The genomic bases of morphological divergence and reproductive isolation driven by ecological speciation in *Senecio* (Asteraceae)

Mark A. Chapman[*,§], Simon J. Hiscock[†] and Dmitry A. Filatov[*]

[*] Department of Plant Sciences, University of Oxford, South Parks Rd, Oxford, OX1 3RB, United Kingdom

[§] Centre for Biological Sciences, University of Southampton, Life Sciences Building 85, Highfield Campus, Southampton, SO17 1BJ, United Kingdom

[†] School of Biological Sciences, University of Bristol, Woodland Road, Bristol, BS8 1UG, United Kingdom

Short title: Genomics of ecological speciation

Corresponding author: Dmitry A. Filatov, Department of Plant Sciences, University of Oxford, South Parks Rd, Oxford, OX1 3RB, United Kingdom; E-mail: dmitry.filatov@plants.ox.ac.uk; Tel: +44 (0)1865 275051; Fax: +44 (0)1865 275074

ABSTRACT

Ecological speciation, driven by adaptation to contrasting environments, provides an attractive opportunity to study the formation of distinct species, and the role of selection and genomic divergence in this process. Here we focus on a particularly clear-cut case of ecological speciation to reveal the genomic bases of reproductive isolation and morphological differences between closely related *Senecio* species, whose recent divergence within the last ~200,000 years was likely driven by the uplift of Mt. Etna (Sicily). These species form a hybrid zone, yet remain morphologically and ecologically distinct, despite active gene exchange. Here we report a high-density genetic map of the *Senecio* genome and map hybrid breakdown to one large and several small quantitative trait loci (QTL). Loci under diversifying selection cluster in three 5 cM regions which are characterised by a significant increase in relative (F_{ST}), but not absolute (d_{XY}), interspecific differentiation. They also correspond to some of the regions of greatest marker density, possibly corresponding to 'cold-spots' of recombination, such as centromeres or chromosomal inversions. Morphological QTL for leaf and floral traits overlap these clusters. We also detected three genomic regions with significant transmission ratio distortion (TRD), possibly indicating accumulation of intrinsic genetic incompatibilities between these recently diverged species. One of the TRD regions overlapped with a cluster of high species differentiation, and another overlaps the large QTL for hybrid breakdown, indicating that divergence of these species may have occurred due to a complex interplay of ecological divergence and accumulation of intrinsic genetic incompatibilities.

Key words: Adaptation, Genomic Divergence, Hybrid Breakdown, *Senecio*, Speciation

INTRODUCTION

Interspecific hybridisation is prevalent in plants (Stebbins, 1959; Stace, 1975; Arnold, 1992; Rieseberg & Carney, 1998), and to a lesser, but not insignificant, degree in animals (Mallet, 2005), such that it appears that many species are able to co-exist with congeners, despite some level of hybridisation. The origin and continued existence of a species in the face of gene flow from others is particularly intriguing to students of evolutionary biology (Coyne & Orr, 2004; Lexer & Widmer, 2008; Wolf *et al.*, 2010; Feder *et al.*, 2012).

Natural hybrid zones usually comprise two species that produce hybrid offspring with some degree of fertility and fitness but do not become homogenised. Hybrid zones often span an ecological gradient, which has led to the suggestion that selection is dependent on the environment (Endler, 1977), and indeed hybrids might be more fit than parental types in an intermediate habitat (Moore, 1977). In the majority of hybrid zones, however, a narrow region of hybridisation and evidence for reduced fitness of hybrids suggests that these are 'tension zones'; maintained by a balance between gene flow into the zone and selection against hybrids (Barton & Hewitt, 1989). Although tension zones do not require any ecological gradient or barrier to be stable, they may coincide with an ecological gradient because the movement of the tension zone can be 'trapped' by an ecological gradient, resulting in a complex interplay of genetic (sub-)species incompatibilities and ecological adaptations to distinct environments in the same hybrid zone (Bierne *et al.*, 2011). Clearly, an understanding of the genomics of species that form hybrid zones will reveal insights into the evolutionary interplay between speciation, reproductive isolation and adaptation (Payseur, 2010).

For species that arise without allopatric separation it is predicted that the early events in speciation are restricted to small genomic regions (potentially even single genes) which diverge

due to selection (Wu, 2001; Nosil *et al.*, 2009). For example if speciation was driven by ecology, alleles at the 'speciation loci' will confer a selective advantage in the species' home environment (Nosil, 2012), and this could lead to the formation of an environmentally dependent hybrid zone. Under this model, genetic divergence accumulates in these speciation loci because of reduced effective gene flow, which also hampers gene flow at linked loci (a process often called 'divergence hitch-hiking', DH; Charlesworth *et al.*, 1997; Via, 2012). The speciation loci are therefore predicted to be found in a region (or regions) of the genome characterised by an increased divergence between species ('speciation island'; Turner *et al.*, 2005), but theory predicts that such 'islands' should not be much larger than ~1cM (Feder & Nosil, 2010).

Investigations of hybridising species often demonstrate that interspecific divergence is not homogeneous (Peichel *et al.*, 2001; Yatabe *et al.*, 2007; Nosil *et al.*, 2008; Kane *et al.*, 2009; Kulathinal *et al.*, 2009; Renaut *et al.*, 2012; Roesti *et al.*, 2012; Stölting *et al.*, 2013). If subsequent genetic mapping reveals that these loci cluster, then it is sometimes inferred that these islands are demarcating the region(s) of the genome that harbour the genes that were initially involved in the speciation process (Turner *et al.*, 2005; Nadeau *et al.*, 2012). If the metric used to calculate divergence is a relative measure (e.g. F_{ST}) then some regions of the genome containing clusters of 'high divergence' loci may instead be regions of low intra-species diversity and not high absolute divergence (Noor & Bennett, 2009; Cruickshank & Hahn, 2014).

To further our understanding of genomic divergence during speciation driven by adaptation to contrasting environments, we took advantage of a particularly attractive study system. The two plant species concerned are *Senecio aethnensis* Jan ex DC., a species endemic to Mt. Etna and only found at altitudes above ca. 2,000 m, and *S. chrysanthemifolius* Poir. a species only found at lower altitudes, less than ca. 1,000 m. The species are morphologically distinct in terms of their leaf dissection and the size of their inflorescences, with *S. aethnensis* having non-dissected leaves and

large inflorescences, while leaves in *S. chrysanthemifolius* are highly dissected and inflorescences are smaller (Brennan et al. 2009). The difference in these traits is likely to reflect adaptations to high and low altitudes.

Both species are strongly self-incompatible with no pre-zygotic barriers to hybridization (Chapman *et al.*, 2005) bar the presence of shared self-incompatibility alleles between individuals (Brennan *et al.*, 2013). Thus, at intermediate altitudes, extensive hybrid swarms of phenotypically and genotypically variable individuals are found, indicative of on-going gene flow (James & Abbott, 2005; Brennan *et al.*, 2009). Although hybrids are locally common in the Mt. Etna hybrid zone, cline analysis reveals both intrinsic and extrinsic selection are acting to maintain the structure of the hybrid zone (Brennan *et al.*, 2009). This study demonstrated that there is selection against hybrid genotypes within the hybrid zone as well as divergent selection acting on aspects of the species' leaf and floral morphology (Brennan *et al.*, 2009). A recently published linkage map from a cross between these species identified significant transmission ratio distortion (TRD) in 27% of 127 mapped AFLP and SSR markers (Brennan et al. 2014).

Senecio aethnensis and *S. chrysanthemifolius* are thought to have diverged very recently (<150,000 years before present) and this divergence took place in the face of gene flow (Chapman et al. 2013; Osborne et al 2013). Based on previous transcriptome sequencing of multiple individuals of both species, divergence between this species pair is heterogeneous across the genome and only a small portion (< 3%) of loci exhibited potential hallmarks of divergent selection (Chapman *et al.*, 2013). To expand upon these findings we were interested in testing if the high divergence 'outliers' form discrete islands of increased differentiation, and if there is overlap between quantitative trait loci (QTL) for morphological species differences and any outlier loci. In our mapping population we also mapped hybrid breakdown, a situation in which a subset of the individuals have reduced fitness due to incompatibilities between the parental genomes (Sweigart

& Willis, 2012) which has been reported for other pairs or taxa that have diverged due to ecological factors (e.g. Fuller, 2008; Renaut & Bernatchez, 2011). Our results reveal complex interactions between morphological, genetic and transcriptomic changes associated with species divergence. Taken together, the results are consistent with the notion that divergence of the two *Senecio* species was driven by adaptation to contrasting environments of high and low altitude during the uplift of Mount Etna.

MATERIALS AND METHODS

For this study an F2 population between *S. aethnensis* and *S. chrysanthemifolius* was generated and a subset of the progeny was randomly selected for transcriptome sequencing. From this sequence data, segregating alleles (single nucleotide polymorphisms, SNPs) were identified and used in the construction of a genetic map. In addition, gene expression was quantified for each individual and compared between the healthy and unhealthy F2s to determine which genes and pathways were significantly affected due to hybrid breakdown. Further information is supplied in Appendix S1.

F2 population generation and sequencing: An F2 population (*S. aethnensis* x *S. chrysanthemifolius*) was generated and grown alongside several individuals of the parental species (see Appendix S1). In total 147 F2 plants were grown, 64 of which were selected for RNA-seq analysis. F2 hybrid breakdown was observed in this cross, with 69 plants developing signs of necrosis or stunted growth (collectively termed 'necrotic'; Fig. 1 A-D) and 78 healthy plants (Fig. 1 E-F). The 64 sequenced plants comprised 48 healthy and 16 necrotic F2s and apart from this designation were selected randomly. RNA was extracted from leaf tissue of 3 week old seedlings using an RNeasy kit (Qiagen, Manchester, UK) with additional DNase treatment (Qiagen). From each an individual RNA-seq library was constructed, which were then run (as a multiplex) across

two lanes of a HiSeq2000 at the Wellcome Trust Centre for Human Genetics. Reads have been submitted to the Short Read Archive (www.ncbi.nlm.nih.gov/sra) under BioProject PRJNA270017.

Sequence Data: Reads were trimmed and filtered and then mapped back onto the reference transcriptome constructed in Chapman *et al.* (2013) using CLC Genomics Workbench (CLC-GW; ver. 5; http://www.clcbio.com). Mapping was carried out with the following settings: mismatch cost 3, insertion cost 3, deletion cost 3, minimum identity 0.95. Reads from the two parental individuals were also mapped back. Mappings were exported as bam files, parsed through *SAMtools* (Li *et al.*, 2009) with settings: minimum base quality 20, minimum mapping quality 5, minimum read depth 5, to convert to fasta files, and aligned in Proseq (Filatov, 2009).

Expression analysis was carried out in CLC-GW with the following settings: maximum number of mismatches 2, minimum length fraction 0.8, minimum similarity fraction 0.9, maximum number of hits for a read 10. The number of reads mapped per transcript was converted to RPKM (Mortazavi *et al.*, 2008; Reads per kilobase per million mapped reads; i.e. standardised to the length of the transcript and depth of sequencing; Appendix S1). Loci with significant differential expression were identified by carrying out locus-by-locus *t*-tests and applying a 5% FDR.

Linkage Mapping: Because we were interested in the genomic distribution of outlier loci relative to non-outliers (see also 'QTL mapping', below) we did not seek to map as many loci as possible from the cross, and instead focussed on the loci with the highest quality reliable genotypic information. We used the following criteria: (1) Parental individuals must be homozygous for different alleles at all SNPs within a locus, (2) At least five SNPs must differentiate the parental alleles, and (3) All 64 F2 individuals must have a genotype (see Appendix S1). Loci with > 15% missing data were removed leaving 966 loci, 112 of which were previously identified (Chapman *et al.*, 2013) as high or low differentiation outlier loci. Mapping was carried out in MapDisto (Lorieux,

2012), initially grouping the markers using a LOD (logarithm of odds) of 10 and an RF (recombination frequency) of 0.2, unlinked loci were removed, and map orders were then compared using 'order' and 'ripple' using the *seriation II* algorithm and *Sum of Adjacent Recombination Frequencies* criteria. The linkage map has been uploaded to Dryad (http://datadryad.org) under doi:10.5061/dryad.gb6g6.

Although a small subset of markers exhibited TRD (i.e. significant χ^2 test based on expected 1:2:1 genotype ratios) these markers were not removed (TRD has been reported in another cross between these species; Brennan *et al*. 2014) and instead we checked that these deviations were not a result of marker scoring errors. DNAs extracted from the two parental individuals and/or a subset of 16 F2 individuals were PCR amplified and Sanger-sequenced or genotyped at the particular loci to ensure that the observed segregation distortion was not an artefact of non-homozygous parental genotypes or biased allelic expression in the F2 (Appendix S1).

Genome-wide analysis of polymorphism and divergence: To investigate polymorphism and divergence along LGs we plotted genetic differentiation (F_{ST} and d_{XY}) and nucleotide diversity (per species per nucleotide π) from the 20 transcriptome data in Chapman *et al.* (2013), as well as calculated a moving average for these statistics based on a bin size of 10 cM and a step of 5 cM.

