplotype was
239 2.6-5.4 times larger during recent human evolution ($s \sim 0.0044 - 0.0075$). These
240 results suggest that the Neandertal OAS haplotype has not experienced constant
241 directional selection in humans.

242 Our results show that the Neandertal haplotype at OAS is associated with several
243 regulatory variants that reduce expression of *OAS3* in response to infection, as
244 well as encode alternate isoforms of *OAS1* and *OAS2*. These dramatic functional

245 implications of the Neandertal OAS haplotype support our case for adaptive
246 introgression at OAS. Yet, because distinct functional polymorphisms segregate
247 together in the same haplotype, inferring the exact variant(s) targeted by positive
248 selection remains a daunting task. We speculate, however, that the direct target of
249 selection is likely to have been the splice variant identified in *OAS1*.

250 The Neandertal haplotype carries the ancestral allele (G) of the OAS1 splice
251 variant (rs10774671), which is common both inside and outside of Africa.
252 However, outside of Africa, the only haplotypes carrying this ancestral splice site
253 are closely related to the Neandertal haplotype, with a few exceptions being rare
254 recombined haplotypes (~2% of all haplotypes with the ancestral allele). This
255 pattern reflects the possibility that Neandertal introgression, in effect, served as a
256 means to resurrect the ancestral splice site from local extinction outside of Africa,
257 probably following the out-of-Africa exodus. The Neandertal-introgressed allele
258 encodes a protein variant (p46) that is associated with higher enzymatic
259 activity(Bonnevie-Nielsen et al., 2005). The adaptive potential of this variant is
260 supported by the observation that this variant (or other variants in strong LD with
261 it) was shown to be associated with: *(i)* reduced infection and replication rates of
262 West Nile virus ((Lim et al., 2009), but see (Bigham et al., 2011)), *(ii)* improved
263 resistance to hepatitis C virus (HCV) infection(Awady et al., 2011; Kwon et al.,
264 2012), and *(iii)* variable symptomology of Tick-Borne Encephalitis (TBE) Virus-
265 Induced Disease (homozygous individuals for the Neandertal haplotype show the
266 most severe symptoms of TBE). Strikingly, West Nile, hepatitis C and TBE are

267 all members of the *Flaviviridae* family, suggesting that Flaviviruses might have
268 been the main drivers of selection in *OAS1*.

269 The differential responses of homozygous carriers of the Neandertal OAS
270 haplotype to the different viruses described above suggest that the Neandertal
271 haplotype is not uniformly beneficial in humans. Thus, it is plausible that alleles at
272 OAS, and particularly the *OAS1* splice variant, might be evolving under balancing
273 selection. This hypothesis is supported by the observation that the *OAS1* splice
274 variant (rs10774671) is found at high frequency worldwide (0.11-0.7), and that
275 *OAS1* is among the most diverse genes in both humans and non-human primates.
276 Indeed, a recent analysis of genome-wide sequence data from a total of 55
277 individuals from four non-human ape species, chimpanzee (*Pan troglodytes*
278 *ellioti*), bonobo (*Pan paniscus*), gorilla (*Gorilla gorilla gorilla*), and orangutan
279 (*Pongo abelii*), identified *OAS1* as in the top 1% of genes showing the largest
280 levels of nucleotide diversity among ape species, consistent with a scenario of
281 long-term balancing selection (*OAS2* and *OAS3* are ranked in the 60th and 36th
282 percentile of the genome-wide distribution, respectively). Further supporting the
283 idea of balancing selection on the introgressed haplotypes, our functional data
284 suggest that the Neandertal haplotype contributes a range of gene expression
285 responses in a cell-type and stimulus-specific manner.

286 In conclusion, our study demonstrates that the frequency and haplotype
287 distribution of Neandertal-like sites can be used in a neutral simulation framework
288 that accounts for local genomic context to investigate the history of selection at a

289 candidate locus for which genome-wide tests of selection provide ambiguous
290 results. When combined with functional data, our results provide the strongest
291 evidence to date in support of adaptive introgression in the OAS region. More
292 generally, our study raises the possibility that adaptive introgression might not
293 necessarily occur to select newly introduced variants but rather as a means to
294 resurrect adaptive variation into modern humans who had lost it due to
295 demographic events.

296 **Materials and Methods**

297 *1. Genome alignments and identification of Neandertal-like sites*

298 Human/chimpanzee ancestral states were computed by parsimony using
299 alignments from the UCSC Genome Browser for the human reference (hg19) and
300 three outgroups chimpanzee (panTro2), orangutan (ponAbe2), and rhesus macaque
301 (rheMac2)(Kent et al., 2002). Ancestral state was assumed to be the chimpanzee
302 allele (if available) if its state was confirmed by matching either orangutan or
303 macaque. All sites with no inferred ancestral state were removed from our
304 analysis.

305 We filtered the Altai Neandertal genome(Prüfer et al., 2014) using the map35_50
306 set of minimum filters provided at
307 (https://bioinf.eva.mpg.de/altai_minimal_filters/). We combined this filtered
308 dataset with the CEU (Utah residents (CEPH) with Northern and Western
309 European Ancestry) and YRI (Yoruba in Ibadan, Nigeria) samples from the 1000

310 Genomes Project Phase 3(Consortium, 2015), which we downloaded from
311 (https://mathgen.stats.ox.ac.uk/impute/1000GP%20Phase%203%20haplotypes%2
312 06%20October%202014.html).

313 We first extracted all variants that were biallelic in our human (CEU+YRI),
314 Neandertal, and chimpanzee alignments. We considered as Neandertal-like sites
315 (NLS) only those variants where the African sample (YRI) had a derived allele
316 frequency of zero and both CEU and Neandertal carried the derived
317 allele. Finally, we required the derived allele to be present in CEU in at least two
318 copies in order to calculate our haplotype-based test statistic ($H_{D/A}$).

319 *2. Demographic model and neutral coalescent simulations*

320 We performed coalescent simulations of the demographic history of the European,
321 African, and Neandertal populations applied by Vernot and Akey(2015) based on
322 previously inferred demographic models (Gravel et al., 2011; Tennessen et al.,
323 2012) with some minor adjustments, including removing the East Asian
324 component of the model and allowing for Neandertal admixture to occur over a
325 time span from 500 to 1000 years (Figure S1). Simulations were performed with
326 macs(Chen et al., 2009) in order to explicitly simulate the genetic map
327 (downloaded with the 1000 Genomes samples at the link above) of the 2Mb region
328 centered on OAS (chr12:112400000-114400000). All simulations were performed
329 with the parameters specified in Figure S1, assuming 25 years per generation and a
330 mutation rate of 2.5×10^{-8} per bp per generation. A sample macs command is
331 given at the end of this section.

332 Data were thinned in a manner similar to Sankararaman et al. 2014 (2014) to
333 account for imperfect SNP ascertainment in the 1000 Genomes dataset, such that
334 SNPs with minor allele count of 1, 2, 3, 4, 5, 6, 7, 8, 9, and >=10 were accepted
335 with probabilities 0.25, 0.5, 0.75, 0.8, 0.9, 0.95, 0.96, 0.97, 0.98, and 0.99,
336 respectively. Additionally, we only kept SNPs that were polymorphic in the
337 simulated CEU sample. Finally, we performed an additional thinning of SNPs
338 with uniform probability of 0.05 of removal to account for slightly elevated SNP
339 density in the simulated data. The resulting simulated datasets had an average
340 SNP density of 4.9 SNPs per kb compared to 3.6 in the real data. This is a slightly
341 larger than ideal difference in SNP density, but we note that neither derived allele
342 frequency, nor our primary haplotype-based test statistic (described below) should
343 be particularly sensitive to SNP density**.** In fact, Figure S6 illustrates that our
344 statistic is conservative with respect to SNP density.

345 To estimate the probabilities of adaptive introgression conditional on the simulated
346 locus containing a NLS (SNPs in which the non-African and Neandertal sample
347 contain a derived allele that is absent from the African sample) we randomly
348 selected a single NLS from the central fifth of the chromosome (positions 0.4-0.6
349 in the unit scaled locus) in which we calculated our test statistics. If an NLS was
350 found, its position on the chromosome was recorded and test statistics were
351 calculated. If no NLS was found, we repeated the simulation step. This approach
352 allowed us to estimate under our demographic model, the proportion of
353 independent loci that will have no obvious signal of introgression under neutrality
354 as the proportion of successful simulation runs to the total number of attempts. We

obtained an approximate 21 percent chance of observing Neandertal introgression at the OAS locus under a neutral model. Additionally, we repeated this simulation approach, setting Neandertal introgression to zero, in a set of 10,000 simulations in order to estimate the likelihood of observing a NLS in the case of incomplete lineage sorting (ILS), to estimate the contribution of ILS to our null distribution. We observed a roughly 0.6% chance that a NLS is produced by ILS at the OAS locus, suggesting that our null distribution is not significantly biased by ILS.

