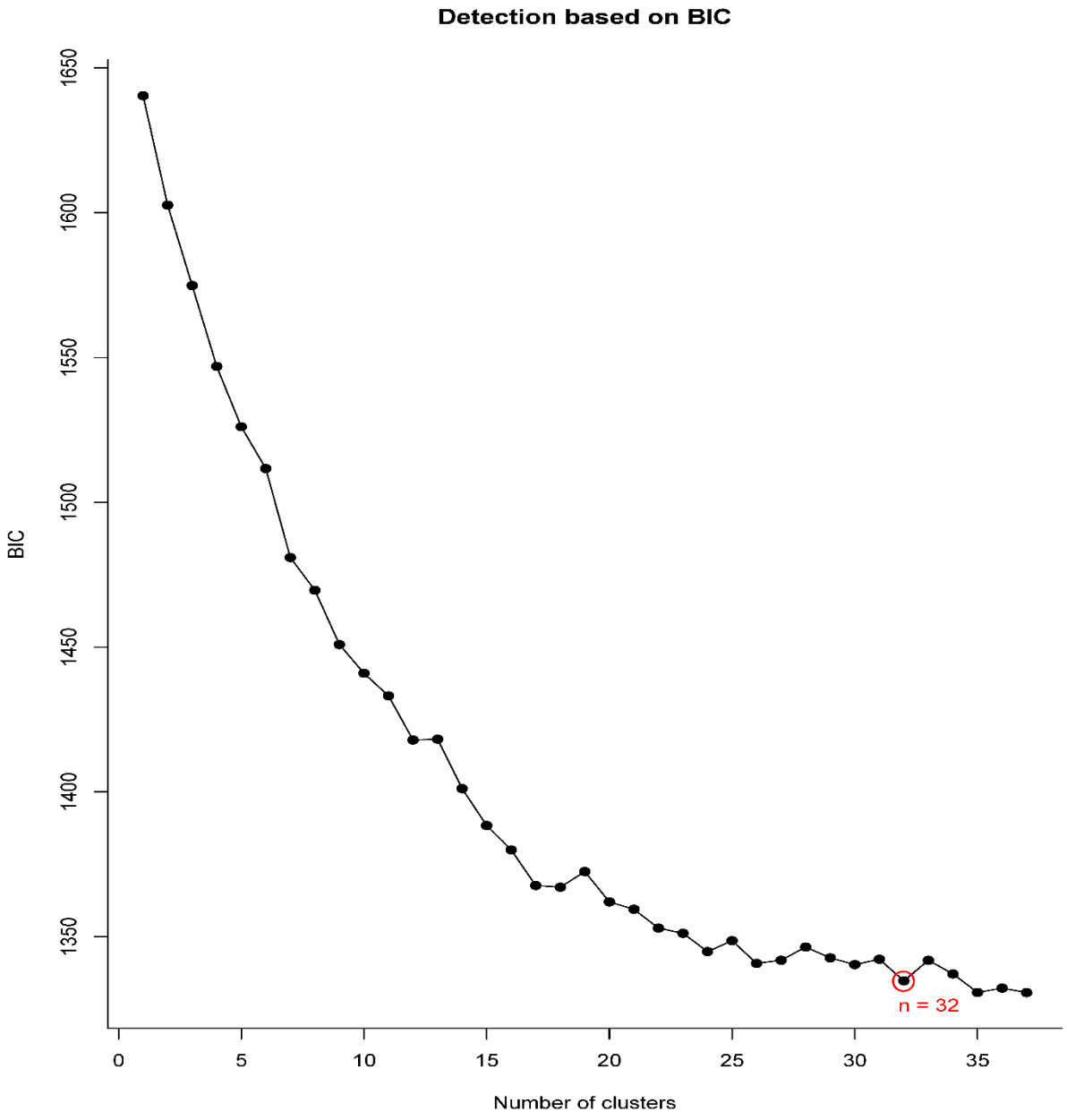
Table S1. Identifying positive selection and the geographic distribution of MHC supertypes.

