

501 **Supporting information**

502 Table S1. The result of BUSCO analysis to evaluate the correctness of gene annotations from sfC
503 and sfR genome assemblies generated in this study (New assembly) and in our previous study
504 (Gouin et al.).

505

	sfC		sfR	
	New assembly	Gouin et al	New assembly	Gouin et al
506 Complete	1,542	1,434	1,562	1,551
Complete and single-copy	1,514	1,240	1,525	1,514
507 Complete and duplicated	28	194	37	37
Fragmented	52	164	35	67
508 Missing	64	60	61	40
Total	1,658	1,658	1,658	1,658

509

510

511

512

513

514

Table S2. Genes within genetic outliers of differentiation identified from the mapping against sfC reference genome

gene ID	Gene name
SFRUCORN0000000043	DDX20_DANRE Probable ATP-dependent RNA helicase DDX20
SFRUCORN0000000044	CG5027-PA
SFRUCORN00000000779	CG1461-PB
SFRUCORN00000000822	KCNT1_CHICK Potassium channel subfamily T member 1
SFRUCORN00000000824	unknown
SFRUCORN00000000825	unknown
SFRUCORN00000000871	DOS_DROME Protein daughter of sevenless
SFRUCORN00000000893	p77
SFRUCORN00000000897	POLO_DROME Serine/threonine-protein kinase polo
SFRUCORN00000000898	unknown
SFRUCORN00000000905	CG8320-PA
SFRUCORN00000000910	PGRP2_HOLDI Peptidoglycan-recognition protein 2
SFRUCORN00000000920	unknown
SFRUCORN00000000923	CG9302-PA
SFRUCORN00000000924	tpi
SFRUCORN00000000950	DYC1_CAEEL Dystrophin-like protein 1
SFRUCORN00000000988	PO11_BRACO Retrovirus-related Pol polyprotein from type-1 retrotransposable element R1 (Fragment)
SFRUCORN00000000997	unknown
SFRUCORN00000000998	unknown
SFRUCORN0000001000	unknown
SFRUCORN0000001068	Su(P)-PA
SFRUCORN0000001072	ACSA_DROME Acetyl-coenzyme A synthetase
SFRUCORN0000001110	SNF4Agamma-PA
SFRUCORN0000001526	Luciferase2
SFRUCORN0000001527	HTP
SFRUCORN0000001531	ADD_ECOUT Adenosine deaminase
SFRUCORN0000001641	NESD_DROME Protein nessun dorma
SFRUCORN0000001873	ACXD-PA
SFRUCORN0000001924	unknown
SFRUCORN0000002508	DDR2_HUMAN Discoidin domain-containing receptor 2
SFRUCORN0000002510	Rrp6-PC
SFRUCORN0000002511	Syx6-PD
SFRUCORN0000002524	GLT25_AEDAE Glycosyltransferase 25 family member
SFRUCORN0000002529	unknown
SFRUCORN0000002537	sl-PA
SFRUCORN0000002538	PLCB3_RAT 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-3
SFRUCORN0000002539	babo-PA
SFRUCORN0000002552	unknown
SFRUCORN0000002553	unknown
SFRUCORN0000002557	DNAJ
SFRUCORN0000002593	unknown
SFRUCORN0000002605	unknown
SFRUCORN0000002607	unknown
SFRUCORN0000002623	ap1
SFRUCORN0000002656	HEAT6_DANRE HEAT repeat-containing protein 6
SFRUCORN0000002658	PPIL6_HUMAN Peptidyl-prolyl cis-trans isomerase-like 6

