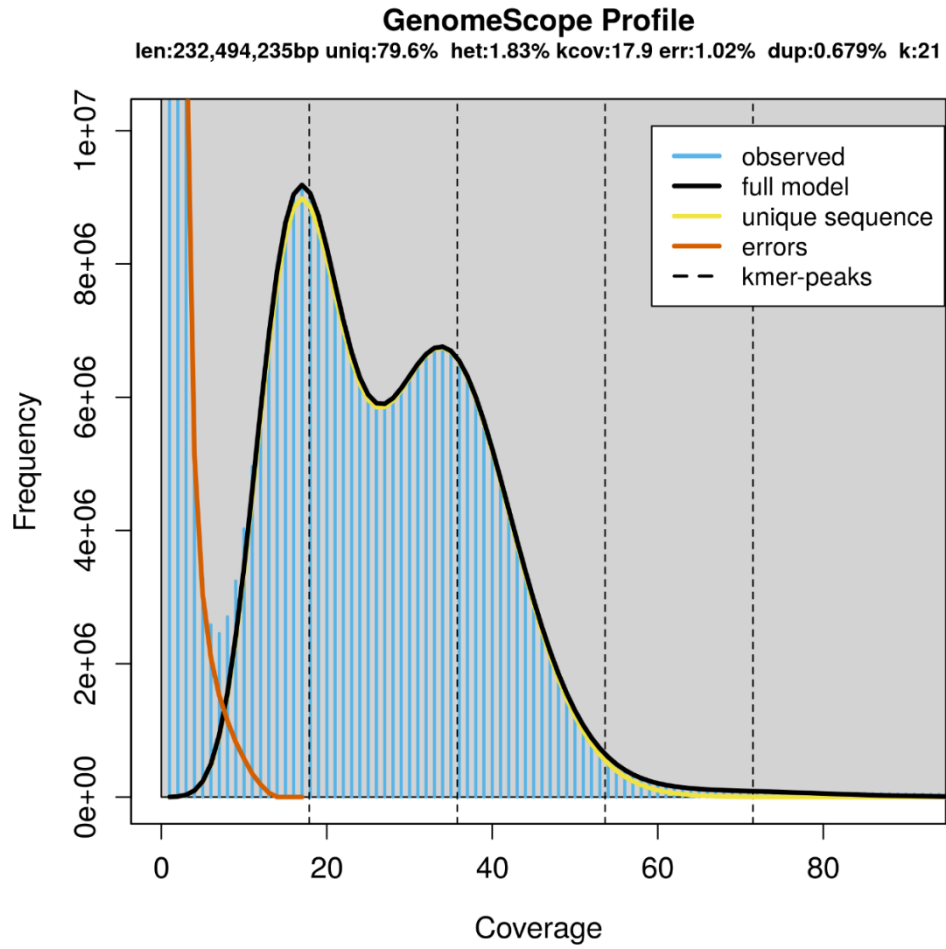


Martijn J.T.N. Timmermans, Amrita Srivathsan, Steve Collins, Rudolf Meier, Alfried P. Vogler.  
Mimicry diversification in *Papilio dardanus* via a genomic inversion in the regulatory region of  
*engrailed-inverted*. Proceedings of the Royal Society B. DOI: 10.1098/rspb.2020.0443