Sample macs command:

```
macs 416 2000000 -t 0.000731 -R oas_recrates.txt -I 3 216 2 198 0 -n 1
58.0027359781 -n 2 0.205198358413 -n 3 70.0410396717 -eg 0 1 482.67144247 -
eg 1e-08 3 570.175408787 -em 2e-08 1 3 0.409323665059 -em 3e-08 3 1
0.409323665059 -eg 0.00699726402189 3 16.9958723831 -en 0.00699727402189
1 1.98002735978 -en 0.03146374829 3 0.759313629055 -en 0.0561902522706 3
1.15461721321 -em 0.0561902622706 1 3 4.386 -em 0.0561902722706 3 1 4.386
-em 0.056088019083 3 2 17.3006865925 -em 0.0569023606299 3 2 0.0 -ej
0.08129305892 3 1 -en 0.08129306892 1 1.98002735978 -en 0.202462380301 1
1.0 -ej 0.89014331235 2 1 -h 1e3 1e3
```

3. Frequency and haplotype-based tests of neutrality

We examined the consistency of genetic variation with our neutral model using several approaches. First, we examined the likelihood of observing (Neandertal) allele frequencies as high as the OAS locus. Under neutrality, allele frequency is

376 not dependent upon recombination rate, therefore, we can estimate the likelihood
377 of our observed NLS frequency in the OAS region. For example, mean allele
378 frequency of all NLS observed between chr12:113344000-113450000 is 0.35
379 (Figure 2A). Frequencies this great or greater were observed in our simulations
380 less than one percent of the time (0.0085), suggesting that this haplotype
381 frequency is unlikely under neutrality.

382 Additionally, we wanted to examine if the haplotypes carrying NLS at the OAS
383 locus are longer than expected under neutrality when conditioning on the observed
384 frequencies and the underlying genetic map, which would provide an additional
385 signature of selection on introgressed haplotypes. For this purpose, we modified a
386 simple haplotype statistic *H*(Schlamp et al., 2016), which measures the average
387 length of pairwise homozygosity tracts in base pairs – a quantity that is very
388 straightforward to interpret. As selective sweeps are expected to create long
389 haplotypes around the selected site, the *H* statistic should be higher in samples
390 containing positively selected haplotypes compared to samples containing
391 neutrally evolving haplotypes, when frequency and recombination are properly
392 controlled, similar to other statistics based on haplotype lengths, such as *EHH,*
393 *iHS,* and *nSL* (Sabeti, P.C. et al., 2002; Voight et al., 2006; Ferrer-Admetlla et al.,
394 2014). However, in contrast to these other statistics, *H* does not require
395 specification of analysis parameters such as minimum haplotype homozygosity
396 levels below which haplotypes are no longer extended.

397 Under adaptive introgression, we specifically expect the introgressed haplotypes to
398 be longer than the ancestral haplotypes. We therefore defined our test statistic,
399 $H_{D/A}$, as:

400 $$H_{D/A} = \ln(H_D/H_A);$$

401 where H_D and H_A are H calculated across haplotypes carrying derived NLS allele
402 versus the ancestral allele.

403 We calculated this statistic for every simulated NLS and for every NLS in our true
404 sample. We compared our simulation results to the observed data in two
405 ways. First, for each observed SNP we gathered all simulated NLS within 0.025
406 frequency and 20 kb. We then calculate the empirical likelihood of observing
407 each observed NLS under our neutral demographic model (Figure 2A). All NLS
408 within the bounds of the three OAS genes are tightly linked, therefore, the peak
409 signal at the center of this locus suggests that the probability of observing these
410 $H_{D/A}$ values are less than one percent.

411 However, to examine a single hypothesis test of the null hypothesis, that the true
412 Neandertal haplotype is not significantly longer than expected under neutrality, we
413 performed a randomization test. First, we calculated mean $H_{D/A}$ across this central
414 locus (chr12:113344000-113450000). Next, we randomly chose one control SNP
415 per true NLS (using the same matching conditions above) and calculated the mean
416 of the control set. We repeated this procedure 100,000 times to create a normally

417 distributed null distribution, which we then used to calculate a Z-score (4.5) and
418 corresponding P (3.35×10^{-6}) for our one-sided hypothesis test (Figure 2B).

419 Finally, we asked how unusual our observation of NLS frequency and $H_{D/A}$ scores
420 were across all simulations. Across approximately one million simulations
421 (1,000,079), we only observed 413 simulations that produced an NLS with a
422 frequency and $H_{D/A}$ score as great or greater than the highest value observed
423 among the true NLS (Figure 2C). In other words, the likelihood of our true sample
424 under the neutral demographic scenario is less than 5×10^{-4}.

425 *4. Analysis of ancient Eurasian data*

426 We utilized supplementary data table 3 from Mathieson et al. (2015). This table
427 includes maximum likelihood allele frequency estimates for three ancient
428 population samples (HG- Hunter-gatherer, EF- Early farmer, SA- Steppe ancestry)
429 and four present day European samples from the 1,000 Genomes Project (see
430 "Genome-wide scan for selection" section of methods in(Mathieson et al., 2015)).
431 We intersect this table with allele frequencies for 1,000 Genomes Yorubans (YRI)
432 and the Altai Neandertal genotypes and only analyze sites for which we have data
433 for all samples (1,004,612 SNPs).

434 To calculate the expected allele frequency in modern samples under drift, we used
435 the estimated proportions (m) of (HG, EF, SA) in each of the four present-day
436 samples estimated by Mathieson et al. (2015): CEU = (0.196, 0.257, 0.547), GBR

437 = (0.362, 0.229, 0.409), IBS = (0, 0.686, 0.314) and TSI = (0, 0.645, 0.355). We
438 calculated the expected frequency E[*p*] of site as:

$$E[p] = \sum_{i}^{\{HG,EF,SA\}} (p_i \, x \, m_i)$$

439

440 Next, we calculated the absolute difference between observed and expected allele
441 frequency in all four present-day European samples at all available sites. To test
442 the null hypothesis that OAS NLS have not changed in frequency more than
443 expected under neutrality at 11 SNPs in the central OAS region
444 (chr12:113200000-113600000), we first calculated the fraction of all autosomal
445 SNPs in the dataset at similar present-day frequency (within 1 percent in the
446 folded frequency spectrum) of each OAS NLS. Finally, we calculated the fraction
447 of autosomal SNPs with an absolute observed minus expected frequency
448 difference greater than or equal to the OAS NLS. These results are given in Table
449 1 and illustrated in Figure 2A.

450 This test does not explicitly incorporate variance in estimated ancient allele
451 frequency. However, any bias in ancient allele frequency estimation should be
452 distributed randomly across the genome. Therefore, our comparison to a genome-
453 wide distribution of SNPs at similar present-day frequency should incorporate
454 most of this error. Nonetheless, the selection test performed by Mathieson et
455 al.(2015) does incorporate such error, so we can also look to the Ps from that test

456 to ensure consistency with our results. In fact, when considering OAS as a single
457 locus, these results are highly significant and consistent with ours (see Table S1).

458 *5. Estimation of selection coefficients*

459 To estimate the selection coefficient s under constant positive selection for a given
460 starting frequency (x_0), final frequency (x_1), and number of generations between
461 these estimates (Δt), we assumed a model of standard logistic growth of a
462 codominant allele:

463 $x_1 = x_0/[x_0+(1- x_0)\exp(-s\Delta t)]$.

464 This equation can be easily solved to obtain s, given x_0 , x_1, and Δt.

465 *6. Sample collection*

466 Buffy coats from 99 healthy European-descent donors were obtained from Indiana
467 Blood Center (Indianapolis, IN, USA). Only individuals self-reported as currently
468 healthy and not under medication were included in the study. The project was
469 approved by the ethics committee at the CHU Sainte-Justine (protocol #4022). The
470 individuals recruited in this study were males aged 18 to 55 years old.

471 *7. DNA Extraction and genotyping*

472 DNA from each of the blood donors was extracted using the Gentra Pure Gene
473 blood kit (Qiagen). Genotyping of each individual was then performed by
474 Illumina's HumanOmni5Exome bead Chip array and complemented with imputed

475 data from the 1000 Genomes data using Impute2(Howie et al., 2009). Here, we
476 only focused on genetic diversity surrounding the OAS region –
477 chr12:113229549-113574044 (~ 344Kb) spanning from the beginning of *RPH3A*
478 to the end of *RASAL1* – for a total of 673 SNPs with a MAF above 10%.

