View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Central Archive at the University of Reading

Genomic sequencing indicates non-random mating of Venturia inaequalis in a mixed cultivar orchard

Article

Accepted Version

Passey, T. A. J., Armitage, A. D., Sobczyk, M. K., Shaw, M. W. and Xu, X. (2020) Genomic sequencing indicates non-random mating of Venturia inaequalis in a mixed cultivar orchard. Plant Pathology, 69 (4). pp. 669-676. ISSN 0032-0862 doi: https://doi.org/10.1111/ppa.13150 Available at http://centaur.reading.ac.uk/88904/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1111/ppa.13150

Publisher: Wiley-Blackwell

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in

University of Reading

the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

1 **Genomic sequencing indicates non-random mating of *Venturia inaequalis* in a**

2 **mixed cultivar orchard**

3

4 Authors: Thomas A. J. Passey[ab]*, Andrew D. Armitage[a], Maria K. Sobczyk[a], Michael W. Shaw[b],

5 Xiangming Xu[a]

6 *[a]NIAB EMR, New Road, East Malling, Kent, ME19 6BJ, UK; [b]School of Agriculture, Policy and*

7 *Development, University of Reading, Reading, RG6 6AR, UK*

8 E-mail: tom.passey@emr.ac.uk (Corresponding author)

9

10 Running head: Non-random mating of apple scab fungus

11

12

13 ## Abstract

14 Apple scab is one of the most economically important diseases of apples worldwide. The
15 disease is caused by the haploid ascomycete *Venturia inaequalis*. Growing apples in cultivar
16 mixtures may reduce disease severity. To determine how the pathogen population structure
17 is affected by host mixtures we studied 24 *V. inaequalis* isolates sampled from three
18 different apple cultivars ("Bramley", "Cox" and "Worcester") growing in a mixed orchard
19 approximately 50 years old. The isolates were aligned against a reference genome and single
20 nucleotide polymorphisms (SNPs) were called between the isolates. The populations isolated
21 from Bramley and Worcester were distinct, while Cox isolates were an admixture. This
22 supports previous tests of the ability of isolates to cross-infect hosts, and molecular
23 comparisons using simple sequence repeats (SSRs). Genotype specific allele (GSA) loci were
24 not distributed randomly across contigs in proportion to contig length, but were clustered.
25 Clustered GSA were observed in almost all contigs. This indicates population differentiation
26 across the whole genome, presumably due to lack of crossing-over events between Bramley
27 and Worcester isolates. This lack is probably due to physical separation effects: sexual
28 mating is more likely to take place and succeed between isolates from lesions on the same
29 leaf than from contact between independently infected leaves in leaf litter on the orchard
30 floor. This would especially be the case if sexual reproduction is initiated before leaf-fall.

31 *Keywords:* Apple scab; apple black spot; ascospore production; super-race; host mixture

32

33

34 ## Introduction

35 The ascomycete *Venturia inaequalis* is the causal agent of apple scab, one of the most
36 important diseases of apples worldwide (MacHardy, 1996). The primary inoculum of the
37 pathogen is predominantly from sexually produced ascospores released from overwintered
38 leaf litter, although, probably depending on climate, some may be from overwintering
39 asexual conidia (Holb *et al.*, 2004, 2005; Passey *et al.*, 2017). If not adequately managed,
40 rounds of secondary infections from conidia can result in large numbers of unmarketable
41 fruit due to unsightly lesions and regular high incidence can lead to premature leaf fall,
42 reduced cumulative growth and very low yields (MacHardy, 1996). Sufficient control to
43 achieve high quality scab-free fruit requires optimum use of numerous fungicide spray
44 rounds through orchard monitoring and disease forecasting systems.

45 An alternative, or supplementary, method of disease control is the use of mixing together
46 cultivars of a crop with differing resistance factors (Wolfe, 1985; Mundt, 2002). The potential
47 for cultivar mixtures to reduce scab development in apple orchards was first assessed by
48 simulation (Blaise & Gessler, 1994)and then tested with a field trial (Bousset *et al.*, 1997).
49 The field trial mixed susceptible and R-gene carrying resistant cultivars and provided
50 evidence that the number of scab lesions per shoot was lower on cv. Golden Delicious and
51 Elstar in mixtures than when these cultivars were in monoculture. The potential of mixtures
52 for scab management was further investigated with a combination of the susceptible cultivar
53 Smoothee and the *Rvi6* R-gene carrying cultivar Baujade (Didelot *et al.*, 2007). Scab leaf
54 incidence was reduced on cv. Smoothee in mixture by 7-21% compared with the mean of
55 monoculture plots, while scab leaf severity was reduced by 15–35%, depending on mixture
56 type and annual epidemic severity. Parisi et al. (2013) investigated scab levels in a mixed
57 orchard of cv. Melrouge, a low susceptibility cultivar, and Pitchounette, a resistant cultivar

58 again carrying the *Rvi6* gene. In 2008 they found 9% of scabbed fruits at harvest in a mixed
59 orchard compared with a mean of 15% in pure stands. In the following year, conditions led
60 to much greater incidence; this was slightly reduced in the mixed orchard, 76% compared
61 with 82% in the pure stands.

62 These studies involved mixing a susceptible cultivar with an R-gene carrying cultivar;
63 however, susceptible cultivars are also known to have differential resistance to apple scab
64 (Sierotzki *et al.*, 1994; Koch *et al.*, 2000). Barbara et al. (2008) showed that isolates of scab
65 sampled from different susceptible cultivars growing within the same orchard could not
66 necessarily infect all other cultivars present. Laboratory crossing between such isolates led
67 to ascospore progenies containing individuals that could infect the whole range of cultivars
68 present in the orchard. Using simple sequence repeat (SSR) markers to look for changes in
69 the *V. inaequalis* populations on the different cultivars indicated that the genetic distances
70 among the populations did not reduce over about a decade, indicating lack of recombination
71 between isolates from different cultivars (Passey *et al.*, 2016). Several hypotheses could
72 explain this. First, sexual reproduction could occur entirely or primarily between isolates
73 infecting the same leaf, because they will be both physically close and having a long time for
74 hyphae to meet and mate. Second, the mating between strains on fallen leaves is possibly
75 more likely among leaves from neighbouring trees in the same row (i.e. same genotypes)
76 than between leave s from different rows.. Third, but less likely, interactions between
77 nearby genetic elements could render recombinants unfit. Finally, chromosomal
78 rearrangements could prevent viable crossing over between certain genotypes, associated
79 by chance with host specialisation.

80 We have obtained further genomic data to confirm this inference of non-random mating.
81 Specifically, we sequenced *V. inaequalis* isolates from different apple cultivars within a single

82 mixed orchard to identify single nucleotide polymorphisms (SNPs) present for subsequent
83 investigation of population differentiation. Of these SNPs, we identified those which had the
84 same allele among isolates from a particular host cultivar, but had a different allele in all
85 isolates from another cultivar; we refer to the alleles at these SNP loci as genotype (cultivar)
86 specific alleles (GSA). We tested whether these GSA loci on each contig were randomly
87 distributed in respect to isolate origin (i.e. cultivars) to infer the nature of mating among
88 isolates from different cultivars.
89

90 ## Materials and methods

91 ### Isolates and DNA extraction

92 Ash Farm, Worcestershire, UK has a 6-row mixed orchard of *Malus* x *domestica* cv. Bramley's
93 Seedling (Bramley), cv. Cox's Orange Pippin (Cox) and cv. Worcester Pearmain (Worcester)
94 on a non-dwarfing rootstock. Each cultivar has two rows with no cultivar being in
95 consecutive rows - Worcester, Cox, Bramley, Cox, Bramley, Worcester. This orchard is ca. 45-
96 50 years old. It has never been sprayed and has not recently been pruned. Scab lesions were
97 sampled from this orchard in 2005 and single spore isolates obtained with no two isolates
98 from the same shoot (Xu *et al.*, 2013). In previous work DNA was extracted from freeze-dried
99 mycelia of single spore isolates for comparison of scab populations on the different cultivars
100 using SSR markers (Passey *et al.*, 2016). The eight isolates from each of the three cultivars
101 with the highest DNA concentrations [quantified and quality-checked using a Nanodrop 1000
102 spectrophotometer (Thermo Scientific)] were selected for Next Generation Sequencing
103 (NGS) on the Illumina MiSeq platform (Supplementary table A).