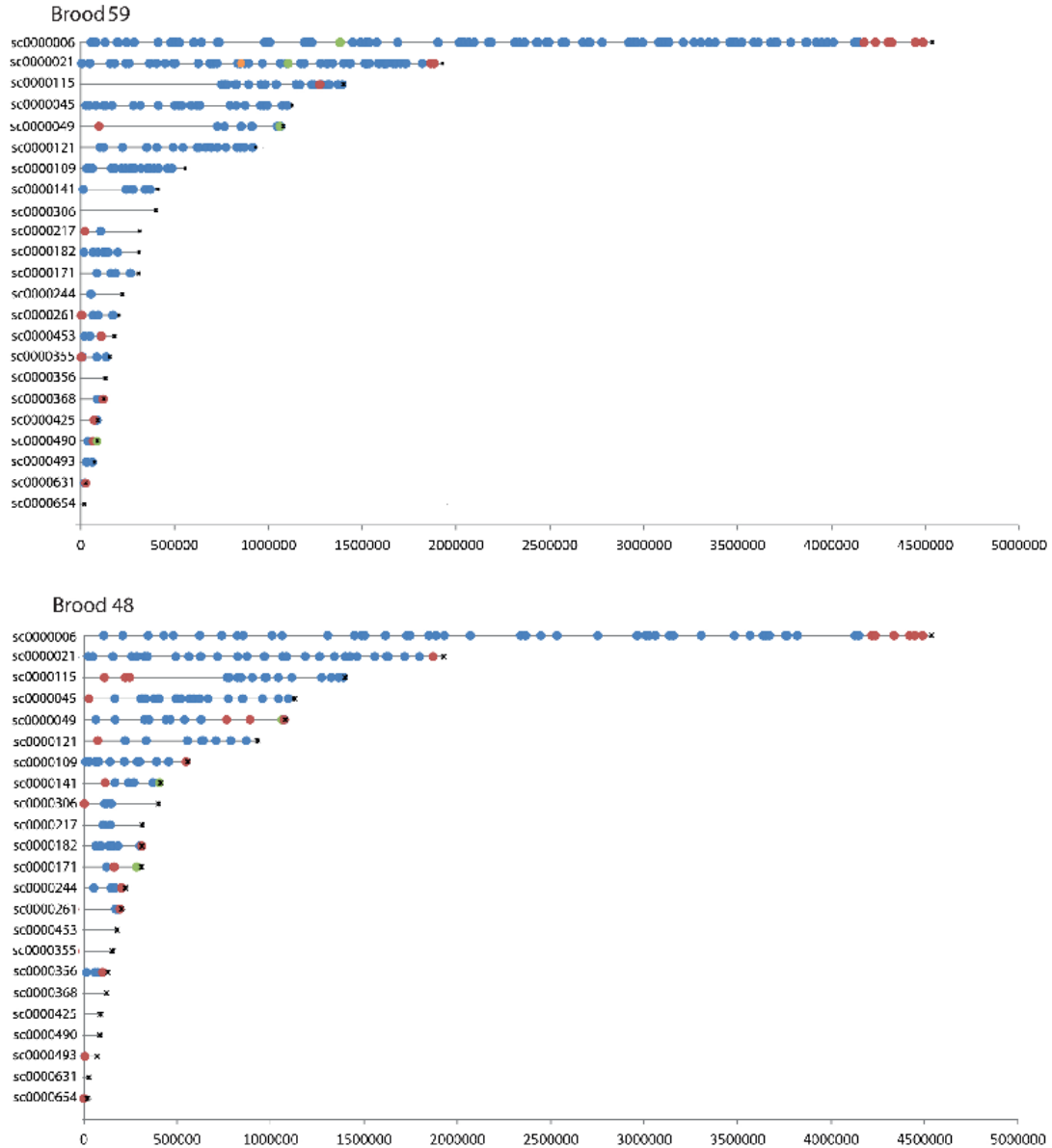
**Electronic supplementary material**

**Figure S1-S8**

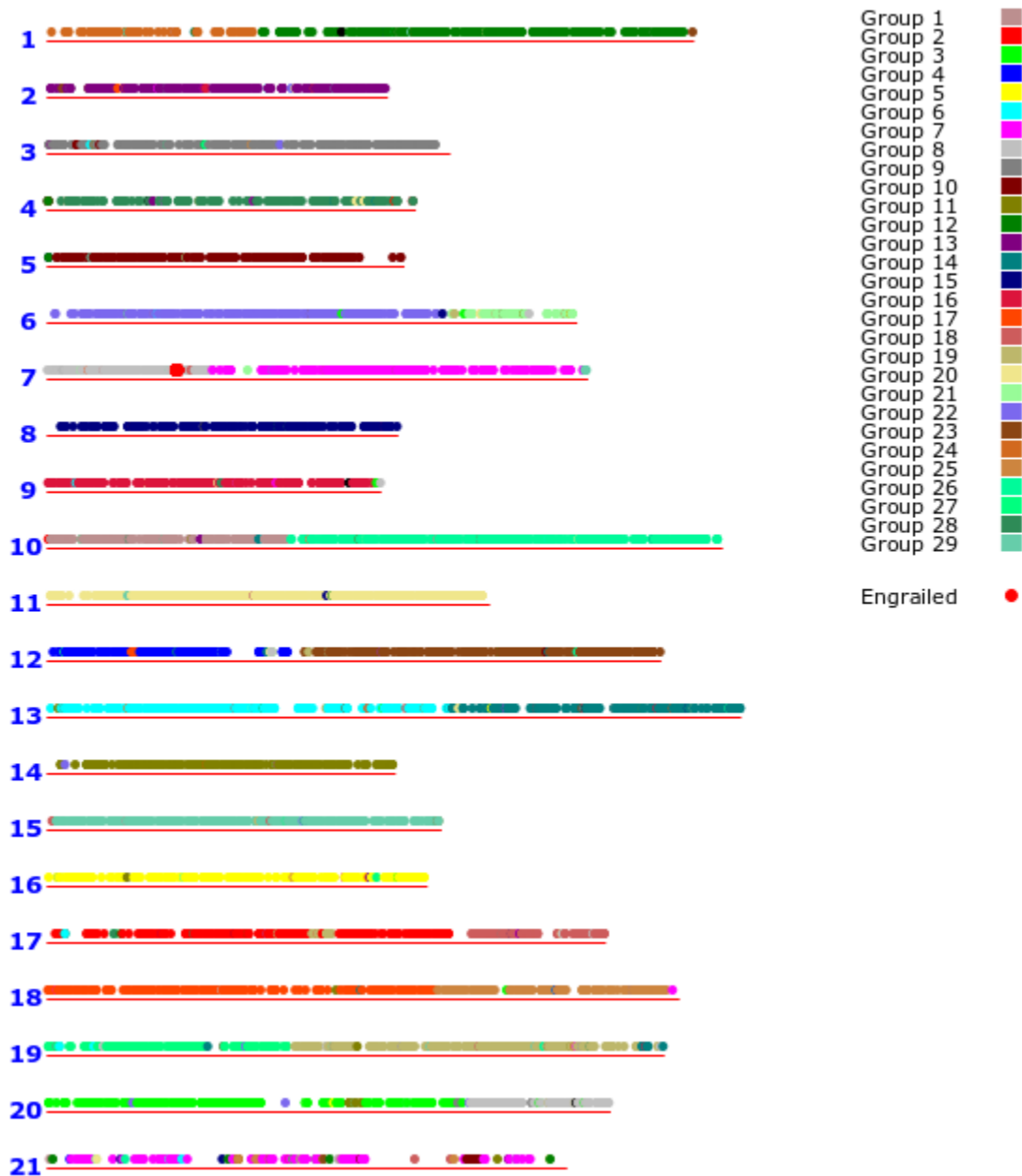
**Table S1-4**



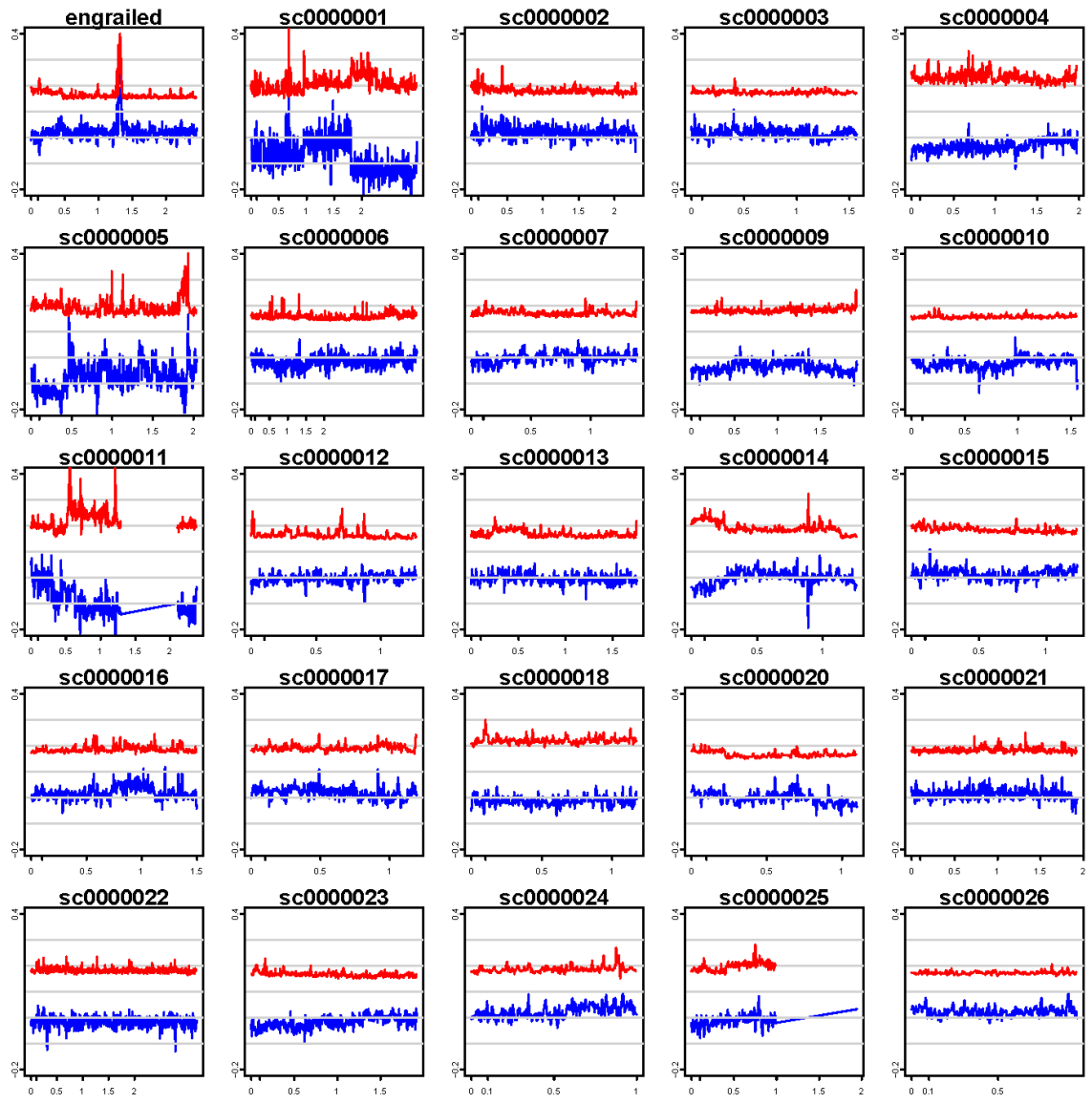
**Figure S1: Estimate of genome size using K-mer count.** Profile was obtained using GenomeScope.

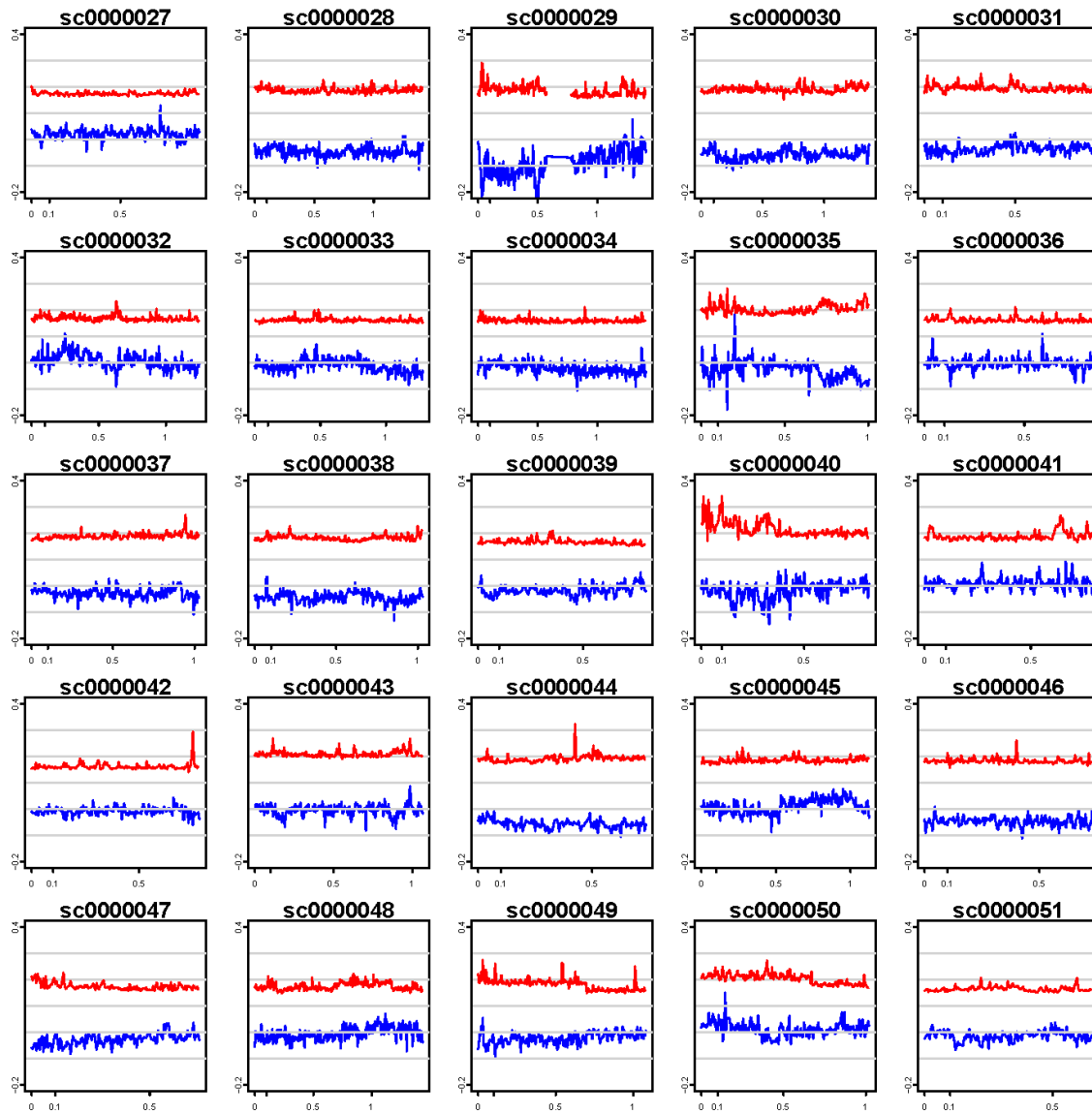


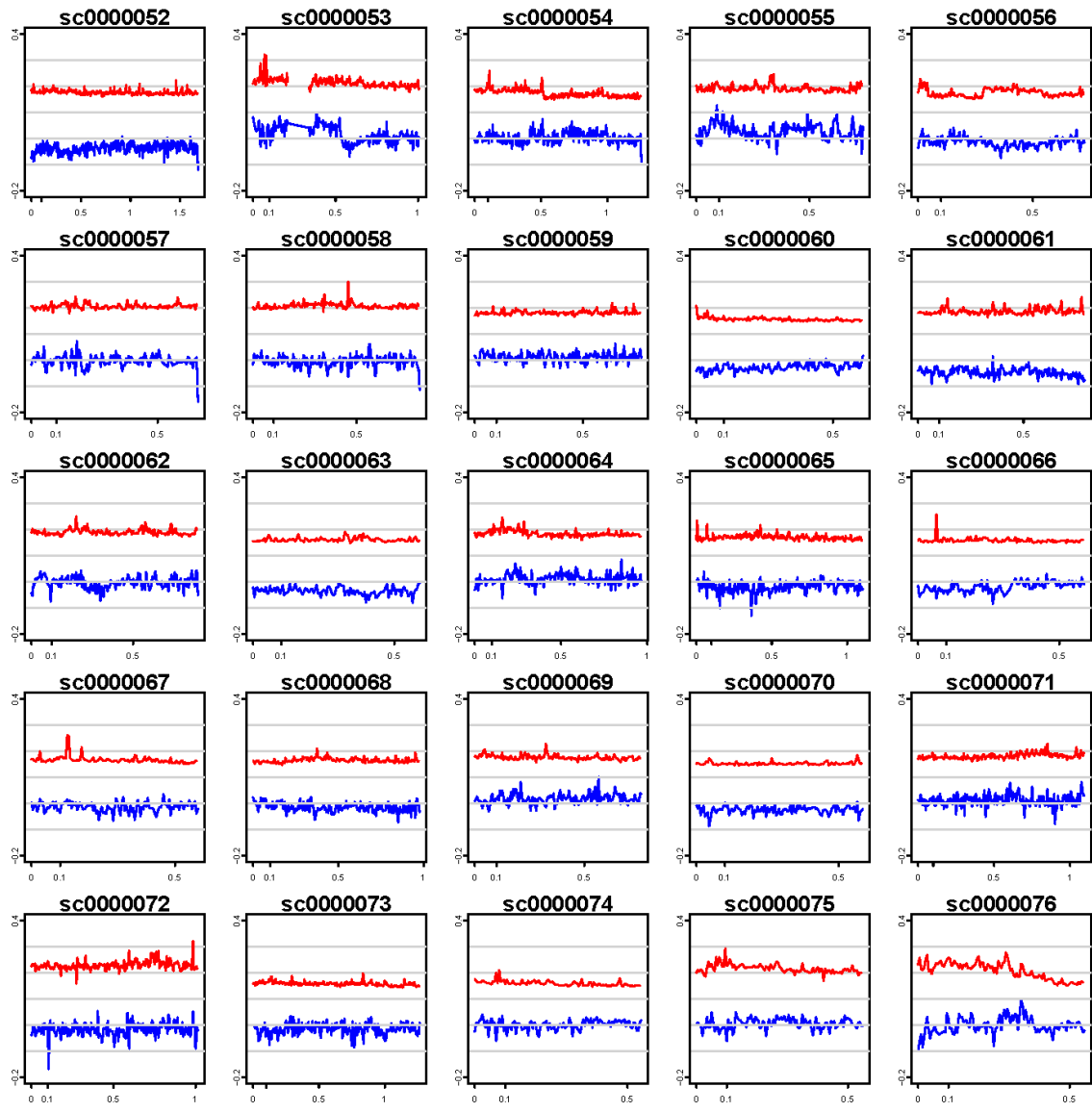
**Figure S2: Scaffolds with deviating SNP patterns for the two pedigree broods.** Top: Brood 59, Bottom: Brood 48. Vertical axis: scaffolds flagged up as being misassembled. Horizontal axis: scaffold length (bp). Colored dots represent SNPs. Blue dots: SNPs showing most frequent pattern. Other colors: SNPs deviating from the most frequent pattern.