SFRUCORN0000002659 ABC
SFRUCORN0000003444 unknown
SFRUCORN0000003793 CG30345-PB
SFRUCORN0000004172 unknown
SFRUCORN0000004173 unknown
SFRUCORN0000004226 unknown
SFRUCORN0000004229 Indy-PD
SFRUCORN0000004474 MET16_CHICK U6 small nuclear RNA (adenine-(43)-N(6))-methyltransferase
SFRUCORN0000004475 Ref1-PA
SFRUCORN0000004483 PGS2_CHICK Decorin
SFRUCORN0000004632 HTP
SFRUCORN0000004665 Gug
SFRUCORN0000005138 unknown
SFRUCORN0000005308 CG7888-PA
SFRUCORN0000005346 FRYL_HUMAN Protein furry homolog-like
SFRUCORN0000005351 unknown
SFRUCORN0000005366 NPRL2_DROME GATOR complex protein NPRL2
SFRUCORN0000005390 VP13A_DICDI Putative vacuolar protein sorting-associated protein 13A
SFRUCORN0000005413 unknown
SFRUCORN0000005438 SdhC-PB
SFRUCORN0000005440 unknown
SFRUCORN0000005461 CBPA3_HUMAN Mast cell carboxypeptidase A
SFRUCORN0000006007 unknown
SFRUCORN0000006164 CG14253-PF
SFRUCORN0000006654 unknown
SFRUCORN0000006896 unknown
SFRUCORN0000007069 ZN862_HUMAN Zinc finger protein 862
SFRUCORN0000007071 unknown
SFRUCORN0000007189 unknown
SFRUCORN0000007281 unknown
SFRUCORN0000007308 unknown
SFRUCORN0000007350 Jarid2-PD
SFRUCORN0000007353 unknown
SFRUCORN0000007444 unknown
SFRUCORN0000007447 unknown
SFRUCORN0000007448 unknown
SFRUCORN0000007451 unknown
SFRUCORN0000007832 futsch
SFRUCORN0000007843 ATP5J_DROME ATP synthase-coupling factor 6, mitochondrial
SFRUCORN0000008195 Dhc98D-PC
SFRUCORN0000008867 unknown
SFRUCORN0000009055 unknown
SFRUCORN0000009271 unknown
SFRUCORN0000009280 unknown
SFRUCORN0000009284 unknown
SFRUCORN0000009383 LPH_HUMAN Lactase-phlorizin hydrolase
SFRUCORN0000009862 unknown
SFRUCORN0000010450 Iris-PA

SFRUCORN0000010451 HTP
SFRUCORN0000010513 unknown
SFRUCORN0000010557 unknown
SFRUCORN0000011018 unknown
SFRUCORN0000011084 unknown
SFRUCORN0000011087 unknown
SFRUCORN0000011834 CG15440-PA
SFRUCORN0000012563 Jarid2-PD
SFRUCORN0000012602 l(1)G0289-PC
SFRUCORN0000012864 CG13293-PC
SFRUCORN0000012917 CG10467-PA
SFRUCORN0000013102 GR146
SFRUCORN0000013104 GR140
SFRUCORN0000013185 unknown
SFRUCORN0000013320 CG12084-PB
SFRUCORN0000013321 unknown
SFRUCORN0000013329 BAG6B_XENLA Large proline-rich protein bag6-B
SFRUCORN0000013361 Socs16D-PA
SFRUCORN0000013362 unknown
SFRUCORN0000013389 fest-PB
SFRUCORN0000013390 TSN
SFRUCORN0000013619 unknown
SFRUCORN0000013643 unknown
SFRUCORN0000013644 unknown
SFRUCORN0000013645 unknown
SFRUCORN0000013647 unknown
SFRUCORN0000013695 unknown
SFRUCORN0000013697 HTP
SFRUCORN0000013726 PARP_SARPE Poly [ADP-ribose] polymerase
SFRUCORN0000013727 PELET_DROME Transposable element P transposase
SFRUCORN0000013895 CG5002-PA
SFRUCORN0000014152 CHIT2_DROME Probable chitinase 2
SFRUCORN0000014377 unknown
SFRUCORN0000014518 unknown
SFRUCORN0000014639 unknown
SFRUCORN0000015156 unknown
SFRUCORN0000015548 unknown
SFRUCORN0000015549 HTP
SFRUCORN0000015551 unknown
SFRUCORN0000015621 Litaf
SFRUCORN0000015622 parvin-PB
SFRUCORN0000015644 CG1246-PD
SFRUCORN0000015649 DYH6_HUMAN Dynein heavy chain 6, axonemal
SFRUCORN0000015652 unknown
SFRUCORN0000015661 DopEcR-PC
SFRUCORN0000015676 unknown
SFRUCORN0000015840 unknown
SFRUCORN0000015934 unknown