104 **Library preparation and sequencing**

105 Paired-end genomic libraries were prepared using NEXTflex Rapid DNA-Seq library prep kit
106 Version 14.02 (Bioo Scientific) following the manufacturer's protocol modified by using
107 Illumina adapters rather than the NEXTflex Barcodes. Libraries were validated using a
108 Fragment Analyzer (Advanced Analytical Technologies) confirming a high proportion of
109 library DNA fragments between 600 and 900 bp long. Library concentrations were quantified
110 using a Qubit 2.0 (Invitrogen/Thermo Fisher), standardised to 9 nM before pooling and then
111 diluted to 4 nM (libraries of 5 isolates). Denatured, pooled libraries at 20 pM were
112 sequenced using 300 bp reads on an Illumina MiSeq.

113 **Alignment of MiSeq reads to reference genome and SNP calling**

114 MiSeq reads for all of the isolates were trimmed to remove adaptors and poor quality data
115 from the sequences using fastq-mcf v1.04.636 (Aronesty, 2013). Read depth was calculated
116 per bp using aligned sequence reads using the SAMtools v.1.3.1 depth function, and Median
117 read coverage determined from these values. Alignment of the trimmed reads of the isolates
118 to the reference genome of isolate 05/172 (Passey *et al.*, 2018) was performed with Bowtie2
119 (Langmead & Salzberg, 2012). After removing multimapping and discordant reads from the
120 isolates with SAMtools v.1.3.1, SNPs were called with GATK v.3.6 (Van der Auwera *et al.*,
121 2013) and then filtered to retain only high-quality, biallelic SNPs using the vcffilter function
122 from vcflib and setting the following filtering options: Minimum quality (phred-scaled
123 probability of a SNP occurring at the site) of 40, minimum MQ (phred-scaled probability that
124 the read is mapped to the correct location) of 30, minimum depth of 10 and minimum GQ
125 (phred-scaled probability that the sample genotype being called is correct, given that there is
126 a SNP at that site) of 30. VCFTools (Danecek et al. 2011) was used to remove indels and
127 missing data for genetic analyses.

128 **Determining genetic structure**

129 Previous work comparing isolates from this orchard using AFLP and SSR screening clearly
130 showed a difference between isolates from different cultivars, particularly between Bramley
131 and Worcester (Xu *et al.*, 2013). To confirm this differentiation was true of isolates genome-
132 wide, identity-by-state (IBS) was calculated based on the percentage similarity of shared
133 alleles between samples to produce a SNP matrix, visualised using R as a heatmap, based on
134 all SNPs after removal of missing data but without quality filtering. A neighbour joining (NJ)
135 tree based on 1000 bootstrap replicates was produced using the ape package in R and
136 visualised using Figtree ver.1.4.3, to show unrooted phylogeny of the isolates after both
137 removal of missing data and quality filtering.

138 For isolates originating from the three different cultivars we ran pairwise searches (i.e.
139 isolates from Bramley and Cox; Bramley and Worcester; Cox and Worcester) for those SNPs
140 where isolates from a single host cultivar shared the same allele at a locus, but the allele
141 differed from those in isolates from other populations (i.e. GSAs), using a custom Python
142 script. The number and positions of all SNP loci, GSA loci, GSA loci in genes and
143 nonsynonymous GSA loci in genes were recorded for each contig.

144 **Aggregation of GSA loci**

145 Those remaining GSA loci after removal of missing data and quality filtering were analysed
146 for aggregation. The number of GSA loci between Bramley and Cox, and between Cox and
147 Bramley, was small. Thus subsequent aggregation analysis of GSA loci was only applied to
148 the GSA loci distinguishing Bramley and Worcester isolates.

149 ***Number of GSAs within a contig:*** We wanted to assess whether the GSA loci in the Bramley
150 and Worcester isolates were randomly distributed among and within the contigs, given the
151 number of total SNP loci in each contig.

152 We ran a permutation test to test for aggregation of GSA loci. Specifically, we tested
153 whether the observed variance in the number of GSA loci between contigs was greater than
154 expected under the assumption of random positioning of GSA loci, conditioned on the total
155 number of SNP loci in the Bramley and Worcester isolates in each contig. We excluded
156 contigs with fewer than 100 SNP loci within the Bramley and Worcester isolates from the
157 permutation test. Such contigs are likely to be either highly conserved regions of DNA and
158 therefore unrepresentative, or poorly sequenced (leading to SNPs being removed during
159 filtering).

160 Each permutation consisted of the following steps: (i) the observed number of SNP loci of
161 all types in the first contig were randomly sampled from the entire set of SNP loci in the
162 Bramley and Worcester isolates (initial source of SNP loci), without replacement; (ii) the
163 number of GSA loci in this random sample of SNP loci was counted; (iii) the sampled SNP loci
164 were removed from the initial source of SNP loci (i.e. sampling without replacement) to form
165 the new source for subsequent sampling; (iv) the above three steps were repeated on the
166 next contig until random samples for all contigs had been constructed; (v) finally, variance in
167 the number of GSA loci on each contig was calculated. A total 999 permutations were
168 conducted to generate a frequency distribution of variance in the number of GSA loci
169 expected under the assumption of random distribution of GSA loci among contigs. The
170 observed variance in the number of GSA loci among the contigs was then compared with the
171 distribution of 999 simulated variances to estimate a p-value.

172 ***Distribution of GSA loci within a contig***: If GSA loci have inter-dependent functions -
173 whether to do with host specificity or not - we would expect selection to maintain them as
174 blocks. The selection required would be weaker in relation to a given degree of crossing if
175 the loci were close together (namely crossing-over events among these loci are less

176 frequently, hence more progeny ascospores will inherit the full set of loci). We used a
177 permutation test to assess whether individual GSA loci were randomly distributed within a
178 contig conditioned on the number of SNP loci observed on each contig. Non-GSA SNP loci
179 were coded '0' while GSA loci were coded '1' as for a run test (Sprent & Smeeton, 2007). As
180 before, we excluded contigs with less than 100 SNPs from the permutation analysis. For each
181 permutation, the observed number of GSA loci on each contig was randomly distributed
182 among the positions of all the SNP loci on the contig. Next, the number of consecutive 1s
183 (i.e., GSAs) was calculated for two consecutive 1s up to eleven consecutive 1s. Only non-
184 overlapping consecutive 1s were counted; thus, for instance, '1111' had two counts of '11'
185 and one count each of '111' and '1111', rather than three counts of '11', two counts of '111'
186 and one count of '1111'. This process of redistributing GSAs and counting consecutive GSAs
187 was repeated 1000 times for each contig. The maximum number of two consecutive 1s up to
188 eleven consecutive 1s from the 1000 permutations was calculated and compared to the
189 observed value for each contig.