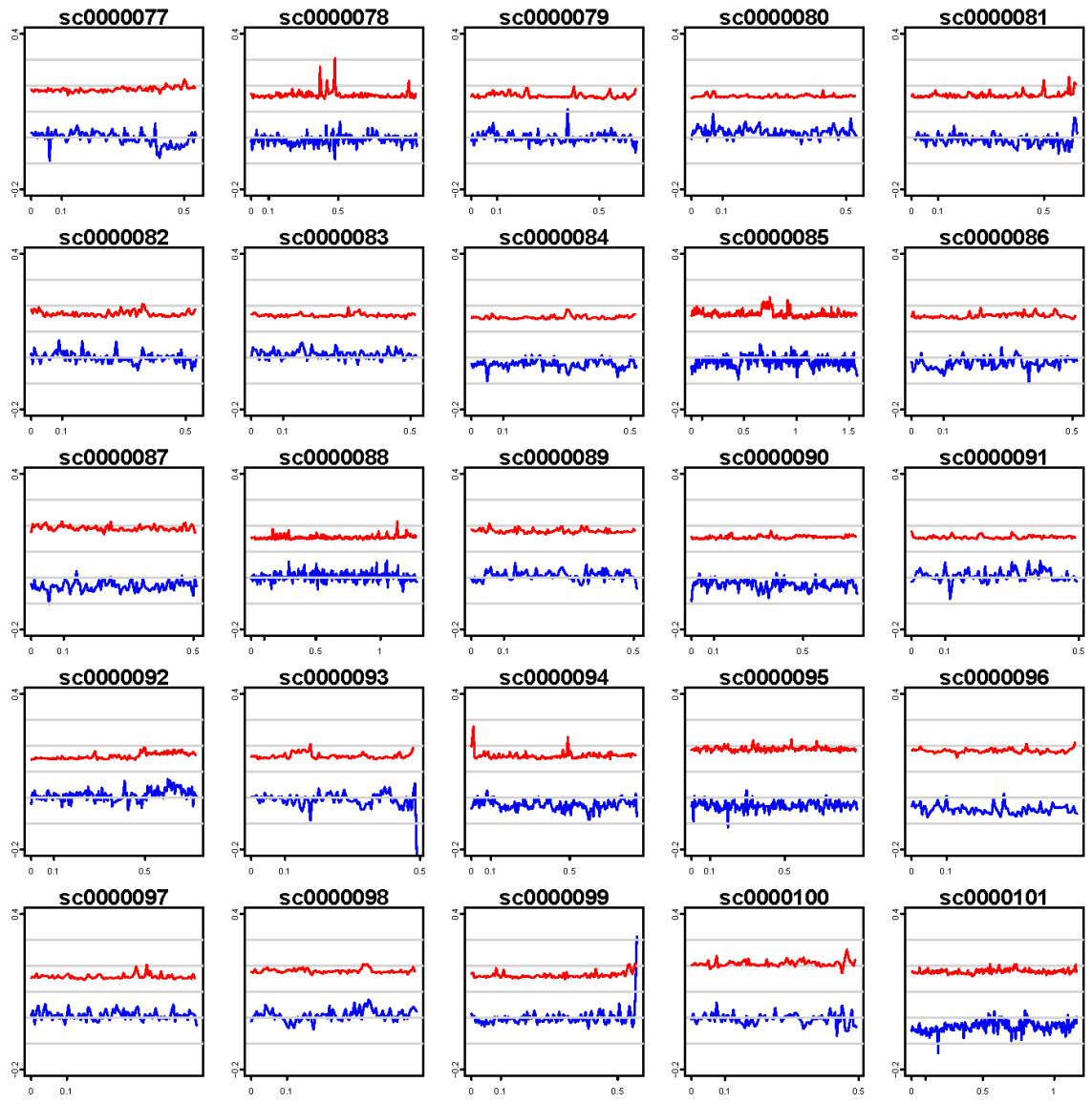


**Figure S3: Comparison between *P. dardanus* draft assembly and *Heliconius melpomene* genome.** BLAST matches between *P. dardanus* and *H. melpomene* are visualised on the *H. melpomene* genome map. *H. melpomene* 21 chromosomes are indicated with red lines. Colored circles indicate positions of *P. dardanus* predicted proteins that show homology to a *H. melpomene* gene. Predicted proteins from similar linkage groups ('chromosome bins') are given identical colors. The transcription factor *engrailed* is part of linkage group 8. In most cases the *P. dardanus* scaffolds within a chromosome bin also mapped to a single chromosome in *H. melpomene*. One exception is *P. dardanus* linkage group 7, which showed similarity to *H. melpomene* chromosome 7 and its Z chromosome (= *H. melpomene* chromosome 21).