We used the HyPhy package (Pond *et al.* 2005) on the Datamonkey webserver (Delport *et al.* 2010a) for model selection to detect sites under selection. The model selection tool (Delport *et al.* 2010b) was used to identify the optimal nucleotide substitution model for further analyses. Recombination was taken into account in the implementation of three separate models of codon-based positive selection: single likelihood ancestor counting (SLAC) (Pond& Frost 2005), fixed effects likelihood (FEL) (Pond& Frost 2005), and mixed effect model of evolution (MEME) (Murrell *et al.* 2012). We adopted a conservative approach that amino acid sites identified by two or more models were retained as sites under positive selection for further analyses. Amino acid sites under positive selection as identified above were used for cluster analysis to define MHC alleles supertypes, following Doytchinova and Flower (Doytchinova& Flower 2005). Secondly, all other amino acid sites (i.e. those that were not found to evolve under positive selection) were excluded during supertype definition. Each retained site was characterized according to five physiochemical descriptor variables: z1 (hydrophobicity), z2 (steric bulk), z3 (polarity), z4 and z5 (electronic effects) (Sandberg *et al.* 1998). Thirdly, discriminant analysis of principle components (DAPC) was implemented to define MHC alleles gene clusters using adegenet 1.4-0 package in R (Jombart 2008; Jombart *et al.* 2010). This analysis implements a k-means clustering algorithm based on Bayesian Information Criterion (BIC); and we chose the optimal number of supertype clusters by BIC values that decreased by a negligible amount (Jombart 2008). DAPC was then performed on retained principal components to assign MHC alleles to a supertype.

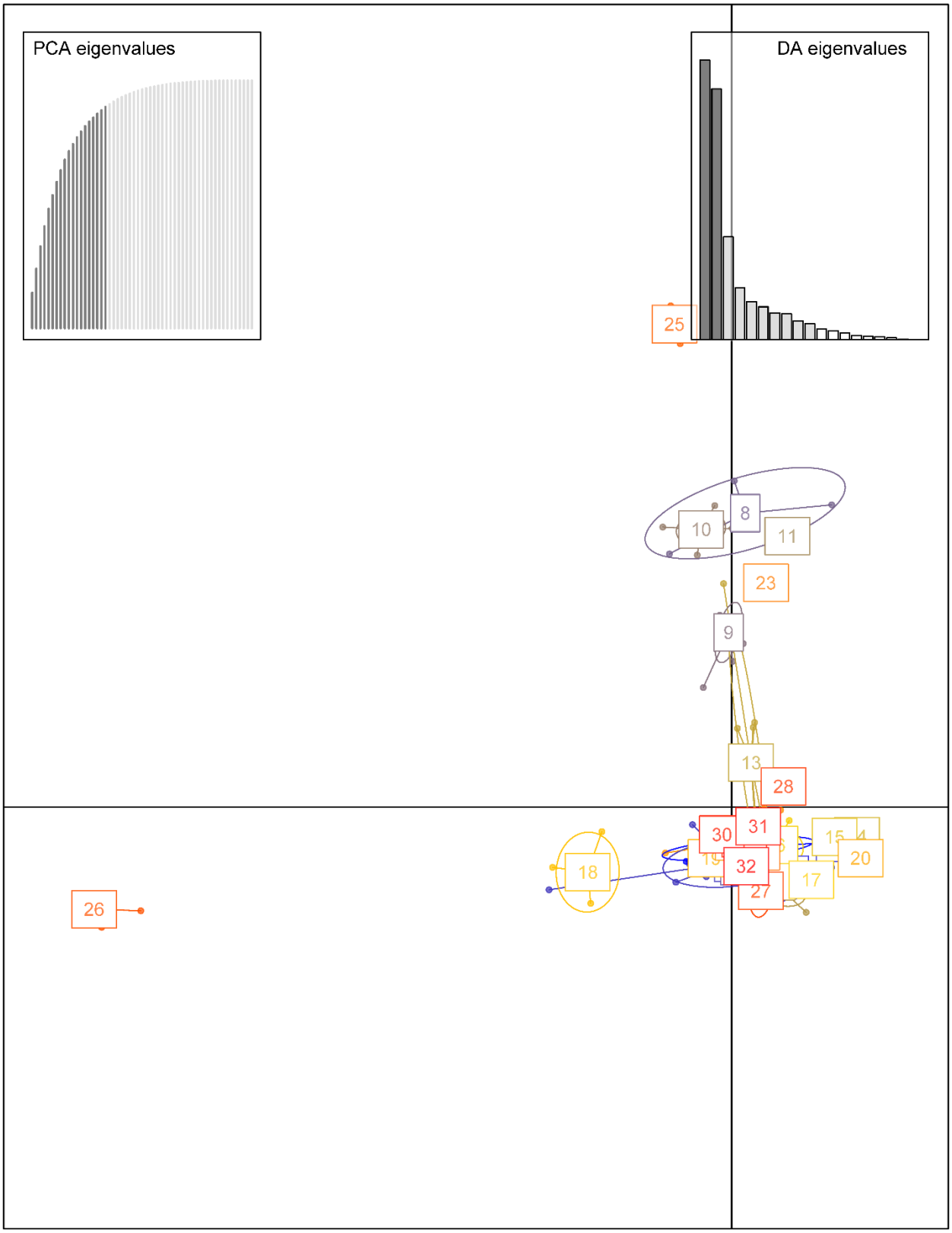
Table S1-1. Amino acid sites under selection.