190

191 Results

192 Genetic structure

193 Isolates 05/036, 05/057 (both isolated from Bramley) and 05/118 (isolated from Cox) had
194 insufficient sequencing coverage and had to be removed from the analyses (Supplementary
195 Table A). Isolates from Bramley grouped separately in neighbour-joining cluster diagrams
196 from isolates from Worcester while isolates from Cox appear to be a mixture, thus
197 supporting previous findings with SSRs (Figures 1 and 2). Due to stringent filtering of SNPs to
198 allow analyses with only the best quality data, the NJ tree (Figure 2) shows some very closely

199 related isolates as identical, however when looking at unfiltered SNPs there are no isolates
200 that have an identical set of SNPs.
201 No GSA loci were observed between populations from Cox (7 samples) and Worcester (8
202 samples), while 160 GSA loci (0.03% of all SNP loci) were found between Cox and Bramley (6
203 samples) populations, and 7168 (1.15% of all SNP loci) between Bramley and Worcester
204 (Table 1). Of the GSA loci between populations from Bramley and Worcester, 3821 (53%)
205 were in the regions of predicted genes, of which 1019 (27%) were nonsynonymous. These
206 proportions are similar to those found among all SNPs: 50% in genes with 28% of these non-
207 synonymous. The breakdown of total SNP and GSA loci for individual contigs is given in
208 Supplementary table B.

209 **Aggregation of GSAs**

210 ***Number of GSAs within a contig between Bramley and Worcester*:** After removal of contigs
211 with less than 100 SNP loci, 99.96% of SNP loci remained, across 92 contigs and covering
212 90.4% of the total genome length. The GSA loci were extremely aggregated in particular
213 contigs (Permutation test $P < 0.001$; Fig. 3).

214 ***Distribution of GSAs within a contig*:** In all contigs with more than 100 SNP loci, the
215 observed number of consecutive GSA loci was much greater than the maximum of the
216 corresponding values in all the 1000 permutations (Supplementary Table C). For example, for
217 contig 47 (Supplementary information D) in 1000 permutations, two consecutive GSAs
218 occurred a maximum 17 times (122 observed), three consecutive GSAs a maximum three
219 times (77 observed) and just once four consecutive GSA loci (53 observed); but the observed
220 contig had a single run of 52 consecutive GSA loci.
221 Across all contigs, the most consecutive GSA loci observed in 1000 permutations was six,
222 on contig 65. In the data, the longest consecutive run of GSAs loci observed was on contig 8,

223 with 121 consecutive GSA loci. Contig 8 was 1.5 Mb long (2.0% of the genome) and has a
224 total of 13,278 SNPs within the combined Bramley and Worcester populations (2.1% of all
225 SNPs across the genome), of which 339 were GSA loci (4.7% of all GSAs across the genome).
226 However, 70% of contig 8 GSA loci were between positions 770024 and 781706, comprising
227 only 0.8% of the total contig length.

228

229 ## Discussion

230 We compared the DNA sequences of isolates of *V. inaequalis* taken from three different
231 cultivars within a single 50 year-old mixed orchard. Isolates from Bramley and Worcester
232 were clearly distinguishable by a large number of SNPs found on many contigs. This supports
233 previous evidence for separation between isolates derived from Bramley and Worcester,
234 based on artificial inoculation studies (Barbara *et al.*, 2008) and molecular comparisons using
235 SSRs (Xu *et al.*, 2013). However, few GSAs separated Cox from either Worcester or Bramley.
236 This also supports previously published evidence. Although the sample sizes in this study
237 were small and therefore some loci with apparent GSA will not actually host GSAs, the
238 permutation tests allow for this, and the lack of GSA between Cox and the other varieties
239 acts as a control for the level of this effect.

240 The mixed orchard where scab isolates were sampled has not received any sanitation or
241 leaf degradation management. Thus, there have been plenty of fallen leaves from all
242 cultivars on which sexual processes of *V. inaequalis* are believed to take place (MacHardy,
243 1996). Any explanation for the observed genetic structure needs to be consistent with the
244 fact that the populations studied here have been in spatial proximity for at least 50 rounds
245 of sexual reproduction. There are two broad classes of explanation. First, host specificity

246 could act as a barrier to gene flow between the populations by selecting for specific alleles
247 and combinations of alleles. Second, the populations on each host may not in practice
248 interbreed with populations from other hosts, preserving both chance founder effects and
249 GSA not necessarily associated with host specificity. This could be for physical or genetic
250 reasons. These mechanisms for intra-species differentiation as observed in the present study
251 are similar to speciation in the *Venturia* genus on different host species (Le Cam *et al.*, 2002;
252 Giraud *et al.*, 2010)**.**

253 We consider first whether host specificity can explain the results. Previously, we showed
254 that fungal population differences in the same mixed orchard decreased over time between
255 Cox and Bramley, but increased in the same period between Bramley and Worcester (Passey
256 et al., 2016). *In vitro* crossing has shown that 1)there are no physiological barriers to sexual
257 recombination between isolates from Worcester and Bramley; and 2) that recombination of
258 virulence factors producing isolates virulent on both Worcester and Bramley did take place
259 (Barbara *et al.*, 2008).

260 GSA distinguishing Worcester and Bramley were spatially aggregated within all contigs.
261 This would be consistent with recombination in the presence of selection for particular
262 alleles at loci required for infection of a particular variety; aggregation would arise from
263 linkage drag around these alleles. However, an over-whelming majority of contigs, covering
264 >90% of the genome length, contain multiple GSA between Bramley and Worcester. If these
265 were all caused by strong selection against hybrids - on both Worcester and Bramley but not
266 Cox - there need to be multiple regions involved in host specialisation on all chromosomes.
267 The nature of the specialisation would also have to be such as to prevent the offspring of
268 Cox-Worcester or Cox-Bramley mating from infecting Cox, so that little or no indirect gene
269 flow occurred between scab populations on Worcester and Bramley. Although all three

270 cultivars are commonly regarded as susceptible to scab, partial resistance evidently exists in
271 all three. However, in the same orchard, seven out of 53 viable ascospores from three
272 crosses between isolates from Cox, Bramley and Worcester could already infect both
273 Bramley and Worcester with no obvious difference in lesion development (Barbara *et al.*,
274 2008). It is hard to reconcile this with a considerable fitness cost across ca. 90% of the whole
275 genome.

276 A small proportion of overwintering inoculum could be conidia (Passey et al., 2017).
277 However, this would not maintain GSA if sexual reproduction were able to move alleles from
278 one variety to another. In addition, it is unlikely that asexual overwintering is restricted to a
279 completely separate sub-group of the *V. inaequalis* population. It is more likely that some
280 individuals overwinter as mycelia due to largely chance factors.

281 A final possibility is that chromosomal rearrangements, such as translocations or
282 inversions, have led to lethality of some recombinants. If several epistatic loci are required
283 for virulence on Worcester, rearrangements holding them together would be favoured. This
284 is hard to test without a fully assembled genome. It would mean that the results here are
285 not general, but depend on specific chromosomal rearrangements in these *V. inaequalis*
286 populations.

287 Taken together, these arguments suggest that rather than being caused by selection in
288 the presence of gene-flow, the genetic structure is consistent with an absence of (or very
289 infrequent) recombination between isolates on different host cultivars. This is possible if
290 sexual mating is more likely to succeed between isolates infecting the same leaf than
291 between isolates from different leaves after leaf-fall. If sex is initiated before leaf-fall, only
292 isolates with similar genetic backgrounds for virulence have the opportunity to mate. Even if
293 sexual mating is initiated on fallen leaves, the chance of successful mating is still likely to be

294 much higher between isolates on the same leaf than isolates on different leaves, since
295 sufficient close contact time to initiate the mating process will be much more likely. Further
296 research is necessary to investigate when mating is initiated relative to leaf fall and the
297 minimum contact time between isolates from different leaves required to initiate the mating
298 process. If the hypothesis of minimal sexual mating among isolates that infected different
299 leaves is correct, there are significant implications for understanding pathogen evolution and
300 practical disease management, such as the timing of end-of-season control measures to
301 reduce sexual reproduction and therefore primary inoculum the following season.

