

Supplemental Information for the manuscript: Population genomics of the invasive Argentine ant (*Linepithema humile*) - adaptive evolution in the introduced supercolonies despite low genetic diversity

Ida Holmberg, Jenni Paviala, Jes S   Pedersen, Heikki Helanter  , Lumi Viljakainen

S10. 20 kbp genomic windows with the highest and lowest Tajima's D values of each supercolony and their overlaps.

Please note that Catalonia was excluded from this step.

LOWEST 1%		SuperExactTest						
Sample/Samples	20 kbp window	Degree	Observed.Overlap	Expected.Overlap	FE	P.value	BH-adj. Pval	
Native	NW_012158064.1_10000, NW_012158064.1_110000	1	95	NA	NA	NA	NA	
Chile	NW_012158181.1_190000, NW_012158181.1_290000	1	95	NA	NA	NA	NA	
Native & Chile	NW_012160615.1_290000, NW_012160615.1_390000	2	4	0.8223	4.8643	0.0092	0.0123	
European Main	NW_012159963.1_70000, NW_012159963.1_170000	1	94	NA	NA	NA	NA	
Native & European Main	NW_012160182.1_70000, NW_012160182.1_170000	2	3	0.8137	3.6870	0.0479	0.0479	
European Main & Chile	NW_012159963.1_70000, NW_012159963.1_170000	2	24	0.8137	29.4961	0.0000	0.0000	
All (Native, European Main & Chile)	NW_012160641.1_170000, NW_012160641.1_390000	3	2	0.0070	283.9645	0.0000	0.0000	

LOWEST 5%		SuperExactTest						
Sample/Samples	20 kbp windows	Degree	Observed.Overlap	Expected.Overlap	FE	P.value	BH-adj. Pval	
Native	NW_012158064.1_10000, NW_012158064.1_110000	1	471	NA	NA	NA	NA	
Chile	NW_012158052.1_70000, NW_012158052.1_170000	1	474	NA	NA	NA	NA	
Native & Chile	NW_012158181.1_70000, NW_012158181.1_170000	2	57	20.3421	2.8021	0.0000	0.0000	
European Main	NW_012158052.1_70000, NW_012158052.1_170000	1	470	NA	NA	NA	NA	
Native & European Main	NW_012159245.1_70000, NW_012159245.1_170000	2	58	20.1704	2.8755	0.0000	0.0000	
European Main & Chile	NW_012158052.1_70000, NW_012158052.1_170000	2	189	20.2989	9.3109	0.0000	0.0000	
All (Native, European Main & Chile)	NW_012159245.1_70000, NW_012159245.1_170000	3	31	0.8711	35.5855	0.0000	0.0000	

HIGHEST 1%		SuperExactTest						
Sample/Samples	20 kbp windows	Degree	Observed.Overlap	Expected.Overlap	FE	P.value	BH-adj. Pval	
Native	NW_012158181.1_110000, NW_012158181.1_210000	1	95	NA	NA	NA	NA	
Chile	NW_012158220.1_90000, NW_012158220.1_190000	1	95	NA	NA	NA	NA	
Native & Chile	NW_012160012.1_30000, NW_012160012.1_130000	2	17	0.8223	20.6731	0.0000	0.0000	
European Main	NW_012158052.1_130000, NW_012158052.1_230000	1	94	NA	NA	NA	NA	
Native & European Main	NW_012160012.1_30000, NW_012160012.1_130000	2	7	0.8137	8.6030	0.0000	0.0000	
European Main & Chile	NW_012160012.1_30000, NW_012160012.1_130000	2	8	0.8137	9.8320	0.0000	0.0000	
All (Native, European Main & Chile)	NW_012160012.1_30000, NW_012160012.1_130000	3	5	0.0070	709.9112	0.0000	0.0000	