QTL mapping: All healthy F2 plants and several individuals of the parental species were phenotyped for a number of characters that differentiate the parental species (Table 1). Traits were measured on the day of anthesis of the first capitulum. Capitulum and floret traits were measured on the first capitulum. Leaf traits were measured from scanned images of the mid-stem leaf using ImageJ (Rasband, 1997-2013). Traits that deviated from normality in the F2 population (indicated in Table 1) were square-root transformed prior to QTL analysis. QTL analysis was carried out in Windows QTL Cartographer (Wang *et al.*, 2005) after reducing the genotypes to only loci

that mapped to unique positions (i.e. where more than one locus mapped to the same position redundant loci were removed), and only including morphological phenotypes for the 48 healthy plants. Morphological traits were mapped using Composite Interval Mapping (CIM) and healthy vs. necrotic was mapped as a binary trait using Interval Mapping (IM), with significance thresholds for all traits determined with 1000 permutations. QTL positions and their 1- and 2-LOD intervals were plotted onto the LGs using MapChart (Voorips, 2002).

Our previous study (Chapman *et al.*, 2013) identified 90 loci that were significantly differentially expressed (DE) between the parental species. Because of DE, these could not be mapped in the F2 (too much missing data), therefore the expression of these loci (calculated above) in the 48 WT F2 was mapped using CIM to identify eQTL (necrotic plants were not included because gene expression was greatly modified in these plants; see below). Again, significant thresholds were determined with 1000 permutations and were mapped onto LGs using MapChart.

RESULTS

Mapping population and linkage map: While the F1 individuals of the cross between *S. aethnensis* and *S. chrysanthemifolius* were all of similar phenotype, intermediate between the parental species (data not shown), the F2 population included a wider variation in phenotypic traits, with some individuals resembling one or the other parental species (e.g. Fig. S1). Curiously, 69 out of 147 F2 plants showed a range of stunted and necrotic phenotypes that either died or never reached maturity (Fig. 1). The necrotic plants exhibited phenotypes uncharacteristic of the parental species including mis-shaped or curled leaves, thick 'brittle' leaves, and necrosis, which may represent a result of hybrid incompatibility between the parental genomes. The genetic bases of this hybrid breakdown are analysed below.

To build a high density map with genic markers we conducted transcriptome sequencing (RNA-seq) for parents and 64 F2 individuals, including 48 progeny with normal phenotype and 16 progeny with necrotic phenotypes. The number of reads per individual was at least 5 million 100b paired-end reads (Table S1). Following single nucleotide polymorphism (SNP) calling and stringent filtering (see methods) we arrived at 966 high quality loci segregating in the mapping population. Genetic mapping resolved eleven linkage groups (LGs) and eight unlinked loci. One of the eleven groups comprised only four loci and was loosely linked to one of the larger groups (using relaxed 'group' settings: LOD 5.0, RF 0.2) and so was included in this LG. The unlinked loci were dropped and the final map comprised 958 loci across 10 LGs, corresponding to the haploid chromosome number of these species. The LGs varied in length from 44.7 to 112.0 cM and included between 34 and 164 loci (Fig. S2). Total map length was 794.6 cM, corresponding to a recombination rate of 0.306 cM/Mb based on the estimated genome size of *S. chrysanthemifolius* (2.6×10^9 bp; Bennett & Smith, 1976).

Marker clustering and transmission ratio distortion: Loci were not evenly distributed along the LGs, with clusters of co-mapping genes clearly visible on several linkage groups; in the most extreme cases, fifty loci or more co-mapped to the exact same positions on LGs 1, 2, 3, 7 and 9 (Fig. 2; Fig. S2). Whilst transmission ratio distortion (TRD; real or erroneous; see below) could lead to the incorrect linkage of loci, four of these five regions of the genome were not characterised by significant TRD, with no evidence for deviation from a 1:2:1 genotype ratio based on χ^2 tests. It is possible that these regions could be marking reduced recombination, leading to a genetic map for a relatively large genomic region collapsing into a single point on the map (see Discussion).

Markers in the fifth largest cluster (on LG3) and in two other regions of the genome not exhibiting marker clustering (on LGs 2 and 4; asterisks in Fig. S2) were, however, characterised by TRD (χ^2 test, $P < 0.001$; Fig. 3). In fact all 134 markers exhibiting TRD at $P < 0.001$ (14.0% of all markers)

reside in these three regions of the linkage map (Fig. S2). The TRD regions on LG2 and 3 are characterised by an excess of *S. chrysanthemifolius*-inherited alleles and a deficit of *S. aethnensis*-inherited alleles, respectively (Fig. 3). The third TRD region (on LG4) showed extreme bias in segregation, with no F2 individuals homozygous for *S. chrysanthemifolius* allele at 21 loci. If we relax the level of significance for TRD to $P < 0.05$ we find 64 additional TRD loci. Most of these (38 of the 64) are in the vicinity of the TRD loci found at $P < 0.001$, with the remainder found in other regions of the linkage map.

Because it is possible that biased allelic expression in heterozygotes could lead to genotyping errors from RNA-seq data resembling this pattern, we genotyped 16 markers from these three TRD regions using DNAs from a subset of the F2 individuals that were used in the RNA-seq genotyping (Table S2). This confirmed the RNA-seq based genotypes for all tested markers; hence these regions of TRD are 'real' and not an artefact of scoring errors or biased allelic expression, for at least the majority of markers. We also confirmed the absence of *S. chrysanthemifolius* homozygous genotypes for the region of LG4; for ten of twelve markers the DNA-based genotypes of the parental and F1 RNA-seq genotypes were identical confirming that the observed lack of homozygotes for the *S. chrysanthemifolius* allele was not an artefact of homozygosity of an F1 parent for *S. aethnensis* alleles at these loci.

Genomic distribution of polymorphism and differentiation: Based on the data obtained from the transcriptome sequencing of ten individuals of each species (Chapman *et al.*, 2013) we compared overall patterns of polymorphism and divergence among LGs. Per locus F_{ST} differed significantly among LGs (one-way ANOVA; $F = 12.66$, 9 d.f., $P = 4.756 \times 10^{-19}$; Fig. S3a), but d_{XY} did not (one-way ANOVA; $F = 0.8701$, 9 d.f., $P = 0.5516$; Fig. S3b). Average heterozygosity (π) differed significantly among LGs and species (two-way ANOVA, LG: $F = 5.57$, 9 d.f., $P = 1.32 \times 10^{-7}$; species: $F = 137.3$, 1 d.f., $P = 1.16 \times 10^{-30}$; interaction: NS; Fig. S3c), consistent with our previous finding that *S.*

249 *chrysanthemifolius* contained lower levels of genetic diversity than *S. aethnensis* (Chapman *et al.*,
250 2013).

251 Population differentiation (F_{ST}, d_{XY}) and genetic diversity (π) were highly heterogeneous across the
252 linkage groups. This variation was clearly structured into regions of high and low differentiation,
253 and/or polymorphism (Fig. 4; Fig. S4). If we classify a 10 cM bin with an F_{ST}, d_{XY} or π within the
254 highest (F_{ST}, d_{XY}) or lowest (π) 10% of the distribution as a high differentiation or low
255 polymorphism region, respectively (dashed horizontal lines in Fig. 4), nine regions (where a region
256 comprises one or more contiguous bins) of the genome (comprising 17 bins) showed high
257 differentiation (F_{ST} and/or d_{XY}). Five and seven regions for *S. aethnensis* and *S. chrysanthemifolius*,
258 respectively (nine bins each) showed low polymorphism.

259 **Clustering of high differentiation outlier markers:** Genetic map allows us to test whether the loci
260 highly differentiated between the species cluster in the genome, forming 'speciation islands'
261 described in other species (Turner *et al.*, 2005; Hohenlohe *et al.*, 2010; Scascitelli *et al.*, 2010;
262 Nadeau *et al.*, 2012; Renaut *et al.*, 2012). This linkage map included 112 outliers identified with
263 the Bayescan program (Foll & Gaggiotti, 2008) previously (Chapman *et al.*, 2013). Positive outliers
264 (n = 81) are identified on the basis of greater than expected divergence between *S. aethnensis* and
265 *S. chrysanthemifolius* (i.e. few shared alleles). Negative outliers (n = 31), on the other hand, have
266 less divergence (more shared alleles). Clusters of positive outliers were visible in the genetic map
267 (red bars in Fig. 4), indeed more than half of the outlier markers (48/81) were found in just three
268 of the 5 cM bins. However, correcting for the total number of loci in each bin, the evidence for
269 clustering was equivocal. Using a 5 cM bin there was some evidence for a significant difference
270 between the observed and expected number of outliers, given the total number of genes in the
271 bin (P = 0.045). With a 10 cM bin size, however, the distribution of outliers was not significantly
272 different to that expected (P = 0.168).

These three outlier clusters, which might have initially passed as speciation islands, are therefore more likely marking regions of high total marker density. The marker clusters (both outliers and non-outliers) (Fig. 4) could represent gene-rich regions of the genome, but more likely they correspond to regions of the genome exhibiting reduced recombination (i.e. higher gene density per cM). To investigate this further we assessed intraspecific polymorphism and interspecific divergence of loci in these three genomic regions containing clusters of markers (Fig. 5). Loci in these three regions exhibited reduced polymorphism (π) relative to the remainder of the genome (*S. aethnensis* 0.0033 ± 0.0031 [SD] within clusters vs. 0.0044 ± 0.0023 outside the clusters, Mann-Whitney U-test $P = 6.08 \times 10^{-9}$; *S. chrysanthemifolius* 0.0021 ± 0.0021 within clusters vs. 0.0032 ± 0.0022 outside the clusters, Mann-Whitney U-test $P = 5.39 \times 10^{-10}$; Fig. 5) a hallmark of reduced recombination (Begun & Aquadro, 1992).

Despite F_{ST} in these three regions being higher than the rest of the genome (0.521 ± 0.267 vs. 0.315 ± 0.208, Mann-Whitney U-test $P = 2.90 \times 10^{-18}$; Fig. 5), absolute divergence, d_{XY}, was not (0.0052 ± 0.0027 vs. 0.0048 ± 0.0023, Mann-Whitney U-test $P = 0.147$; Fig. 5). This strongly suggests that because F_{ST} is a relative measure of differentiation, contrasting inter- and intra-population diversity, the low diversity (π) in these regions is elevating F_{ST} even though absolute divergence is no higher than the rest of the genome.

Looking throughout the genome as a whole, F_{ST} is strongly negatively correlated with distance from the nearest outlier locus (F_{ST}: Spearman's rank correlation coefficient (r_s) = -0.352, $P = 1.321 \times 10^{-24}$), whereas d_{XY} is only weakly correlated (r_s = -0.094, P = 0.0078; Fig. S5a,b; LG8 was excluded because no positive outliers reside on this LG). This negative correlation between distance from outliers and F_{ST} is mirrored by a positive correlation between distance and genetic diversity (π_{aet}: r_s = 0.151, $P = 8.399 \times 10^{-6}$; π_{chr}: r_s = 0.123, P = 0.0003; Fig. S5c).

296 **QTL analysis of interspecific differences:** Individuals of the parental species were significantly
297 different in ten of the 13 studied morphological traits (Table 1). The analysis resolved QTL for all
298 but one trait (D2F). Morphological QTL were found on only four of the 10 LGs, with one to five QTL
299 per trait (Fig. 6). Effect sizes of the QTL were typically moderate (PVE 20-35%; Table 2).
300 Considerable overlap/clustering of QTL was uncovered; however in some cases this is likely due to
301 the traits being directly related (e.g. overlap of QTL for leaf perimeter and leaf dissection on LGs 5
302 and 6). In other cases QTL may co-localise, e.g. floral and leaf traits on LGs 1, 6 and 7, as they
303 represent specific aspects of 'plant size', therefore plants with large leaves also have large
304 inflorescences, for example. Positive outliers from Chapman *et al.* (2013) were found to lie under
305 morphological QTL (based on 1-LOD QTL interval) more often than expected by chance (χ^2 test; P =
306 3.22×10^{-7}; Fig. 6).

307 Ninety loci were previously shown to be differentially expressed (DE) between the two species
308 (Chapman *et al.*, 2013). For 67 we were able to map expression QTL (eQTL) affecting differential
309 expression (average 1.82 eQTL per DE locus; total 122 eQTL; Fig. S6). These eQTL overlapped the
310 morphological and necrotic QTL (Fig. S6) and the positive outliers more than would be expected by
311 chance (Fisher's exact test P = 0.0045, P = 0.0043 and P = 0.0003, respectively).

312 **QTL analysis of hybrid breakdown:** The necrotic plants (Fig. 1) did not reach maturity and may
313 represent the consequences of hybrid breakdown. Genetic mapping resolved one large QTL for
314 hybrid breakdown, explaining 67.5% of the phenotypic variance, plus six smaller QTL (Fig. 6; Table
315 2). Minor necrosis QTL on LGs 1 and 7 were found to co-localise with other morphological QTL
316 (Fig. 6).