302 The number of GSA loci separating the isolates from Cox and isolates of the other two
303 cultivars was very low. This is consistent with the isolates of Cox being a mixture, with some
304 isolates grouping more closely with isolates from Bramley and some more closely with
305 isolates from Worcester (Fig 2). This hypothesis is supported both by artificial inoculation
306 studies (Barbara *et al.*, 2008) and by population comparisons based on SSR markers (Xu *et*
307 *al.*, 2013; Passey *et al.*, 2016). However, it means that, unless these two sub-populations are
308 separated in some way, isolates virulent on both Bramley and Worcester via the Cox isolates
309 as a mating partner and with the alternate allele at the genome specific allele should arise by
310 recombination, giving rise to "super-race" isolates virulent on all the cultivars in the mixture.
311 It is an important consideration in establishing a mixed cultivar orchard that the cultivars
312 have differing resistance factors. Although we have no evidence for a breakdown in the
313 difference between Bramley and Worcester in this orchard, this may be due to specific
314 factors due to the particular varieties mixed. In particular, the maintenance of separate Cox-
315 Bramley and Cox-Worcester groupings in the Cox population could be explained by
316 chromosomal rearrangements, with the loci conferring virulence lying within an inverted or
317 translocated region, which would give rise to lethal recombinants in crosses heterogeneous

318 for the rearranged regions. The proportion of the genome involved in, or affected by,
319 rearrangements would be extensive but not unprecedented (Raeside *et al.*, 2014; Shi-Kunne
320 *et al.*, 2018; Olarte *et al.*, 2019). As pointed out above, chromosomal rearrangements would,
321 of course, also be another explanation for the aggregation of GSAs. In general, therefore,
322 despite the present results, it would be prudent not to have a potential intermediate
323 cultivar.

324 Conidia are likely to have a role as part of the primary inoculum in some regions (Holb *et*
325 *al.*, 2004, 2005; Passey *et al.*, 2017). If climate of the study orchard does not favour
326 ascospore production and so only asexual clonal races of the pathogen are present, this
327 could also explain the differentiation between Worcester and Bramley, even with an
328 intermediate cultivar, Cox, present. However, this is highly unlikely. Although no
329 pseudothecia have ever been recovered in areas of Israel that lack the necessary lower
330 winter temperatures (Boehm *et al.*, 2003), there is no indication that this happens in
331 orchards in temperate regions and all the evidence is that ascospores are, at least, an
332 important source of inoculum in these areas.

333 If mating among scab isolates within an orchard is not random because the opposite mating
334 types have to be able to infect the same host, this would reduce the rate at which virulence
335 factors against different resistance genes can recombine. This could explain how populations
336 on different cultivars within the same orchard remain different after ca. 50 years of sexual
337 reproduction, though less obviously how the Cox population remains sub-divided.

338 The implementation of orchards with mixed cultivars of differing resistance, shown to
339 decrease levels of apple scab compared to monoculture (Didelot *et al.*, 2007; Parisi *et al.*,
340 2013), is a more attractive option if the risk of "super-races", combining virulence on all
341 components of a mixture, emerging is much less than predicted on the assumption of

342 random mating across cultivars. The risk is further reduced if orchards are replaced after
343 short periods: current commercial apple orchards are replaced after only 15-20 years.

344 In summary, the findings presented here add to the evidence that isolates from specific
345 cultivars within a mixed orchard do not mate at random. We suggest that the most likely
346 explanation is a low frequency of mating between isolates initially infecting different leaves,
347 though other explanations are possible. This separation of sub-populations suggests that the
348 risk of super-races in mixed orchards may be low enough for mixtures that reduce apple scab
349 to remain viable for the lifetime of commercial orchards (15-20 years). A similar conclusion
350 would follow for other pathosystems requiring hyphal mating on living tissue.

351

352 ## Acknowledgements

353 The authors would like to thank Tony Roberts for sample collection and single spore
354 isolations and Dr Richard Harrison for his ideas and advice. This research was funded by the
355 British Department for Environment, Food and Rural Affairs (Defra) and the National
356 Association of Cider Makers (NACM).

357

358 ## Data availability

359 The data that support the findings of this study are available from the corresponding author
360 upon reasonable request.

361

362 ## References

363

364 Aronesty E, 2013. Comparison of Sequencing Utility Programs. *The Open Bioinformatics*
365 *Journal* **7**, 1–8.
366 Van der Auwera GA, Carneiro MO, Hartl C *et al.*, 2013. From FastQ data to high-confidence

367 variant calls: The genome analysis toolkit best practices pipeline. In: *Current Protocols in*
368 *Bioinformatics*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 11.10.1-11.10.33.
369 Barbara DJ, Roberts AL, Xu X-M, 2008. Virulence characteristics of apple scab (Venturia
370 inaequalis) isolates from monoculture and mixed orchards. *Plant Pathology* **57**, 552–
371 561.
372 Blaise P, Gessler C, 1994. Cultivar mixtures in apple orchards as a mean to control apple
373 scab? *Norwegian Journal of Agricultural Sciences* **17**, 105–112.
374 Boehm E, Freeman S, Shabi E, Michailides T, 2003. Microsatellite primers indicate the
375 presence of asexual populations of Venturia inaequalis in coastal Israeli apple orchards.
376 *Phytoparasitica* **31**, 236–251.
377 Bousset L, Blaise P, Kellerhals M, Gessler C, 1997. Mixtures of apple cultivars in orchards:
378 effect on the scab epidemics. *IOBC/WPRS Bulletin* **20**, 42–48.
379 Le Cam B, Parisi L, Arene L, 2002. Evidence of two formae speciales in Venturia inaequalis,
380 responsible for apple and pyracantha scab. *Phytopathology* **92**, 314–20.
381 Didelot F, Brun L, Parisi L, 2007. Effects of cultivar mixtures on scab control in apple
382 orchards. *Plant Pathology* **56**, 1014–1022.
383 Giraud T, Gladieux P, Gavrilets S, 2010. Linking the emergence of fungal plant diseases with
384 ecological speciation. *Trends in ecology & evolution* **25**, 387–95.
385 Holb I, Heijne B, Jeger M, 2004. Overwintering of conidia of Venturia inaequalis and the
386 contribution to early epidemics of apple scab. *Plant Disease* **88**, 751–757.
387 Holb IJ, Heijne B, Jeger MJ, 2005. The widespread occurrence of overwintered conidial
388 inoculum of Venturia inaequalis on shoots and buds in organic and integrated apple
389 orchards across the Netherlands. *European Journal of Plant Pathology* **111**, 157–168.
390 Koch T, Kellerhals M, Gessler C, 2000. Virulence pattern of Venturia inaequalis field isolates
391 and corresponding differential resistance in Malus x domestica. *Journal of*
392 *Phytopathology* **148**, 357–364.
393 Langmead B, Salzberg SL, 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*
394 **9**, 357–359.
395 MacHardy WE, 1996. *Apple Scab: Biology, Epidemiology, and Management*. St Paul, MN:
396 American Phytopathological Society.
397 Mundt CC, 2002. Use of multiline cultivars and cultivar mixtures for disease management.
398 *Annual Review of Phytopathology* **40**, 381–410.
399 Olarte RA, Menke J, Zhang Y *et al.*, 2019. Chromosome rearrangements shape the
400 diversification of secondary metabolism in the cyclosporin producing fungus
401 Tolypocladium inflatum. *BMC Genomics* **20**, 120.
402 Parisi L, Gros C, Combe F, Parveaud C-E, Gomez C, Brun L, 2013. Impact of a cultivar mixture
403 on scab, powdery mildew and rosy aphid in an organic apple orchard. *Crop Protection*
404 **43**, 207–212.
405 Passey TAJ, Armitage AD, Xu X, 2018. Annotated Draft Genome Sequence of the Apple Scab
406 Pathogen Venturia inaequalis. *Microbiology Resource Announcements* **7**.
407 Passey TAJ, Robinson JD, Shaw MW, Xu X-M, 2017. The relative importance of conidia and
408 ascospores as primary inoculum of Venturia inaequalis in a southeast England orchard.
409 *Plant Pathology*.
410 Passey TAJ, Shaw MW, Xu X-M, 2016. Differentiation in populations of the apple scab fungus
411 Venturia inaequalis on cultivars in a mixed orchard remain over time. *Plant Pathology*