HIGHEST 5%		SuperExactTest						
Sample/Samples	20 kbp windows	Degree	Observed.Overlap	Expected.Overlap	FE	P.value	BH-adj. Pval	
Native	NW_012158181.1_110000, NW_012158181.1_210000	1	471	NA	NA	NA	NA	
Chile	NW_012158220.1_90000, NW_012158220.1_190000	1	474	NA	NA	NA	NA	
Native & Chile	NW_012158820.1_10000, NW_012158820.1_20000	2	55	20.3421	2.7038	0.0000	0.0000	
European Main	NW_012158052.1_130000, NW_012158052.1_230000	1	470	NA	NA	NA	NA	
Native & European Main	NW_012158820.1_10000, NW_012158820.1_20000	2	39	20.1704	1.9335	0.0001	0.0001	
European Main & Chile	NW_012158820.1_10000, NW_012158820.1_20000	2	59	20.2989	2.9066	0.0000	0.0000	
All (Native, European Main & Chile)	NW_012158820.1_10000, NW_012158820.1_20000	3	19	0.8711	21.8105	0.0000	0.0000	

S10. Genes with the highest and lowest Tajima's D values of each supercolony and their overlaps.

Please note that Catalonia was excluded from this step.

LOWEST 1%		SuperExactTest						
Sample/Samples	Genes	Degree	Observed.Overlap	Expected.Overlap	FE	P.value	BH-adj. Pval	
Native	LOC105667306, LOC1056	1	110	NA	NA	NA	NA	NA
Chile	LOC105667323, LOC1056	1	110	NA	NA	NA	NA	NA
Native & Chile	LOC105673307, LOC1056	2	3	1.1025	2.7211	0.0982	0.0982	
European Main	LOC105667304, LOC1056	1	109	NA	NA	NA	NA	NA
Native & European Main	LOC105669939, LOC1056	2	7	1.0925	6.4074	0.0001	0.0002	
European Main & Chile	LOC105667346, LOC1056	2	24	1.0925	21.9683	0.0000	0.0000	
All (Native, European Main & Chile)		3	0	0.0109	0.0000	1.0000	1.0000	

LOWEST 5%		SuperExactTest						
Sample/Samples	Genes	Degree	Observed.Overlap	Expected.Overlap	FE	P.value	BH-adj. Pval	
Native	LOC105667287, LOC1056	1	547	NA	NA	NA	NA	NA
Chile	LOC105667306, LOC1056	1	548	NA	NA	NA	NA	NA
Native & Chile	LOC105667306, LOC1056	2	63	27.3126	2.3066	0.0000	0.0000	
European Main	LOC105667304, LOC1056	1	545	NA	NA	NA	NA	NA
Native & European Main	LOC105667304, LOC1056	2	56	27.1631	2.0616	0.0000	0.0000	
European Main & Chile	LOC105667346, LOC1056	2	139	27.2128	5.1079	0.0000	0.0000	
All (Native, European Main & Chile)	LOC105668109, LOC1056	3	24	1.3563	17.6952	0.0000	0.0000	

HIGHEST 1%		SuperExactTest						
Sample/Samples	Genes	Degree	Observed.Overlap	Expected.Overlap	FE	P.value	BH-adj. Pval	
Native	LOC105667314, LOC1056	1	110	NA	NA	NA	NA	NA
Chile	LOC105667324, LOC1056	1	110	NA	NA	NA	NA	NA
Native & Chile	LOC105667482, LOC1056	2	28	1.1025	25.3967	0.0000	0.0000	
European Main	LOC105667325, LOC1056	1	109	NA	NA	NA	NA	NA
Native & European Main	LOC105667482, LOC1056	2	24	1.0925	21.9683	0.0000	0.0000	
European Main & Chile	LOC105667482, LOC1056	2	32	1.0925	29.2911	0.0000	0.0000	
All (Native, European Main & Chile)	LOC105667482, LOC1056	3	17	0.0109	1552.5518	0.0000	0.0000	

HIGHEST 5%		SuperExactTest						
Sample/Samples	Genes	Degree	Observed.Overlap	Expected.Overlap	FE	P.value	BH-adj. Pval	
Native	LOC105667294, LOC1056	1	547	NA	NA	NA	NA	NA
Chile	LOC105667317, LOC1056	1	548	NA	NA	NA	NA	NA
Native & Chile	LOC105667324, LOC1056	2	111	27.3126	4.0641	0.0000	0.0000	
European Main	LOC105667295, LOC1056	1	545	NA	NA	NA	NA	NA
Native & European Main	LOC105667404, LOC1056	2	104	27.1631	3.8287	0.0000	0.0000	
European Main & Chile	LOC105667402, LOC1056	2	111	27.2128	4.0790	0.0000	0.0000	
All (Native, European Main & Chile)	LOC105667404, LOC1056	3	68	1.3563	50.1365	0.0000	0.0000	