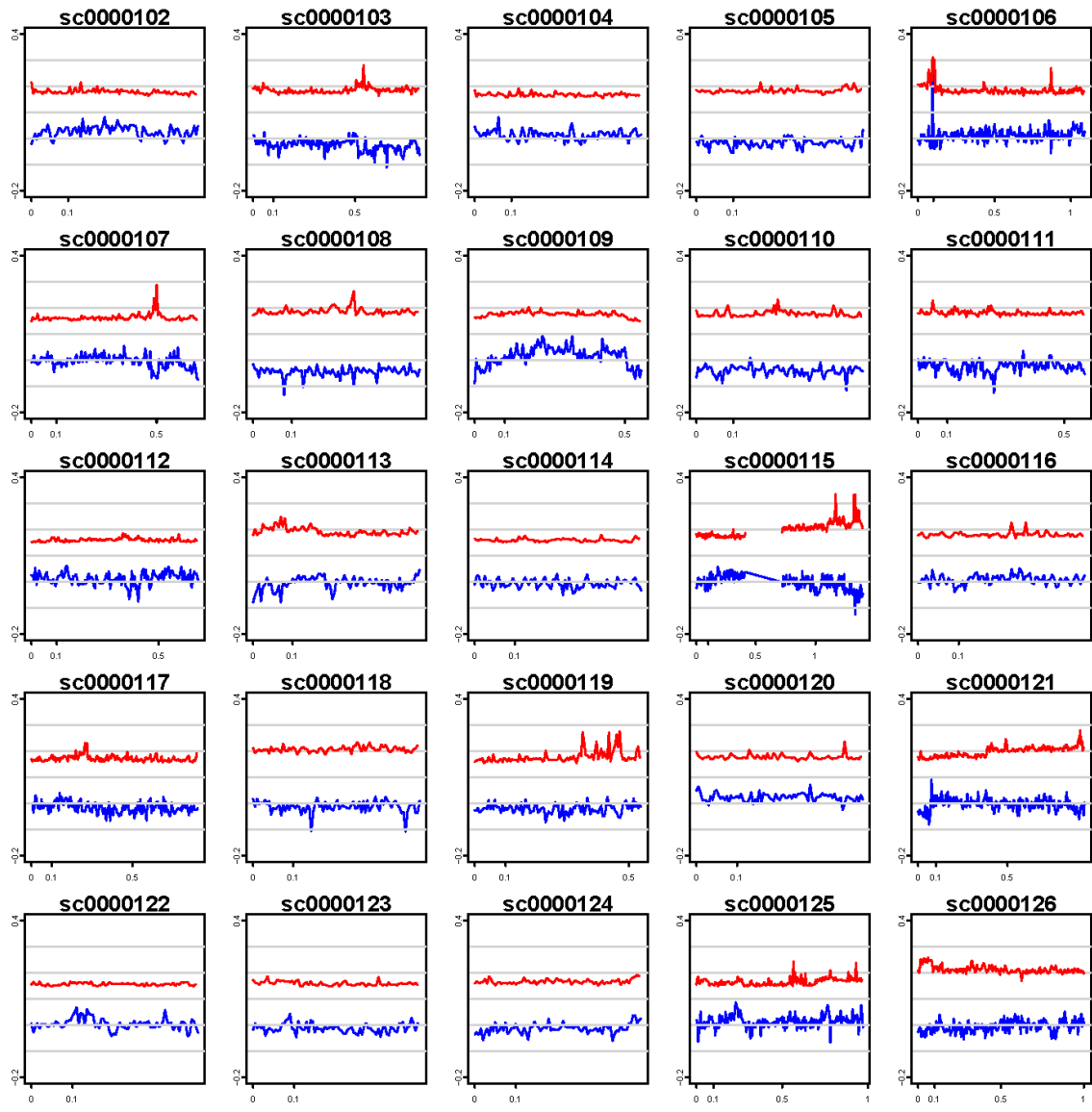


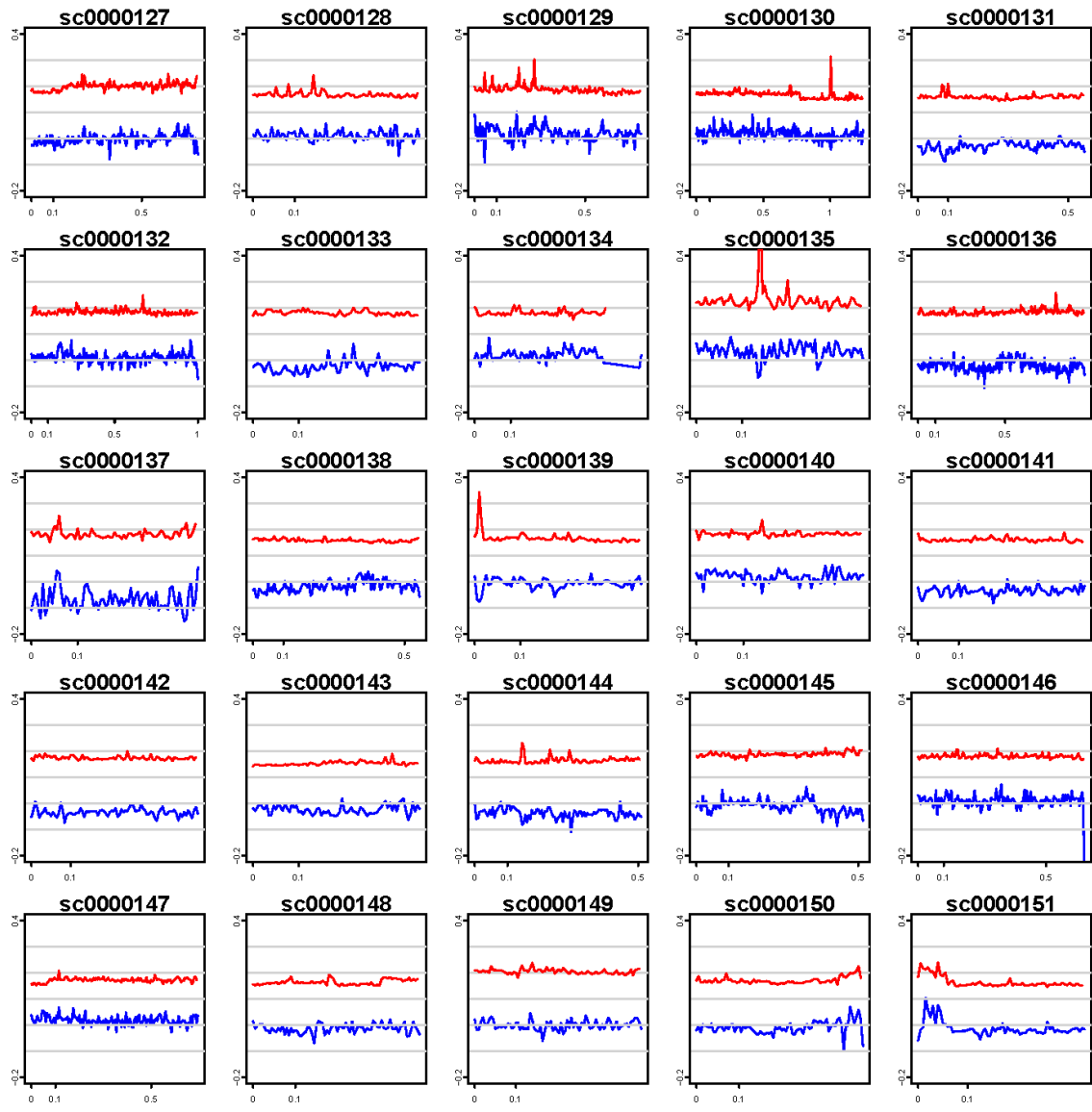


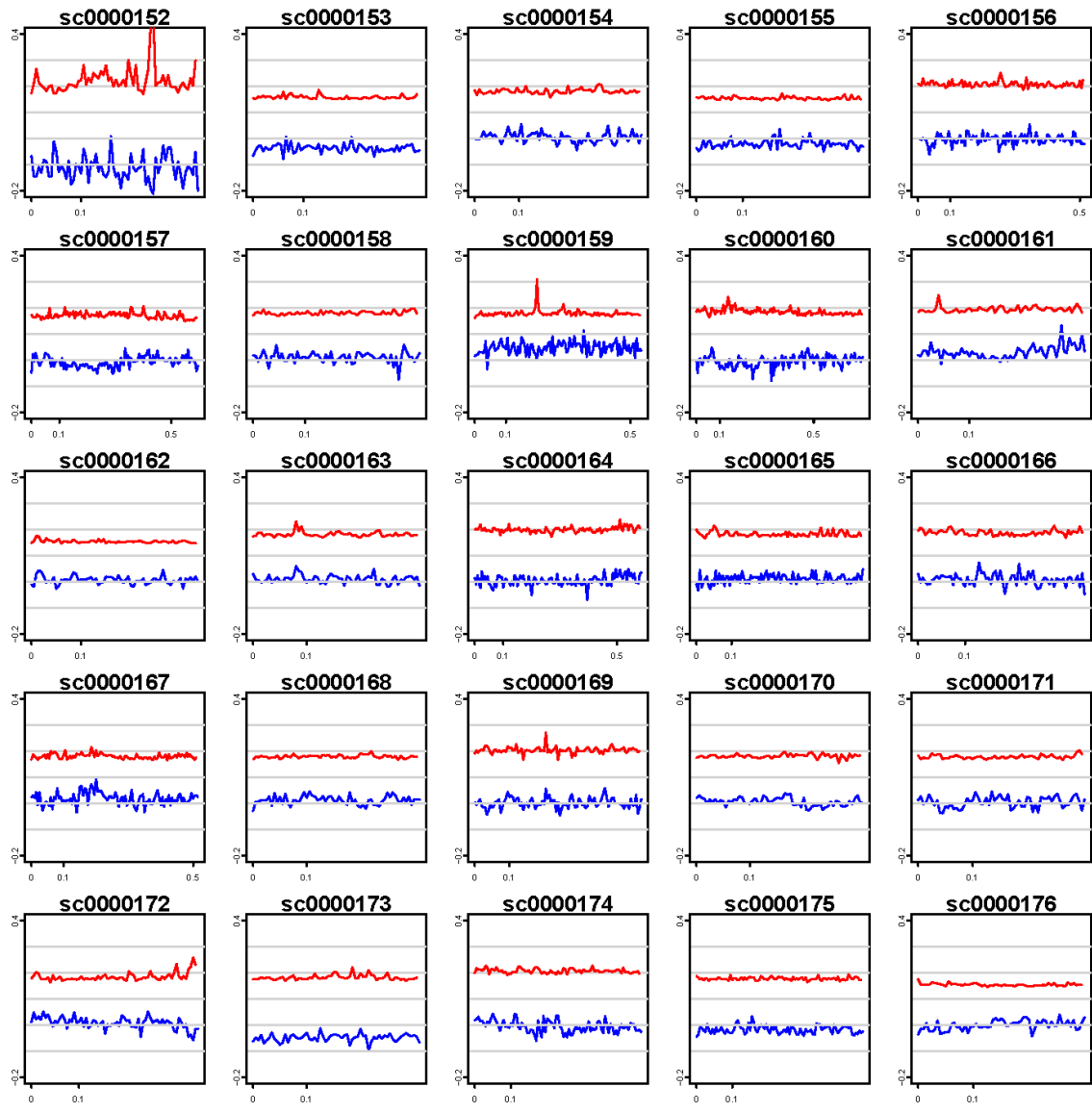


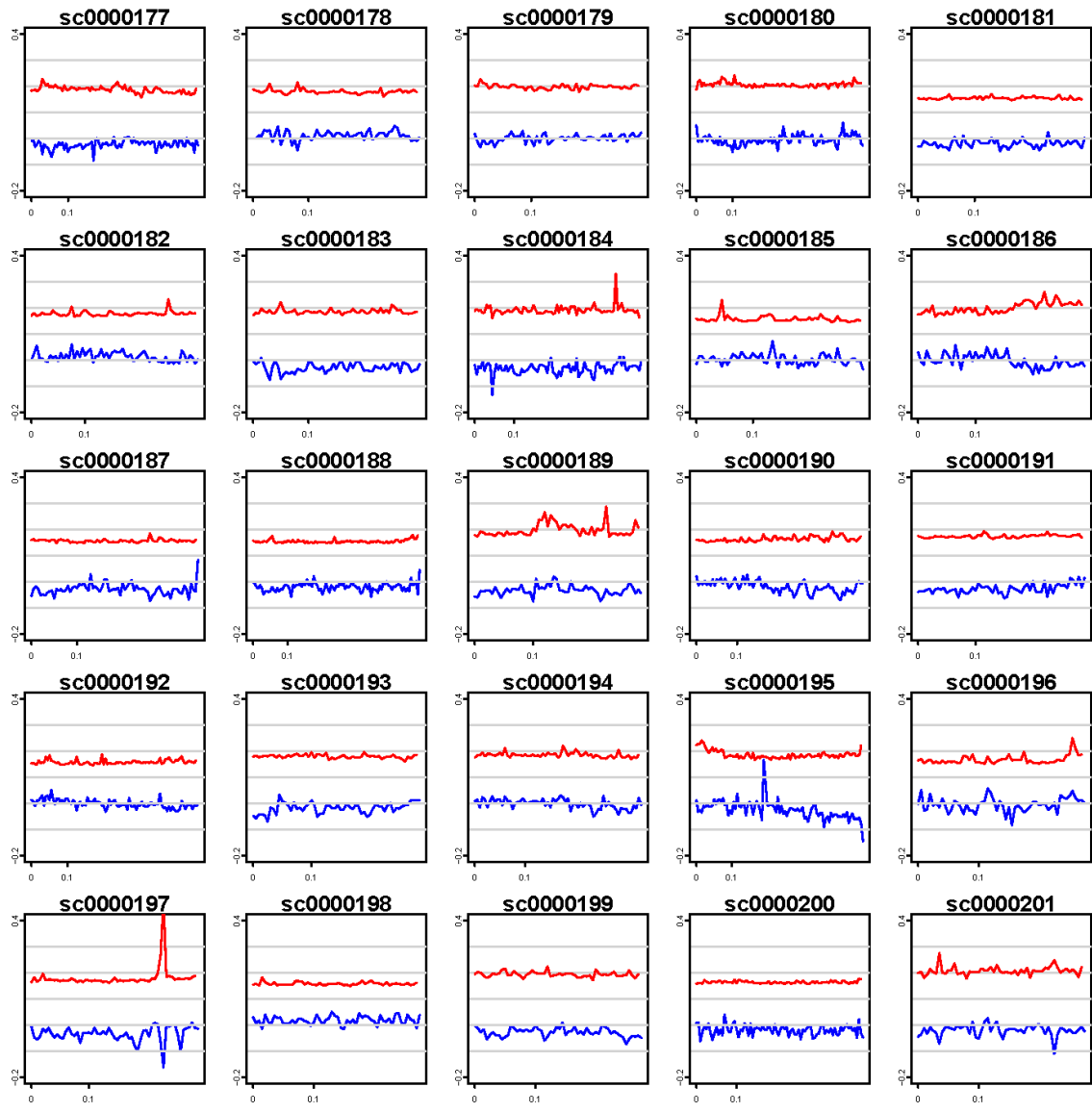


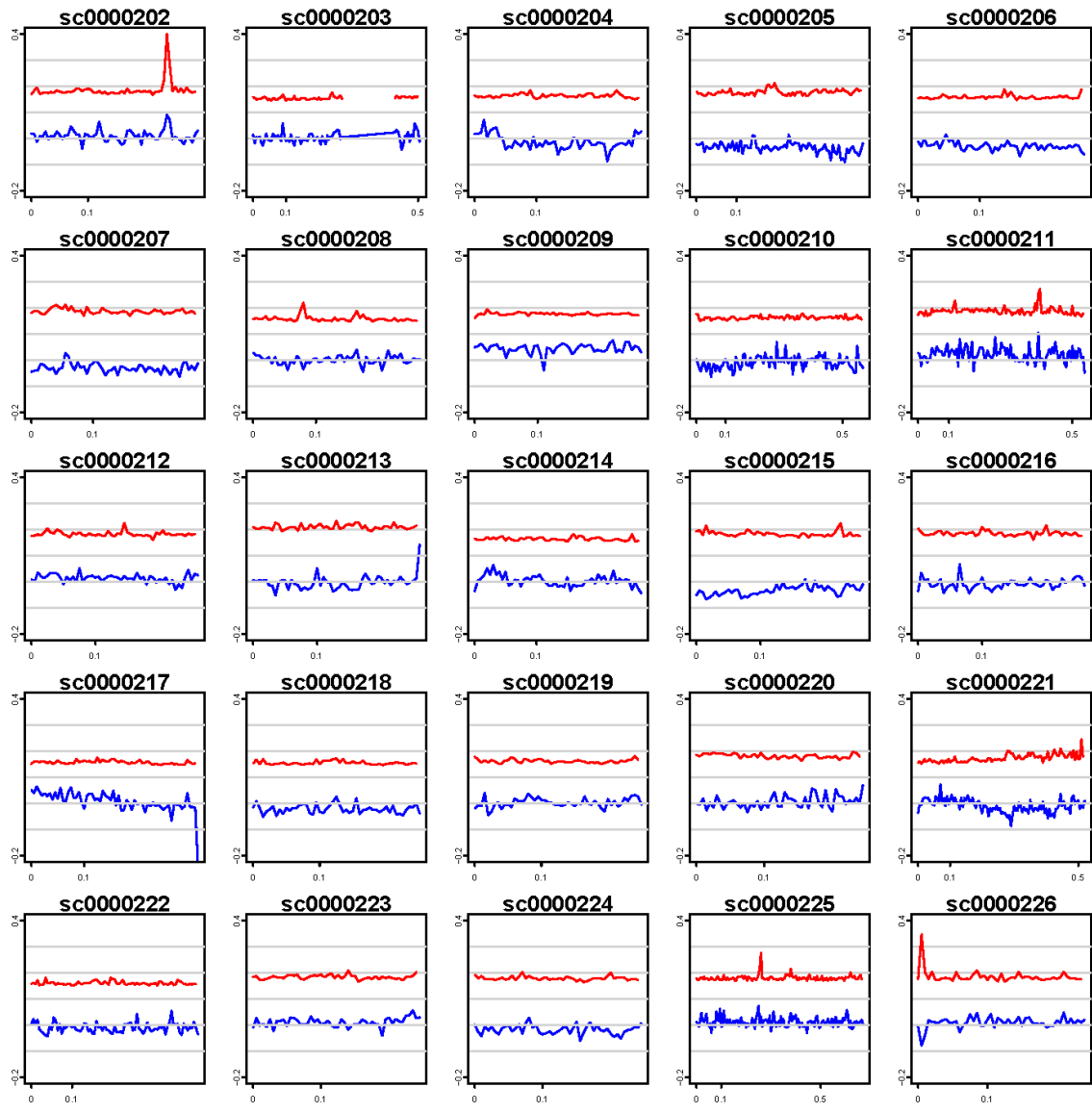