SFRUCORN0000015937 Iris-PA
SFRUCORN0000015938 HTP
SFRUCORN0000016603 unknown
SFRUCORN0000017016 unknown
SFRUCORN0000017116 unknown
SFRUCORN0000017788 unknown
SFRUCORN0000018022 f-cup-PB
SFRUCORN0000018552 unknown
SFRUCORN0000018597 TM209_XENTR Transmembrane protein 209
SFRUCORN0000018598 PCH2_DANRE Pachytene checkpoint protein 2 homolog
SFRUCORN0000018599 Gbeta5-PA
SFRUCORN0000018601 Bsg-PB
SFRUCORN0000018602 ERMP1_MOUSE Endoplasmic reticulum metalloproteinase 1
SFRUCORN0000018603 EDC3_DANRE Enhancer of mRNA-decapping protein 3
SFRUCORN0000018652 CG1561-PB
SFRUCORN0000018654 unknown
SFRUCORN0000018760 unknown
SFRUCORN0000018765 unknown
SFRUCORN0000018790 unknown
SFRUCORN0000018791 PELET_DROME Transposable element P transposase
SFRUCORN0000019197 CG6142-PA
SFRUCORN0000019207 Iris-PA
SFRUCORN0000019269 Gba1b-PC
SFRUCORN0000019382 unknown
SFRUCORN0000019566 DYHC_TRIGR Dynein beta chain, ciliary
SFRUCORN0000019904 unknown
SFRUCORN0000019909 unknown
SFRUCORN0000020413 LDH_DROME L-lactate dehydrogenase
SFRUCORN0000021306 HTP
SFRUCORN0000021307 Iris-PA
SFRUCORN0000021333 HTP
SFRUCORN0000021418 unknown
SFRUCORN0000021509 LRRX1_DICDI Putative leucine-rich repeat-containing protein DDB_G0290503
SFRUCORN0000021681 unknown
SFRUCORN0000021682 unknown
SFRUCORN0000021727 unknown
SFRUCORN0000021788 unknown
SFRUCORN0000021960 CG5589-PA
SFRUCORN0000022083 FEN1_DROVI Flap endonuclease 1
SFRUCORN0000022294 DNJ60_DROME DnaJ-like protein 60
SFRUCORN0000022384 unknown
SFRUCORN0000022405 unknown
SFRUCORN0000022410 Nost-PD
SFRUCORN0000022561 unknown
SFRUCORN0000022670 RTBS_DROME Probable RNA-directed DNA polymerase from transposon BS
SFRUCORN0000022876 Cat
SFRUCORN0000023028 CG2065-PB
SFRUCORN0000023154 unknown

SFRUCORN0000023935 unknown
SFRUCORN0000024133 CG17292-PD
SFRUCORN0000024207 TT21B_HUMAN Tetratricopeptide repeat protein 21B
SFRUCORN0000024251 EIF1A_XENLA Probable RNA-binding protein EIF1AD
SFRUCORN0000024469 PIF1_BDEBA ATP-dependent DNA helicase pif1
SFRUCORN0000024982 CYP340AA2P
SFRUCORN0000024985 unknown
SFRUCORN0000025268 unknown
SFRUCORN0000025550 unknown
SFRUCORN0000025554 S35B1_DROME Solute carrier family 35 member B1 homolog
SFRUCORN0000025824 unknown
SFRUCORN0000025825 Gyf-PF
SFRUCORN0000025836 unknown
SFRUCORN0000025865 ppk13-PA
SFRUCORN0000026045 b6-PA
SFRUCORN0000026401 unknown
SFRUCORN0000026481 ENV_NPVLD Envelope fusion protein
SFRUCORN0000026484 HTP
SFRUCORN0000026485 unknown
SFRUCORN0000027627 SP27A_DROME Serine protease inhibitor 27A
SFRUCORN0000027628 PGBD3_HUMAN PiggyBac transposable element-derived protein 3
SFRUCORN0000027968 HARB1_MOUSE Putative nuclease HARB11
SFRUCORN0000027969 unknown
SFRUCORN0000028024 unknown
SFRUCORN0000028058 ATC1_ANOGA Calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum type
SFRUCORN0000028059 unknown
SFRUCORN0000028060 NEUFC_DROME Neuferricin homolog
SFRUCORN0000028061 acyl
SFRUCORN0000028062 GGNB2_DROME Gametogenetin-binding protein 2-like
SFRUCORN0000028066 AL7A1_CAEEL Putative aldehyde dehydrogenase family 7 member A1 homolog
SFRUCORN0000028067 LR74A_HUMAN Leucine-rich repeat-containing protein 74A
SFRUCORN0000028451 Iris-PA
SFRUCORN0000028452 HTP
SFRUCORN0000028562 unknown
SFRUCORN0000028627 unknown
SFRUCORN0000028630 ZMPB_STRPN Zinc metalloprotease ZmpB
SFRUCORN0000028686 IR2
SFRUCORN0000028687 unknown
SFRUCORN0000028867 nmo-PC
SFRUCORN0000029853 Jon65Aiii-PA
SFRUCORN0000029855 p77
SFRUCORN0000029973 unknown
SFRUCORN0000030044 Polycalin
SFRUCORN0000030072 unknown
SFRUCORN0000030284 unknown
SFRUCORN0000030285 Iris-PA
SFRUCORN0000030543 unknown
SFRUCORN0000030680 PO11_BRACO Retrovirus-related Pol polyprotein from type-1 retrotransposable element R1 (Fragment)