479 *8. Isolation of monocytes and differentiation of macrophages*

480 Blood mononuclear cells were isolated by Ficoll-Paque centrifugation. Monocytes
481 were purified from peripheral blood mononuclear cells (PBMCs) by positive
482 selection with magnetic CD14 MicroBeads (Miltenyi Biotech) using the
483 autoMACS Pro Separator. All samples had purity levels above 90%, as measured
484 by flow cytometry using an antibody against CD14 (BD Biosciences). Monocytes
485 were then cultured for 7 days in RPMI-1640 (Fisher) supplemented with 10%
486 heat-inactivated FBS (FBS premium, US origin, Wisent), L-glutamine (Fisher)
487 and M-CSF (20ng/mL; R&D systems). Cell cultures were fed every 2 days with
488 complete medium supplemented with the cytokines previously mentioned. Before
489 infection, we systematically verified that the differentiated macrophages presented
490 the expected phenotype for non-activated macrophages (CD1a+, CD14+, CD83−,
491 and HLA-DRlow (BD Biosciences)).

492 *9. Bacterial preparation and infection of macrophages*

493 The day prior to infection, aliquots of *Salmonella typhimurium* (Keller strain) were
494 thawed and bacteria were grown overnight in Tryptic Soy Broth (TSB) medium.
495 Bacterial culture was diluted to mid-log phase prior to infection and supernatant

496 density was checked at OD600. Monocyte-derived macrophages were then
497 infected with *Salmonella typhimurium* at a multiplicity of infection (MOI) of 10:1.
498 A control group of non-infected macrophages was treated the same way but using
499 medium without bacteria. After 2 hours in contact with the bacteria, macrophages
500 were washed and cultured for another hour in the presence of 50 mg/ml of
501 gentamycin in order to kill all extracellular bacteria present in the medium. The
502 cells were then washed a second time and cultured in complete medium with 3
503 mg/ml gentamycin for an additional 2 hours, the time point to which we refer in
504 the main text.

505 *10. Infection/stimulation of PBMC*

506 PBMCs from a subset of 30 individuals used to derive macrophages were cultured
507 in RPMI-1640 (Fisher) supplemented with 10% heat-inactivated FBS (FBS
508 premium, US origin, Wisent) and 1% L-glutamine (Fisher). The 30 individuals
509 were chosen based on their genotype for kgp4570197, a SNP which derived allele
510 is of Neandertal origin and that we used as a proxy to identify individuals
511 harbouring the Neandertal haplotype in the OAS region. From the 30 individuals,
512 and based on this SNP, 9 individuals were homozygous for the Neandertal
513 haplotype, 9 were heterozygous, and 12 homozygous for the modern human
514 sequence.

515 For each of the tested individuals, PBMCs (1 million per condition) were
516 stimulated/infected with one of the following viral-associated immune challenges:
517 polyI:C (10 µg/ml, TLR3 agonist), gardiquimod (0.5µg/ml, TLR7 and TLR8

518 agonist), Influenza PR8 WT (multiplicity of infection (MOI) of 0.05:1), Herpes
519 simplex virus (HSV) 1 ($1.55x10^2$ CPE), and HSV2 ($19.5x10^4$ CPE). PBMCs were
520 stimulated/infected for 4 hours with TLR ligands and Influenza, and 6h with
521 HSV1 and HSV2. A control group of non-infected PBMC was treated the same
522 way but with only medium.

523 *11. RNA extraction, RNA-seq library preparation, and sequencing*

524 Total RNA was extracted from the non-infected and infected/stimulated cells
525 using the miRNeasy kit (Qiagen). RNA quantity was evaluated
526 spectrophotometrically, and the quality was assessed with the Agilent 2100
527 Bioanalyzer (Agilent Technologies). Only samples with no evidence of RNA
528 degradation (RNA integrity number > 8) were kept for further experiments. RNA-
529 sequencing libraries were prepared using the Illumina TruSeq protocol. Once
530 prepared, indexed cDNA libraries were pooled (6 libraries per pool) in equimolar
531 amounts and sequenced with single-end 100bp reads on an Illumina HiSeq2500.
532 Results based on the entire dataset are described elsewhere (Nédélec *et al*., under
533 revision). Here, we only studied transcript-level and gene-level expression
534 estimates for *OAS1*, *OAS2* and *OAS3*.

535 *12. Quantifying gene expression values from RNA-seq data*

536 Adaptor sequences and low quality score bases (Phred score < 20) were first
537 trimmed using Trim Galore (version 0.2.7). The resulting reads were then mapped
538 to the human genome reference sequence (Ensembl GRCh37 release 65) using

539 TopHat (version 2.0.6) and using a hg19 transcript annotation database
540 downloaded from UCSC. Gene-level expression estimates were calculated using
541 featureCounts (version 1.4.6-p3) and transcript-level expression values were
542 obtained using RSEM under default parameters.

543 *13. Quantitative real time PCR*

544 For the PBMC samples we measured the expression levels of OAS and interferon
545 genes using real time PCR. 100ng of high-quality RNA was reverse-transcribed
546 into cDNA using the qScript cDNA SuperMix (Quanta Biosciences). Quantitative
547 real time PCR was performed using 96.96 Dynamic Array™ IFCs and the
548 BioMark™ HD System from Fluidigm. For the TaqMan gene assays, we used the
549 following TaqMan Gene Expression Assay (Applied BioSystems) to quantify the
550 expression levels of interferon genes: *IFNA1* (Hs03044218), *IFNA6*
551 (Hs00819627), and *IFNG* (Hs00989291). To quantify the overall expression levels
552 of OAS genes, we used probes that capture all common isoforms of *OAS1*
553 (Hs00973635), *OAS2* (Hs00942643), and OAS3 (Hs00196324). Custom-made
554 probes were designed to specifically target the short-isoform of *OAS2* (Forward
555 Primer Sequence CTGCAGGAACCCGAACAGTT; Reverse Primer Sequence
556 ACTCATGGCCTAGAGGTTGCA; Reporter Sequence
557 AGAGAAAAGCCAAAGAA). As housekeeping genes we used: *GAPDH*
558 (Hs02758991), *GUSB* (Hs99999908), *HPRT1* (Hs99999909), and *POLR2A*
559 (Hs00172187). The results reported in the manuscript used *POLR2A* as a reference
560 but all conclusions remain unchanged when using any of the other housekeeping
561 genes.

562 We start by doing a preamplification of the cDNA using the PreAmp Master Mix
563 (Fluidigm). Preamplified cDNA was then diluted 2X on a solution of 10 mM Tris–
564 HCl (pH 8.0) and 0.1 mM EDTA. In order to prepare samples for loading into the
565 integrated fluid circuit (IFC), a mix was prepared consisting of 360 μL TaqMan
566 Fast Advanced Master Mix (Applied BioSystems) and 36 μL 20× GE Sample
567 Loading Reagent (Fluidigm). 2.75 μL of this mix was dispensed to each well of a
568 96-well assay plate and mixed with 2.25 μL of preamplified cDNA. Following
569 priming of the IFC in the IFC Controller HX, 5 μL of the mixture of cDNA and
570 loading reagent were dispensed in each of the sample inlet of the 96.96 IFC. For
571 the TaqMan gene assays, 5 μL of mixes consisting of 2.5 μL 20× TaqMan Gene
572 Expression Assay (Applied BioSystems) and 2.5 μL 2X Assay Loading Reagent
573 (Fluidigm) were dispensed to each detector inlet of the 96.96 IFC. After loading
574 the assays and samples into the IFC in the IFC Controller HX, the IFC was
575 transferred to the BioMark HD and PCR was performed using the thermal protocol
576 GE 96 × 96 Fast v1.pcl. This protocol consists of a Thermal Mix of 70 °C, 30 min;
577 25 °C, 10 min, Hot Start at 95 °C, 1 min, PCR Cycle of 35 cycles of (96 °C, 5 s;
578 60 °C, 20 s). Data was analysed using Fluidigm Real-Time PCR Analysis software
579 using the Linear (Derivative) Baseline Correction Method and the Auto
580 (Detectors) Ct Threshold Method.

581 To quantify the expression levels of the *OAS1* isoform associated with the derived
582 allele at the splicing variant rs10774671 we used SybrGreen and the following
583 forward (GCTGAGGCCTGGCTGAATTA), and reverse
584 (CCACTTGTTAGCTGATGTCCTTGA) primers. PCR was performed using the

585 thermal protocol 50 °C, 2 min; 95 °C, 10 min, PCR Cycle of 40 cycles of (95 °C,
586 15 s; 60 °C, 1 min). A melting curve was also performed to check for non-specific
587 amplification.