317 The marker loci closest to the peak of the LG4 major QTL were homozygous for the allele inherited
318 from the *S. aethnensis* parent in all necrotic F2 plants and were heterozygous in all normal F2
319 plants. In this 14 cM region, homozygous alleles inherited from *S. chrysanthemifolius* were never

found in the 64 F2. Loci under the minor QTL did not show such strict differentiation between the normal and necrotic plants, with the exception of loci under one of the LG7 QTL (30 markers on the full map, from cM 91 – 108) that were never homozygous for the *S. aethnensis* parent allele in the 16 necrotic plants.

Gene expression in the F2 population: Comparisons were made between RNAs extracted from the leaves of wild-type F2 and necrotic F2 seedlings for all loci in the reference transcriptome. Of 18,797 loci, 18,687 were expressed in at least one individual, and of these, 2,301 (12.3%) were differentially expressed (DE) based on *t*-tests and a 5% FDR. The majority of DE loci (1709/2301; 74.3%) were under-expressed in the necrotic F2 relative to the wild-type F2. Positive outlier loci (n = 199) and differentially expressed loci (i.e. *S. aethnensis* vs. *S. chrysanthemifolius*; n = 90) identified in Chapman *et al.* (2013) were no more likely to be differentially expressed in the F2 than non-outliers and non-differentially expressed loci (χ^2 test; $P = 0.882$ and $P = 0.821$, respectively). There was also no difference in interspecific divergence (F_{ST} between *S. aethnensis* and *S. chrysanthemifolius*) between the DE (wild-type F2 vs. necrotic F2) and non-DE loci (0.202 ± 0.191 and 0.195 ± 0.191, respectively; Mann-Whitney $U = 4.67 \times 10^6$, $z = -1.501$, $P = 0.134$) suggesting that genes mis-expressed in the necrotic hybrids are not those that are more (or less) divergent between the species.

Several gene ontology (GO) terms were significantly over-represented (FDR < 0.05) in the loci DE between wild-type and necrotic F2 plants, many relating to basic physiological and developmental processes. After reducing the GO terms to the most specific terms (i.e. removing terms which contained more than one over-represented sub-term) there was evidence for genes involved in translation (e.g. ribosomal structural constituents [$P = 4.2 \times 10^{-27}$] and ribosome biogenesis [$P = 5.7 \times 10^{-3}$]), RNA methylation ($P = 2.3 \times 10^{-7}$), and development of leaves, roots and floral organs, being over-represented (Table S3). After separation into the loci over-expressed (n = 592 loci) and

under-expressed (n = 1709 loci) in the necrotic plants different pathways/functions appear affected by up- and down-regulation. In the up-regulated loci (necrotic expression > healthy expression) there is significant (FDR < 0.05) over-representation of just a few GO terms, for example 'protein targeting to the vacuole' ($P = 4.0 \times 10^{-5}$) and 'involved in the peroxisome' ($P = 1.5 \times 10^{-4}$). In contrast, much more drastic changes are seen in the down-regulated loci (Table S3). Loci involved in the structural constituents of the ribosome were highly over-represented ($P = 2.9 \times 10^{-34}$), with > 6% of the DE loci and < 1% of the non-DE loci being categorised by this GO term. Other related terms were highly over-represented (e.g. 'RNA methylation', 'translation') suggesting massive disruption of translation in the necrotic plants. Similarly high over-representation was found for multiple GO categories involved in basic growth and development pathways, e.g. stomatal morphogenesis ($P = 5.6 \times 10^{-9}$), cell wall organisation ($P = 3.1 \times 10^{-7}$), leaf morphogenesis ($P = 3.4 \times 10^{-3}$), and floral organ development ($P = 8.3 \times 10^{-3}$). One hundred and thirty four of the 2,301 DE loci were mapped in the F2 and were not found to coincide with morphological or hybrid breakdown QTL any more than by chance (1-LOD QTL interval; χ^2 test; $P = 0.937$).

DISCUSSION

Combining genome scans for outlier loci, with linkage mapping and QTL analysis can be particularly informative with regards the genomics of adaptation and/or speciation (Scotti-Saintagne *et al.*, 2004; Scascitelli *et al.*, 2010; Michel *et al.*, 2011; Via *et al.*, 2012). The genetic and phenotypic divergence of high altitude *S. aethnensis* and low altitude *S. chrysanthemifolius* is a likely example of ecological speciation. The two species appear to have diverged with gene flow relatively recently (Chapman *et al.*, 2013; Muir *et al.*, 2013; Osborne *et al.*, 2013); they are adapted to contrasting environments of high and low altitudes, and demonstrate a number of morphologically divergent traits evolving under divergent selection (Brennan *et al.*, 2009). Contemporary gene flow is evidenced by an extensive hybrid zone with many late generation

hybrids (Brennan *et al.*, 2009), in line with the observation that there is almost no intrinsic barrier to heterospecific mating (Chapman *et al.*, 2005) which is promoted by the presence of self-incompatibility (Brennan *et al.*, 2013). This study system therefore represents the ideal situation to understand the interplay between adaptation, reproductive isolation and the origin of new species.

To investigate these processes, we carried out linkage mapping of an *S. aethnensis* x *S. chrysanthemifolius* F2 population and identified the map positions of highly-differentiated (at the sequence and expression level) outlier loci. Additionally we mapped QTL for the morphological and gene expression differences between these two species and identified QTL for hybrid breakdown. Finally, we interrogated the leaf transcriptomes of the F2 plants to identify differentially expressed (DE) loci, and pathways that are significantly affected in the normal vs. necrotic plants. As far as we are aware we have, for the first time in a plant study system, brought together these complementary investigations to gain an in-depth understanding of the genomics of species divergence and reproductive isolation.

The *Senecio* genetic map: This study saw the generation of a linkage map for *Senecio* comprising nearly 1,000 markers spread across ten linkage groups, the haploid chromosome number. Average distance between markers was just 0.87 cM (or 2.41 cM if we consider only one marker at each genomic location). It therefore represents an important resource for the study of this group of species.

In line with the findings of Brennan *et al*. (2014), we have found evidence for strong transmission ratio distortion (TRD) in 14.0% of markers in the F2 of the crosses between these *Senecio* species. We checked for genotyping errors that could have arisen from biased allelic expression and in just two such markers did the DNA and RNA genotypes not tally perfectly. This TRD could have

391 resulted from the build-up of genomic incompatibilities between the species, and might therefore
392 play a role in preventing gene flow – a possibility we discuss below.

393 **Are there speciation islands in *Senecio*?:** In many pairs of species, sub-species or ecotypes, a
394 heterogeneous pattern of genomic divergence has been uncovered and where these loci have
395 been mapped, the most highly divergent alleles are often found to cluster (Turner *et al.*, 2005;
396 Hohenlohe *et al.*, 2010; Scascitelli *et al.*, 2010; Nadeau *et al.*, 2012; Renaut *et al.*, 2012). This has
397 been used as evidence to support the hypotheses that (1) only a small portion of the genome may
398 confer the differences between species (Wu, 2001), and (2) these speciation loci are nested in
399 small genomic regions ('islands') of increased interspecific divergence (Feder *et al*., 2012). These
400 islands are predicted to arise due to locally reduced effective gene flow because of linkage to the
401 locus under divergent selection (Divergence Hitch-hiking, DH; Via, 2012).

402 We used this framework to investigate whether speciation islands are present in the *Senecio*
403 genome because we previously demonstrated that average genomic differentiation between
404 these species is quite low (mean F_{ST} across loci = 0.196; mean d_{XY} across loci = 0.0035) but highly
405 heterogeneous (Chapman *et al.*, 2013). In the current study, plotting interspecific divergence onto
406 the genetic map reveals that the genomes of these species are highly heterogeneous with respect
407 to interspecific relative (F_{ST}) and absolute (d_{XY}) divergence (Figs. 2 and 4). Further, the distribution
408 of 81 highly differentiated loci placed on the map is suggestive of putative speciation islands, with
409 48 of the high divergence outliers (59.3%) clustering in only three 5 cM regions of the genome
410 (Fig. 4). This therefore represents the pattern that would be expected under DH. Statistically,
411 however this clustering was at best marginally significant if one takes into account marker density
412 (P = 0.045 and 0.168 for 5 and 10 cM bin sizes, respectively). The observed clustering of the
413 outliers is thus mainly due to clustering of many loci in the regions where the genetic map is
414 unresolved (Figs. 2 and 4, Fig. S2).

These gene clusters may correspond to gene-rich regions, but locally reduced recombination, for example in the proximity of centromeres (Round *et al.*, 1997) appears a more likely cause. Upon analysis of the three genomic regions containing high numbers of outliers we find no evidence for higher than average absolute inter-species divergence (d_{XY} in Fig. 5). Instead these regions are characterised by significantly reduced intra-species diversity (π in Fig. 5). This is consistent with recent re-analysis of speciation islands in which low polymorphism and not high absolute divergence was typically found for genomic regions previously identified as speciation islands (Cruickshank & Hahn, 2014). However, we do find regions where both F_{ST} and d_{XY} are elevated (e.g. on LGs 2 and 6), indicating that 'speciation islands' with reduced interspecific gene flow may be present in the *Senecio* genome.

We also found a strong negative correlation between F_{ST} and distance from the closest outlier locus, but only a weak correlation between d_{XY} and distance (Fig. S5), suggesting that F_{ST} is not independent of surrounding loci and adds to the growing evidence that F_{ST} can be artificially elevated in some regions of the genome due to reduced intraspecific genetic diversity (π) (Cruickshank & Hahn, 2014) as was resolved here (Fig. 7). We cannot therefore conclude that strong divergent selection has given rise to the apparent speciation island-like clustering; this will only be resolved by combining denser linkage maps with physical maps currently being created as part of the *Senecio* genome project (T. Batstone, M.A. Chapman, O.G. Osborne, D.A. Filatov, R.J. Abbott, and S.J. Hiscock, in preparation).

Recombination may be reduced not only because of the presence of centromeres, but also due to heterozygosity in F1s for inversions that differentiate the parental taxa. Inversions are known to be the hotspots for accumulation of species divergence and have long been suspected to be involved in speciation (Noor *et al.*, 2001; Rieseberg, 2001; Navarro & Barton, 2003). The relatively recent speciation time estimated for these two species (Chapman *et al.*, 2013; Osborne *et al.*,

2013), however, makes it unlikely that these species already differ by several chromosomal rearrangements, given a maximum rate of chromosomal rearrangements in plants estimated to be 7-10 per million years (Burke *et al.*, 2004; Barb *et al.*, 2014). Still, inversions are known to segregate within populations of the same species (e.g. Craddock & Carson, 1989; Corbett-Detig & Hartl, 2012; Charron *et al.*, 2014) and can be quickly fixed within one or both species if associated with adaptive alleles (Kirkpatrick & Barton, 2006).

Another possible explanation for supressed recombination (and consequentially marker clustering) is that some of these regions are involved in hybrid breakdown. Indeed, one cluster of outlier markers did show TRD (on LG3; Fig. S2), with homozygous alleles from *S. aethnensis* found at low frequency. Interestingly, Brennan *et al*. (2014) also identified a region of their linkage map with the same pattern of selection against homozygous *S. aethnensis* alleles. Unfortunately, it is not possible to tell whether they identified the same or a different genomic region compared to our map. Another region, on our LG4, was devoid of *S. chrysanthemifolius* homozygous genotypes and harboured the QTL for hybrid breakdown. This could have important implications for gene flow in the *Senecio* hybrid zone, if certain allelic combinations in hybrids are unfit.

Sequence divergence and phenotypic QTL: It is interesting that two of the clusters of outliers (on LGs 1 and 6) were in regions of the genome harbouring morphological QTL, and outlier loci in general were more likely to be found to coincide with morphological QTL than expected by chance ($P = 3.22 \times 10^{-7}$). Furthermore, eQTL for loci differentially expressed (DE) between the parental species also tended to cluster with morphological QTL ($P = 0.0045$). Brennan *et al.* (2009) studied morphological clines in the *Senecio* hybrid zone and concluded that groups of morphological traits (both leaf and floral) were under environmental-based selection. Given this, we would expect that the regions of the genome that harbour the loci controlling these traits would show increased interspecific divergence and this is precisely what we see – both at the DNA sequence and gene

expression levels. Genetic divergence builds up in regions of the genome that control the locally adaptive and ecologically conditioned phenotypes. Similarly, in whitefish and sticklebacks, outlier markers were also found to be present in regions of the genome with morphological QTL for traits that were putatively under selection (Hohenlohe *et al.*, 2010; Renaut *et al.*, 2012).