412 **65**, 1133–1141.
413 Raeside C, Gaffé J, Deatherage DE *et al.*, 2014. Large chromosomal rearrangements during a
414 long-term evolution experiment with Escherichia coli. *mBio* **5**, e01377-14.
415 Shi-Kunne X, Faino L, van den Berg GCM, Thomma BPHJ, Seidl MF, 2018. Evolution within the
416 fungal genus *Verticillium* is characterized by chromosomal rearrangement and gene
417 loss. *Environmental Microbiology* **20**, 1362–1373.
418 Sierotzki H, Eggenschwiler M, Boillat O, McDermott JM, Gessler C, 1994. Detection of
419 variation in virulence toward susceptible apple cultivars in natural populations of
420 Venturia inaequalis. *Phytopathology* **84**, 1005–1009.
421 Sprent P, Smeeton NC, 2007. *Applied nonparametric statistical methods.* Chapman &
422 Hall/CRC.
423 Wolfe MS, 1985. The current status and prospects of multiline cultivars and variety mixtures
424 for disease resistance. *Annual Review of Phytopathology* **23**, 251–273.
425 Xu X, Harvey N, Roberts A, Barbara D, 2013. Population variation of apple scab (Venturia
426 inaequalis) within mixed orchards in the UK. *European Journal of Plant Pathology* **135**,
427 97–104.
428

429 **Supplementary table A.** *Venturia inaequalis* isolates from Ash Farm, UK, a mixed cultivar
430 orchard of Bramley, Cox and Worcester. Genomes of all isolates were sequenced on the
431 Illumina platform.
432 **Supplementary table B.** The whole genome sequence (WGS) of 14 *V. inaequalis* isolates (6
433 from Bramley, 8 from Worcester) were aligned to the 05/172 reference genome (one of the
434 8 Worcester isolates). The total number of SNPs between the 14 isolates were called and
435 then the number of genotype specific polymorphic alleles (GSAs) (i.e. isolates from Bramley
436 shared the same allele at a locus but this differed to that shared among the Worcester
437 isolates) in the whole genome, those just in genes and those nonsynonymous.
438 **Supplementary table C.** Number in brackets is the maximum number of consecutive
439 genotype specific allelic (GSA) loci in a contig from 1000 permutations of the allocation of
440 GSA to positions occupied by a single nucleotide polymorphism. A lack of number on top
441 row for each contig indicates there was no occurrence in the permutation for that number
442 of consecutive GSA loci. The main entry (second row for each contig) is the observed
443 number of consecutive GSA on each contig. Only contigs with any GSA loci presented.
444 **Supplementary information D.** An example input data of genotype specific allele (GSA) (1)
445 and non-GSA SNPs (0) of contig 47 for GSA redistribution analysis of GSAs within a contig.
446 Each number represent a SNP within the 14 isolates (6 from Bramley, 8 from Worcester)
447 sequenced and aligned to the assembled genome of *Venturia inaequalis* isolate 05/172.
448

449 **Figure 1.** Heatmap to represent clustering of 21 *V. inaequalis* isolates from three different
450 apple cultivars, Bramley (B), Cox (C) and Worcester (W), present in the same orchard. Data
451 from identity-by-state (IBS) calculated on the percentage similarity of shared alleles
452 between samples to produce a SNP matrix (the darker the shading the more alike the
453 isolates). Due to rounding percentage values, several isolates showed "100%" IBS;
454 examining unfiltered SNPs showed that there are no isolates that have an identical set of
455 SNPs.

456 **Figure 2.** A Neighbour joining (NJ) tree, in polar format, showing clustering of *Venturia*
457 *inaequalis* isolates from three different apple cultivars, Bramley, Cox and Worcester,
458 present in the same orchard.

459 **Figure 3.** Plots of standard deviation of (A) the number of Bramley-Worcester genotype
460 specific allele (GSA) loci and (B) the number of Bramley-Worcester GSA as the percentage of
461 total number of Cox-Bramley-Worcester GSA on each contig under the assumption of
462 random distribution of all the GSA loci across all contigs. The 1000 values consisted of 999
463 from permutations and the observed values.

Supplementary table A. *Venturia inaequalis* isolates from Ash Farm, UK, a mixed cultivar orchard of Bramley, Cox and Worcester. Genomes of all isolates were sequenced on the Illumina MiSeq platform

Isolate ID	Host cultivar	MiSeq run	Amount of data from MiSeq run (bp)	Coverage[c]
05/007	Bramley	2	13136208	31
05/024	Bramley	4	12453319	33
05/025	Bramley	3	14340023	33
05/030	Bramley	4	11430665	28
05/036	Bramley	5	2029822	3[d]
05/044	Bramley	2	19637369	45
05/049	Worcester[b]	1	13700931	39
05/057	Bramley	1	13336096	10[d]
05/083	Cox	2	12003095	29

05/096	Cox	3	13063352	4631
05/097	Cox	5	12123271	27
05/098	Cox	1	17982080	52
05/101	Cox	4	13983601	37
05/106	Cox	4	9259589	24
05/118	Cox	3	11758642	9[d]
05/119	Cox	1	14438772	38
05/172[a]	Worcester	2	12672586	32
05/173	Worcester	5	19115153	48
05/182	Worcester	2	15752464	40
05/190	Worcester	5	13003150	33
05/196	Worcester	3	11826771	33
05/197	Worcester	4	13898993	36
05/199	Bramley[b]	1	21335557	54
05/202	Worcester	3	14365407	37

[a]The genome of isolate 05/172 was also sequenced by the PacBio platform and used to assemble the reference genome used in this publication.
[b]Isolates were mislabelled during library prep; this was picked up and checked during data analysis and as such the isolate codes for this publication have been switched; i.e. the isolate in 05/199 in this publication is actually the isolate 05/049 from Bramley and *vice versa*.
[c]Coverage showing median read coverage per bp.
[d]Because of low coverage, these isolates were in population analysis.

Supplementary table B. The whole genome sequence (WGS) of 14 *V. inaequalis* isolates (6 from Bramley, 8 from Worcester), sequenced on the Illumina Miseq platform, were aligned to the 05/172 reference genome (one of the 8 Worcester isolates). The total number of SNPs between the 14 isolates were called and then the number of genotype specific polymorphic alleles (GSAs) (i.e. isolates from Bramley shared the same allele at a locus but this differed to that shared among the Worcester isolates) in the whole genome, those just in genes and those nonsynonymous

Contig	Length (bp)	Total SNPs in B/W populations	Total GSA loci in genome	GSA loci in genes	Genotype specific nonsynonymous polymorphic loci in genes
1	3847617	39750	88	24	6
2	2883036	35428	367	162	53
3	2469270	25959	294	156	39
4	1643167	18383	88	31	10
5	1572910	16462	483	204	55
6	1553562	13901	48	21	6
7	1545189	15053	290	158	27
8	1540187	13278	339	307	79
9	1520579	11419	249	120	36
10	1471990	11801	18	4	0
11	1469107	18420	209	105	29
12	1466925	16430	79	32	12