SFRUCORN0000030685 unknown
SFRUCORN0000030735 Trh-PA
SFRUCORN0000030847 unknown
SFRUCORN0000031008 vib-PC
SFRUCORN0000031191 unknown
SFRUCORN0000031523 unknown
SFRUCORN0000031525 S7A14_DANRE Probable cationic amino acid transporter
SFRUCORN0000031736 HTP
SFRUCORN0000031761 CG5180-PC
SFRUCORN0000031823 unknown
SFRUCORN0000031840 HTP
SFRUCORN0000031898 unknown
SFRUCORN0000031978 unknown
SFRUCORN0000031979 unknown
SFRUCORN0000032339 unknown
SFRUCORN0000032340 unknown
SFRUCORN0000032613 unknown
SFRUCORN0000032783 unknown
SFRUCORN0000032786 ZBED1_HUMAN Zinc finger BED domain-containing protein 1
SFRUCORN0000032787 unknown
SFRUCORN0000032788 unknown
SFRUCORN0000032812 unknown
SFRUCORN0000032857 RL32_SPOFR 60S ribosomal protein L32
SFRUCORN0000032858 HTP
SFRUCORN0000032859 unknown
SFRUCORN0000032917 unknown
SFRUCORN0000032921 CG14688-PB
SFRUCORN0000032972 unknown
SFRUCORN0000032973 unknown
SFRUCORN0000033235 unknown
SFRUCORN0000033239 PGBD4_HUMAN PiggyBac transposable element-derived protein 4
SFRUCORN0000033242 unknown
SFRUCORN0000033374 LRC71_MOUSE Leucine-rich repeat-containing protein 71
SFRUCORN0000033375 unknown
SFRUCORN0000033376 CG30100-PB
SFRUCORN0000033602 HTP
SFRUCORN0000033968 EXOC2_DROME Exocyst complex component 2

Table S3. Genes within genetic outliers of differentiation identified from the mapping against sfR reference genome

gene ID	Gene name
SFRURICE0000000400	POE_DROME Protein purity of essence
SFRURICE0000000401	unknown
SFRURICE0000000460	unknown
SFRURICE0000000510	unknown
SFRURICE0000001614	unknown
SFRURICE0000001897	unknown
SFRURICE0000002022	unknown
SFRURICE0000002025	unknown
SFRURICE0000002026	unknown
SFRURICE0000002045	unknown
SFRURICE0000002049	unknown
SFRURICE0000002058	SNF4Agamma-PA
SFRURICE0000002073	CG41520-PA
SFRURICE0000002095	ACSA_DROME Acetyl-coenzyme A synthetase
SFRURICE0000002102	Su(P)-PA
SFRURICE0000002114	cxe022a
SFRURICE0000002115	CG2930-PD
SFRURICE0000002172	unknown
SFRURICE0000002176	TCF25_MOUSE Transcription factor 25
SFRURICE0000002182	S15A1_RAT Solute carrier family 15 member 1
SFRURICE0000002185	unknown
SFRURICE0000002211	DYC1_CAEEL Dystrophin-like protein 1
SFRURICE0000002214	unknown
SFRURICE0000002240	tpi
SFRURICE0000002241	CG9302-PA
SFRURICE0000002252	CG8314-PA
SFRURICE0000002253	PGRP2_HOLDI Peptidoglycan-recognition protein 2
SFRURICE0000002255	unknown
SFRURICE0000002256	CG8320-PA
SFRURICE0000002271	POLO_DROME Serine/threonine-protein kinase polo
SFRURICE0000002272	SPRR3_RABIT Small proline-rich protein 3
SFRURICE0000002297	unknown
SFRURICE0000002316	CG13917-PD
SFRURICE0000002330	unknown
SFRURICE0000002334	unknown
SFRURICE0000002335	KCNT1_CHICK Potassium channel subfamily T member 1
SFRURICE0000002339	unknown
SFRURICE0000002340	Iris-PA
SFRURICE0000002358	unknown
SFRURICE0000002625	CG31100-PB
SFRURICE0000003127	SP047
SFRURICE0000003128	GAGJ_DROME Nucleic-acid-binding protein from mobile element jockey
SFRURICE0000003833	unknown
SFRURICE0000003911	CG31183-PB