588 *14. Genotype–Phenotype Association Analysis*

589 eQTL, asQTL were performed against *OAS1*, *OAS2* and *OAS3*. We examined
590 associations between SNP genotypes and the phenotype of interest using a linear
591 regression model, in which phenotype was regressed against genotype. In
592 particular, expression levels were considered as the phenotype when searching for
593 eQTL and the percentage usage of each isoform in each gene when mapping
594 asQTL. To avoid low power caused by rare variants, only SNPs in the OAS region
595 with a minor allele frequency of 10% across all individuals were tested (i.e., 673
596 SNPs within the region chr12:113229549-113574044). In all cases, we assumed
597 that alleles affected the phenotype in an additive manner. For the eQTL and
598 asQTL analyses on macrophages we mapped Salmonella-infected, and non-
599 infected samples separately. For the non-infected and infected/stimulated PBMCs
600 we only tested expression levels against the SNPs identified as eQTL or asQTL in
601 the macrophage data (specifically, the SNPs for which boxplots are shown in
602 Figure 3).

603 **Author Contributions:** Conception and design: AJS, PWM, LBB; Acquisition of
604 data: AJS, AD, YN, VY, LBB; Analysis and interpretation of data: AJS, AD, YN,
605 PWM, LBB; Contributed unpublished, essential data, or reagents: CA, JET;
606 Drafting or revising the article: AJS, PWM, LBB.

Acknowledgements: We thank all members from the Barreiro and Messer lab for helpful discussions and comments on the manuscript. We thank Dr Silvia Vidal for the gift of the Influenza PR8 WT used in this study, Dr Hugo Soudeyns and Doris Ransy for advice with the viral infections, and Marc Montero and George (PJ) Perry for sharing their analysis on nucleotide diversity levels of OAS genes in non-human primate species.

Funding Statement: This study was funded by grants from the Canadian Institutes of Health Research (301538 and 232519), the Human Frontiers Science Program (CDA-00025/2012) and the Canada Research Chairs Program (950-228993) (to L.B.B.). Y.N. was supported by a fellowship from the *Réseau de Médecine Génétique Appliquée* (RMGA).

References:

Awady, El, M.K., Anany, M.A., Esmat, G., Zayed, N., Tabll, A.A., Helmy, A., Zayady, El, A.R., Abdalla, M.S., Sharada, H.M., Raziky, El, M., Akel, El, W., Abdalla, S., Bader El Din, N.G., 2011. Single nucleotide polymorphism at exon 7 splice acceptor site of OAS1 gene determines response of hepatitis C virus patients to interferon therapy. Journal of Gastroenterology and Hepatology. 26, 843–850.

Bigham, A.W., Buckingham, K.J., Husain, S., Emond, M.J., Bofferding, K.M., Gildersleeve, H., Rutherford, A., Astakhova, N.M., Perelygin, A.A., Busch, M.P., Murray, K.O., Sejvar, J.J., Green, S., Kriesel, J., Brinton, M.A., Bamshad, M., 2011. Host Genetic Risk Factors for West Nile Virus Infection and Disease Progression. PLoS ONE. 6, e24745.

Bonnevie-Nielsen, V., Field, L.L., Lu, S., Zheng, D.-J., Li, M., Martensen, P.M., Nielsen, T.B., Beck-Nielsen, H., Lau, Y.L., Pociot, F., 2005. 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. The American Journal of Human Genetics. 76, 623–633.

Chen, G.K., Marjoram, P., Wall, J.D., 2009. Fast and flexible simulation of DNA

638 sequence data. Genome Research. 19, 136–142.
639 Consortium, T.1.G.P., 2015. A global reference for human genetic variation.
640 Nature. 526, 68–74.
641 Deschamps, M., Laval, G., Fagny, M., Itan, Y., Abel, L., Casanova, J.-L., Patin,
642 E., Quintana-Murci, L., 2016. Genomic Signatures of Selective Pressures and
643 Introgression from Archaic Hominins at Human Innate Immunity Genes.
644 American Journal of Human Genetics. 98, 5–21.
645 Ferrer-Admetlla, A., Liang, M., Korneliussen, T., Nielsen, R., 2014. On detecting
646 incomplete soft or hard selective sweeps using haplotype structure. Molecular
647 Biology and Evolution. 31, 1275–1291.
648 Fu, Q., Li, H., Moorjani, P., Jay, F., Slepchenko, S.M., Bondarev, A.A., Johnson,
649 P.L.F., Aximu-Petri, A., Prüfer, K., de Filippo, C., Meyer, M., Zwyns, N.,
650 Salazar-García, D.C., Kuzmin, Y.V., Keates, S.G., Kosintsev, P.A., Razhev,
651 D.I., Richards, M.P., Peristov, N.V., Lachmann, M., Douka, K., Higham,
652 T.F.G., Slatkin, M., Hublin, J.-J., Reich, D., Kelso, J., Viola, T.B., Pääbo, S.,
653 2014. Genome sequence of a 45,000-year-old modern human from western
654 Siberia. Nature. 514, 445–449.
655 Gravel, S., Henn, B.M., Gutenkunst, R.N., Indap, A.R., Marth, G.T., Clark, A.G.,
656 Yu, F., Gibbs, R.A., Bustamante, C.D., Project, 1.G., 2011. Demographic
657 history and rare allele sharing among human populations. Proceedings of the
658 National Academy of Sciences of the United States of America. 108, 11983–
659 11988.
660 Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M.,
661 Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., 2010. A draft sequence of the
662 Neandertal genome. Science. 328, 710–722.
663 Gulig, P.A., Doyle, T.J., Clare-Salzler, M.J., Maiese, R.L., Matsui, H., 1997.
664 Systemic infection of mice by wild-type but not Spv- Salmonella typhimurium
665 is enhanced by neutralization of gamma interferon and tumor necrosis factor
666 alpha. Infection and Immunity. 65, 5191–5197.
667 Howie, B.N., Donnelly, P., Marchini, J., 2009. A Flexible and Accurate Genotype
668 Imputation Method for the Next Generation of Genome-Wide Association
669 Studies. PLoS Genetics. 5, e1000529.
670 Hublin, J.J., 2009. The origin of Neandertals. Proceedings of the National
671 Academy of Sciences. 106, 16022–16027.
672 Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M.,
673 Haussler, D., 2002. The human genome browser at UCSC. Genome Research.
674 12, 996–1006.
675 Kwon, Y.-C., Kang, J.-I., Hwang, S.B., Ahn, B.-Y., 2012. The ribonuclease l-
676 dependent antiviral roles of human 2′,5′-oligoadenylate synthetase family
677 members against hepatitis C virus. FEBS letters. 587, 156–164.
678 LaRock, D.L., Chaudhary, A., Miller, S.I., 2015. Salmonellae interactions with
679 host processes. Nature Reviews Microbiology. 13, 191–205.
680 Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K.,
681 Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., Berger, B.,
682 Economou, C., Bollongino, R., Fu, Q., Bos, K.I., Nordenfelt, S., Li, H., de
683 Filippo, C., Prüfer, K., Sawyer, S., Posth, C., Haak, W., Hallgren, F.,

Fornander, E., Rohland, N., Delsate, D., Francken, M., Guinet, J.-M., Wahl, J., Ayodo, G., Babiker, H.A., Bailliet, G., Balanovska, E., Balanovsky, O., Barrantes, R., Bedoya, G., Ben-Ami, H., Bene, J., Berrada, F., Bravi, C.M., Brisighelli, F., Busby, G.B.J., Cali, F., Churnosov, M., Cole, D.E.C., Corach, D., Damba, L., van Driem, G., Dryomov, S., Dugoujon, J.-M., Fedorova, S.A., Romero, I.G., Gubina, M., Hammer, M., Henn, B.M., Hervig, T., Hodoglugil, U., Jha, A.R., Karachanak-Yankova, S., Khusainova, R., Khusnutdinova, E., Kittles, R., Kivisild, T., Klitz, W., Kucinskas, V., Kushniarevich, A., Laredj, L., Litvinov, S., Loukidis, T., Mahley, R.W., Melegh, B., Metspalu, E., Molina, J., Mountain, J., Näkkäläjärvi, K., Nesheva, D., Nyambo, T., Osipova, L., Parik, J., Platonov, F., Posukh, O., Romano, V., Rothhammer, F., Rudan, I., Ruizbakiev, R., Sahakyan, H., Sajantila, A., Salas, A., Starikovskaya, E.B., Tarekegn, A., Toncheva, D., Turdikulova, S., Uktveryte, I., Utevska, O., Vasquez, R., Villena, M., Voevoda, M., Winkler, C.A., Yepiskoposyan, L., Zalloua, P., Zemunik, T., Cooper, A., Capelli, C., Thomas, M.G., Ruiz-Linares, A., Tishkoff, S.A., Singh, L., Thangaraj, K., Villems, R., Comas, D., Sukernik, R., Metspalu, M., Meyer, M., Eichler, E.E., Burger, J., Slatkin, M., Pääbo, S., Kelso, J., Reich, D., Krause, J., 2014. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature. 513, 409–413.