The genetics of hybrid breakdown: The presence of a major QTL for hybrid breakdown that did not overlap any outlier markers suggests that, whilst there is selection for divergent morphologies, hybrid breakdown does not appear to be increasing genomic divergence to the same extent. However it is important to remember that 'Bateson-Dobzhansky-Muller' (BDM) hybrid incompatibility is dependent on alleles at more than one locus, with alleles at each locus not intrinsically unfit. Instead it requires that alleles derived at one locus in one species are incompatible with alleles derived at an alternate locus in the other species, which is what we observed. Alleles at the LG4 QTL inherited from the *S. aethnensis* parent when in the homozygous state were incompatible with *S. chrysanthemifolius* alleles at the LG7 (cM 91 – 108) minor QTL. Alternatively, it is possible that the alleles at the major QTL are interacting with a cytoplasmic locus to result in hybrid breakdown. In the F2 we generated, all plants have the same (*S. chrysanthemifolius*-derived) cytoplasm and the necrotic plants were all homozygous for *S. aethnensis* alleles at the major necrosis QTL. It is noteworthy that Hegarty et al. (2009) crossed these two species to generate F1 to F5 generations of hybrids. A strong drop in germination success was noted in the F2 and F3, but there were no noted examples of necrosis such as we uncovered here.

Associated with hybrid breakdown were differences in gene expression in ca. 12% of the transcriptome when comparing the normal and necrotic F2 plants. This stands in stark contrast to the comparative analysis of the transcriptomes of the parental species (Chapman *et al.*, 2013) in which only 90 of >18,000 loci (ca. 0.5%) were significantly differentially expressed (DE). Although

direct comparisons cannot be made because the current and previous study used different tissues, it seems there was a much greater effect of hybrid breakdown on gene expression than there was based on interspecific differences between healthy parental taxa. A study of hybrid breakdown in lake whitefish revealed an intriguingly similar pattern with <0.1% of loci DE between the parental morphs, however 39% of transcripts were DE in a comparison of the transcriptomes of phenotypically abnormal and normal backcross hybrids (Renaut & Bernatchez, 2011). Our analysis was restricted to seedling leaf tissue to reduce the effect of divergent leaf morphology on gene expression in more mature plants, and therefore we are only able to identify DE leaf-expressed loci, and we might find that different numbers or subsets of loci are DE in other tissues.

A number of gene ontology (GO) categories associated with basic developmental processes were significantly over-represented in the loci differentially expressed (DE) between necrotic and normal F2 plants, suggesting severe impairment of normal growth and development at the gene expression level. The most drastic changes though were associated with ribosomes and translation; for example, the GO terms 'structural constituent of ribosome', 'RNA methylation' (a hallmark of translation regulation; Motorin & Helm, 2011), and 'translation' were all highly over-represented in the loci that were under-expressed in the necrotic F2s relative to the normal F2s (all FDR $P < 0.0001$; Table S3).

Several studies of hybrid breakdown in plants have revealed that genes involved in pathogen- or stress-response are often mis-expressed in the unfit hybrids resulting in necrotic phenotypes (e.g. (Bomblies *et al.*, 2007; Jeuken *et al.*, 2009; Yamamoto *et al.*, 2010). Although we did not see GO terms explicitly involved in pathogen-response being over-represented in our DE loci, homologues of some genes known to play a role in pathogen-response were DE. Of six candidate pathogen-response genes listed as DE in necrotic *Arabidopsis* hybrids (Bomblies *et al.*, 2007), homologues of three (*enhanced disease susceptibility 1*, *phytoalexin deficient 4*, and *pathogenesis-related protein*

511 *1*) were significantly DE between necrotic and normal *Senecio* F2s. In addition, the Cellular
512 Compartment GO term 'peroxisome' was over-represented in the DE loci; one function of which
513 includes response to fungal pathogens, suggesting that mis-expression of pathogen-response
514 genes might underlie, at least in part, the genetics of reproductive isolation (Lipka *et al.*, 2005).

515 **CONCLUSIONS**

516 This study set out to genetically map the distribution of markers that exhibited putative adaptive
517 divergence and to investigate the genetic basis of morphological divergence and hybrid
518 breakdown. We found that more than half of the markers with previous evidence of adaptive
519 divergence clustered into three regions of the genome; however, the clustering appears to be due
520 to a higher density of genes in regions with reduced recombination, rather than accumulation of
521 divergence outliers in 'speciation islands'. Whether these gene clusters represent inversions
522 distinguishing the two species, or centromeric regions (Gimenez *et al.*, 2013; reviewed in Faria &
523 Navarro, 2010), requires further investigation using our emerging genome sequences for the two
524 species. Interestingly, we also discovered clusters of markers showing significant TRD, indicative of
525 accumulation of intrinsic genetic incompatibilities between these relatively recently diverged
526 (<200,000 years) species. One region of TRD corresponded to a QTL for hybrid breakdown, and the
527 distribution of genotypes at this locus showed evidence for incompatibility with another region of
528 the genome.

529 Divergence of these species may thus have occurred due to a complex interplay of ecological
530 divergence and accumulation of intrinsic genetic incompatibilities. The relative importance of
531 these forces may be evaluated in future studies devoted to dissection of species incompatibilities
532 and genetic analysis of quantitative traits involved in adaptation to contrasting environments of
533 high and low altitudes.

ACKNOWLEDGEMENTS

We are grateful to Owen Osborne for discussions throughout this project and to Richard Abbott and Adrian Brennan for sharing their recently accepted manuscript. We thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics for the generation of the Sequencing data. This study was funded by UK Natural Environment Research Council grant to DAF and SJH (NE/G017646/1).

COMPETING INTERESTS

The authors have no competing interests

REFERENCES

Arnold, M. L. 1992. Natural hybridization as an evolutionary process. *Annu Rev Ecol Syst* **23**: 237-261.

Barb, J. G., Bowers, J. E., Renaut, S., Rey, J. I., Knapp, S. J., Rieseberg, L. H. & Burke, J. M. 2014. Chromosomal Evolution and Patterns of Introgression in *Helianthus*. *Genetics* **197**: 969-979.

Barton, N. H. & Hewitt, G. M. 1989. Adaptation, speciation and hybrid zones. *Nature* **341**: 497-503.

Begun, D. J. & Aquadro, C. F. 1992. Levels of naturally-occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. *Nature* **356**: 519-520.

Bennett, M. D. & Smith, J. B. 1976. Nuclear DNA amounts in angiosperms. *Philos Trans R Soc Lond B* **274**: 227-274.

Bierne, N., Welch, J., Loire, E., Bonhomme, F. & David, P. 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol Ecol* **20**: 2044-2072.

Bomblies, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J. L. *et al.* 2007. Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol* **5**: e236.

Brennan, A. C., Bridle, J. R., Wang, A.-L., Hiscock, S. J. & Abbott, R. J. 2009. Adaptation and selection in the *Senecio* (Asteraceae) hybrid zone on Mount Etna, Sicily. *New Phytol***183**: 702-717.

Brennan, A. C., Harris, S. A. & Hiscock, S. J. 2013. The population genetics of sporophytic self-incompatibility in three hybridizing Senecio (Asteraceae) species with contrasting population histories. *Evolution (N.Y.)* **67**: 1347-1367.

Brennan, A. C., Hiscock, S. J. & Abbott, R. J. 2014. Interspecific crossing and genetic mapping reveal intrinsic genomic incompatibility between two Senecio species that form a hybrid zone on Mount Etna, Sicily. Heredity 113: 195-204.

Burke, J. M., Lai, Z., Salmaso, M., Nakazato, T., Tang, S., Heesacker, A., *et al.* 2004. Comparative mapping and rapid karyotypic evolution in the genus *Helianthus*. *Genetics* **167**: 449-457.

Chapman, M. A., Forbes, D. G. & Abbott, R. J. 2005. Pollen competition among two species of *Senecio* (Asteraceae) that form a hybrid zone on Mt. Etna, Sicily. *Am J Bot* **92**: 730-735.

569 Chapman, M. A., Hiscock, S. J. & Filatov, D. A. 2013. Genomic divergence during speciation driven by
570 adaptation to altitude. *Mol Biol Evol* **30**: 2553-2567.
571 Charlesworth, B., Nordborg, M. & Charlesworth, D. 1997. The effects of local selection, balanced
572 polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided
573 populations. *Genet Res***70**: 155-174.
574 Charron, G., Leducq, J. B. & Landry, C. R. 2014. Chromosomal variation segregates within incipient species
575 and correlates with reproductive isolation. *Mol Ecol* **23**: 4362-4372.
576 Corbett-Detig, R. B. & Hartl, D. L. 2012. Population Genomics of Inversion Polymorphisms in *Drosophila*
577 *melanogaster*. *PLoS Genet* **8**.
578 Coyne, J. A. & Orr, H. A. 2004. *Speciation*. Sinauer, Sunderland, Massachusetts, USA.
579 Craddock, E. M. & Carson, H. L. 1989. Chromosomal inversion patterning and population differentiation in a
580 young insular species, *Drosophila silvestris*. *Proc Natl Acad Sci U S A* **86**: 4798-4802.
581 Cruickshank, T. E. & Hahn, M. W. 2014. Reanalysis suggests that genomic islands of speciation are due to
582 reduced diversity, not reduced gene flow. *Mol Ecol* **23**: 3133-3157.
583 Endler, J. A. 1977. *Geographic Variation, Speciation, and Clines.* Princeton University Press, Princeton, USA.
584 Faria, R. & Navarro, A. 2010. Chromosomal speciation revisited: rearranging theory with pieces of evidence.
585 *Trends Ecol Evol* **25**: 660-669.
586 Feder, J. L., Egan, S. P. & Nosil, P. 2012. The genomics of speciation-with-gene-flow. *Trends Genet* **28**: 342-
587 350.
588 Feder, J. L. & Nosil, P. 2010. The efficacy of divergence hitchhiking in generating genomic islands during
589 ecological speciation. *Evolution (N.Y.)* **64**: 1729-1747.
590 Filatov, D. A. 2009. Processing and population genetic analysis of multigenic datasets with ProSeq3
591 software. *Bioinformatics* **25**: 3189-3190.
592 Foll, M. & Gaggiotti, O. 2008. A genome-scan method to identify selected loci appropriate for both
593 dominant and codominant markers: A Bayesian perspective. *Genetics* **180**: 977-993.
594 Fuller, R. C. 2008. Genetic incompatibilities in killifish and the role of environment. *Evolution (N.Y.)* **62**:
595 3056-3068.
596 Gimenez, M. D., White, T. A., Hauffe, H. C., Panithanarak, T. & Searle, J. B. 2013. Understanding the basis of
597 diminished gene flow between hybridizing chromosome races of the house mouse. *Evolution (N.Y.)*
598 **67**: 1446-1462.
599 Hegarty, M. J., Barker, G. L., Brennan, A. C., Edwards, K. J., Abbott, R. J. & Hiscock, S. J. 2009. Extreme
600 changes to gene expression associated with homoploid hybrid speciation. *Mol Ecol* **18**: 877-889.
601 Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A. & Cresko, W. A. 2010. Population
602 genomics of parallel adaptation in Threespine Stickleback using sequenced RAD tags. *PLoS Genet* **6**:
603 e1000862.
604 James, J. K. & Abbott, R. J. 2005. Recent, allopatric, homoploid hybrid speciation: The origin of *Senecio*
605 *squalidus* (Asteraceae) in the British Isles from a hybrid zone on Mount Etna, Sicily. *Evolution (N.Y.)*
606 **59**: 2533-2547.
607 Jeuken, M. J. W., Zhang, N. W., McHale, L. K., Pelgrom, K., den Boer, E., Lindhout, P., *et al.* 2009. Rin4
608 causes hybrid necrosis and race-specific resistance in an interspecific lettuce hybrid. *Plant Cell* **21**:
609 3368-3378.
610 Kane, N. C., King, M. G., Barker, M. S., Raduski, A., Karrenberg, S., Yatabe, Y., *et al.* 2009. Comparative
611 genomic and population genetic analyses indicate highly porous genomes and high levels of gene
612 flow between divergent *Helianthus* species. *Evolution (N.Y.)* **63**: 2061-2075.
613 Kirkpatrick, M. & Barton, N. 2006. Chromosome inversions, local adaptation and speciation. *Genetics* **173**:
614 419-434.
615 Kulathinal, R. J., Stevison, L. S. & Noor, M. A. F. 2009. The genomics of speciation in *Drosophila*: Diversity,
616 divergence, and introgression estimated using low-coverage genome sequencing. *PLoS Genet* **5**:
617 e1000550.
618 Lexer, C. & Widmer, A. 2008. The genic view of plant speciation: recent progress and emerging questions.
619 *Philos Trans R Soc Lond B* **363**: 3023-3036.
620 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., *et al.* 2009. The Sequence
621 Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.