Codon	SLAC dN-dS	SLAC *p-value*	FEL dN/dS	FEL *p-value*	MEME beta2	MEME *p-value*
3	16.2549	0.005487	2.982	0.005839	93.7674	0
4			2.57E+15	0.01505	10.5835	0.002412
6	56.3055	9.77E-08	29.85	1.73E-09	46.8084	3.22E-15
30	15.2679	0.025918	6.958	0.008364	3.47552	0.013052
33	12.6698	0.018404	4.901	0.00671	2.93245	0.01046
34			Infinite	0.029542	0.764468	0.043773
38	10.1497	0.022595	6.746	0.01247	1.98425	0.020077
40	6.98243	0.047484	Infinite	0.018142	1.6012	0.028351
41	30.2026	0.007931	1.78E+14	0.000344	19.2414	7.70E-07
51	26.1261	6.06E-08	Infinite	5.56E-09	4.50712	1.72E-08
52	9.55965	0.002284	Infinite	0.000233	1.59781	0.000493
62	41.8675	1.11E-06	6.522	3.01E-06	36.1454	1.18E-09
65	15.3834	0.00658	5.66	0.00215	3.33372	0.00391
66	37.0685	0.000752	2.63E+16	1.39E-05	11.4339	2.43E-05
70	11.0301	0.002495	5.332	0.00554	15.9012	0.000805
92	48.5137	7.14E-07	4.876	6.61E-08	121.815	0

K-means clustering



Discriminant analysis of principal components (DAPC) scatterplots of the MHCsupertypes


PCA eigenvalues
DA eigenvalues
25
8
10
11
23
9
13
28
30
31
32
27
18
20
17
26

Reference

Delport W, Poon AFY, Frost SDW, Pond SLK (2010a) Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* **26**, 2455-2457.

Delport W, Scheffler K, Botha G, *et al.* (2010b) CodonTest: Modeling Amino Acid Substitution Preferences in Coding Sequences. *Plos Computational Biology* **6**.

Doytchinova IA, Flower DR (2005) In silico identification of supertypes for class II MHCs. *Journal of Immunology* **174**, 7085-7095.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *Bmc Genetics* **11**.

Murrell B, Wertheim JO, Moola S, *et al.* (2012) Detecting Individual Sites Subject to Episodic Diversifying Selection. *Plos Genetics* **8**.

Pond SLK, Frost SD (2005) Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular Biology and Evolution* **22**, 1208-1222.

Pond SLK, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676-679.

Sandberg M, Eriksson L, Jonsson J, Sjostrom M, Wold S (1998) New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. Journal of Medicinal Chemistry 41, 2481-2491.

Table S6. Tests of recombination at MHC I for *P. nigromaculatus.* Significance level $P = 0.05$ for evidence of recombination. These four putative recombinants were excluded from supertype analysis.

Allele	RDP	GeneConv	BootScan	MaxChi	Chimaera	SiScan	3Seq	LARD	PhylPro
83	—	—	—	0.01	0.01	< 0.001	—	< 0.001	—
156	0.02	—	—	0.001	—	—	—	< 0.001	< 0.001
146	—	—	—	0.036	0.041	< 0.001	—	0.002	< 0.001
200	—	—	—	0.001	0.005	< 0.001	—	0.02	0.01

— No recombination event identified by the specified method ($P > 0.05$)