S10. Lists of genes under selection that putatively function in the following categories:

Targets of positive selection (low Tajima's D)

Genes with putative neurological/neuron-related function		Gene ontology (NCBI):			
ID	Name	Function	Process	Component	Model
LOC105669987	BTB/POZ domain-containing protein 1-like	enables protein binding	involved_in neurogenesis	located_in P-body located_in cytosol	protein coding
LOC105670081	LIM/homeobox protein Awh-like	enables DNA-binding transcription f	involved_in neuron differentiationinv	located_in nucleus	protein coding
LOC105670079	LIM/homeobox protein Awh	enables DNA-binding transcription f	involved_in neuron differentiationinv	located_in nucleus	protein coding
LOC105670993	enhancer of split mbeta protein-like	enables DNA-binding transcription f	involved_in anterior/posterior patterr	located_in nucleus	protein coding
LOC105672811	inhibitory POU protein	enables DNA-binding transcription f	involved_in regulation of transcriptioni	-	protein coding
LOC105678045	gamma-aminobutyric acid receptor subunit beta	enables GABA-A receptor activityyen	involved_in chemical synaptic transm	located_in membranelocated_in neuron pr	protein coding
LOC105678369	neurotrimin-like	enables axon guidance receptor acti	involved_in homophilic cell adhesion	located_in axon located_in neuronal cell bo	protein coding
LOC105668854	gamma-1-syntrophin	enables protein bindingenables stru	part_of dystrophin-associated glycop	-	protein coding
LOC105671064	synaptic vesicle glycoprotein 2B	enables transmembrane transporte	involved_in transmembrane transpor	located_in membrane	protein coding
LOC105677918	protein still life, isoform SIF type 1	enables guanyl-nucleotide exchange	involved_in activation of GTPase activ	located_in membranelocated_in synapse	protein coding
LOC105669672	homeobox protein extradenticle	enables DNA-binding transcription f	involved_in animal organ morphogen	located_in nucleus	protein coding
LOC105668761	patronin	enables calmodulin bindingenables	involved_in cytoplasmic microtubule	located_in microtubule minus-end	protein coding
LOC105678066	glutamate receptor 1-like	enables glutamate-gated receptor a	involved_in modulation of chemical s	located_in plasma membrane located_in pc	protein coding
LOC105679459	enhancer of split m7 protein-like	enables DNA-binding transcription f	transcription by RNA polymerase II	located_in nucleus	protein coding

Targets of positive selection (low Tajima's D)

Chemical signalling	
ID	Name
LOC105670843	odorant receptor 13a-like
LOC105670807	odorant receptor 22c-like
LOC105668181	acyl-CoA Delta(11) desaturase-like
LOC105673531	fatty acid synthase-like
LOC105667306	lipase 3-like
LOC105669334	fatty acid synthase-like
LOC105668015	acyl-CoA Delta(11) desaturase-like
LOC105671493	acyl-CoA:lysophosphatidylglycerol acyltrans
LOC105676392	putative gustatory receptor 28a
LOC105675232	uncharacterized
LOC105673563	endocuticle structural protein SgAbd-6-like

Targets of balancing selection (high Tajima's D values)

Genes with putative neurological/neuron-related function		Gene ontology:			
ID	Name	Function	Process	Component	Model
LOC105675099	synaptic vesicle glycoprotein 2B-like	enables transmembrane transporte	involved_in transmembrane transpor	located_in membrane	protein coding
LOC105667794	glycine receptor subunit alpha-2	enables chloride channel activityyen	involved_in chemical synaptic transm	located_in membrane located_in neuron pr	protein coding
			-	-	

Targets of balancing selection (high Tajima's D values)