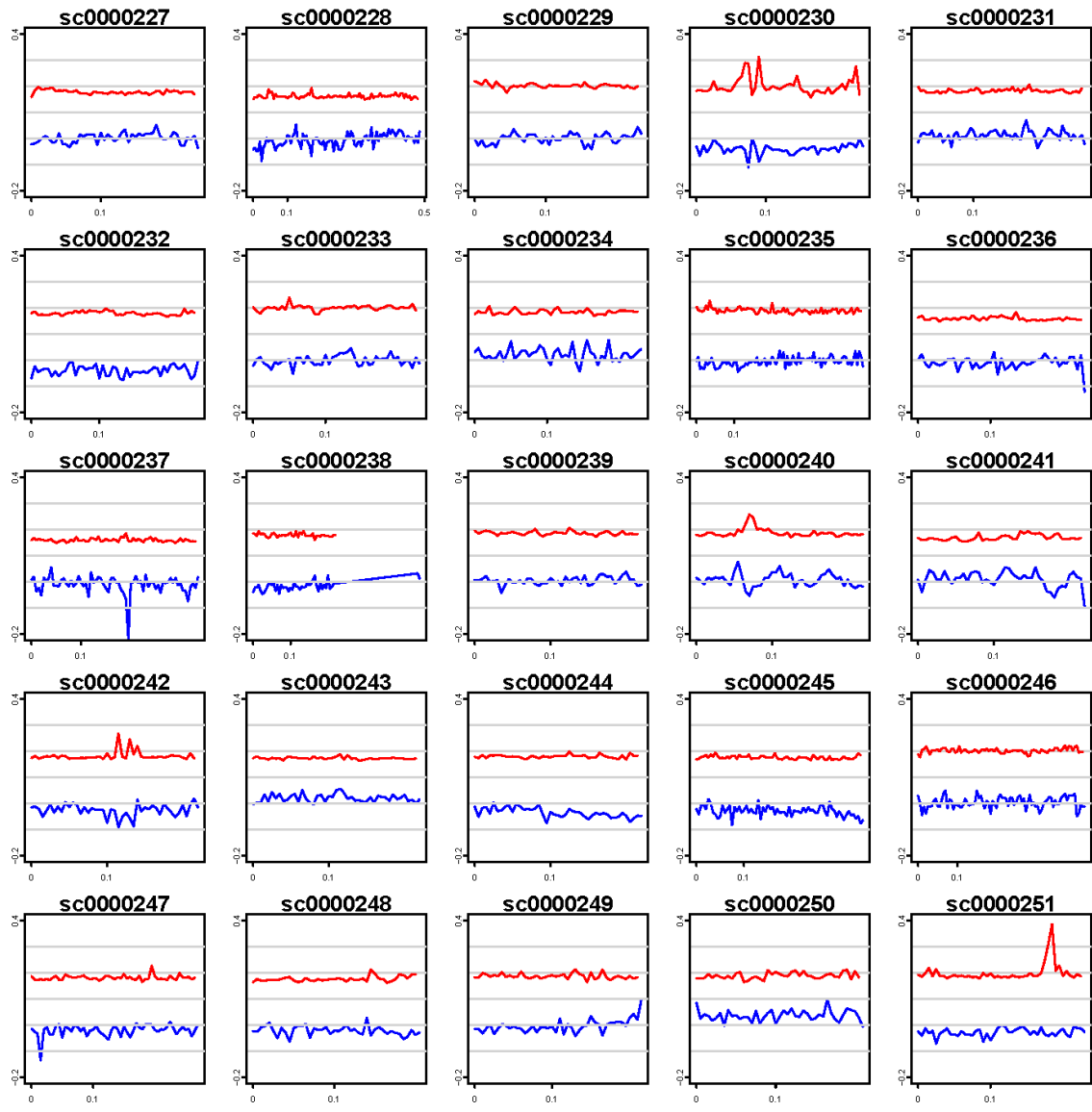


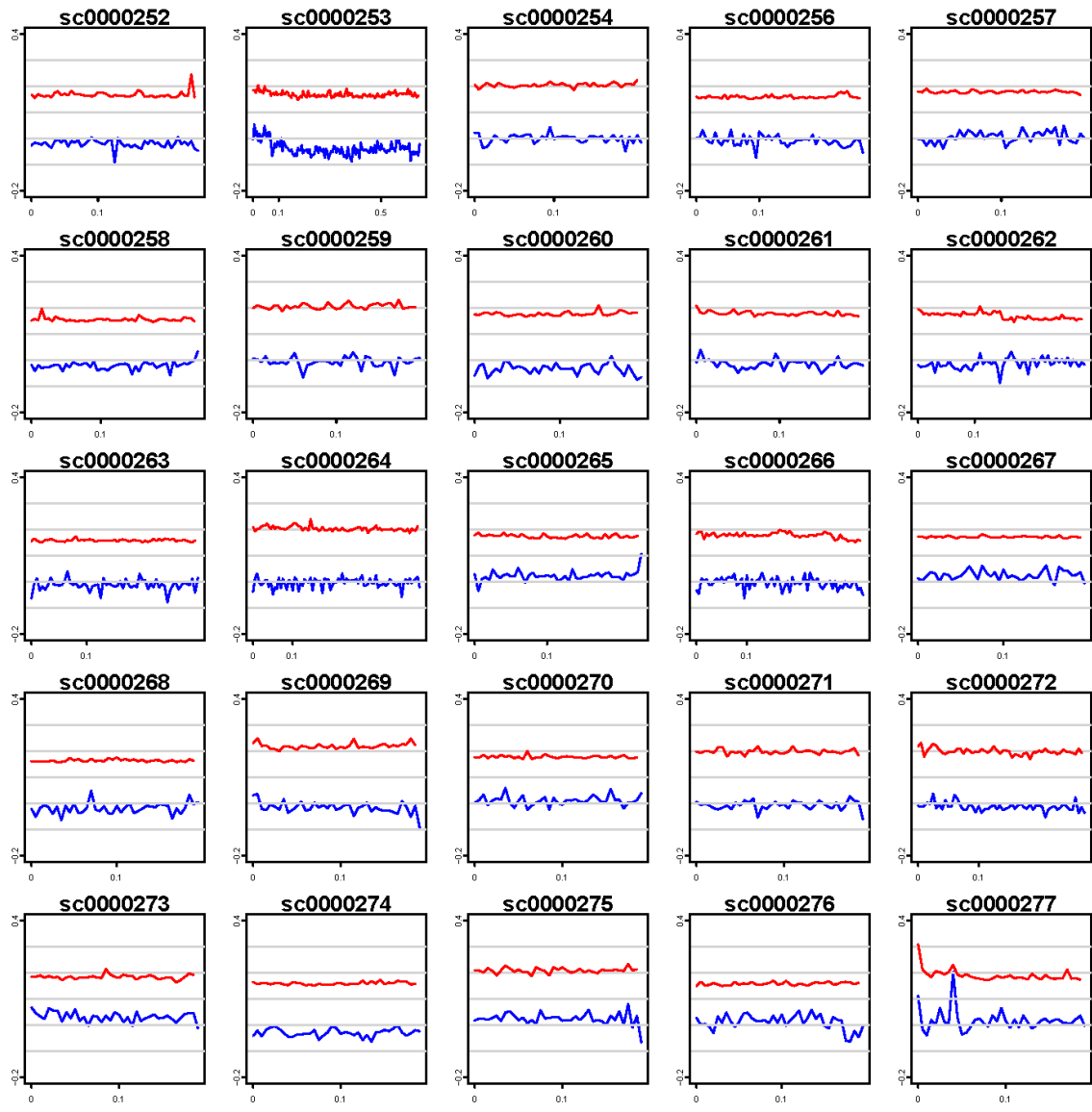


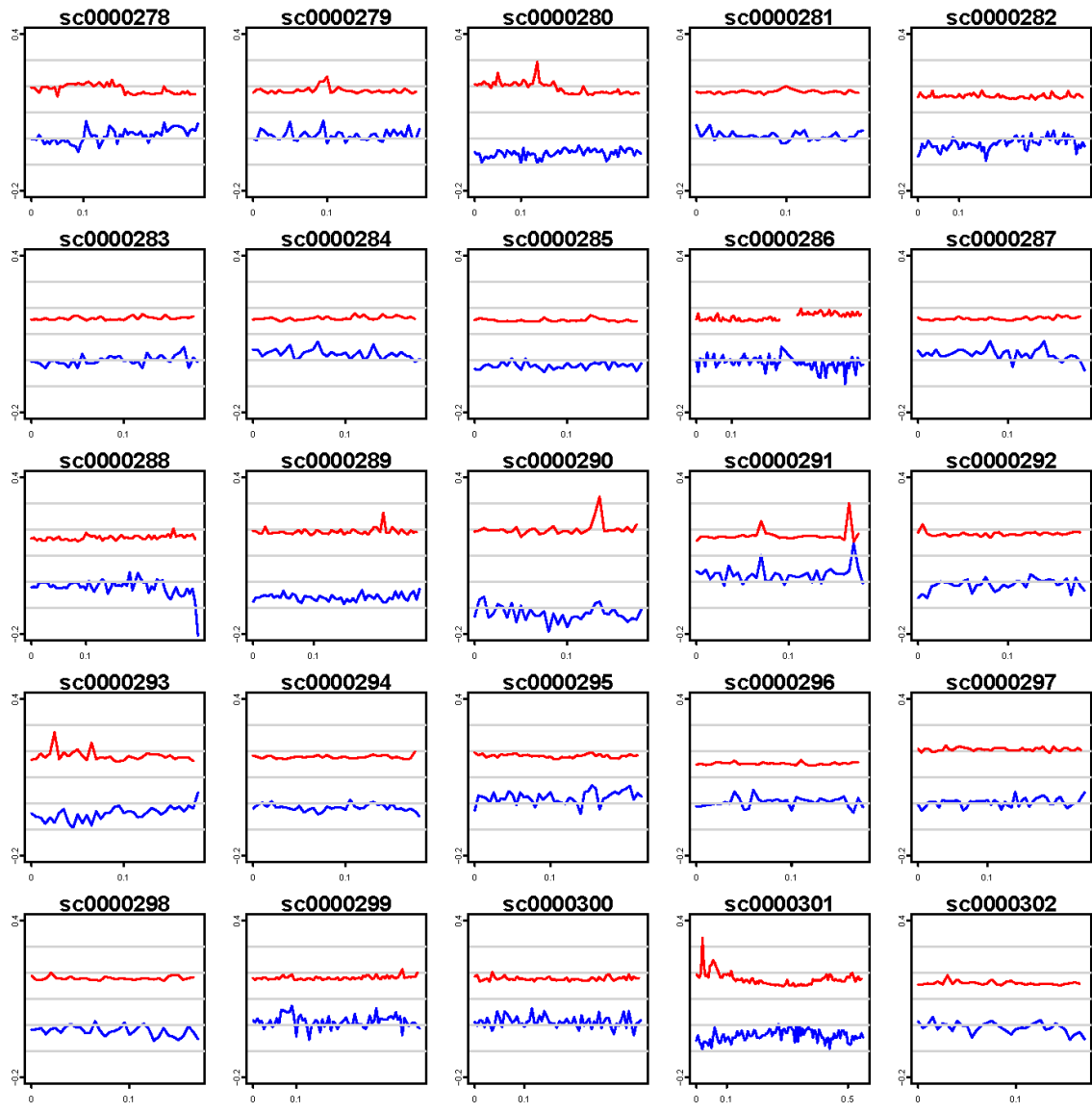




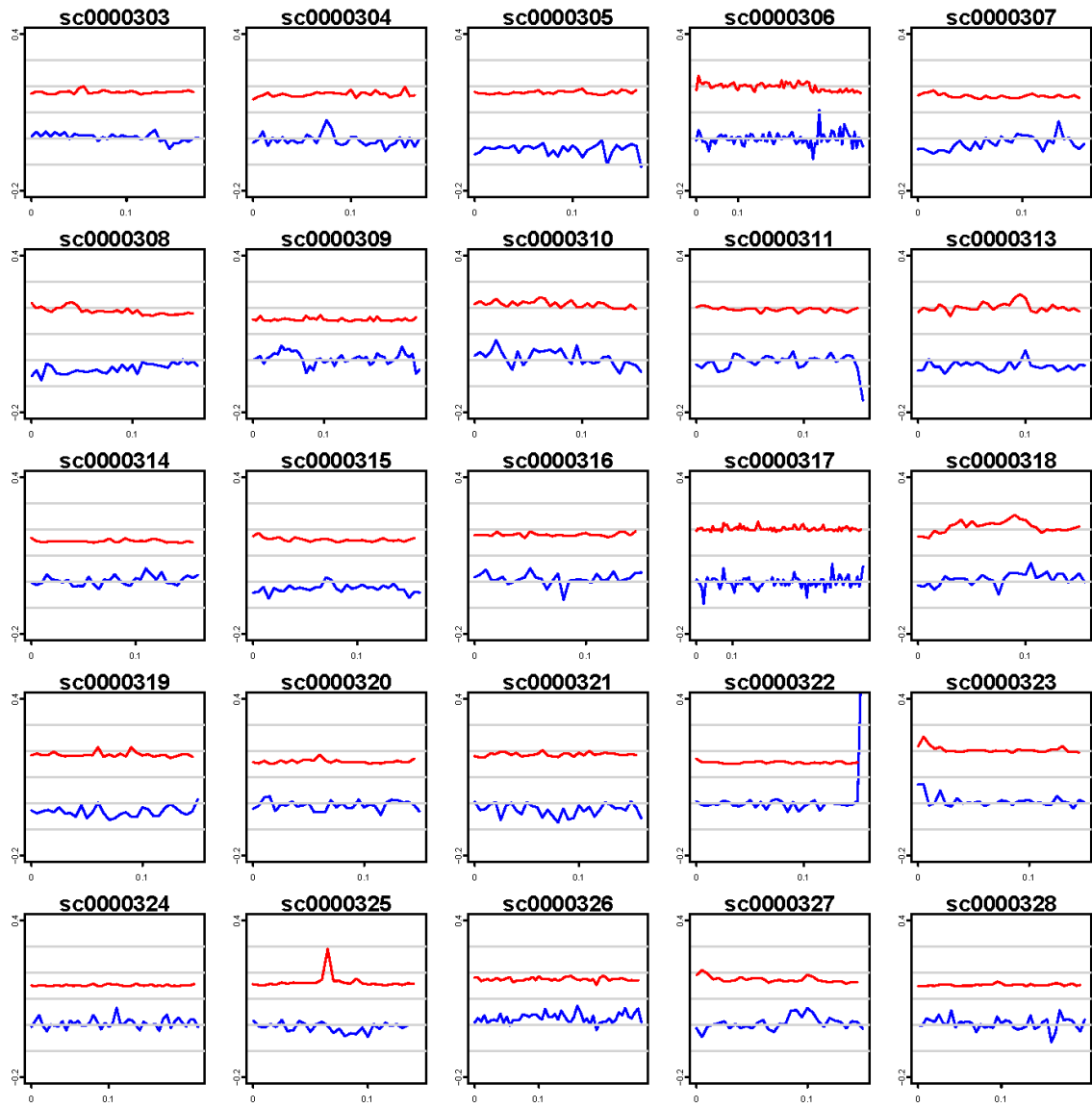


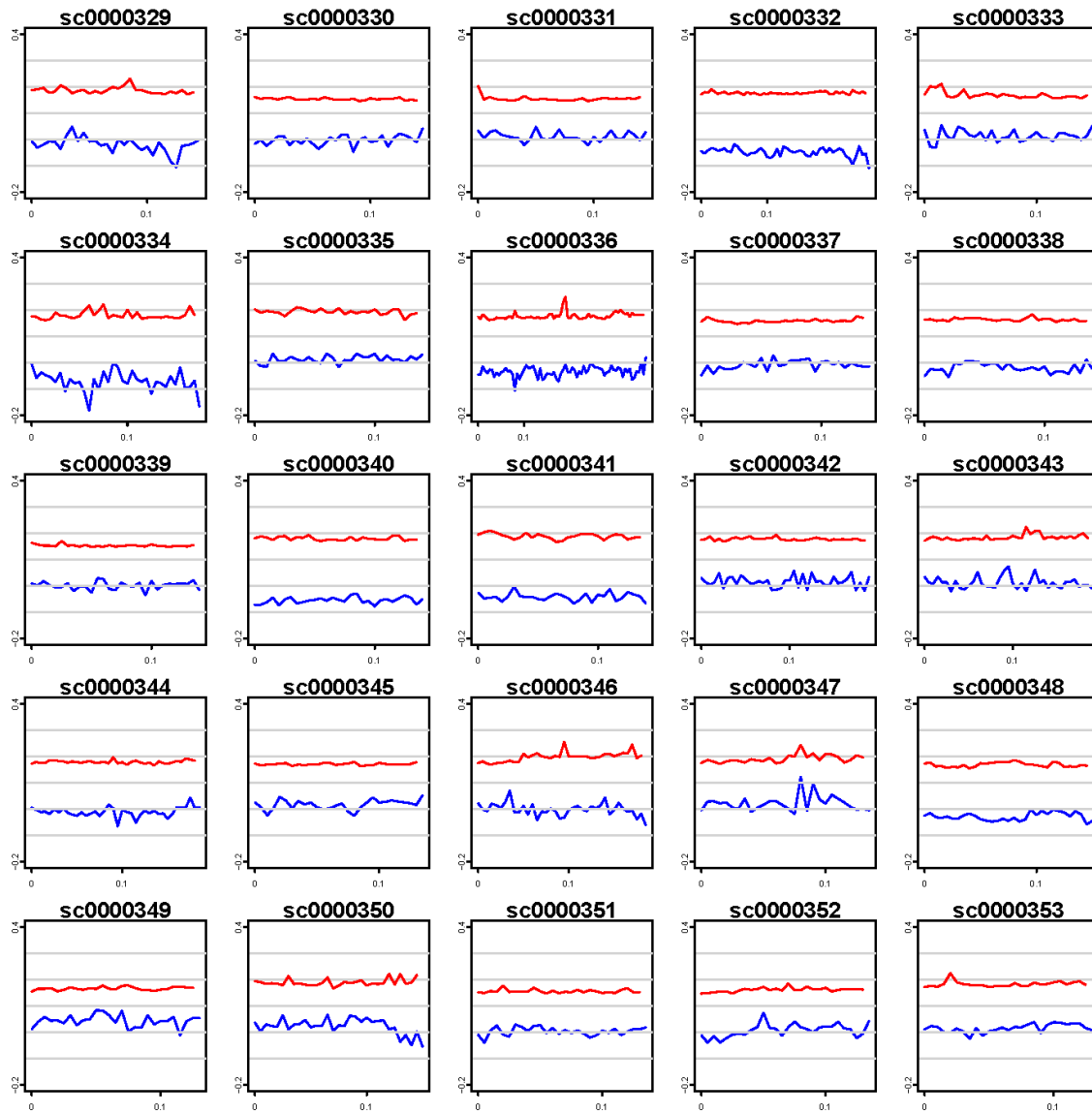


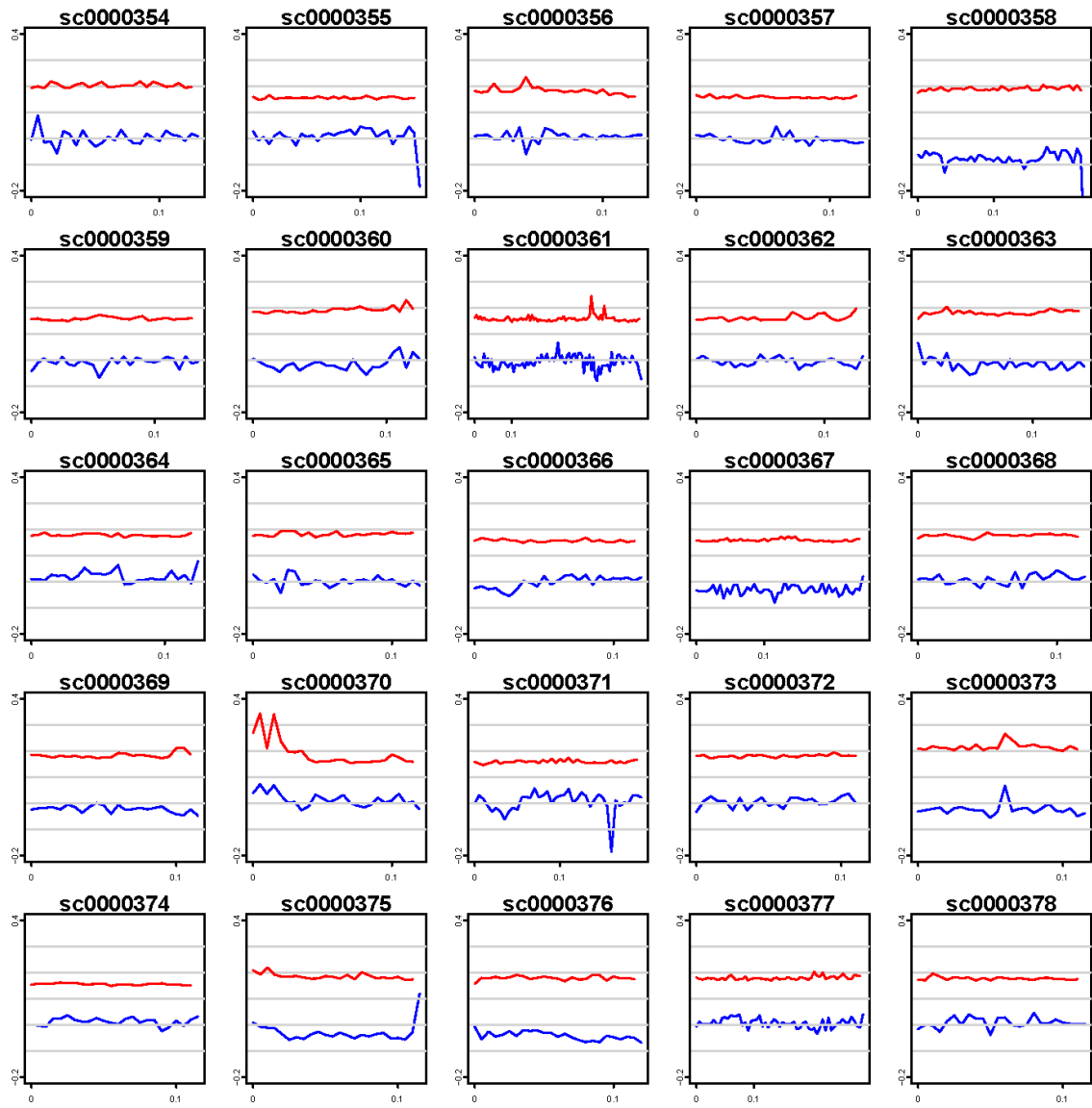