Lipka, V., Dittgen, J., Bednarek, P., Bhat, R., Wiermer, M., Stein, M., *et al.* 2005. Pre- and postinvasion defenses both contribute to nonhost resistance in *Arabidopsis*. *Science* **310**: 1180-1183.
Lorieux, M. 2012. MapDisto: fast and efficient computation of genetic linkage maps. *Molecular Breeding* **30**: 1231-1235.
Mallet, J. 2005. Hybridization as an invasion of the genorne. *Trends Ecol Evol* **20**: 229-237.
Michel, A. P., Sim, S., Powell, T. H. Q., Taylor, M. S., Nosil, P. & Feder, J. L. 2011. Widespread genomic divergence during sympatric speciation. *Proc Natl Acad Sci U S A* **107**: 9724-9729.
Moore, W. S. 1977. An evaluation of narrow hybrid zones in vertebrates. *Q Rev Biol***52**: 263-277.
Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**: 621-628.
Motorin, Y. & Helm, M. 2011. RNA nucleotide methylation. *WIREs RNA* **2**: 611-631.
Muir, G., Osborne, O. G., Sarasa, J., Hiscock, S. J. & Filatov, D. A. 2013. Recent ecological selection on regulatory divergence is shaping clinal variation in *Senecio* on Mount Etna. *Evolution (N.Y.)* **67**: 3032-3042.
Nadeau, N. J., Whibley, A., Jones, R. T., Davey, J. W., Dasmahapatra, K. K., Baxter, S. W., *et al.* 2012. Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. *Philos Trans R Soc Lond B* **367**: 343-353.
Navarro, A. & Barton, N. H. 2003. Accumulating postzygotic isolation genes in parapatry: A new twist on chromosomal speciation. *Evolution (N.Y.)* **57**: 447-459.
Noor, M. A. F. & Bennett, S. M. 2009. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity (Edinb)* **103**: 439-444.
Noor, M. A. F., Grams, K. L., Bertucci, L. A. & Reiland, J. 2001. Chromosomal inversions and the reproductive isolation of species. *Proc Natl Acad Sci U S A* **98**: 12084-12088.
Nosil, P. 2012. *Ecological Speciation*. Oxford University Press, Oxford.
Nosil, P., Egan, S. P. & Funk, D. J. 2008. Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. *Evolution (N.Y.)* **62**: 316-336.
Nosil, P., Funk, D. J. & Ortiz-Barrientos, D. 2009. Divergent selection and heterogeneous genomic divergence. *Mol Ecol* **18**: 375-402.
Osborne, O. G., Batstone, T. E., Hiscock, S. J. & Filatov, D. A. 2013. Rapid speciation with gene flow following the formation of Mount Etna, revealed by transcriptome sequencing of endemic Etnean *Senecio* species. *Genome Biol Evol* **5**: 1704-15.
Payseur, B. A. 2010. Using differential introgression in hybrid zones to identify genomic regions involved in speciation. *Mol Ecol Resour* **10**: 806-820.
Peichel, C. L., Nereng, K. S., Ohgi, K. A., Cole, B. L. E., Colosimo, P. F., Buerkle, C. A., *et al.* 2001. The genetic architecture of divergence between threespine stickleback species. *Nature* **414**: 901-905.
Rasband, W. S. (1997-2013) ImageJ (imagej.nih.gov/ij/). pp. U.S. National Institutes of Health, Bethesda, Maryland, USA,.
Renaut, S. & Bernatchez, L. 2011. Transcriptome-wide signature of hybrid breakdown associated with intrinsic reproductive isolation in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Heredity (Edinb)* **106**: 1003-1011.
Renaut, S., Maillet, N., Normandeau, E., Sauvage, C., Derome, N., Rogers, S. M. *et al.* 2012. Genome-wide patterns of divergence during speciation: the lake whitefish case study. *Philos Trans R Soc Lond B* **367**: 354-363.
Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. *Trends Ecol Evol* **16**: 351-358.
Rieseberg, L. H. & Carney, S. E. 1998. Tansley review no. 102: Plant hybridisation. *New Phytol* **140**: 599-624.
Roesti, M., Hendry, A. P., Salzburger, W. & Berner, D. 2012. Genome divergence during evolutionary diversification as revealed in replicate lake–stream stickleback population pairs. *Mol Ecol* **21**: 2852-2862.
Round, E. K., Flowers, S. K. & Richards, E. J. 1997. *Arabidopsis thaliana* Centromere Regions: Genetic Map Positions and Repetitive DNA Structure. *Genome Res* **7**: 1045-1053.

673 Scascitelli, M., Whitney, K. D., Randell, R. A., King, M., Buerkle, C. A. & Rieseberg, L. H. 2010. Genome scan
674 of hybridizing sunflowers from Texas (*Helianthus annuus* and *H. debilis*) reveals asymmetric
675 patterns of introgression and small islands of genomic differentiation. *Mol Ecol* **19**: 521-541.
676 Scotti-Saintagne, C., Mariette, S., Porth, I., Goicoechea, P. G., Barreneche, T., Bodenes, K., *et al.* 2004.
677 Genome scanning for interspecific differentiation between two closely related oak species *Quercus*
678 *robur* L. and *Q. petraea* (Matt.) Liebl. *Genetics* **168**: 1615-1626.
679 Stace, C. A. 1975. *Hybridisation and the Flora of the British Isles*. Academic Press, London and New York.
680 Stebbins, G. L. 1959. The role of hybridisation in evolution. *Proc Am Philos Soc* **103**: 231-251.
681 Stölting, K. N., Nipper, R., Lindtke, D., Caseys, C., Waeber, S., Castiglione, S. *et al.* 2013. Genomic scan for
682 single nucleotide polymorphisms reveals patterns of divergence and gene flow between
683 ecologically divergent species. *Mol Ecol* **22**: 842-855.
684 Sweigart, A. L. & Willis, J. H. 2012. Molecular evolution and genetics of postzygotic reproductive isolation in
685 plants. *F1000 Biol Rep* **4**: 23.
686 Turner, T. L., Hahn, M. W. & Nuzhdin, S. V. 2005. Genomic islands of speciation in *Anopheles gambiae*. *PLoS*
687 *Biol* **3**: 1572-1578.
688 Via, S. 2012. Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-
689 gene-flow. *Philos Trans R Soc Lond B* **367**: 451-460.
690 Via, S., Conte, G., Mason-Foley, C. & Mills, K. 2012. Localizing FST outliers on a QTL map reveals evidence
691 for large genomic regions of reduced gene exchange during speciation-with-gene-flow. *Mol Ecol* **21**:
692 5546-5560.
693 Voorips, R. E. 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. *J Hered*
694 **93**: 77-78.
695 Wang, S., Basten, C. J. & Zeng, Z.-B. 2005. Windows QTL Cartographer 2.5. Department of Statistics, North
696 Carolina State University, Raleigh, NC. (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm).
697 Wolf, J. B. W., Lindell, J. & Backstrom, N. 2010. Speciation genetics: current status and evolving approaches.
698 *Philos Trans R Soc Lond B* **365**: 1717-1733.
699 Wu, C.-I. 2001. The genic view of the process of speciation. *Journal of Evolutionary Biology* **14**: 851-865.
700 Yamamoto, E., Takashi, T., Morinaka, Y., Lin, S., Wu, J., Matsumoto, T.,*et al.* 2010. Gain of deleterious
701 function causes an autoimmune response and Bateson–Dobzhansky–Muller incompatibility in rice.
702 *Molecular Genetics and Genomics* **283**: 305-315.
703 Yatabe, Y., Kane, N. C., Scotti-Saintagne, C. & Rieseberg, L. H. 2007. Rampant gene exchange across a strong
704 reproductive barrier between the annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics*
705 **175**: 1883-1893.

706

707 **FIGURE LEGENDS**

708 **Figure 1** – Plants in the F2 segregated for necrotic phenotypes (A-D) and healthy phenotypes (E,F).

709 **Figure 2 -** Representative linkage groups (LG) from the interspecific mapping population. Positive

710 and negative BayeScan outlier loci are indicated in red and blue, respectively. Distances are in

711 Kosambi cM. Asterisks indicate loci showing significant transmission ratio distortion ($P < 0.001$).

712 The complete map is displayed in Fig. S2.

713 **Figure 3 –** Genotype frequencies along the three LGs that showed TRD. Genotypes are AA (blue),
714 *S. aethnensis* homozygote; CC (red), *S. chrysanthemifolius* homozygote; AC (purple) heterozygote.
715 Grey boxes indicate the markers which display significant distortion (χ^2 test, $P < 0.001$).

716 **Figure 4** - $F_{ST,}$ d_{XY} (A), and π (B) are not homogeneous across linkage groupsin this *Senecio* species
717 pair. LGs 1-10 from are shown left to right with the scale in map distance. (A) Sliding window
718 (width 10 cM, step 5 cM) analysis of F_{ST} (green line) and d_{XY} (purple line) across the ten linkage
719 groups (left hand scale). The dashed lines indicate the upper 10% of the distribution. Red and blue
720 bars indicate the number (right hand scale) of positive (red) and negative (blue) outliers per 5 cM
721 bin; (B) Sliding window (width 10 cM, step 5 cM) analysis of π in *S. aethnensis* (blue) and *S.*
722 *chrysanthemifolius* (red) across the ten linkage groups. The blue and red dashed line indicates the
723 lower 10% of the distribution for each species. Linkage groups 1 to 10 are indicated in each panel.

724 **Figure 5 –** Population genetic data for markers in the three regions of the genome (on LGs 1, 3 and
725 6) exhibiting putative outlier clustering and the rest of the genome.

726 **Figure 6** – QTL for morphological differences and hybrid breakdown. The seven LGs with QTL are
727 shown. Markers mapping to the same position were excluded prior to QTL analysis. Traits (see
728 table 1) are colour-coded (green: leaf, blue: inflorescence, red: plant height; black, necrosis). One-
729 LOD (boxes) and 2-LOD QTL intervals (whiskers) are indicated.

730 **Figure S1** – Parental and F2 leaf silhouettes. Two *S. chrysanthemifolius* leaves are shown on the
731 left and two *S. aethnensis* leaves are shown on the right. Other leaves are from the healthy F2
732 plants.

733 **Figure S2** - Linkage map of the cross between *S. aethnensis* Pro2-21 and *S. chrysanthemifolius*
734 Ran1-5. Positive and negative BayeScan outlier loci are indicated in red and blue, respectively.

735 Distances are in Kosambi cM. Asterisks indicate loci showing significant transmission ratio
736 distortion (P < 0.001).

737 **Figure S3** – Among linkage group (LG) differences in F_{ST} (A) and π (B). F_{ST} varied significantly among
738 LGs whether or not the outlier markers were included in the analysis whereas d_{XY} did not. π varied
739 among LGs and species.

740 **Figure S4** - Per locus F_{ST} (top panel) and π (bottom panel; blue = *S. aethnensis*, red = *S.*
741 *chrysanthemifolius*) across all linkage groups.

742 **Figure S5** – Distance from nearest outlier locus was significantly negatively correlated with F_{ST} (P =
743 1.321×10^{-24}) and d_{XY} (P = 0.0078), and positively correlated with nucleotide diversity (π; P = 8.399
744 $\times 10^{-6}$ and P = 0.0003 for *S. aethnensis* and *S. chrysanthemifolius*, respectively).

745 **Figure S6** – eQTL (red) for 67 loci differentially-expressed between *S. aethnensis* and *S.*
746 *chrysanthemifolius* superimposed on the morphological/hybrid breakdown QTL map (Fig. 6).

747 **Table S1 (.xls) –** Numbers of sequence reads produced per individual.

748 **Table S2 (.xls) –** Primers used in the targeted genotyping and sequencing.

749 **Table S3 (.xls) –** GO terms over-represented in the differentially-expressed loci relative to the non-
750 differentially-expressed loci. (Top) all 2,301 loci, (Middle) 592 loci over-expressed in the necrotic
751 plants and (Bottom) 1,709 loci under-expressed in the necrotic plants.