Chemical signalling	
ID	Name
LOC105668241	endochitinase
LOC105670810	odorant receptor 4-like
LOC105670754	odorant receptor 22c-like
LOC105670819	odorant receptor 45b-like
LOC105676832	fatty acid synthase-like
LOC105679973	fatty acid binding protein 1-B.1-like
LOC105672373	sterol O-acyltransferase 2-like
LOC105673304	putative odorant receptor 85d
LOC105668074	fatty acyl-CoA reductase 1-like
LOC105674860	odorant receptor 63a-like
LOC105673490	odorant receptor coreceptor
LOC105680075	fatty acid synthase-like
LOC105669369	gustatory receptor for sugar taste 64a-like
LOC105674865	odorant receptor 67a-like

Gene ontology

function	process	composition	Model
enables odorant binding	enables olfa involved_in detection of chemical stimulus	involvement located_in plasma membrane	protein coding
enables odorant binding	enables olfa involved_in detection of chemical stimulus	involvement located_in plasma membrane	protein coding
enables iron ion binding	enables ste involved_in unsaturated fatty acid biosynthetic process	located_in endoplasmic reticulum	protein coding
-	-	-	protein coding
enables hydrolase activity, acting on	involved_in lipid metabolic process	-	protein coding
-	-	-	protein coding
-	-	-	pseudo
enables acyltransferase activity	located_in endomembrane system	-	protein coding
involved_in sensory perception of taste	-	located_in axon	located_in cytoplasm protein coding
enables GPI-anchor transamidase activity	involved_in attachment of GPI anchor to protein	part_of GPI-anchor transamidation	protein coding
enables structural constituent of chitin	-	located_in chitin-based extracellular matrix	protein coding

Targets of positive selection (low Tajima's D)**Genes with immune/oxidative-stress -related function**

ID	Name
LOC105668581	protein toll-like

Gene ontology

function	process
-	-

Targets of balancing selection (high Tajima's D values)**Genes with immune/oxidative-stress -related function**

ID	Name
LOC105675515	peroxidase
LOC105677448	myeloid differentiation primary response protein 1

Gene ontology

function	process
enables heme binding	enables peroxidase involved_in response to oxidative stress
enables TIR domain binding	involved_in MyD88-dependent signaling