Lim, J.K., Lisco, A., McDermott, D.H., Huynh, L., Ward, J.M., Johnson, B., Johnson, H., Pape, J., Foster, G.A., Krysztof, D., Follmann, D., Stramer, S.L., Margolis, L.B., Murphy, P.M., 2009. Genetic Variation in OAS1 Is a Risk Factor for Initial Infection with West Nile Virus in Man. PLoS Pathogens. 5, e1000321.

Llorente, M.G., Jones, E.R., Eriksson, A., Siska, V., Arthur, K.W., Arthur, J.W., Curtis, M.C., Stock, J.T., Coltorti, M., Pieruccini, P., Stretton, S., Brock, F., Higham, T., Park, Y., Hofreiter, M., G, B.D., Bhak, J., Pinhasi, R., Manica, A., 2015. Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent. Science.

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., Sirak, K., Gamba, C., Jones, E.R., Llamas, B., Dryomov, S., Pickrell, J., Arsuaga, J.L., de Castro, J.M.B., Carbonell, E., Gerritsen, F., Khokhlov, A., Kuznetsov, P., Lozano, M., Meller, H., Mochalov, O., Moiseyev, V., Guerra, M.A.R., Roodenberg, J., Vergès, J.M., Krause, J., Cooper, A., Alt, K.W., Brown, D., Anthony, D., Lalueza-Fox, C., Haak, W., Pinhasi, R., Reich, D., 2015. Genome-wide patterns of selection in 230 ancient Eurasians. Nature.

Mendez, F.L., Watkins, J.C., Hammer, M.F., 2012. A Haplotype at STAT2 Introgressed from Neanderthals and Serves as a Candidate of Positive Selection in Papua New Guinea. The American Journal of Human Genetics. 91, 265–274.

Mendez, F.L., Watkins, J.C., Hammer, M.F., 2013. Neandertal origin of genetic variation at the cluster of OAS immunity genes. Molecular Biology and Evolution. 30, 798–801.

Meyer, M., Arsuaga, J.L., de Filippo, C., Nagel, S., Aximu-Petri, A., Nickel, B.,

730 Martínez, I., Gracia, A., de Castro, J.M.B., Carbonell, E., Viola, B., Kelso, J.,
731 Prüfer, K., Pääbo, S., 2016. Nuclear DNA sequences from the Middle
732 Pleistocene Sima de los Huesos hominins. Nature.
733 Meyer, M., Kircher, M., Gansauge, M.T., Li, H., Racimo, F., Mallick, S.,
734 Schraiber, J.G., Jay, F., Prufer, K., de Filippo, C., 2012. A high-coverage
735 genome sequence from an archaic denisovan individual. Science. 338, 222–
736 226.
737 Nauciel, C., Espinasse-Maes, F., 1992. Role of Gamma Interferon and Tumor-
738 Necrosis-Factor-Alpha in Resistance to Salmonella-Typhimurium Infection.
739 Infection and Immunity. 60, 450–454.
740 Player, M.R., Torrence, P.F., 1998. The 2–5 A system: Modulation of viral and
741 cellular processes through acceleration of RNA degradation. Pharmacology &
742 Therapeutics. 78, 55–113.
743 Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S.,
744 Heinze, A., Renaud, G., Sudmant, P.H., de Filippo, C., Li, H., Mallick, S.,
745 Dannemann, M., Fu, Q., Kircher, M., Kuhlwilm, M., Lachmann, M., Meyer,
746 M., Ongyerth, M., Siebauer, M., Theunert, C., Tandon, A., Moorjani, P.,
747 Pickrell, J., Mullikin, J.C., Vohr, S.H., Green, R.E., Hellmann, I., Johnson,
748 P.L.F., Blanche, H., Cann, H., Kitzman, J.O., Shendure, J., Eichler, E.E., Lein,
749 E.S., Bakken, T.E., Golovanova, L.V., Doronichev, V.B., Shunkov, M.V.,
750 Derevianko, A.P., Viola, B., Slatkin, M., Reich, D., Kelso, J., Pääbo, S., 2014.
751 The complete genome sequence of a Neanderthal from the Altai Mountains.
752 Nature. 505, 43–49.
753 Racimo, F., Marnetto, D., Huerta-Sanchez, E., 2016. The landscape of uniquely
754 shared archaic alleles in present-day human populations. bioRxiv. 045237.
755 Racimo, F., Sankararaman, S., Nielsen, R., Huerta-Sanchez, E., 2015. Evidence
756 for archaic adaptive introgression in humans. Nature Reviews Genetics.
757 Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola,
758 B., Briggs, A.W., Stenzel, U., Johnson, P.L.F., 2010. Genetic history of an
759 archaic hominin group from Denisova Cave in Siberia. Nature. 468, 1053–
760 1060.
761 Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z.P., Richter, D.J., Schaffner,
762 S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J., 2002.
763 Detecting recent positive selection in the human genome from haplotype
764 structure. Nature. 419, 832–837.
765 Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., Kelso, J., Pääbo, S.,
766 Patterson, N., Reich, D., 2014. The genomic landscape of Neanderthal
767 ancestry in present-day humans. Nature. 507, 354–357.
768 Sankararaman, S., Mallick, S., Patterson, N., Reich, D., 2016. The Combined
769 Landscape of Denisovan and Neanderthal Ancestry in Present-Day Humans.
770 Current Biology.
771 Sawyer, S., Renaud, G., Viola, B., Hublin, J.-J., Gansauge, M.-T., Shunkov, M.V.,
772 Derevianko, A.P., Prüfer, K., Kelso, J., Pääbo, S., 2015. Nuclear and
773 mitochondrial DNA sequences from two Denisovan individuals. Proceedings
774 of the National Academy of Sciences. 201519905.
775 Schlamp, F., Made, J., Stambler, R., Chesebrough, L., Boyko, A.R., Messer, P.W.,

2016. Evaluating the performance of selection scans to detect selective sweeps in domestic dogs. Molecular Ecology. 25, 342–356.

Ségurel, L., Quintana-Murci, L., 2014. Preserving immune diversity through ancient inheritance and admixture. Current opinion in immunology. 30C, 79–84.

Tennessen, J.A., Bigham, A.W., O'Connor, T.D., Fu, W., Kenny, E.E., Gravel, S., McGee, S., Do, R., Liu, X., Jun, G., Kang, H.M., Jordan, D., Leal, S.M., Gabriel, S., Rieder, M.J., Abecasis, G., Altshuler, D., Nickerson, D.A., Boerwinkle, E., Sunyaev, S., Bustamante, C.D., Bamshad, M.J., Akey, J.M., Broad GO, Seattle GO, on behalf of the NHLBI Exome Sequencing Project, 2012. Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. Science. 337, 64–69.

Vernot, B., Akey, J.M., 2014. Resurrecting surviving Neandertal lineages from modern human genomes. Science. 343, 1021.

Vernot, B., Akey, J.M., 2015. Complex History of Admixture between Modern Humans and Neandertals. The American Journal of Human Genetics.

Voight, B.F., Kudaravalli, S., Wen, X., Pritchard, J.K., 2006. A map of recent positive selection in the human genome. PLoS Biology. 4, e72.

Yang, M.A., Malaspinas, A.S., Durand, E.Y., Slatkin, M., 2012. Ancient structure in Africa unlikely to explain Neanderthal and non-African genetic similarity. Molecular Biology and Evolution. 29, 2987–2995.