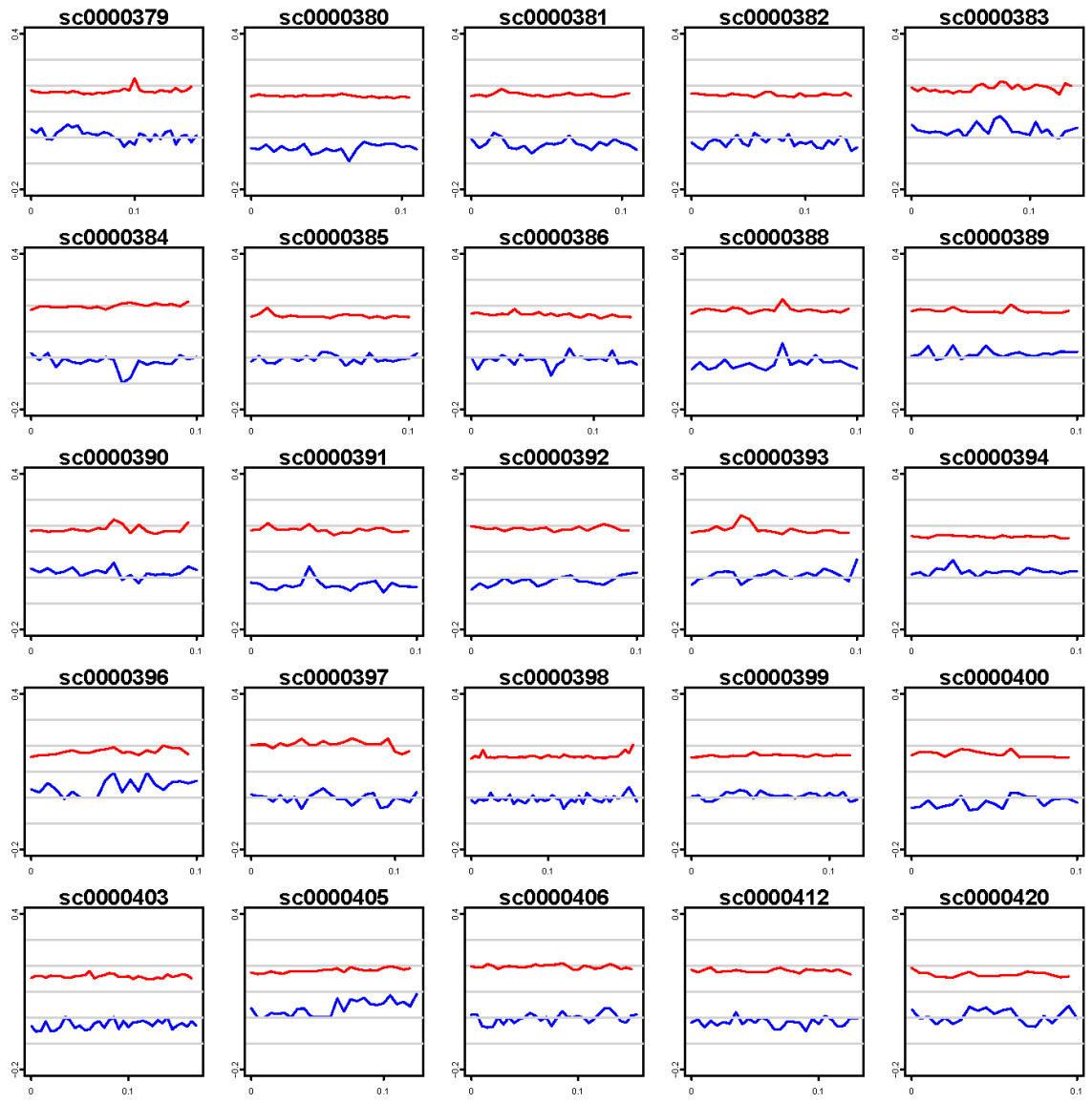


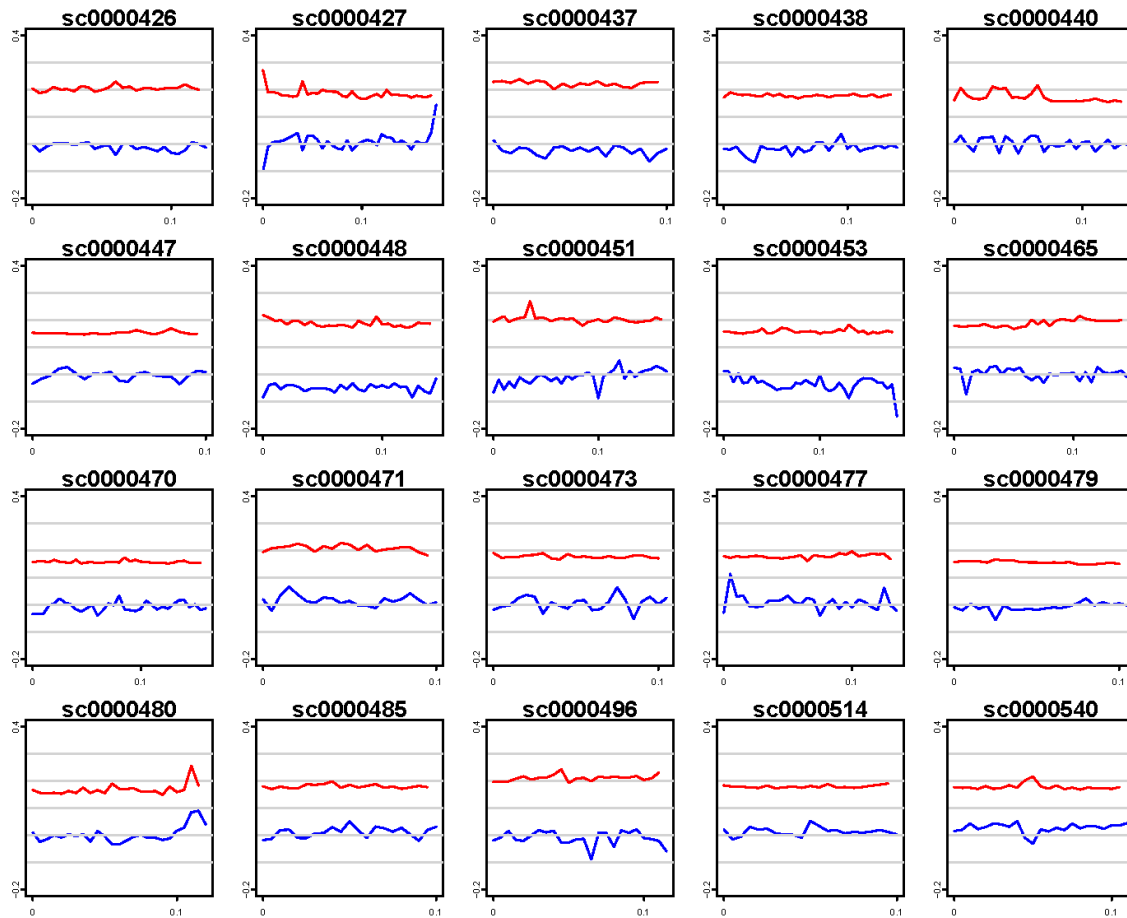




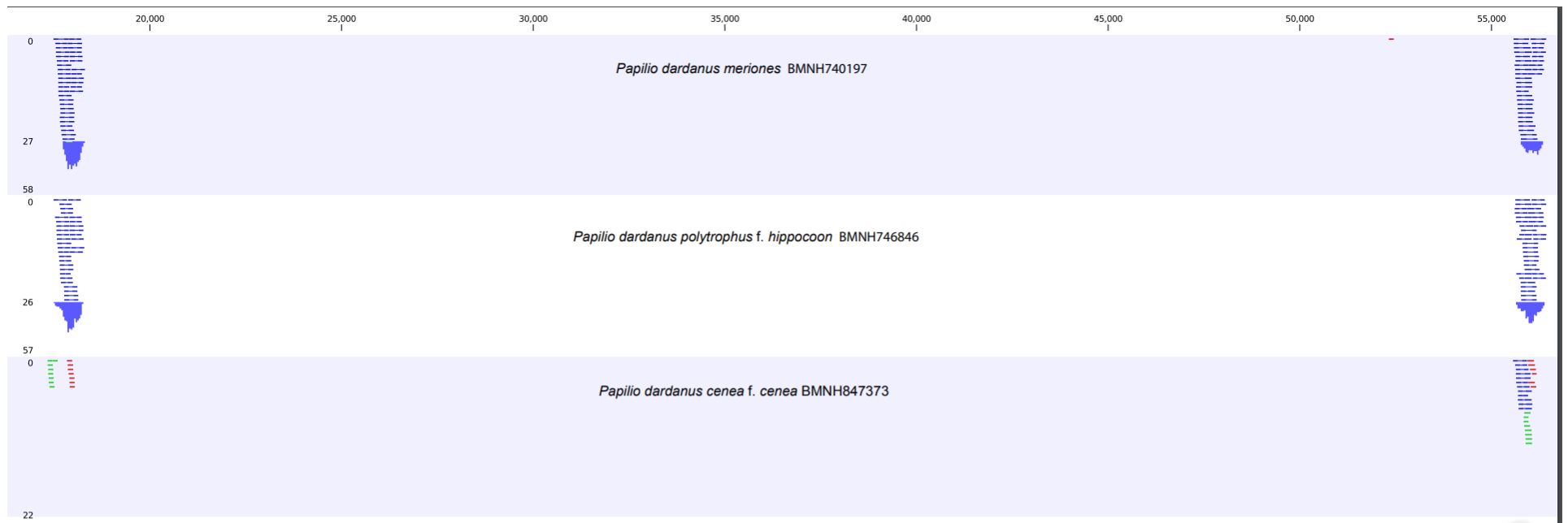








**Figure S4: Linkage disequilibrium (blue) and  $F_{st}$  (red) for the 420 scaffolds > 100 kb.** The first graph (“engrailed”) represents the scaffold containing the engrailed locus. engrailed: LD and  $F_{st}$  for the scaffold containing the *engrailed-inverted* locus. Horizontal scale is in million base pair.