SFRURICE0000004270	RM11_DROME 39S ribosomal protein L11, mitochondrial
SFRURICE0000004307	CG14253-PF
SFRURICE0000004448	p77
SFRURICE0000004470	unknown
SFRURICE0000004500	unknown
SFRURICE0000004929	unknown
SFRURICE0000005012	unknown
SFRURICE0000005014	Jon65Aiii-PA
SFRURICE0000005041	unknown
SFRURICE0000005042	unknown
SFRURICE0000005043	unknown
SFRURICE0000005044	unknown
SFRURICE0000005045	unknown
SFRURICE0000005046	HARB1_RAT Putative nuclease HARB1
SFRURICE0000005109	unknown
SFRURICE0000005147	unknown
SFRURICE0000005148	ZBED1_HUMAN Zinc finger BED domain-containing protein 1
SFRURICE0000005157	unknown
SFRURICE0000005158	unknown
SFRURICE0000005159	unknown
SFRURICE0000005166	unknown
SFRURICE0000005167	unknown
SFRURICE0000005168	unknown
SFRURICE0000005188	I(1)G0289-PC
SFRURICE0000005225	Jarid2-PD
SFRURICE0000005345	unknown
SFRURICE0000005778	Y3556_DROME Uncharacterized protein CG3556
SFRURICE0000005838	unknown
SFRURICE0000005939	GNAL_DROME Guanine nucleotide-binding protein subunit alpha homolog
SFRURICE0000006018	mRpL50-PA
SFRURICE0000006095	bma-PE
SFRURICE0000006119	PGRP
SFRURICE0000006128	NR2E1_XENLA Nuclear receptor subfamily 2 group E member 1
SFRURICE0000006151	CG42258-PC
SFRURICE0000006157	TSN
SFRURICE0000006184	Trh-PA
SFRURICE0000006204	unknown
SFRURICE0000006205	unknown
SFRURICE0000006245	CG4928-PC
SFRURICE0000006256	HARB1_BOVIN Putative nuclease HARB1
SFRURICE0000006261	Ddr-PG
SFRURICE0000006262	unknown
SFRURICE0000006263	Rrp6-PC
SFRURICE0000006322	HTP
SFRURICE0000007150	unknown
SFRURICE0000007496	sl-PA

SFRURICE0000007497	babo-PA
SFRURICE0000007498	unknown
SFRURICE0000007533	Fhos-PG
SFRURICE0000007545	unknown
SFRURICE0000007555	unknown
SFRURICE0000008580	unknown
SFRURICE0000008581	unknown
SFRURICE0000008589	unknown
SFRURICE0000008670	unknown
SFRURICE0000008679	CG9503-PA
SFRURICE0000008873	unknown
SFRURICE0000009171	unknown
SFRURICE0000009172	CATA_LACSK Catalase
SFRURICE0000009286	HTP
SFRURICE0000009527	Lmx
SFRURICE0000009545	unknown
SFRURICE0000009564	CG13293-PC
SFRURICE0000009589	unknown
SFRURICE0000009602	unknown
SFRURICE0000009992	unknown
SFRURICE0000010006	p77
SFRURICE0000010007	Esp-PB
SFRURICE0000010268	ATTY_RAT Tyrosine aminotransferase
SFRURICE0000010479	unknown
SFRURICE0000010572	DNJ60_DROME DnaJ-like protein 60
SFRURICE0000010669	unknown
SFRURICE0000010670	unknown
SFRURICE0000010671	unknown
SFRURICE0000010672	unknown
SFRURICE0000011119	CG30100-PB
SFRURICE0000011120	LRC71_HUMAN Leucine-rich repeat-containing protein 71
SFRURICE0000011121	unknown
SFRURICE0000011470	CYP4L9
SFRURICE0000011557	unknown
SFRURICE0000011575	unknown
SFRURICE0000012296	unknown
SFRURICE0000012297	unknown
SFRURICE0000012298	unknown
SFRURICE0000013185	unknown
SFRURICE0000013186	Gyf-PF
SFRURICE0000013195	unknown
SFRURICE0000013242	CYC
SFRURICE0000013269	COPB2_DROME Coatomer subunit beta'
SFRURICE0000013272	NPRL2_DROME GATOR complex protein NPRL2
SFRURICE0000013514	unknown
SFRURICE0000013769	TIGD4_MOUSE Tigger transposable element-derived protein 4