13	1444683	16911	254	90	29
14	1434827	7843	47	11	7
15	1433712	13375	42	24	6
16	1432488	15795	149	73	19
17	1345551	15547	176	116	27
18	1224983	11025	22	11	5
19	1201024	13765	133	52	13
20	1189902	12346	529	259	55
21	989026	8262	12	5	3
22	960501	13438	94	48	18
23	953805	6543	61	40	3
24	887866	7589	2	1	1
25	878632	7455	87	38	5
26	877845	11043	11	0	0
27	830644	7224	45	30	4
28	802546	8038	59	48	7
29	775226	5108	49	32	3
30	757030	5133	1	0	0
31	757014	4206	42	13	6
32	742314	3749	62	24	4
33	713553	9208	46	5	4
34	702798	6598	10	4	3
35	667479	8316	14	11	3
36	664565	4321	0	0	0
37	642819	6378	170	94	17
38	633838	4775	3	0	0
39	631560	8285	399	158	36
40	616300	5535	8	0	0
41	615051	4108	0	0	0
42	612523	2330	0	0	0
43	606529	8417	25	13	8
44	599691	7232	81	45	25
45	572125	7163	237	171	62
46	570550	7497	257	151	28
47	548817	8366	270	193	33
48	544245	5585	231	141	34
49	531805	7819	42	36	9
50	510422	1584	0	0	0
51	486582	2049	2	1	1
52	483498	4553	1	0	0
53	467987	1618	0	0	0
54	448816	6344	175	98	48
55	403330	2362	1	0	0
56	400139	3177	18	11	5
57	386935	1347	1	0	0
58	359892	4218	107	61	20

59	347592	2998	1	0	0
60	334733	2927	0	0	0
61	334267	2826	1	0	0
62	316014	1835	0	0	0
63	307156	4892	4	0	0
64	259082	2355	1	0	0
65	255655	2879	235	152	37
66	249385	2827	2	1	1
67	246485	1439	77	37	13
68	243293	2172	112	93	17
70	239636	796	0	0	0
71	239618	1123	2	1	1
72	238937	2896	53	16	1
73	236519	2721	0	0	0
74	227977	1006	2	2	2
75	223988	1210	0	0	0
76	222240	1059	29	15	11
77	205344	1583	2	0	0
79	194393	1952	0	0	0
80	193792	10	0	0	0
81	190172	79	0	0	0
82	189157	1573	4	4	0
83	188550	1504	21	6	4
85	173860	1172	0	0	0
86	165300	764	0	0	0
88	156979	1207	1	0	0
89	148886	320	0	0	0
90	145307	1200	1	0	0
91	144156	783	6	6	5
92	144025	1500	69	44	12
93	142133	584	0	0	0
94	140254	568	0	0	0
96	133811	28	0	0	0
97	123707	1353	1	1	1
99	110078	497	0	0	0
101	98753	4	0	0	0
102	89478	278	1	1	1
120	60568	634	16	10	10
144	44087	12	0	0	0
145	43653	1	0	0	0
166	34764	21	0	0	0
183	29119	86	0	0	0
190	28319	12	0	0	0
Remaining contigs[a]	6144191	0	0	0	0
Total	72310420	625550	7168	3821	1019

[a]Contigs containing no SNPs, equating to 8.5% of the genome

465

Supplementary table C. The top row for each contig is the maximum number of consecutive genotype specific allelic (GSA) loci in a contig from 1000 permutations of the allocation of GSA to positions occupied by a single nucleotide polymorphism; the main entry in bold (second row for each contig) is the observed number of consecutive GSA on each contig. Only contigs with any GSA loci presented

Contig number	No. of GSAs	Number of consecutive Genotype Specific Alleles 2	3	4	5	6	7	8	9	10	11
1	88	3	1	0	0	0	0	0	0	0	0
		29	***15***	***9***	***7***	***5***	***4***	***2***	***2***	***2***	***2***
2	367	11	1	1	0	0	0	0	0	0	0
		108	***54***	***29***	***15***	***9***	***8***	***5***	***5***	***3***	***2***
3	294	12	1	0	0	0	0	0	0	0	0
		111	***60***	***40***	***29***	***18***	***16***	***12***	***10***	***8***	***7***
4	88	4	1	0	0	0	0	0	0	0	0
		33	***15***	***12***	***5***	***3***	***3***	***3***	***1***	***1***	***1***
5	483	26	4	1	0	0	0	0	0	0	0
		164	***87***	***56***	***38***	***27***	***21***	***18***	***15***	***14***	***12***
6	48	3	1	0	0	0	0	0	0	0	0
		12	***6***	***4***	***4***	***1***	***1***	***1***	***1***	***1***	***1***
7	290	15	2	1	0	0	0	0	0	0	0
		98	***52***	***32***	***26***	***20***	***13***	***10***	***8***	***7***	***6***
8	339	19	2	1	1	0	0	0	0	0	0
		137	***80***	***57***	***44***	***34***	***27***	***24***	***19***	***16***	***13***
9	249	13	2	1	0	0	0	0	0	0	0
		87	***44***	***28***	***17***	***14***	***10***	***9***	***6***	***4***	***4***
10	18	1	0	0	0	0	0	0	0	0	0
		5	***2***	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
11	209	8	2	0	0	0	0	0	0	0	0
		72	***37***	***22***	***12***	***9***	***5***	***4***	***3***	***3***	***3***
12	79	4	0	0	0	0	0	0	0	0	0
		25	***14***	***6***	***4***	***3***	***3***	***1***	***1***	***1***	***1***
13	254	11	2	0	0	0	0	0	0	0	0
		90	***49***	***31***	***20***	***16***	***13***	***11***	***9***	***7***	***6***
14	47	3	0	0	0	0	0	0	0	0	0
		9	***6***	***2***	***1***	***0***	***0***	***0***	***0***	***0***	***0***
15	42	2	0	0	0	0	0	0	0	0	0
		17	***9***	***5***	***3***	***2***	***2***	***1***	***1***	***1***	***1***
16	149	5	2	0	0	0	0	0	0	0	0
		60	***38***	***27***	***20***	***18***	***14***	***12***	***11***	***8***	***7***
17	176	7	2	0	0	0	0	0	0	0	0
		60	***33***	***23***	***13***	***12***	***8***	***8***	***7***	***5***	***4***
18	22	2	0	0	0	0	0	0	0	0	0
		6	***3***	***1***	***1***	***1***	***1***	***0***	***0***	***0***	***0***

19	133	5	1	0	0	0	0	0	0	0	0
		44	***19***	***14***	***8***	***6***	***4***	***4***	***3***	***2***	***1***
20	529	37	6	1	1	0	0	0	0	0	0
		189	***90***	***55***	***35***	***26***	***18***	***13***	***10***	***9***	***7***
21	12	1	0	0	0	0	0	0	0	0	0
		2	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
22	94	5	1	0	0	0	0	0	0	0	0
		25	***11***	***8***	***4***	***2***	***2***	***2***	***1***	***1***	***1***
23	61	3	1	0	0	0	0	0	0	0	0
		20	***11***	***6***	***3***	***3***	***2***	***2***	***2***	***1***	***1***
24	2	0	0	0	0	0	0	0	0	0	0
		0	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
25	87	5	1	1	1	0	0	0	0	0	0
		32	***18***	***12***	***9***	***6***	***5***	***4***	***4***	***3***	***2***
26	11	1	0	0	0	0	0	0	0	0	0
		3	***1***	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
27	45	4	0	0	0	0	0	0	0	0	0
		17	***8***	***6***	***4***	***1***	***1***	***1***	***0***	***0***	***0***
28	59	4	1	0	0	0	0	0	0	0	0
		23	***12***	***8***	***6***	***4***	***2***	***2***	***2***	***2***	***2***
29	49	3	1	0	0	0	0	0	0	0	0
		17	***8***	***7***	***4***	***2***	***2***	***2***	***1***	***1***	***10***
31	42	4	1	0	0	0	0	0	0	0	0
		10	***3***	***1***	***1***	***1***	***1***	***0***	***0***	***0***	***0***
32	62	5	1	0	0	0	0	0	0	0	0
		25	***12***	***7***	***6***	***5***	***2***	***0***	***0***	***0***	***0***
33	46	3	1	0	0	0	0	0	0	0	0
		19	***12***	***9***	***6***	***4***	***3***	***3***	***2***	***2***	***2***
34	10	1	0	0	0	0	0	0	0	0	0
		2	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
35	14	1	0	0	0	0	0	0	0	0	0
		2	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
37	170	13	3	1	0	0	0	0	0	0	0
		54	***25***	***14***	***6***	***5***	***3***	***3***	***1***	***0***	***0***
38	3	0	0	0	0	0	0	0	0	0	0
		1	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
39	399	30	4	2	1	0	0	0	0	0	0
		148	***88***	***54***	***35***	***27***	***21***	***17***	***15***	***13***	***9***
40	8	1	0	0	0	0	0	0	0	0	0
		3	***2***	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***0***
43	25	2	0	0	0	0	0	0	0	0	0
		8	***3***	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
44	81	5	1	0	0	0	0	0	0	0	0
		24	***11***	***4***	***3***	***2***	***2***	***1***	***1***	***1***	***0***
45	237	15	3	1	1	0	0	0	0	0	0
		99	***59***	***43***	***32***	***24***	***20***	***18***	***13***	***12***	***10***