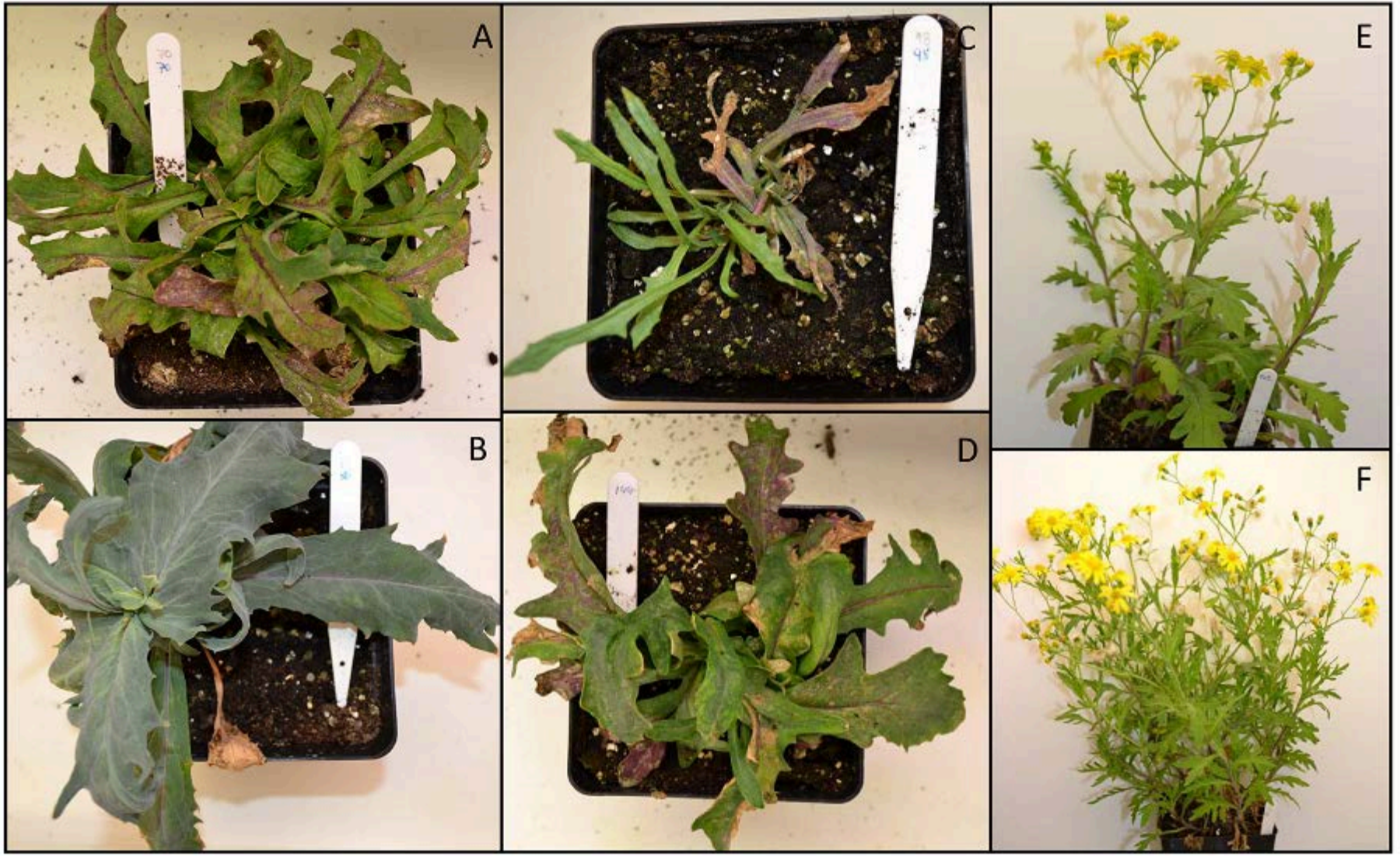
Table 1 – Traits measured in the F2 mapping population. Asterisks denote traits that were square-root transformed before QTL analysis. Acronyms correspond to the trait codes used in the QTL figure (Fig. 2) and table 2. Parental means (untransformed data) ± S.D. and *t*-test for significant differences between *S. chrysanthemifolius* and *S. aethnensis* are given. 'n.s.', non-significant ($P > 0.05$), '-' not measured.

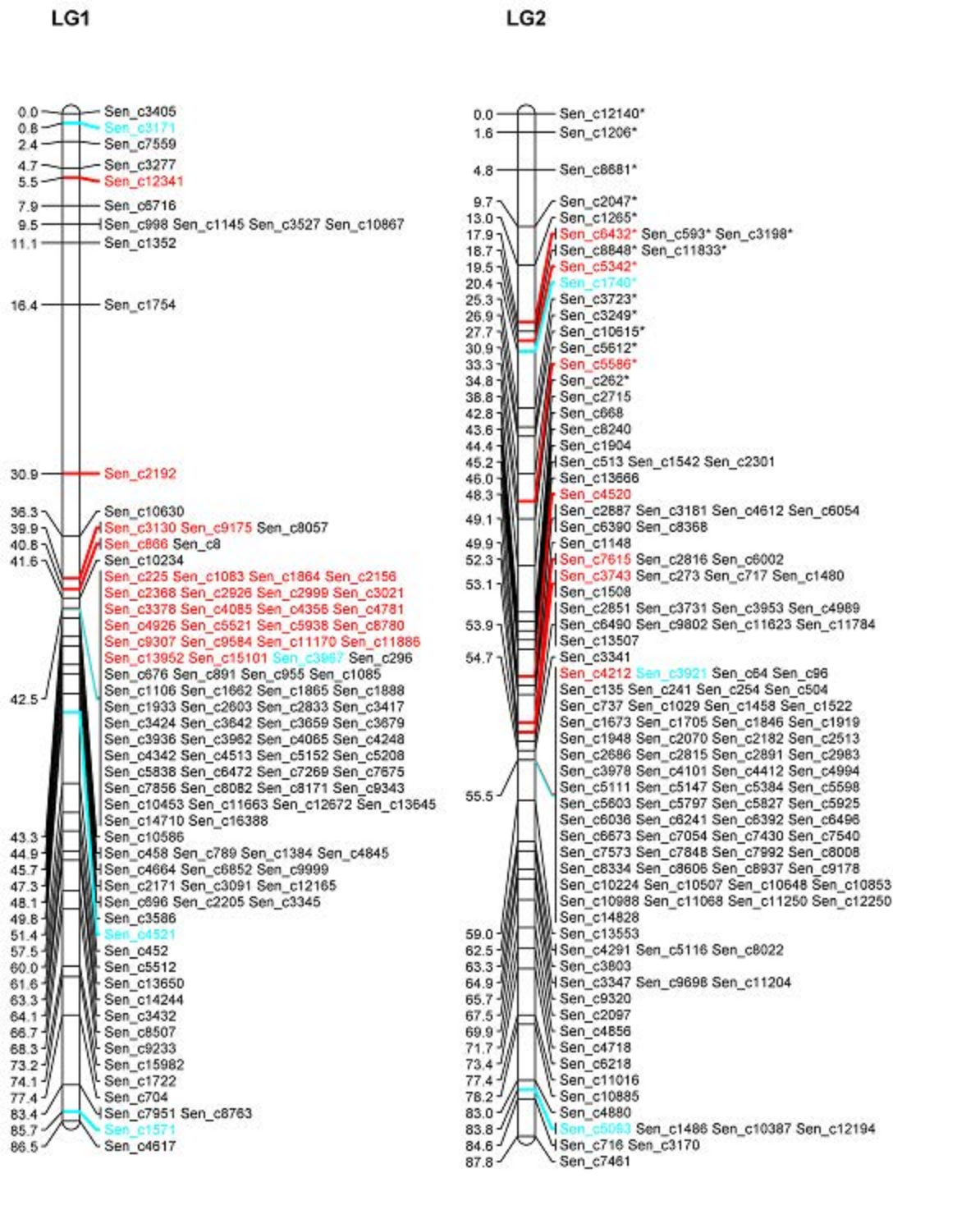
Trait	Description of measurement	Acronym	*S. chrysanthemifolius*	*S. aethnensis*	*t*-test *P* value
Days to flowering	Days from germination to first anthesis	D2F	-	-	-
Plant Height*	Plant height on day of first anthesis (mm)	Ht	321.54 (± 108.79)	285.38 (± 23.47)	n.s.
Capitulum Width	Width of apical capitulum (mm)	CapWid	4.83 (± 0.21)	6.43 (± 0.52)	2.2×10^{-5}
Capitulum Height	Height of apical capitulum (mm)	CapHt	9.50 (± 0.50)	10.69 (± 0.48)	6.9×10^{-5}
Number of ray florets	Count of rays on apical capitulum	nRays	12.15 (± 1.21)	11.00 (± 2.14)	n.s.
Ray floret length	Average length of three ray florets (mm)	RayLen	8.73 (± 1.29)	11.70 (± 2.55)	0.013
Ray floret width	Average width of three ray florets (mm)	RayWid	2.72 (± 0.26)	3.97 (± 0.67)	8.3×10^{-4}
Leaf length	Length of midleaf (mm) using ImageJ	LfLen	1,415.08 (± 180.74)	1,032.63 (± 110.95)	8.8×10^{-6}
Leaf area	Area of midleaf (mm^2) using ImageJ	LfArea	378,748 (± 85,187)	255,720 (± 58,534)	9.5×10^{-4}
Leaf perimeter	Perimeter of midleaf (mm) using ImageJ	LfPerim	17,559 (± 3,654)	2,871 (± 215)	5.2×10^{-9}
Leaf dissection*	LfPerim /√LfArea	LfDiss	28.88 (± 6.52)	5.75 (± 0.55)	1.9×10^{-8}
Leaf standardised perimeter*	LfPerim/ LfLen	LfStdPerim	12.51 (± 2.89)	2.79 (± 0.15)	4.1×10^{-8}
Leaf standardised area	√LfArea/LfLen	LfStdArea	0.44 (± 0.05)	0.49 (± 0.06)	0.044

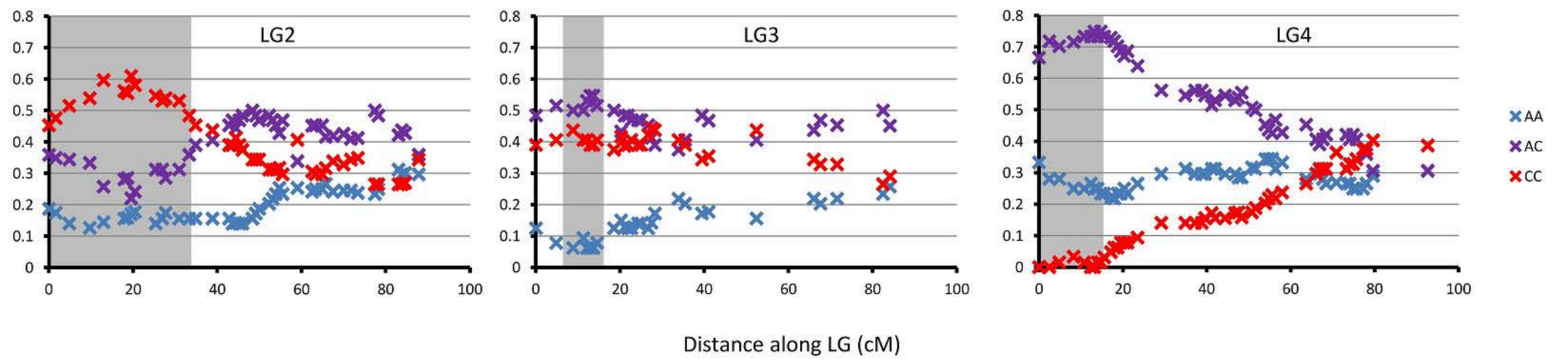
Table 2 – Results of the QTL analysis. Morphological traits (see table 1 for acronyms) were mapped using CIM and necrosis was mapped as a binary trait using IM. LG, linkage group; LOD, logarithm of odds; Add, Additive effect; Dom, dominance effect; PVE, Percentage of variance explained. -2, +2 and -1, +1 indicate the positions (in cM) of the 2-LOD and 1-LOD intervals, respectively.

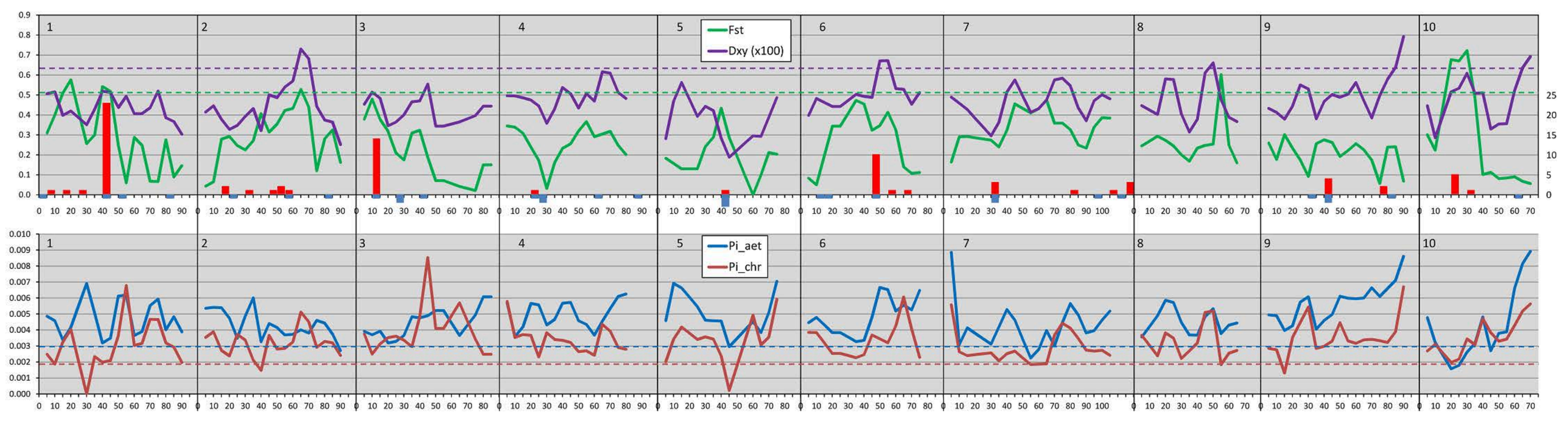
Trait	LG	Peak (cM)	LOD	Add	Dom	-2 LOD	-1 LOD	+1 LOD	+2 LOD	PVE
Ht	7	35.8	4.622	1.61	-0.48	32.5	35.0	36.4	39.7	28.30
Ht	7	65.49	4.107	-1.02	-0.816	64.5	65.05	68	68.8	25.58
CapWid	1	46.18	3.9	0.209	-0.116	24.5	44.5	50.3	55	24.47
CapHt	7	85.34	5.231	0.683	0.687	75	85	86.2	86.5	31.37
nRays	1	8.91	4.731	1.062	0.0396	3	5.5	10.2	26.8	28.85
nRays	1	21.47	5.207	1.104	-0.08	3.3	20.8	22.5	26.5	31.25
RayLen	1	17.14	4.607	-0.75	1.1931	3.5	11	22	23	28.22
RayLen	6	65.23	3.778	-1.05	0.5162	56.2	56.8	65.02	70	23.81
RayWid	1	21.47	11.74	-0.44	0.6582	5.53	11	22	22.5	57.04
LfLen	6	64.23	4.186	-108	77.379	49.2	63.3	70.2	76.03	26.01
LfLen	7	3.33	6.431	208.3	19.498	0	0	6.2	9.8	37.04
LfArea	1	30.58	5.83	0	0	12	27.5	34.55	43.2	34.26
LfArea	5	24.95	4.304	0	0	0	19.9	26.9	42	26.63
LfPerim	1	29.58	3.734	0	0	11	27	40.1	41.5	23.56
LfPerim	5	24.95	5.865	0	0	7.5	16.5	27	28.9	34.43
LfPerim	5	37.01	4.403	0	0	7.3	34.5	40	45	27.15
LfPerim	6	42.83	5.158	0	0	26.8	40.5	46	58	31.00
LfPerim	6	55.28	4.78	0	-432	26.8	50	57.7	58	29.10
LfDiss	5	17.96	4.452	-0.26	0.0911	8	13.5	32.7	43.09	27.41
LfDiss	6	31.79	4.211	-0.27	-0.079	8.7	22.5	39	39.5	26.14
LfDiss	6	75.26	4.459	-0.24	-0.022	59.5	72.5	76.03	76.03	27.45
LfDiss	7	65.49	4.82	-0.26	-0.247	64	65	67	69	29.31
LfStdPerim	7	88.49	4.155	-0.18	-0.144	71	75.5	90	96	25.84
LfStdArea	1	30.58	4.337	-0.04	0.0229	11.7	25	43	43.05	26.81