SFRURICE0000014448	CG31974-PC
SFRURICE0000014634	PGBD4_HUMAN PiggyBac transposable element-derived protein 4
SFRURICE0000015028	unknown
SFRURICE0000015077	unknown
SFRURICE0000015341	POL4_DROME Retrovirus-related Pol polyprotein from transposon 412
SFRURICE0000015610	unknown
SFRURICE0000015795	unknown
SFRURICE0000015895	ARP6_CHICK Actin-related protein 6
SFRURICE0000016153	unknown
SFRURICE0000016269	CG2970-PB
SFRURICE0000016270	CC112_MACFA Coiled-coil domain-containing protein 112
SFRURICE0000016671	ATC1_ANOGA Calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum type
SFRURICE0000016672	unknown
SFRURICE0000016673	TANT_DROME Protein tantalus
SFRURICE0000016674	NEUFC_DROPS Neuferricin homolog
SFRURICE0000016675	acyl
SFRURICE0000016676	GGNB2_DROME Gametogenetin-binding protein 2-like
SFRURICE0000016681	AL7A1_CAEEL Putative aldehyde dehydrogenase family 7 member A1 homolog
SFRURICE0000016683	LR74A_HUMAN Leucine-rich repeat-containing protein 74A
SFRURICE0000016967	unknown
SFRURICE0000017102	unknown
SFRURICE0000017103	unknown
SFRURICE0000017177	HTP
SFRURICE0000017334	CG44247-PB
SFRURICE0000017335	DopEcR-PC
SFRURICE0000017346	DYH6_HUMAN Dynein heavy chain 6, axonemal
SFRURICE0000017348	unknown
SFRURICE0000017354	unknown
SFRURICE0000017373	unknown
SFRURICE0000017374	unknown
SFRURICE0000017385	parvin-PB
SFRURICE0000017391	CG12084-PB
SFRURICE0000017392	unknown
SFRURICE0000017393	unknown
SFRURICE0000017394	unknown
SFRURICE0000017396	GR10
SFRURICE0000017400	BAG6A_XENLA Large proline-rich protein bag6-A
SFRURICE0000017583	dpr20-PA
SFRURICE0000017610	unknown
SFRURICE0000018020	unknown
SFRURICE0000018050	unknown
SFRURICE0000018054	unknown
SFRURICE0000018081	unknown
SFRURICE0000018082	unknown
SFRURICE0000018517	GR146
SFRURICE0000018518	GR140

SFRURICE0000018584	mgI-PE
SFRURICE0000018712	PGS2_CHICK Decorin
SFRURICE0000019020	OSGP2_RAT Probable tRNA N6-adenosine threonylcarbamoyltransferase, mitochondrial
SFRURICE0000019161	unknown
SFRURICE0000019168	futsch
SFRURICE0000019239	unknown
SFRURICE0000019320	DSCAM
SFRURICE0000019609	unknown
SFRURICE0000019614	b6-PA
SFRURICE0000019708	unknown
SFRURICE0000019955	ENV_NPVLD Envelope fusion protein
SFRURICE0000020106	unknown
SFRURICE0000020108	unknown
SFRURICE0000020186	unknown
SFRURICE0000020187	RTJK_DROME RNA-directed DNA polymerase from mobile element jockey
SFRURICE0000020194	unknown
SFRURICE0000020195	Y2R2_DROME Putative 115 kDa protein in type-1 retrotransposable element R1DM
SFRURICE0000020196	unknown
SFRURICE0000020197	unknown
SFRURICE0000020217	unknown
SFRURICE0000020218	unknown
SFRURICE0000020266	CG8177-PA
SFRURICE0000020376	unknown
SFRURICE0000020399	spn
SFRURICE0000020400	unknown
SFRURICE0000020401	HTP
SFRURICE0000020488	CG18547-PA
SFRURICE0000020871	HTP
SFRURICE0000020872	RTXE_DROME Probable RNA-directed DNA polymerase from transposon X-element
SFRURICE0000020907	unknown
SFRURICE0000021181	ENV_NPVLD Envelope fusion protein
SFRURICE0000021402	unknown
SFRURICE0000021473	unknown
SFRURICE0000021518	ppk13-PA
SFRURICE0000021798	unknown
SFRURICE0000021799	unknown
SFRURICE0000021942	unknown
SFRURICE0000021943	CG10904-PB
SFRURICE0000022586	CG6178-PA
SFRURICE0000022670	Sbf-PA
SFRURICE0000022671	unknown
SFRURICE0000022713	unknown
SFRURICE0000022714	CYP4AU2
SFRURICE0000022911	ADCY8_HUMAN Adenylate cyclase type 8
SFRURICE0000023067	unknown
SFRURICE0000023093	p77