46	257	18	3	1	0	0	0	0	0	0	0
		103	***58***	***41***	***26***	***21***	***18***	***14***	***10***	***9***	***9***
47	270	17	3	1	0	0	0	0	0	0	0
		122	***77***	***53***	***39***	***33***	***24***	***21***	***19***	***16***	***11***
48	231	20	5	1	1	0	0	0	0	0	0
		87	***47***	***30***	***23***	***18***	***12***	***7***	***5***	***5***	***4***
49	42	2	1	0	0	0	0	0	0	0	0
		8	***3***	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
51	2	1	0	0	0	0	0	0	0	0	0
		1	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
54	175	14	3	1	0	0	0	0	0	0	0
		70	***45***	***28***	***21***	***16***	***14***	***12***	***11***	***8***	***7***
55	1	0	0	0	0	0	0	0	0	0	0
		0	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
56	18	2	1	0	0	0	0	0	0	0	0
		2	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
58	107	9	2	1	0	0	0	0	0	0	0
		47	***30***	***21***	***16***	***13***	***10***	***9***	***6***	***6***	***6***
63	4	1	0	0	0	0	0	0	0	0	0
		1	***1***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
65	235	31	7	3	1	1	0	0	0	0	0
		65	***24***	***15***	***10***	***4***	***3***	***3***	***1***	***1***	***1***
66	2	0	0	0	0	0	0	0	0	0	0
		0	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
67	77	10	2	1	0	0	0	0	0	0	0
		17	***4***	***2***	***2***	***1***	***1***	***0***	***0***	***0***	***0***
68	112	13	3	1	1	0	0	0	0	0	0
		51	***30***	***21***	***16***	***12***	***9***	***8***	***6***	***5***	***4***
71	2	0	0	0	0	0	0	0	0	0	0
		0	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
72	53	6	1	1	0	0	0	0	0	0	0
		15	***6***	***4***	***2***	***2***	***0***	***0***	***0***	***0***	***0***
74	2	1	0	0	0	0	0	0	0	0	0
		0	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
76	29	4	1	1	0	0	0	0	0	0	0
		13	***9***	***6***	***4***	***4***	***2***	***2***	***2***	***1***	***1***
77	2	1	0	0	0	0	0	0	0	0	0
		1	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
82	4	1	0	0	0	0	0	0	0	0	0
		0	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
83	21	3	1	0	0	0	0	0	0	0	0
		3	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***	***0***
91	6	1	0	0	0	0	0	0	0	0	0
		2	***1***	***1***	***1***	***0***	***0***	***0***	***0***	***0***	***0***
92	69	9	2	1	0	0	0	0	0	0	0
		24	***13***	***7***	***4***	***4***	***3***	***3***	***3***	***1***	***1***

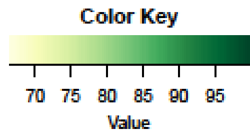
120	16	3	1	0	0	0	0	0	0	0	0
		4	***2***	***1***	***1***	***0***	***0***	***0***	***0***	***0***	***0***

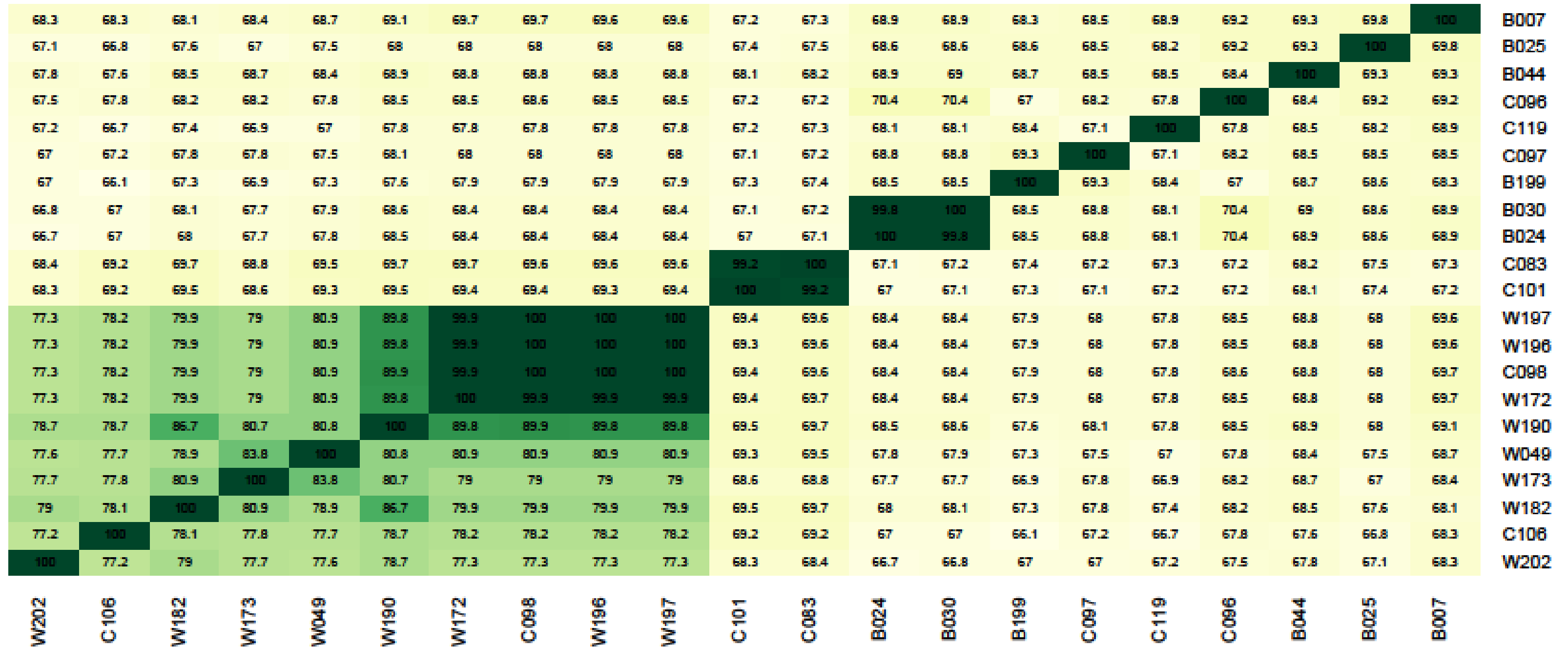
Supplementary information D. An example input data of genotype specific allele (GSA) (1) and non-GSA SNPs (0) of contig 47 for GSA redistribution analysis of GSAs within a contig. Each number represent a SNP within the 14 isolates (6 from Bramley, 8 from Worcester) sequenced and aligned to the assembled genome of *Venturia inaequalis* isolate 05/172.

```
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000001111111111000001111110000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000010100110000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000110010111100000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000111000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000111010111111111111111111111
0011101111111111111111111101111111111111011111101111111111111111111111111111110001111111111111111111
```