		Targets of positive selection (low Tajima's D)							Targets of positive selection (high Tajima's D values)
composition	Model	other transcription factors, signalling genes etc.		Gene ontology			composition	Model	
-	pseudo	ID	Name	function	process				
		LOC105675249	LIM domain only protein 3-like	enables DNA-binding transcription factor binding	involved_in positive regulation of trans	located_in nucleus	protein coding		
		LOC105675250	LIM domain only protein 3-like	enables DNA-binding transcription factor binding	involved_in positive regulation of trans	located_in nucleus	protein coding		
		LOC105677901	homeobox protein Nkx-6.1	enables DNA-binding transcription factor activity	involved_in cell differentiation	located_in nucleus	protein coding		
		LOC105679616	glutamate receptor 3-like	enables glutamate-gated receptor activity	-	located_in membran	protein coding		
		LOC105678675	guanine nucleotide-binding protein	enables G protein-coupled receptor binding	involved_in adenylate cyclase-modulat	part_of heterotrimer	protein coding		
		LOC105678074	putative serine protease K12H4.7	enables dipeptidyl-peptidase activity	involved_in proteolysis	-	protein coding		
		LOC105675233	nesprin-1-like	enables actin filament binding	involved_in nuclear migration	located_in cytoplasm	protein coding		
		LOC105672900	transcription factor Sp9-like	enables DNA-binding transcription factor activity	involved_in regulation of transcription by RNA polymerase II	protein coding			
		LOC105672751	mucin-17-like	enables protein binding	-	-	protein coding		
		LOC105670065	putative serine/threonine-protein kinase	enables ATP binding	involved_in protein phosphorylation	-	protein coding		
		LOC105669914	multiple coagulation factor deficient	-	-	-	protein coding		
		LOC105669693	aryl hydrocarbon receptor nuclear translocator-like	enables DNA-binding transcription factor activity	involved_in regulation of transcription	part_of aryl hydrocarbon receptor	protein coding		
		LOC105669680	glucose transporter type 1	enables hexose transmembrane transporter activity	involved_in monosaccharide transmembrane transport	located_in membran	protein coding		
		LOC105669444	glucose-6-phosphate isomerase	enables calcium channel activity	involved_in calcium ion import across plasma membrane	located_in cytosol	protein coding		
		LOC105667492	CCR4-NOT transcription complex subunit 1	enables molecular adaptor activity	involved_in nuclear-transcribed mRNA processing	part_of CCR4-NOT complex	protein coding		
		LOC105667346	calmodulin, striated muscle-like	enables calcium ion binding	-	-	protein coding		
		Targets of balancing selection (high Tajima's D values)							Targets of balancing selection (low Tajima's D values)
composition	Model	other transcription factors, signalling genes etc.		Gene ontology			composition	Model	Putative endogenous ortholog
located_in extracellular space	protein coding	ID	Name	function	process				ID
located_in plasma membrane	protein coding	LOC105674327	zinc finger protein SNAI3-like	-	-	-	protein coding		LOC105675906
		LOC105680088	galanin receptor type 3-like	enables G protein-coupled receptor activity	involved_in G protein-coupled receptor signaling pathway	located_in membran	protein coding		LOC105679308
		LOC105670608	manganese-transporting ATPase 13	enables ATP binding	involved_in intracellular calcium ion homeostasis	located_in endoplasmic reticulum	protein coding		LOC105679309
		LOC105669174	pro-resilin-like	-	-	located_in extracellular space	protein coding		LOC105679434
		LOC105669160	pro-resilin-like	-	-	located_in extracellular space	protein coding		LOC105680069
		LOC105676333	hemolymph lipopolysaccharide-binding protein	-	-	-	protein coding		LOC105667930
		LOC105669509	arylsulfatase B-like	-	-	-	protein coding		LOC105676718
		LOC105668859	facilitated trehalose transporter 1	enables transmembrane transporter activity	involved_in transmembrane transport	located_in membran	protein coding		LOC105678032
		LOC105673502	nef-associated protein 1-like	enables tRNA (adenine-N6)-methyltransferase activity	-	-	protein coding		LOC105677292
		LOC105668864	calcium and integrin-binding protein	enables calcium ion binding	-	-	protein coding		LOC105679572
		LOC105677495	gamma-interferon-inducible lysosomal protein	enables oxidoreductase activity, acting on a sulfhydryl group	-	-	protein coding		LOC105669297
		LOC105673995	derlin-1	enables misfolded protein binding	involved_in endoplasmic reticulum unfolded protein response	part_of Hrd1p ubiquitin-proteasome pathway	protein coding		LOC105668921
									LOC105671592

selection (low Tajima's D)

selection (high Tajima's D values)

s virus genes

Name	Conserved domains
uncharacterized	RT_pepA17 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=238822 Pao retrotransposon peptidase https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=283123
uncharacterized	Pao retrotransposon peptidase https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=283123 RT_like Superfamily https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=295487
uncharacterized	DUF1759 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=281552 pepsin_retropepsin_like Superfamily https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=299705
uncharacterized	Integrase core domain https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=279057
uncharacterized	Pao retrotransposon peptidase https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=283123 RT_like Superfamily https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=295487
uncharacterized	RT_pepA17 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=238822 DUF1759 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=281552 Pao retrotransposon peptidase https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=283123
uncharacterized	DUF1759 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=281552 pepsin_retropepsin_like Superfamily https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=299705
uncharacterized	DUF1759 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=281552 RT_like Superfamily https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=295487 pepsin_retropepsin_like Superfamily https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=295487
piggyBac transposable el	Transposase IS4
uncharacterized	Integrase core domain https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=279057
uncharacterized	RT_pepA17 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=238822 DUF1759 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=281552 Pao retrotransposon peptidase https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=283123
uncharacterized	RT_pepA17 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=238822 DUF1759 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=281552 Pao retrotransposon peptidase https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=283123
uncharacterized	RT_pepA17 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=238822 DUF1759 https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=281552 Pao retrotransposon peptidase https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=283123

pepsin_retropepsin_like Superfamily <https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=299705> Integrase core domain <https://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=304425>