SFRURICE0000023266	unknown
SFRURICE0000023267	CG5953-PD
SFRURICE0000023268	unknown
SFRURICE0000023269	unknown
SFRURICE0000023270	ARC1_DROME Activity-regulated cytoskeleton associated protein 1
SFRURICE0000023314	unknown
SFRURICE0000023838	unknown
SFRURICE0000023969	unknown
SFRURICE0000023970	unknown
SFRURICE0000023994	unknown
SFRURICE0000023995	unknown
SFRURICE0000023996	unknown
SFRURICE0000023997	unknown
SFRURICE0000023999	unknown
SFRURICE0000024005	unknown
SFRURICE0000024766	Iris-PA
SFRURICE0000025293	GPRK1_DROME G protein-coupled receptor kinase 1
SFRURICE0000025295	PPIL6_HUMAN Peptidyl-prolyl cis-trans isomerase-like 6
SFRURICE0000025296	HEAT6_DANRE HEAT repeat-containing protein 6
SFRURICE0000025880	unknown
SFRURICE0000026019	unknown
SFRURICE0000026534	unknown
SFRURICE0000026715	RTJK_DROFU RNA-directed DNA polymerase from mobile element jockey
SFRURICE0000026944	unknown
SFRURICE0000027555	unknown
SFRURICE0000027556	unknown
SFRURICE0000027889	HTP
SFRURICE0000027962	unknown
SFRURICE0000027963	unknown
SFRURICE0000028011	PI4KIIIalpha-PC
SFRURICE0000028129	unknown
SFRURICE0000028494	cxe001k
SFRURICE0000029158	unknown
SFRURICE0000029226	CYP340AA2P
SFRURICE0000029347	HTP
SFRURICE0000029628	unknown
SFRURICE0000029671	CG9572-PB
SFRURICE0000029672	unknown
SFRURICE0000030208	unknown
SFRURICE0000030209	unknown
SFRURICE0000030239	PB1_MOUSE Protein polybromo-1
SFRURICE0000030298	unknown
SFRURICE0000030473	unknown
SFRURICE0000030476	unknown
SFRURICE0000030477	RTJK_DROME RNA-directed DNA polymerase from mobile element jockey
SFRURICE0000030502	CG9248-PC

SFRURICE0000030959	SYEP_DROME Bifunctional glutamate/proline--tRNA ligase
SFRURICE0000031106	unknown
SFRURICE0000031293	Taz-PD
SFRURICE0000031294	unknown
SFRURICE0000031295	EDC3_DANRE Enhancer of mRNA-decapping protein 3
SFRURICE0000031296	ERMP1_MOUSE Endoplasmic reticulum metalloproteinase 1
SFRURICE0000031297	Bsg-PB
SFRURICE0000031299	Gbeta5-PA
SFRURICE0000031300	PCH2_DANRE Pachytene checkpoint protein 2 homolog
SFRURICE0000031302	CG6479-PA
SFRURICE0000031303	unknown
SFRURICE0000031391	Dop1R1-PC
SFRURICE0000031759	unknown
SFRURICE0000031760	unknown
SFRURICE0000032013	unknown
SFRURICE0000032055	unknown
SFRURICE0000032091	unknown
SFRURICE0000032092	unknown
SFRURICE0000032367	unknown
SFRURICE0000032368	unknown
SFRURICE0000032415	unknown
SFRURICE0000032416	unknown
SFRURICE0000032777	POLR1_ARATH Retrovirus-related Pol polyprotein from transposon RE1

515 Table S4. The number of genes within outliers of genetic differentiation that are potentially
516 associated with interactions with host-plants. refC and refR are the results from the mapping against
517 sfC and sfR reference genome assemblies, respectively.

518	Functions	refC	refR
	Chemosensory	3	3
519	Immunity	1	0
	Oxidative stress	10	9
	Development	4	4
520	P450	1	3
	Circadian Signaling	0	1
521	Esterase	0	2
	Serine Protease	0	1

522

523

524 Figure S1. Distribution of F_{st} calculated from each of eight groups from which hapflk scores were
525 calculated. These groups were generated by a random grouping of scaffolds into eights.

526

527 Figure S2. The read depth (upper) and the alignment rate (lower) of the mappings reads against the
528 reference genomes of sfC and sfR. As 'R1' individual has a particularly lower read depth, we
529 excluded this individual in this paper. R1 has the lowest alignment rate, as well.

530

531 Figure S3. The relationship of exon density with π (A) and F_{st} (B) from the scaffolds without
532 outliers.

533

534 Figure S4. Log-transformed p-values and hapflk scores of outliers of genetic differentiation with
535 higher d_{XY} than genomic average along the sequences in the reference genome of the sfC. The red
536 bars indicate the borders of the outliers and dotted lines show p-value equals to 0.001.

537

538 Figure S5. The distribution of π of corn strain (sfC) and rice strain (sfR) from the outliers and 100kb
539 windows from the whole genome sequences.

540

541 Figure S6. The distribution of d_{XY} from outliers (red) and 10kb windows from whole genome
542 sequences (blue), based on the mapping against sfR reference assembly.

543

544 Figure S7. The analysis of ancestry coefficient. a). The relationship between K and cross entropy.
545 b). The ancestry coefficient when $K = 2$.

546

547 Figure S8. Histogram of F_{st} calculated in 10kb windows based on the mapping against sfR
548 reference assembly. The red vertical bar indicates $F_{st} = 0$.

549

550 Figure S9. A. F_{st} calculated according to the distance from the nearest outlier of genetic
551 differentiation based on the mapping against sfR reference assembly. The solid red curve is fitted
552 smooth-spline with $df = 5$, and the red dotted curves are 95% confidence intervals with 1,000
553 bootstrapping. The vertical dotted line indicates the distance equal to 200kb. B. The barplot shows
554 F_{st} from scaffolds with outliers and that from without outliers.