517 01111111111111111111000000000000000000000000000000000000000000000000000000000000000000000
518 00000000000000001100000000000000000000000000000000000000000000000000000000000000000000000
519 00000001001000000000000000000000000000000000000000000000000000000000000000000000000000100
520 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000011
521 11001001000100011100000001000000000000000000000000000000000000000000000000000000000000000
522 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
523 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
524 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
525 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
526 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
527 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
528 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
529 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
530 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
531 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
532 00000000000000000000000000000000011111110000000000000111111110000000000000111111111111111
533 11111111111111111111111111111111111101110110111111111111111111110000000000000000000000000
534 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
535 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
536 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
537 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
538 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
539 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
540 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
541 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
542 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
543 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
544 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
545 00000000000000000000000000000000000000000000000000000000000000000000000000000100000000000
546 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
547 00000000000000000000000000000000000000001000000011110010000000000000000000000000010000000
548 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
549 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
550 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
551 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
552 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
553 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
554 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
555 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
556 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
557 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
558 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
559 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
560 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
561 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
562 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
563 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
564 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
565 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
566 00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
567
568
569
570
571

572





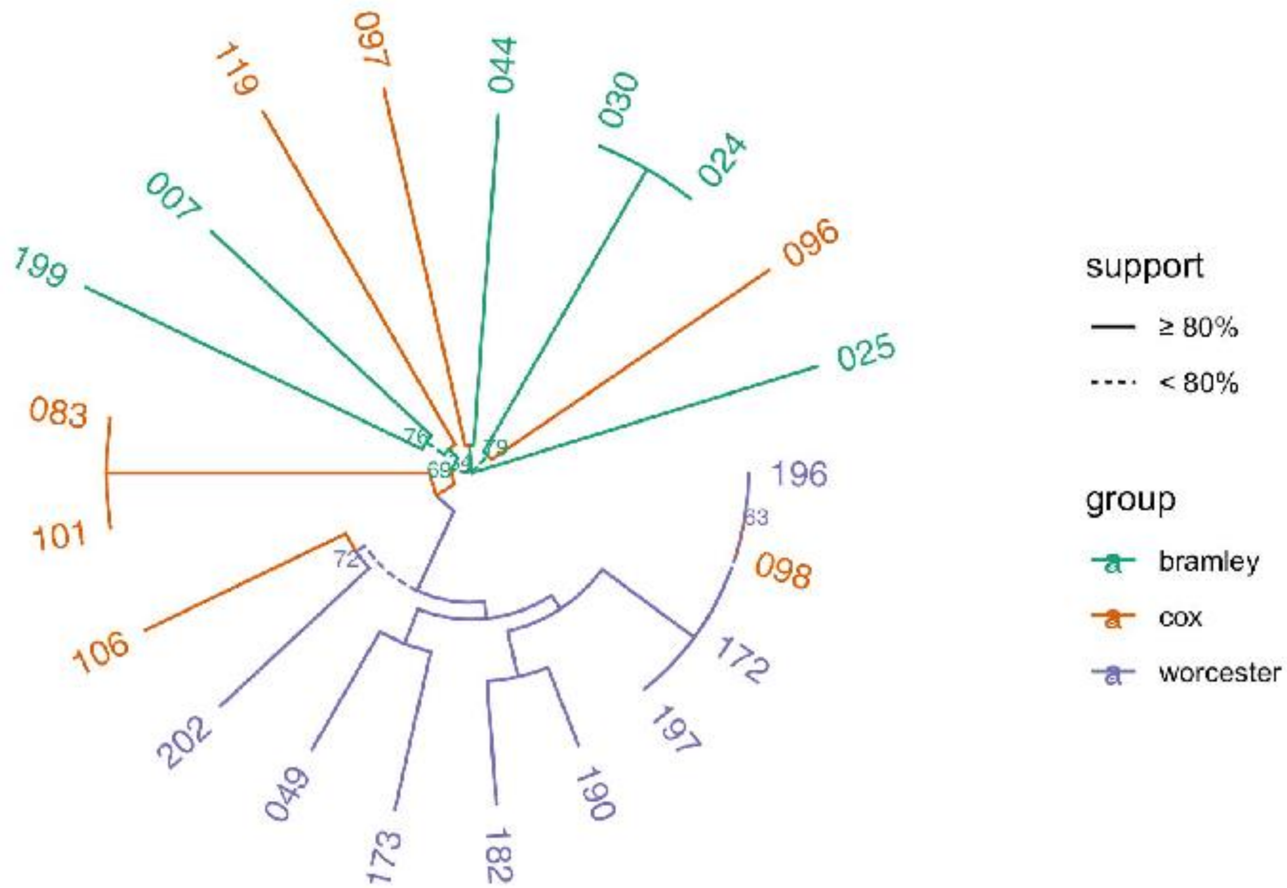
573

574 **Figure 1.** Heatmap to represent clustering of 21 *V. inaequalis* isolates from three different apple cultivars, Bramley (B), Cox (C) and Worcester

575 (W), present in the same orchard. Data from identity-by-state (IBS) calculated on the percentage similarity of shared alleles between samples

576 to produce a SNP matrix (the darker the shading the more alike the isolates). Due to rounding percentage values, several isolates showed

577 "100%" IBS; examining unfiltered SNPs showed that there are no isolates that have an identical set of SNPs.



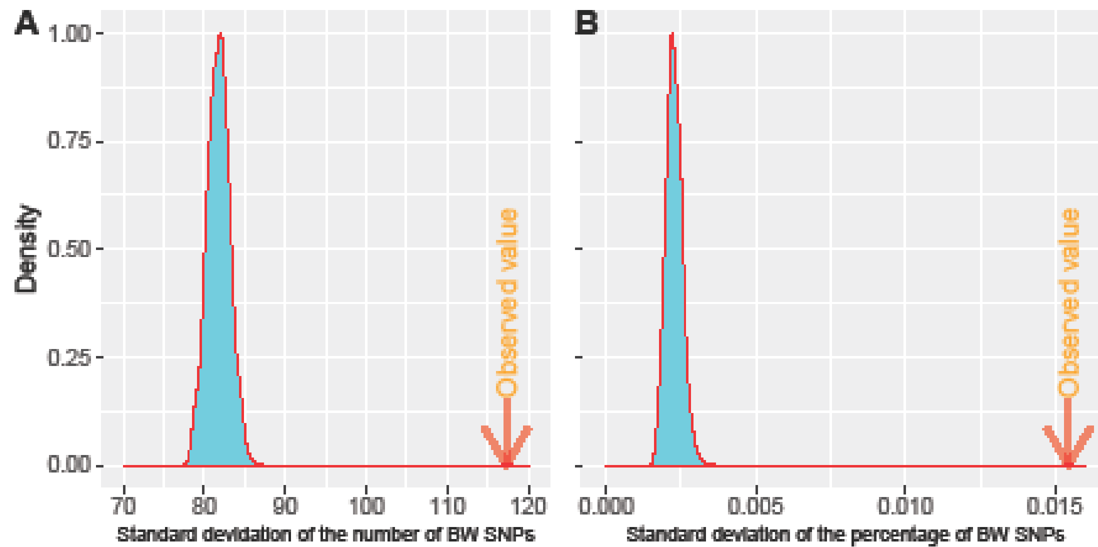
578

579 **Figure 2.** A Neighbour joining (NJ) tree, in polar format, showing clustering of *Venturia inaequalis* isolates from three different apple cultivars,

580 Bramley, Cox and Worcester, present in the same orchard.

581

582



583 **Figure 3.** Plots of standard deviation of (A) the number of Bramley-Worcester genotype specific allele (GSA) loci and (B) the number of

584 Bramley-Worcester GSA as the percentage of total number of Cox-Bramley-Worcester GSA on each contig under the assumption of random

585 distribution of all the GSA loci across all contigs. The 1000 values consisted of 999 from permutations and the observed values.