**Figure S5: Reads mapping near the inversion breakpoint.** Top: *Papilio dardanus meriones* ('male-like'), Middle: *Papilio dardanus* f. *hippococonides*, Bottom: *Papilio dardanus* f. *cenea*. Reads are mapped to the f. *hippococonides* based draft genome sequence. Correctly placed reads ( $\rightarrow \leftarrow$ ) are indicated in blue. Distant and oppositely-oriented placed paired reads ( $\rightarrow \rightarrow$  or  $\leftarrow \leftarrow$ ) are indicated in green and red. Numbers at the left indicate coverage.

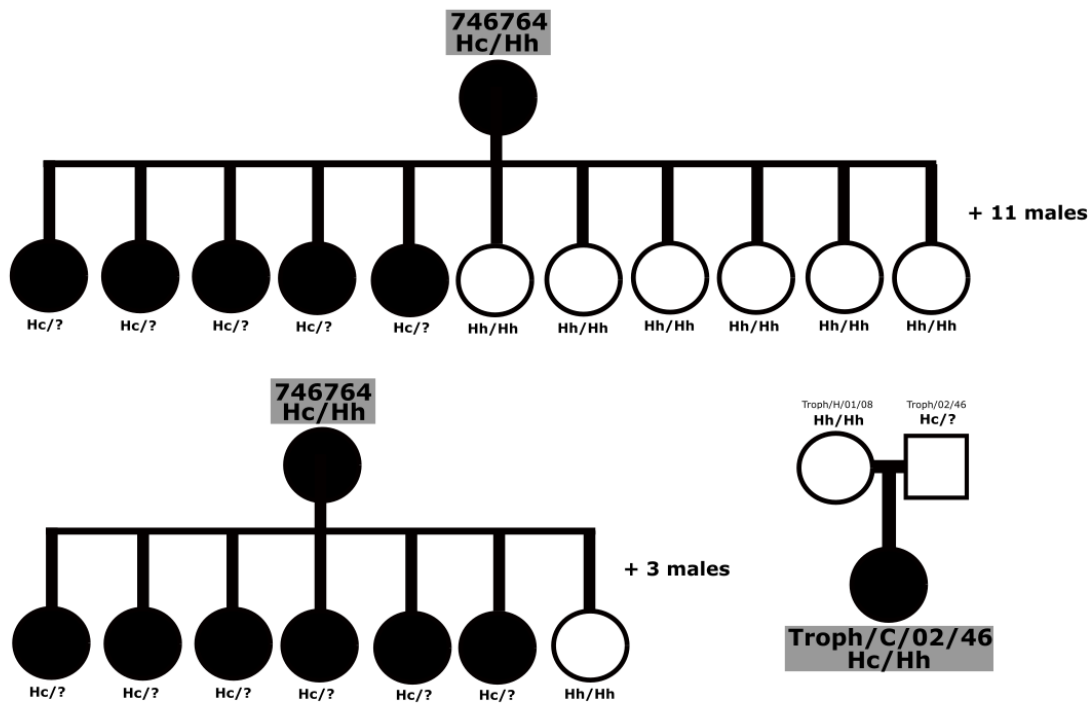
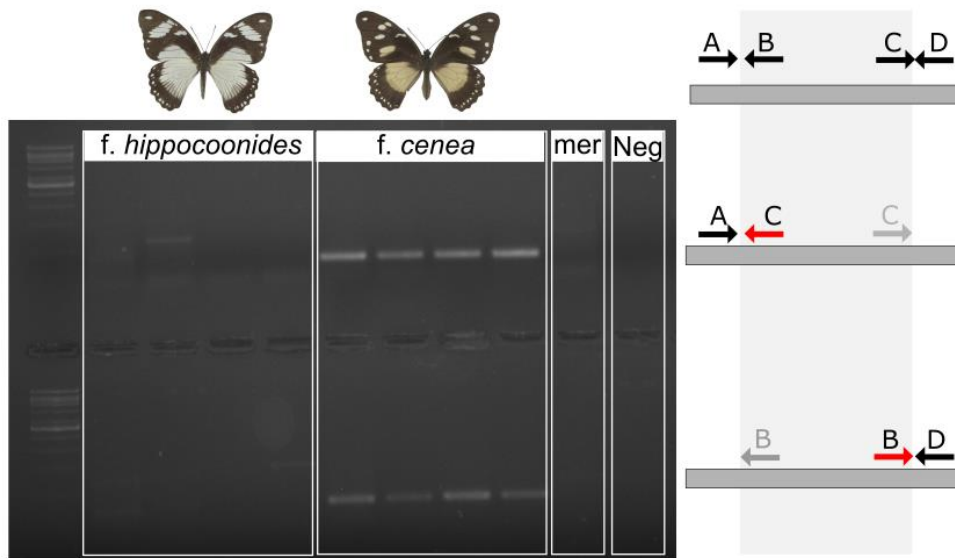


Figure S6: Laboratory cross details on the three heterozygous ( $H_c/H_h$ ) *P. dardanus polytrophus f. cenea* used for whole genome shotgun sequencing.



**Figure S7: PCR validation of inversions using females of two forms (*f. hippocooides* and *f. cenea*).** The *P. dardanus meriones* specimen used for whole genome shotgun sequencing is also included (“mer”). Primer combinations are given on the right.



## microRNA 2768 (miR-2768)

*Papilio dardanus* ATCAGGCGCCGGTGCAATATTTGACCAATTTTATAGCA---AACAATCGCTATAGCAAATGTGAATTGGTTAAGATAATGCATCGCCCGCTTGACT  
*Bombyx mori* (NR\_107338) GTCGGCGACCGGTGAATATATTTGACCAATTTTATAATCGCTCGATCG--TCAATCAAAGTGAATTGGTTAAGATAATGCATCGTTCCTGATC

**Figure S8: Sequence of microRNA 2768 located in the intronic region of *invected*.** miR-2768 is Lepidoptera specific and has been shown to interact with *cubitus interruptus*. Top: Sequence of *Papilio dardanus*, Bottom: Sequence of *Bombyx mori* (Genbank accession number: NR\_107338).

**Table S1. Completeness of the assembly as estimated using BUSCO version 3 compared to various recently published genomes.**

<b>Species</b>	<b>Number of scaffolds</b>	<b>N50</b>	<b>Total Length</b>	<b>BUSCO summary (%)</b>	<b>BUSCO summary (absolute)</b>
<i>Papilio dardanus</i>	<b>7,365</b>	596,599	231,123,043	C:89.3%[D:0.7%],F:7.0%,M:3.7%,n:2442	C: 2180 [D:16], F:172, M: 90, n:2,442
<i>Papilio glaucus</i>	<b>60,470</b>	230,841	374,815,656	C:92.2%[D:0.9%],F:5.6%,M:2.2%,n:2442	C: 2230 [D:22], F:136, M: 54, n:2,442
<i>Papilio machaon</i>	<b>63,187</b>	1,174,287	278,436,446	C:91.5%[D:1.0%],F:5.9%,M:2.6%,n:2442	C: 2234 [D:25], F:143, M: 65, n:2,442
<i>Papilio memnon</i>	<b>11,084</b>	5,456,730	232,544,191	C:94.9%[D:1.1%],F:3.1%,M:2.0%,n:2442	C: 2317 [D:26], F:75, M: 50, n:2,442
<i>Papilio polytes</i>	<b>3,874</b>	3,672,263	227,021,014	C:88.1%[D:0.3%],F:5.9%,M:6.0%,n:2442	C: 2152 [D:8], F:143, M: 147, n:2,442
<i>Papilio xuthus</i>	<b>5,572</b>	6,198,915	243,890,167	C:96.1%[D:1.3%],F:1.2%,M:2.7%,n:2442	C: 2347 [D:31], F:30, M: 65, n:2,442

**Table S2. Information on *Papilio dardanus* samples used in this study.**

	BIOLOGICAL SAMPLE	SOURCE	VOUCHER NUMBER	YEAR COLLECTED
Draft Genome Sequencing	<i>Papilio dardanus tibullus</i> (inbred male)	Laboratory bred specimen	BMNH1043181	X
WGS sequencing	<i>Papilio dardanus polytrophus</i> f. <i>hippocoonides</i>	Kenya, via ABRI	BMNH746826	2002
	<i>Papilio dardanus polytrophus</i> f. <i>hippocoonides</i>	Kenya, via ABRI	BMNH746846	2003
	<i>Papilio dardanus polytrophus</i> f. <i>hippocoonides</i>	Kenya, via ABRI	BMH746848	2003
	<i>Papilio dardanus cenea</i> f. <i>hippocoonides</i>	South-Africa	BMNH847389	2013
	<i>Papilio dardanus polytrophus</i> f. <i>cenea</i>	Kenya, via ABRI	BMNH746453	2002
	<i>Papilio dardanus polytrophus</i> f. <i>cenea</i>	Kenya, via ABRI	BMNH746764	2003
	<i>Papilio dardanus polytrophus</i> f. <i>cenea</i>	Laboratory bred specimen	Troph-c-02-46	X
	<i>Papilio dardanus cenea</i> f. <i>cenea</i>	South-Africa	BMNH847353	2013
	<i>Papilio dardanus meriones</i>	Madagascar	BMNH740197	2003
Validation of inversion via PCR	<i>Papilio dardanus</i> f. <i>hippocoonides</i>	Kenya, via ABRI	BMNH746615	2003
	<i>Papilio dardanus</i> f. <i>hippocoonides</i>	Kenya, via ABRI	BMNH746616	2003
	<i>Papilio dardanus</i> f. <i>hippocoonides</i>	Kenya, via ABRI	BMNH746618	2003
	<i>Papilio dardanus tibullus</i> f. <i>hippocoonides</i>	Laboratory bred specimen	BMNH746650	X
	<i>Papilio dardanus</i> f. <i>cenea</i>	Kenya, via ABRI	BMNH746747	2002
	<i>Papilio dardanus</i> f. <i>cenea</i>	Kenya, via ABRI	BMNH746624	2003
	<i>Papilio dardanus</i> f. <i>cenea</i>	Kenya, via ABRI	BMNH746619	2003
	<i>Papilio dardanus</i> f. <i>cenea</i>	Kenya, via ABRI	BMNH746627	2003
RADseq - Brood 48	Parent Brood 48	Laboratory bred specimen	BMNH1041083	X
	Parent Brood 48	Laboratory bred specimen	BMNH1041077	X
RADseq - Brood 59	Parent Brood 59	Laboratory bred specimen	BMNH1041063	X
	Parent Brood 59	Laboratory bred specimen	BMNH1041081	X

**Table S3: Fisher exact tests for association between inversion and phenotype.**

*P. dardanus f. hippocoonides* vs *P. dardanus f. cenea*

	no inversion	inversion
<i>f. hippocoonides</i> (recessive)	9*	0
<i>f. cenea</i>	0	8**

two tailed p value

<0.0001

\*) 5 Illumina shotgun sequence data + 4 PCR based.

\*\*\*) 4 Illumina shotgun sequence data sequenced + 4 PCR based.

**Table S4: Voucher numbers of specimens from brood 48 and brood 59 that show evidence for genetic recombination near *engrailed*. Positions of SNPs (along merged scaffold) in-between which recombination occurred are given.**

Brood	Voucher number	SNP1	SNP2
48	BMNH1043214	1155503	1644533
	BMNH1043229	1996091	2183720
59	BMNH1043209	195408	308710
	BMNH1043194	777466	777466
	BMNH1043200	1025647	1148751
	BMNH1043155	1742876	1836404
	BMNH1043204	1742876	1836404
	BMNH1043207	1742876	1836404
	BMNH1043244	1742876	1836404