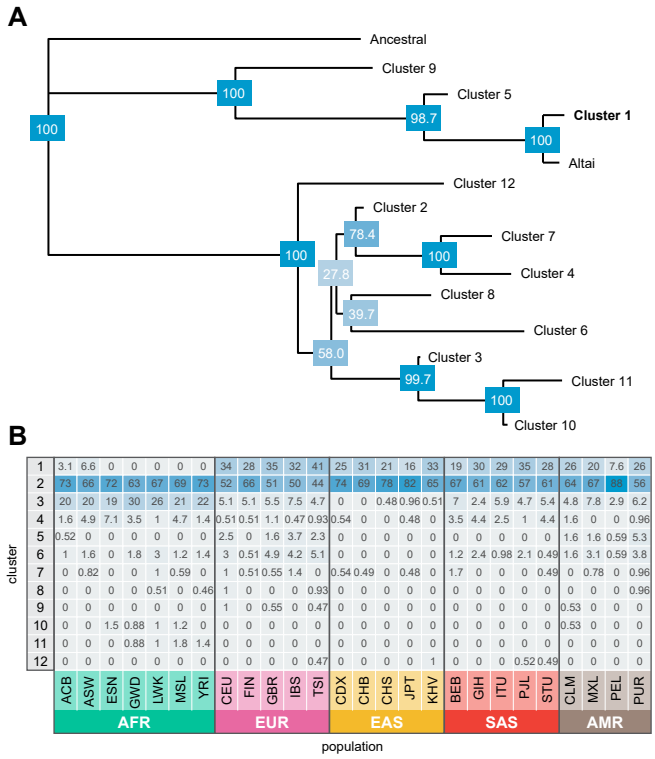


Figure 1: Neandertal introgresssed haplotypes in the OAS region. (A) Neighbor-joining tree of 5008 phased haplotypes spanning chr12: 113344739-113449528 (hg19) from phase 3 of the 1000 genomes project. Haplotypes condensed into 12 core haplotypes based on majority allele in clusters with pairwise differences of 85 or less. The figure illustrates that the Altai haplotype is very similar to "cluster 1" haplotypes found in several human populations. Bootstrap values (1000 replicates) are provided in blue boxes at each node. (B) Frequencies of the 12 core haplotypes within each 1000 genomes project population sample. The most common Neandertal-like haplotype, cluster 1, is found only outside of sub-Saharan African samples, with the exception of recently admixed populations. Population codes can be found at http://www.1000genomes.org/category/population/.

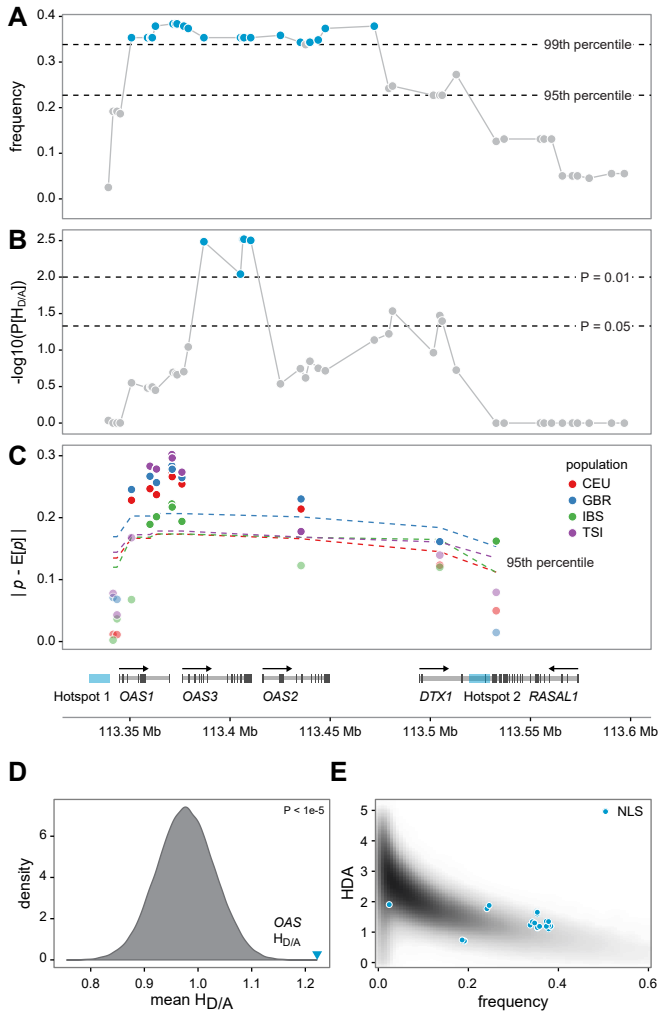


Figure 2: OAS-introgressed haplotypes show multiple signatures of positive selection. (A) Comparison of frequency (y-axis) of NLS in the OAS locus in CEU sample with respect to neutral expectations (dashed lines) based on coalescent simulations. (B) The $H_{D/A}$ values associated with several NLS are significantly larger (-log10 Ps in the y-axis) than expected under a neutral model of evolution. $H_{D/A}$ at each NLS in the OAS region were compared against the distributions of simulated NLS drawn to match local recombination rate and allele frequency. (C) Absolute difference between observed and expected allele frequency in all four present-day European samples (y-axis) based on ancient DNA data. Dashed lines represent the 95th percentile of the expected distribution based on similar deviations calculated on a dataset of approximately one million SNPs scattered around the genome and with comparable present-day frequencies to those found for NLS in the OAS region (D) Comparison of mean $H_{D/A}$ across OAS locus to randomized frequency and genomic position-matched SNPs from simulations. OAS locus is a strong outlier ($p < 10^{-5}$) compared to simulations. (E) Joint density distribution of $H_{D/A}$ and Neandertal-like site derived frequency. Cluster of NLS in OAS are significant outliers to neutral distribution ($p < 5 \times 10^{-4}$).

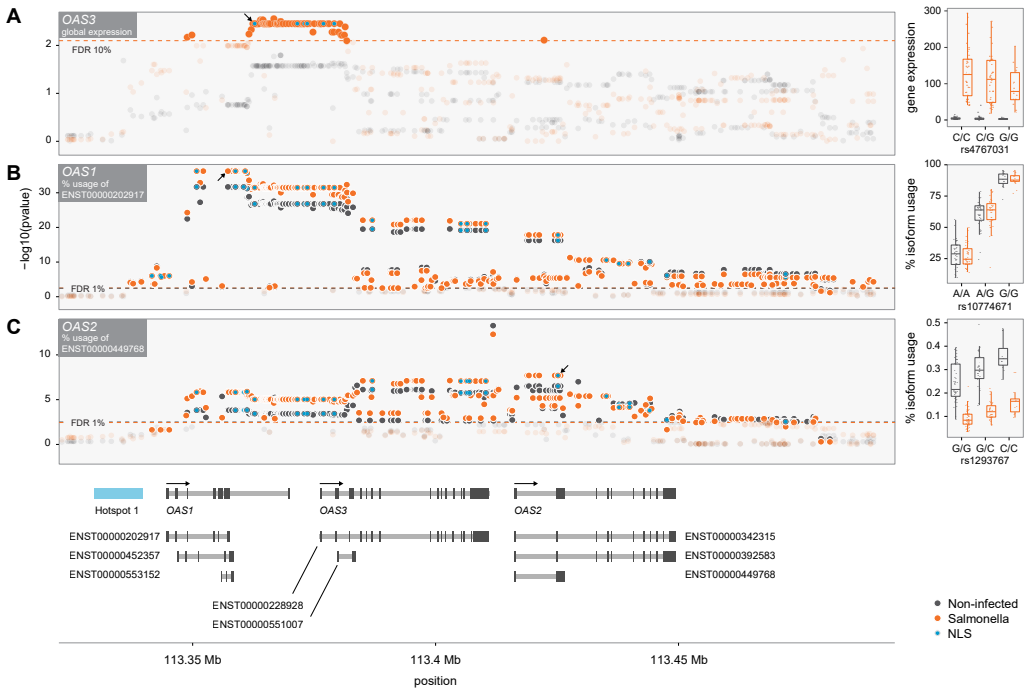


Figure 3: Pervasive impact of the Neandertal haplotype on the regulation of *OAS* genes in primary macrophages. (A) –log 10 Ps (y-axis) for the association between genotypes for SNPs with a MAF > 10% in the OAS region and expression levels of *OAS3* in non-infected (black) and Salmonella-infected macrophages (orange). The dashed line shows the P cutoff corresponding to an FDR of 10%. The right panel shows a boxplot for the association between genotypes at the NLS rs4767031 (x-axis) and the expression levels of *OAS3* (y-axis)(B) –log 10 Ps (y-axis) for the association between genotypes in the OAS regions and the percentage usage of isoform ENST00000202917 (i.e., p46 in the text) in non-infected (black) and Salmonella-infected macrophages (orange). The dashed line shows the P cutoff corresponding to an FDR of 1%. The right panel shows a boxplot for the association between genotypes at the splicing variant rs10774671 (x-axis) and the percentage usage of isoform p46 (y-axis). (C) Similar to (B) but for the percentage usage of isoform ENST00000449768 of OAS2. In all the panels NLS are highlighted by blue dots. The arrows on panel A-C highlight the location of the SNPs for which the boxplots are shown on the right.