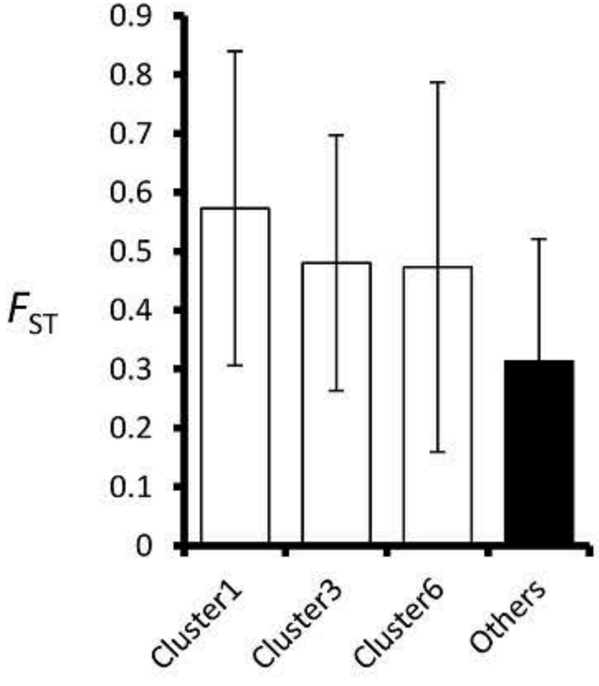
Necrosis	1	16	2.77	-1.33	-0.55	11.8	12	18.5	24.5	18.04
Necrosis	1	32	4.04	-1.12	-0.75	28	29.5	35.5	39.5	25.25
Necrosis	2	63	2.74	-0.84	-1.02	53.5	60	64.8	73	17.89
Necrosis	4	7	15.63	40.32	-73.18	4.7	6.5	11.5	12	67.51
Necrosis	7	81	3.47	-1.08	-0.38	80	80.8	82.3	84.5	22.11
Necrosis	7	103	3.63	-1.20	0.93	91.5	100.8	108.22	108.22	23.00
Necrosis	8	55	4.69	0.02	-1.92	51.5	51.8	56.3	61.1	28.65

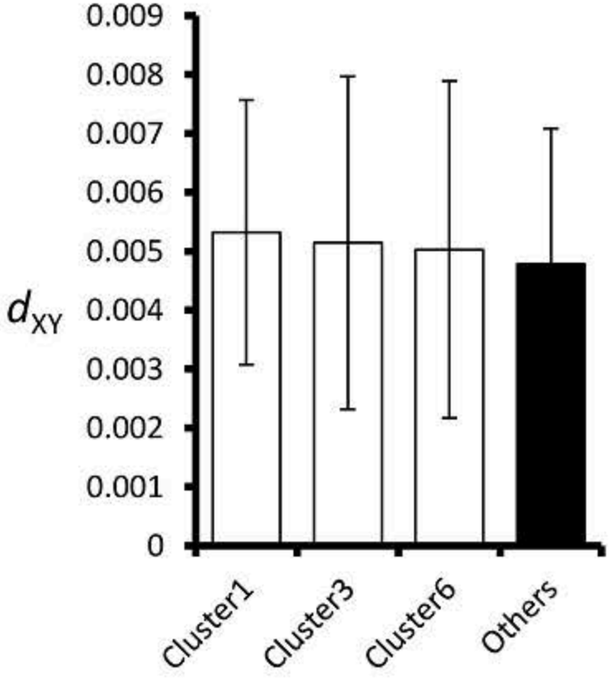

A
B
C
D
E
F

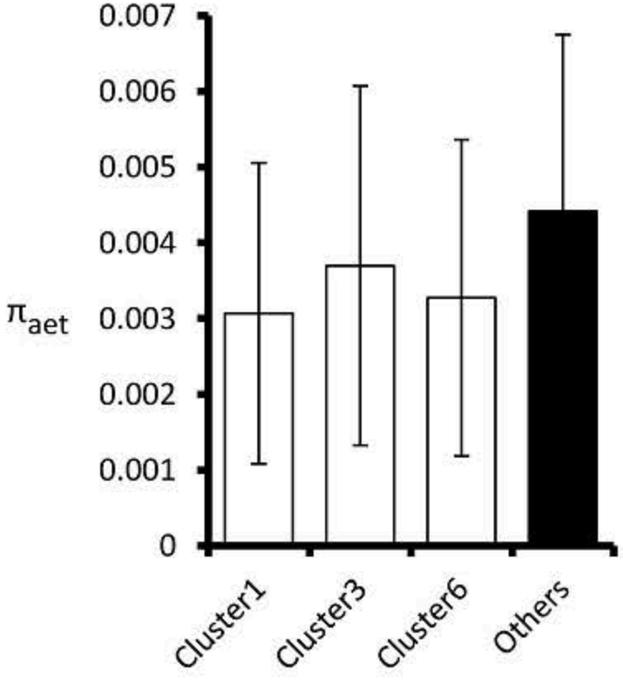

LG1
0.0 Sen_c3405
0.8 Sen_c3171
2.4 Sen_c7559
4.7 Sen_c3277
5.5 Sen_c12341
7.9 Sen_c6716
9.5 Sen_c998 Sen_c1145 Sen_c3527 Sen_c10867
11.1 Sen_c1352
16.4 Sen_c1754
30.9 Sen_c2192
36.3 Sen_c10630
39.9 Sen_c3130 Sen_c9175 Sen_c8057
40.8 Sen_c866 Sen_c8
41.6 Sen_c10234
Sen_c225 Sen_c1083 Sen_c1864 Sen_c2156
Sen_c2368 Sen_c2926 Sen_c2999 Sen_c3021
Sen_c3378 Sen_c4085 Sen_c4356 Sen_c4781
Sen_c4926 Sen_c5521 Sen_c5938 Sen_c8780
Sen_c9307 Sen_c9584 Sen_c11170 Sen_c11886
Sen_c13952 Sen_c15101 Sen_c3967 Sen_c296
Sen_c676 Sen_c891 Sen_c955 Sen_c1085
42.5 Sen_c1106 Sen_c1662 Sen_c1865 Sen_c1888
Sen_c1933 Sen_c2603 Sen_c2833 Sen_c3417
Sen_c3424 Sen_c3642 Sen_c3659 Sen_c3679
Sen_c3936 Sen_c3962 Sen_c4065 Sen_c4248
Sen_c4342 Sen_c4513 Sen_c5152 Sen_c5208
Sen_c5838 Sen_c6472 Sen_c7269 Sen_c7675
Sen_c7856 Sen_c8082 Sen_c8171 Sen_c9343
Sen_c10453 Sen_c11663 Sen_c12672 Sen_c13645
Sen_c14710 Sen_c16388
43.3 Sen_c10586
44.9 Sen_c458 Sen_c789 Sen_c1384 Sen_c4845
45.7 Sen_c4664 Sen_c6852 Sen_c9999
47.3 Sen_c2171 Sen_c3091 Sen_c12165
48.1 Sen_c896 Sen_c2205 Sen_c3345
49.8 Sen_c3586
51.4 Sen_c4521
57.5 Sen_c452
60.0 Sen_c5512
61.6 Sen_c13650
63.3 Sen_c14244
64.1 Sen_c3432
66.7 Sen_c8507
68.3 Sen_c9233
73.2 Sen_c15982
74.1 Sen_c1722
77.4 Sen_c704
83.4 Sen_c7951 Sen_c8763
85.7 Sen_c1571
86.5 Sen_c4617
LG2
0.0 Sen_c12140*
1.6 Sen_c1206*
4.8 Sen_c8681*
9.7 Sen_c2047*
13.0 Sen_c1265*
17.9 Sen_c6432* Sen_c593* Sen_c3198*
18.7 Sen_c8848* Sen_c11833*
19.5 Sen_c5342*
20.4 Sen_c1740*
25.3 Sen_c3723*
26.9 Sen_c3249*
27.7 Sen_c10615*
30.9 Sen_c5612*
33.3 Sen_c5586*
34.8 Sen_c262*
36.8 Sen_c2715
42.8 Sen_c668
43.6 Sen_c8240
44.4 Sen_c1904
45.2 Sen_c513 Sen_c1542 Sen_c2301
46.0 Sen_c13666
48.3 Sen_c4520
Sen_c2887 Sen_c3181 Sen_c4612 Sen_c6054
49.1 Sen_c6390 Sen_c8368
49.9 Sen_c1148
52.3 Sen_c7615 Sen_c2816 Sen_c6002
53.1 Sen_c3743 Sen_c273 Sen_c717 Sen_c1480
Sen_c1508
Sen_c2851 Sen_c3731 Sen_c3953 Sen_c4989
53.9 Sen_c6490 Sen_c9802 Sen_c11623 Sen_c11784
Sen_c13507
54.7 Sen_c3341
Sen_c4212 Sen_c3921 Sen_c64 Sen_c96
Sen_c135 Sen_c241 Sen_c254 Sen_c504
Sen_c737 Sen_c1029 Sen_c1458 Sen_c1522
Sen_c1673 Sen_c1705 Sen_c1846 Sen_c1919
Sen_c1948 Sen_c2070 Sen_c2182 Sen_c2513
Sen_c2686 Sen_c2815 Sen_c2891 Sen_c2983
Sen_c3978 Sen_c4101 Sen_c4412 Sen_c4994
Sen_c5111 Sen_c5147 Sen_c5384 Sen_c5598
55.5 Sen_c5603 Sen_c5797 Sen_c5827 Sen_c5925
Sen_c6036 Sen_c6241 Sen_c6392 Sen_c6496
Sen_c6673 Sen_c7054 Sen_c7430 Sen_c7540
Sen_c7573 Sen_c7848 Sen_c7992 Sen_c8008
Sen_c8334 Sen_c8606 Sen_c8937 Sen_c9178
Sen_c10224 Sen_c10507 Sen_c10648 Sen_c10853
Sen_c10988 Sen_c11068 Sen_c11250 Sen_c12250
Sen_c14828
59.0 Sen_c13553
62.5 Sen_c4291 Sen_c5116 Sen_c8022
63.3 Sen_c3803
64.9 Sen_c3347 Sen_c9698 Sen_c11204
65.7 Sen_c9320
67.5 Sen_c2097
69.9 Sen_c4856
71.7 Sen_c4718
73.4 Sen_c6218
77.4 Sen_c11016
78.2 Sen_c10885
83.0 Sen_c4880
83.8 Sen_c5083 Sen_c1486 Sen_c10387 Sen_c12194
84.6 Sen_c716 Sen_c3170
87.8 Sen_c7461