555

556 Figure S10. The relationship of exon density with π (A) and F_{st} (B), based on the mapping against
557 sfR reference assembly.

558

559 Figure S11. The relationship of exon density with π (A) and F_{st} (B) from the scaffolds without
560 outliers, based on the mapping against sfR reference assembly.

561

562 Figure S12. Log-transformed p-values and hapflk scores of outliers of genetic differentiation with
563 higher d_{XY} than genomic average along the sequences in the reference genome of the sfR. The red
564 bars indicate the borders of the outliers, and dotted lines show p-value equals to 0.001.

565

566 Figure S13. The distribution of π of corn strain (sfC) and rice strain (sfR) from outliers and 100kb
567 windows from the whole genome sequences, based on the mapping against sfR reference assembly.

568

569 Figure S14. The distribution of d_{XY} from outliers (red) and 10kb windows from whole genome
570 sequences (blue), based on the mapping against sfR reference assembly.

571

572 Figure S15. Principal component analysis from the outliers (A) and from the whole genome (B),
573 based on the mapping against sfR reference assembly. The red and blue dots represent sfC and sfR,
574 respectively.

575

576 Figure S16. The analysis of ancestry coefficient, based on the mapping against sfR reference
577 assembly. A. The relationship between K and cross entropy. B. The ancestry coefficient when $K = 2$.

578

579 Figure S17. Principal component analysis from mitochondrial genomes.

580

581 Figure S18. Mitochondrial phylogenetic tree. The mitochondrial sequence of each individual was
582 inferred by mapping against mitochondrial genomes (NCBI accession number: KM362176) and by
583 variant and non-variant calling. Then, multiple sequence alignment was generated together with
584 *Spodoptera litura* (NCBI accession number: KF701043) using prank software. The phylogenetic
585 tree was reconstructed with 1,000 bootstrapping replications using FastME software

586

587 Figure S19. The analysis of ancestry coefficient from the mitochondrial genome. A. The
588 relationship between K and cross entropy. B. The ancestry coefficient when $K = 2$.

589

590 Figure S20. We performed a simple forward simulation using slim software with a wide range of
591 migration rate to test mitochondrial divergence time can explain the level of observed nuclear
592 genetic differentiation ($F_{st} = 0.0176$). The simulation was performed during $5 \times N_e$ generations in
593 100kb sequences. Assuming that the generation time for each generation is 0.1 years (lab condition)
594 and that N_e is 4 million, $5 \times N_e$ generation time corresponds to $5 \times 4 \times 10^6 \times 0.1 = 2 \times 10^6$ years,
595 which is the reported mitochondrial divergence time based on the molecular clock[43]. The
596 mutation rate, recombination rate, N_e of the ancestral population, N_e of two derived populations

597 after the split from the ancestral population, and the generation time after the split are 1.16×10^{-4} ,
598 1.188×10^{-3} , 200, 100, and 500, respectively. For each migration rate, 500 independent simulations
599 were performed, and the calculated F_{st} was averaged. The red horizontal bar indicates the genomic
600 average F_{st} , which is 0.0176. Please note that the used parameters were rescaled by 4000 folds from
601 2.9×10^{-9} and 2.97×10^{-8} , for mutation rate, and recombination rate, respectively. And the used
602 parameters were rescaled by 1/4000 folds from 8×10^6 , 4×10^6 , and 2×10^7 , for N_e of the ancestral
603 population, and N_e of two derived population after the split from the ancestral population, and the
604 number of generation after the split, respectively.

605

606 Figure S21. The read-depth of mappings against FR (NCBI ID: FR.X78688.1), which are expected
607 to be present only in the sfR.

608

609 Figure S22. A possible explanation for the discrepancy of identified strains between the
610 mitochondrial genome and nuclear TPI gene. Names of the leaves of the trees show the strains
611 identified from the mitochondrial genome. For example, sfC.1 and sfC.2 are the individuals
612 identified as sfC according to mitochondrial markers. As noted in the main text of the paper,
613 mitochondrial divergence time predates the averaged nuclear divergence time. The divergence time
614 of TPI gene postdates the averaged nuclear divergence time. The dashed horizontal bars indicate
615 average nuclear differentiation time. (upper) If a selective sweep on TPI gene occurs at the common
616 ancestor of sfR.2 and sfR.3, then sfC.1, sfC.2, and sfR.1 share the common genotypes. (lower) If
617 the selective sweep occurs at the common ancestor of sfC.1, sfC.2, and sfR.1, these three
618 individuals share the common genotypes. In these two cases, sfR.1 will be identified as sfC when
619 TPI gene is used as a marker.