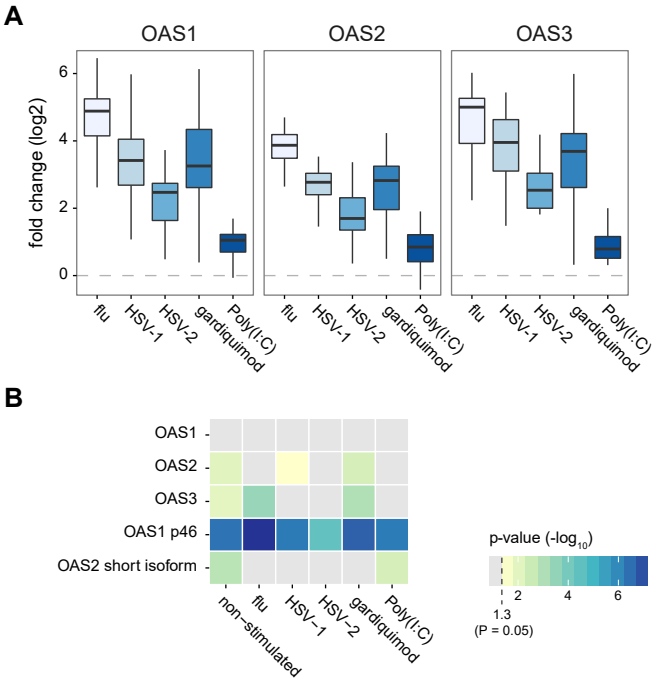
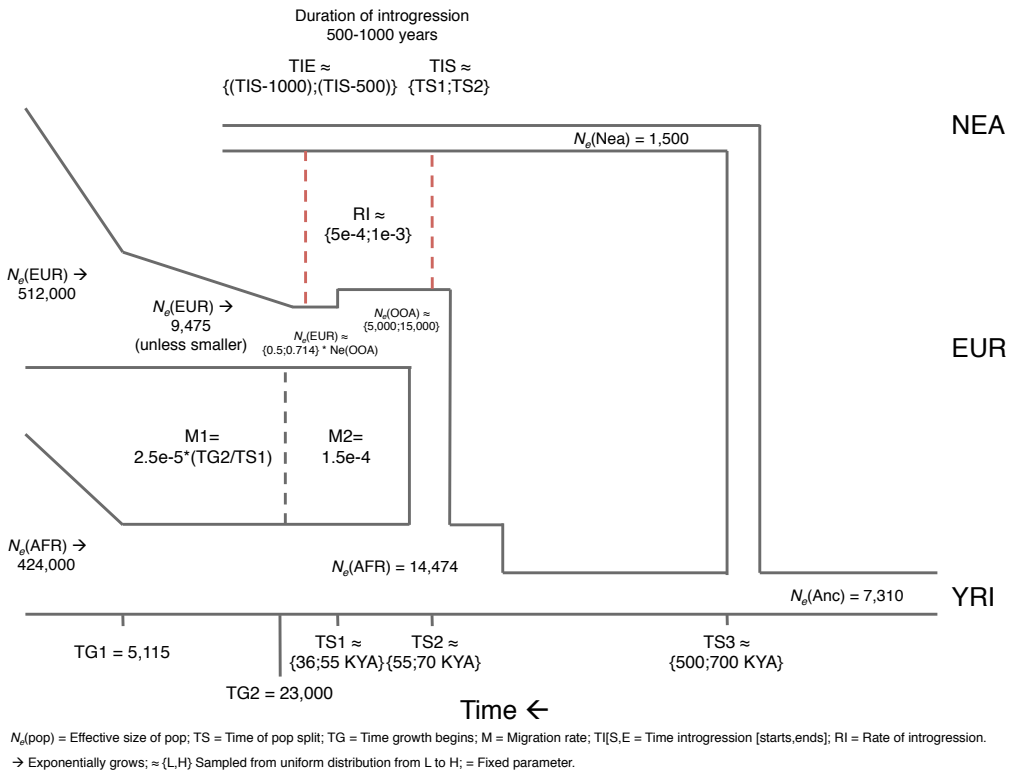
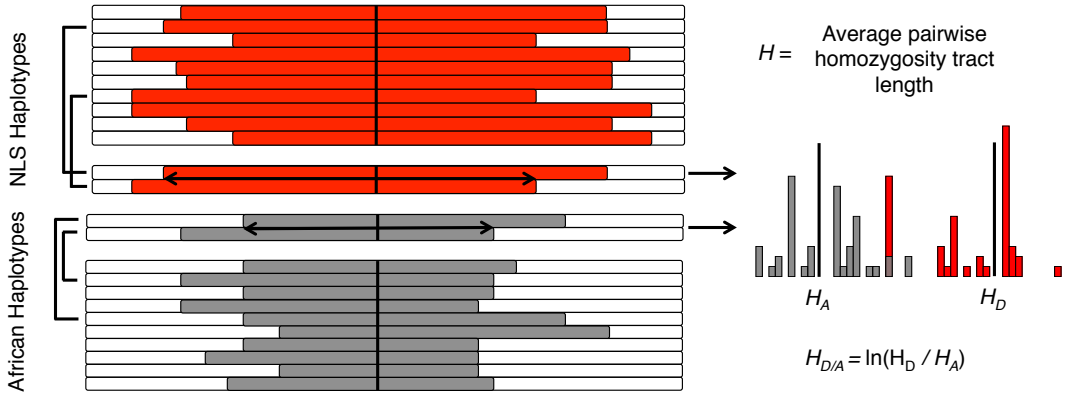


Figure 4: The Neandertal haplotype in the OAS regions has a different impact on the regulation of OAS genes depending on the viral agents PBMCs are exposed to. (A) Log 2 fold induction (y-axis) of *OAS1*, *OAS2* and *OAS3* in response to different viral agents or viral-associated immune stimuli (B) –log10 P for the association between genotype status for the Neandertal haplotype and overall expression levels of OAS genes and the expression of specific isoforms of *OAS1* and *OAS2* (those identified in Figure 3 as associated with NLS).

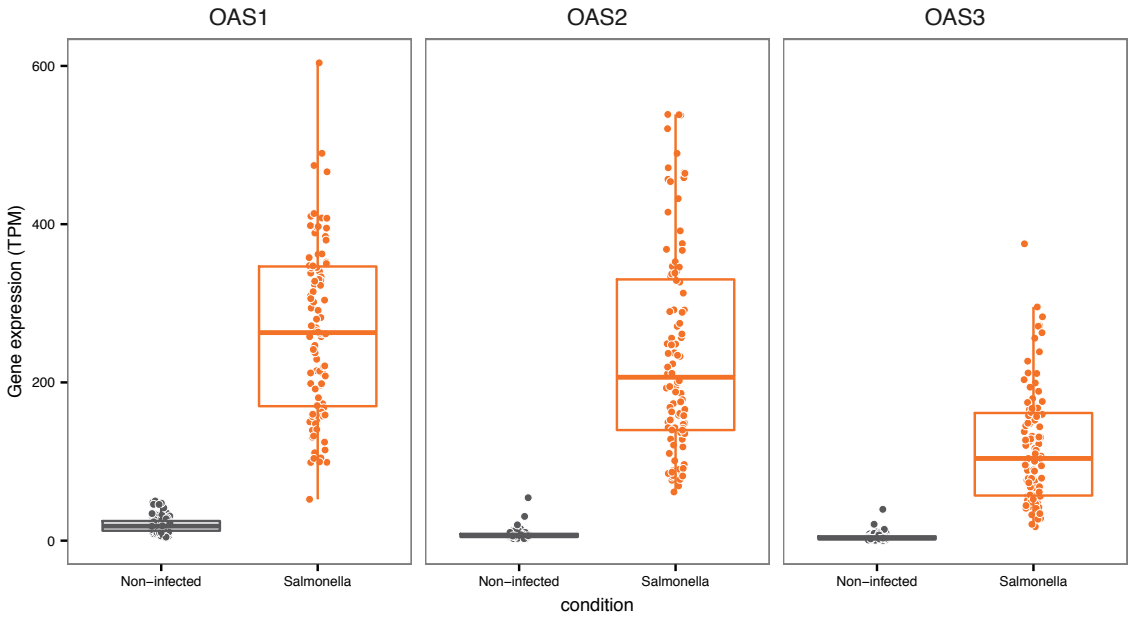
Supplementary Figures:



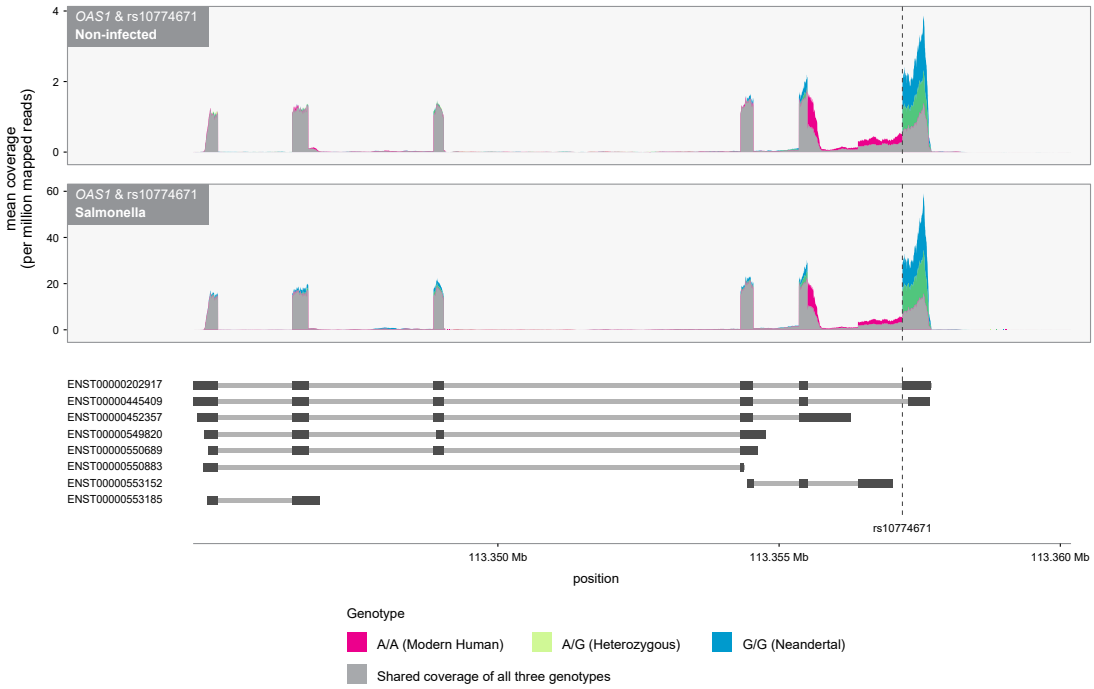
Supplementary Figure 1: Demographic model for neutral coalescent simulations.



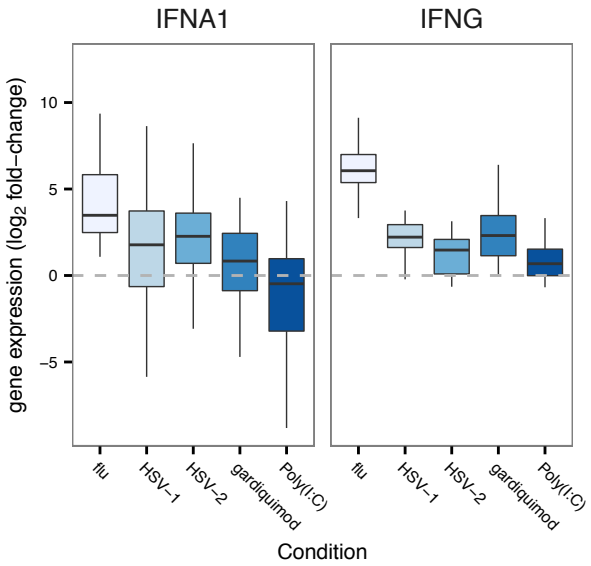
896
897 Supplementary Figure 2: Illustration of $H_{D/A}$ statistic.
898
899
900
901



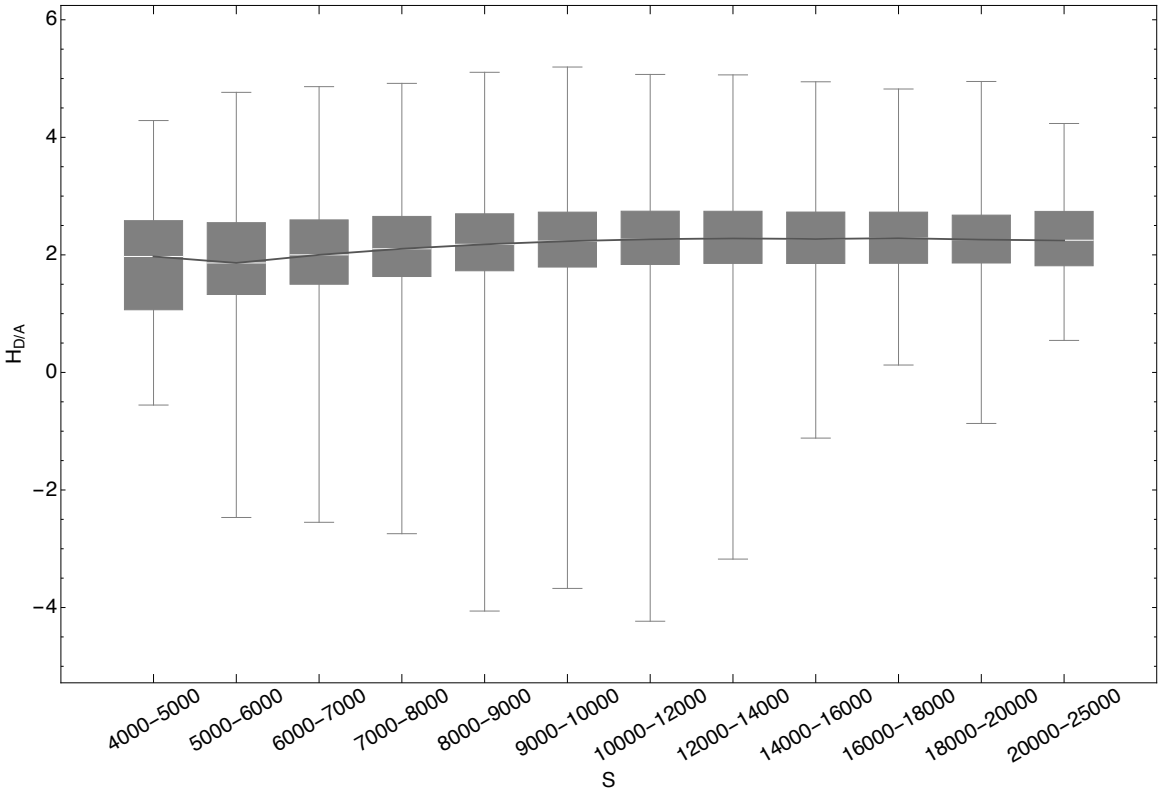
Supplementary Figure 3: Expression levels of OAS genes in primary macrophages before and after infection with *Salmonella*.



Supplementary Figure 4: The splicing variant rs10774671 is a strong asQTL for *OAS1*. Plotted is the normalized average coverage at which each base was sequenced along the genomic regions encoding the gene *OAS1*. Individuals were stratified according to their genotype at rs10774671. Below the figure are gene models from Ensembl database. Individuals carrying the G allele at rs10444671 (i.e., the Neandertal allele) primarily express the transcript ENST00000202917 (referred to as p46 in the text) whereas individuals carrying the A derived allele lose the splice site, which leads to the usage of a distinct isoform.



Supplementary Figure 5: Log 2 fold induction (y-axis) of *IFNA1* (type I interferon) and *IFNG* (type II interferon) in PBMCs upon infection/stimulation with several viral agents.



Supplementary Figure 6: $H_{D/A}$ does not vary substantially with SNP density. $H_{D/A}$ statistic plotted against number of segregating sites (S) per locus across all neutral coalescent simulations. The slight increase in $H_{D/A}$ with elevated S is conservative with respect to our one-tailed test for elevated $H_{D/A}$.

964 **Supplementary Tables:**
965
966

	CEU	GBR	IBS	TSI	Mathieson et al. 2015 P
113341598	0.856	0.407	0.967	0.256	0.233
113343563	0.862	0.428	0.529	0.524	0.335
113350796	0.0123	0.0194	0.395	0.0557	0.0249
113360025	0.00850	0.0114	0.0317	0.00633	0.000277
113363284	0.0126	0.0170	0.0288	0.00732	0.0000288
113370966	0.00654	0.00902	0.0190	0.00476	2.29E-08
113371114	0.00735	0.0102	0.0209	0.00518	0.000000315
113375983	0.00947	0.0140	0.0333	0.00787	0.00459
113435450	0.0167	0.0284	0.132	0.0409	0.0000375
113504725	0.0852	0.0797	0.134	0.0797	0.157
113532885	0.354	0.853	0.0112	0.218	0.0236
p <= 0.01					
0.01<p<=0.05					

967
968 Supplementary Table 1: Proportion of similar frequency genome-wide SNPs with
969 deviations from expected frequency as great or greater than 11 Neandertal-like
970 sites. Fraction of all non-OAS SNPs for which absolute difference in expected
971 (based on aDNA samples) and observed allele frequency that are as great or
972 greater than NLS derived allele frequency difference. Final column reports P from
973 Mathieson et al. 2015 selection test based on same aDNA samples.
974
975