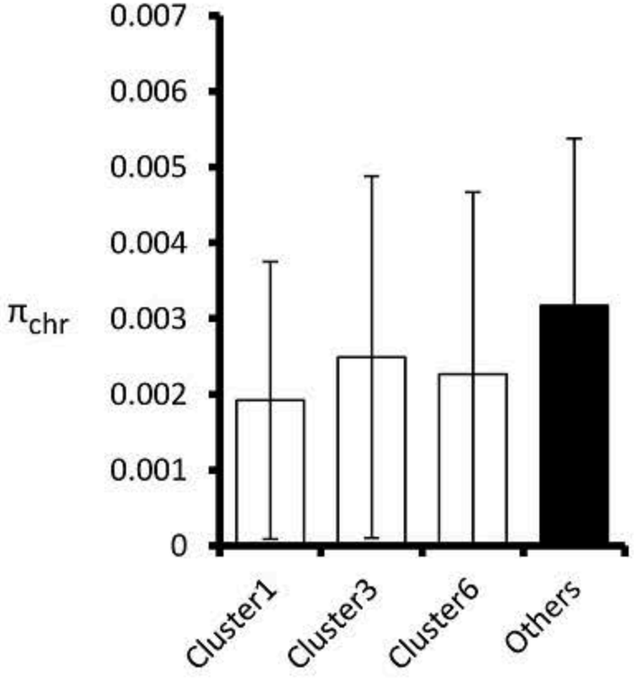

LG2
LG3
LG4
AA
AC
CC
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0
20
40
60
80
100
Distance along LG (cM)

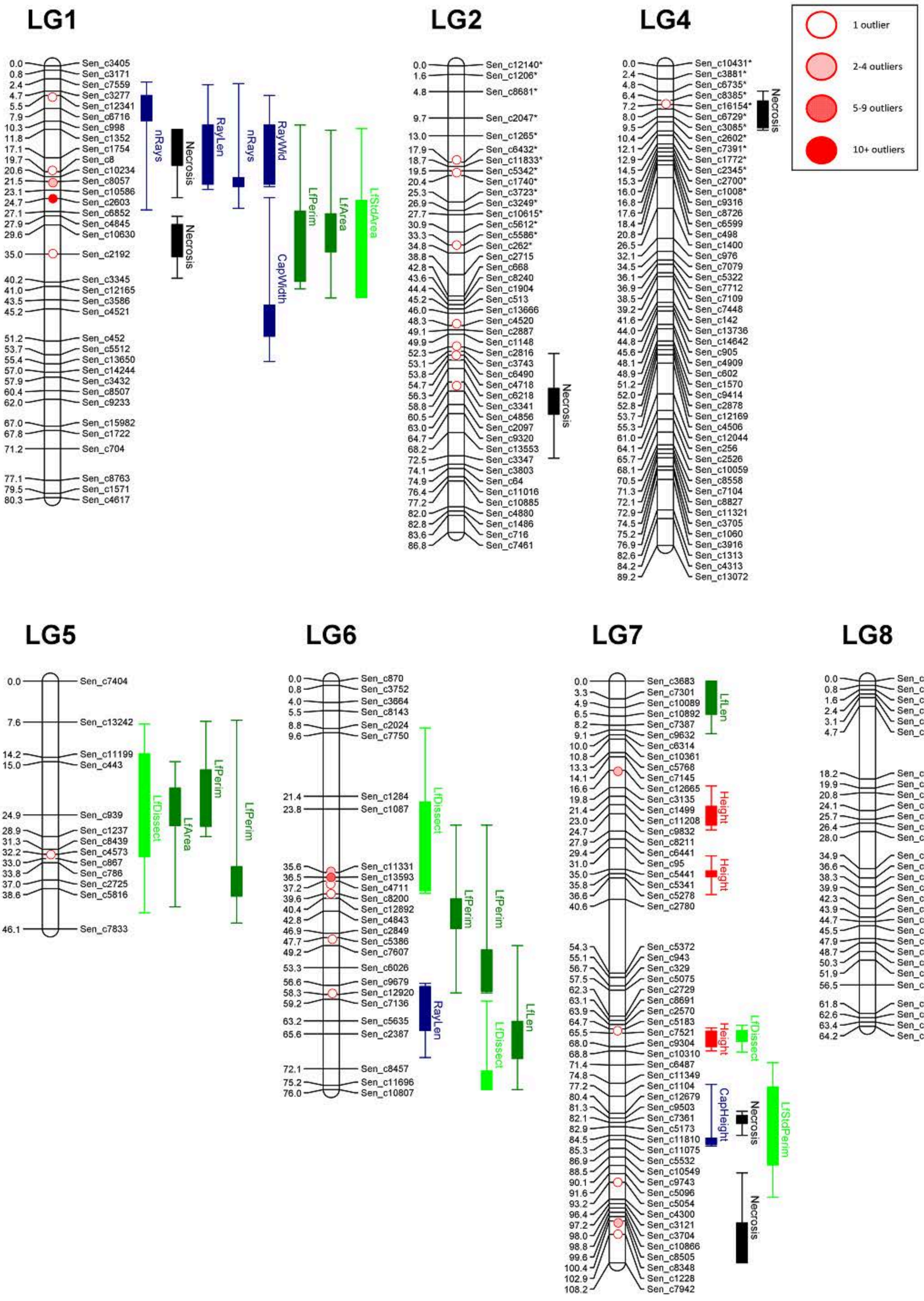

1
2
3
4
5
6
7
8
9
10
Fst
Dxy (x100)
Pi_aet
Pi_chr


F_{ST}
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Cluster1
Cluster3
Cluster6
Others


d_{XY}
0.009
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
Cluster1
Cluster3
Cluster6
Others


π_{aet}
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
Cluster1
Cluster3
Cluster6
Others


π_{chr}
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0
Cluster1
Cluster3
Cluster6
Others


LG1
0.0 Sen_c3405
0.8 Sen_c3171
2.4 Sen_c7559
4.7 Sen_c3277
5.5 Sen_c12341
7.9 Sen_c6716
10.3 Sen_c998
11.8 Sen_c1352
17.1 Sen_c1754
19.7 Sen_c8
20.6 Sen_c10234
21.5 Sen_c8057
23.1 Sen_c10586
24.7 Sen_c2603
27.1 Sen_c6852
27.9 Sen_c4845
29.6 Sen_c10630
35.0 Sen_c2192
40.2 Sen_c3345
41.0 Sen_c12165
43.5 Sen_c3586
45.2 Sen_c4521
51.2 Sen_c452
53.7 Sen_c5512
55.4 Sen_c13650
57.0 Sen_c14244
57.9 Sen_c3432
60.4 Sen_c8507
62.0 Sen_c9233
67.0 Sen_c15982
67.8 Sen_c1722
71.2 Sen_c704
77.1 Sen_c8763
79.5 Sen_c1571
80.3 Sen_c4617
nRays
Necrosis
Necrosis
RayLen
nRays
RayWid
CapWidth
LfPerim
LfArea
LfStdArea
LG2
0.0 Sen_c12140*
1.6 Sen_c1206*
4.8 Sen_c8681*
9.7 Sen_c2047*
13.0 Sen_c1265*
17.9 Sen_c6432*
18.7 Sen_c11833*
19.5 Sen_c5342*
20.4 Sen_c1740*
25.3 Sen_c3723*
26.9 Sen_c3249*
27.7 Sen_c10615*
30.9 Sen_c5612*
33.3 Sen_c5586*
34.8 Sen_c262*
38.8 Sen_c2715
42.8 Sen_c668
43.6 Sen_c8240
44.4 Sen_c1904
45.2 Sen_c513
46.0 Sen_c13666
48.3 Sen_c4520
49.1 Sen_c2887
49.9 Sen_c1148
52.3 Sen_c2816
53.1 Sen_c3743
53.8 Sen_c6490
54.7 Sen_c4718
56.3 Sen_c6218
58.8 Sen_c3341
60.5 Sen_c4856
63.0 Sen_c2097
64.7 Sen_c9320
68.2 Sen_c13553
72.5 Sen_c3347
74.1 Sen_c3803
74.9 Sen_c64
76.4 Sen_c11016
77.2 Sen_c10885
82.0 Sen_c4880
82.8 Sen_c1486
83.6 Sen_c716
86.8 Sen_c7461
Necrosis
LG4
0.0 Sen_c10431*
2.4 Sen_c3881*
4.8 Sen_c6735*
6.4 Sen_c8385*
7.2 Sen_c16154*
8.0 Sen_c6729*
9.5 Sen_c3085*
10.4 Sen_c2602*
12.1 Sen_c7391*
12.9 Sen_c1772*
14.5 Sen_c2345*
15.3 Sen_c2700*
16.0 Sen_c1008*
16.8 Sen_c9316
17.6 Sen_c8726
18.4 Sen_c6599
20.8 Sen_c498
26.5 Sen_c1400
32.1 Sen_c976
34.5 Sen_c7079
36.1 Sen_c5322
36.9 Sen_c7712
38.5 Sen_c7109
39.2 Sen_c7448
41.6 Sen_c142
44.0 Sen_c13736
44.8 Sen_c14642
45.6 Sen_c905
48.1 Sen_c4909
48.9 Sen_c602
51.2 Sen_c1570
52.0 Sen_c9414
52.8 Sen_c2878
53.7 Sen_c12169
55.3 Sen_c4506
61.0 Sen_c12044
64.1 Sen_c256
65.7 Sen_c2526
68.1 Sen_c10059
70.5 Sen_c8558
71.3 Sen_c7104
72.1 Sen_c8827
72.9 Sen_c11321
74.5 Sen_c3705
75.2 Sen_c1060
76.9 Sen_c3916
82.6 Sen_c1313
84.2 Sen_c4313
89.2 Sen_c13072
Necrosis
1 outlier
2-4 outliers
5-9 outliers
10+ outliers
LG5
0.0 Sen_c7404
7.6 Sen_c13242
14.2 Sen_c11199
15.0 Sen_c443
24.9 Sen_c939
28.9 Sen_c1237
31.3 Sen_c8439
32.2 Sen_c4573
33.0 Sen_c867
33.8 Sen_c786
37.0 Sen_c2725
38.6 Sen_c5816
46.1 Sen_c7833
LfDissect
LfArea
LfPerim
LfPerim
LG6
0.0 Sen_c870
0.8 Sen_c3752
4.0 Sen_c3664
5.5 Sen_c8143
8.8 Sen_c2024
9.6 Sen_c7750
21.4 Sen_c1284
23.8 Sen_c1087
35.6 Sen_c11331
36.5 Sen_c13593
37.2 Sen_c4711
39.6 Sen_c8200
40.4 Sen_c12892
42.8 Sen_c4843
46.9 Sen_c2849
47.7 Sen_c5386
49.2 Sen_c7607
53.3 Sen_c6026
56.6 Sen_c9679
58.3 Sen_c12920
59.2 Sen_c7136
63.2 Sen_c5635
65.6 Sen_c2387
72.1 Sen_c8457
75.2 Sen_c11696
76.0 Sen_c10807
LfDissect
LfPerim
RayLen
LfPerim
LfDissect
LfLen
LG7
0.0 Sen_c3683
3.3 Sen_c7301
4.9 Sen_c10089
6.5 Sen_c10892
8.2 Sen_c7387
9.1 Sen_c9632
10.0 Sen_c6314
10.8 Sen_c10361
13.3 Sen_c5768
14.1 Sen_c7145
16.6 Sen_c12865
19.8 Sen_c3135
21.4 Sen_c1499
23.0 Sen_c11208
24.7 Sen_c9832
27.9 Sen_c8211
29.4 Sen_c6441
31.0 Sen_c95
35.0 Sen_c5441
35.8 Sen_c5341
36.6 Sen_c5278
40.6 Sen_c2780
54.3 Sen_c5372
55.1 Sen_c943
56.7 Sen_c329
57.5 Sen_c5075
62.3 Sen_c2729
63.1 Sen_c8691
63.9 Sen_c2570
64.7 Sen_c5183
65.5 Sen_c7521
68.0 Sen_c9304
68.8 Sen_c10310
71.4 Sen_c6487
74.8 Sen_c11349
77.2 Sen_c1104
80.4 Sen_c12679
81.3 Sen_c9503
82.1 Sen_c7361
82.9 Sen_c5173
84.5 Sen_c11810
85.3 Sen_c11075
86.9 Sen_c5532
88.5 Sen_c10549
90.1 Sen_c9743
91.6 Sen_c5096
93.2 Sen_c5054
96.4 Sen_c4300
97.2 Sen_c3121
98.0 Sen_c3704
98.8 Sen_c10866
99.6 Sen_c8505
100.4 Sen_c8348
102.9 Sen_c1228
108.2 Sen_c7942
LfLen
Height
Height
Height
LfDissect
CapHeight
Necrosis
Necrosis
LfStdPerim
LG8
0.0 Sen_c6686
0.8 Sen_c688
1.6 Sen_c10440
2.4 Sen_c5012
3.1 Sen_c1780
4.7 Sen_c6483
18.2 Sen_c2861
19.9 Sen_c4490
20.8 Sen_c3829
24.1 Sen_c4189
25.7 Sen_c5755
26.4 Sen_c2448
28.0 Sen_c1560
34.9 Sen_c11559
36.6 Sen_c2015
38.3 Sen_c14235
39.9 Sen_c4172
42.3 Sen_c9018
43.9 Sen_c957
44.7 Sen_c6632
45.5 Sen_c2599
47.9 Sen_c5745
48.7 Sen_c4360
50.3 Sen_c6787
51.9 Sen_c11089
56.5 Sen_c2677
61.8 Sen_c5740
62.6 Sen_c6326
63.4 Sen_c1730
64.2 Sen_c10162
Necrosis