1 Article (Methods)

2

3 # Reconstruction of haplotype-blocks selected during experimental

4 evolution

5

6 Susanne U. Franssen[1], Nicholas H. Barton[2], Christian Schlötterer*[1]

7

8 [1]Institut für Populationsgenetik, Vetmeduni Vienna, Vienna, Austria

9 [2]Institute of Science and Technology Austria (IST Austria), Klosterneuburg, Austria

10

11 * Corresponding author

12 Email: christian.schloetterer@vetmeduni.ac.at (CS)

13

14

15

16 Key words:

17 haplotype reconstruction; sequencing of pooled individuals (Pool-Seq); experimental

18 evolution; replicated time series data; selection; Evolve and Resequence (E&R)

19

Abstract

The genetic analysis of experimentally evolving populations typically relies on short reads from pooled individuals (Pool-Seq). While this method provides reliable allele frequency estimates, the underlying haplotype structure remains poorly characterized. With small population sizes and adaptive variants that start from low frequencies, the interpretation of selection signatures in most evolve and resequencing studies remains challenging. To facilitate the characterization of selection targets, we propose a new approach that reconstructs selected haplotypes from replicated time series, using Pool-Seq data. We identify selected haplotypes through the correlated frequencies of alleles carried by them. Computer simulations indicate that selected haplotype-blocks of several Mb can be reconstructed with high confidence and low error rates, even when allele frequencies change only by 20% across three replicates. Applying this method to real data from *D. melanogaster* populations adapting to a hot environment, we identify a selected haplotype-block of 6.93 Mb. We confirm the presence of this haplotype-block in evolved populations by experimental haplotyping, demonstrating the power and accuracy of our haplotype reconstruction from Pool-Seq data. We propose that the combination of allele frequency estimates with haplotype information will provide the key to understanding the dynamics of adaptive alleles.

Introduction

The combination of experimental evolution with whole genome resequencing, called "Evolve and Resequence" (E&R) (Turner et al. 2011), is a widely used approach to study genotype-phenotype maps or to characterize adaptive variation (Long et al. 2015; Schlötterer et al. 2015). To date, E&R has been used for diverse taxonomic groups ranging from viruses (Sanjuán et al. 2005), through bacteria (Masri et al. 2015) and yeast (Lang and Desai 2014) to higher eukaryotic systems (*reviewed in* Schlötterer et al. 2015).

48 In microbes, experimental evolution typically starts from a single genotype and
49 adaptation is based on novel mutations. In contrast, multicellular experimental evolution
50 studies use substantially smaller population sizes; adaptation occurs from standing
51 variation, which is shuffled during the experiment by recombination (Teotónio et al.
52 2009; Burke et al. 2010; Burke et al. 2014; Sheng et al. 2015).
53
54 The emerging picture from eukaryotic E&R studies (e.g. Turner et al. 2011; Orozco-
55 terWengel et al. 2012; Turner and Miller 2012; Remolina et al. 2012; Tobler et al. 2014;
56 Martins et al. 2014; Reed et al. 2014; Franssen et al. 2015; Jha et al. 2015) is that very
57 many putative candidate loci are identified, which have more pronounced allele
58 frequency changes than expected under neutrality. Nevertheless, the candidates are too
59 numerous to be direct targets of selection (Nuzhdin and Turner 2013). Consequently the
60 localization of causative variants has been difficult (Schlötterer et al. 2015), with the
61 exception of a few studies, where the underlying trait architecture suggests a simple
62 genetic basis (e.g. Martins et al. 2014). Several computer simulation studies have
63 explored possible improvements to the experimental design of E&R and concluded that
64 the key to reliable identification of targets of selection is to reduce linkage disequilibrium
65 (LD). This can be achieved by several means, including more founding haplotypes, larger
66 population sizes, more generations, and more replicates (Kofler and Schlötterer 2013;
67 Baldwin-Brown et al. 2014). Furthermore, alternating generations with and without
68 selection can also increase recombination thereby reducing LD (Kessner and Novembre
69 2015).
70
71 A particularly challenging case is when selected alleles in the founder population reside
72 only on a single haplotype. Since the number of recombination events during typical E&R
73 studies is relatively small, this results in long-range hitchhiking among all SNPs
74 associated with the selected haplotype. While the general problem had been described
75 before (e.g. Logeswaran and Barton 2011), two studies recently demonstrated this

prevalence of false positives due to long-range hitchhiking in E&R studies (Tobler et al. 2014; Franssen et al. 2015). Tobler et al. (2014) recognized the importance of long-range hitchhiking through the presence of neutral sites among candidate SNPs. Franssen et al. (2015) identified such selected haplotype-blocks by mapping candidate SNPs to founder chromosomes, which showed a strong and consistent allele frequency change of these block specific SNPs across replicates. In total, 17 haplotype-blocks sized up to a few Mb were detected (Franssen et al. 2015). This illustrates the importance of initial associations / linkage between alleles in the founder population. If haplotype-blocks sized up to a few Mb are carrying a selected variant and stay intact over a few tens of generations, hitchhiking variants will be distributed over the entire haplotype-block. The density of singleton markers in the founder population, i.e. markers in LD to the selected allele that are hitchhiking, can vary locally depending on the shape of the underlying genealogy. Thus, the power to identify haplotype-blocks may also vary along the chromosome.

While Franssen et al. (2015) have shown that haplotype-blocks can be identified from a subset of the founder chromosomes, it is clear that the efficiency of this approach depends on the fraction of the founder chromosomes for which sequence information is available. A second, less obvious limitation comes from the presence of identity by descent (IBD) regions in natural populations. Although linkage disequilibrium in natural *D. melanogaster* populations is restricted to 100-200 bp (Mackay et al. 2012; Langley et al. 2012; Franssen et al. 2015), it has also been shown that long stretches of sequence identity between different haplotypes can nevertheless be detected (IBD regions) (Langley et al. 2012; Pool et al. 2012). If a neutral haplotype carries a low frequency IBD region, which is shared with a haplotype carrying a beneficial mutation close to but outside the IBD region, the neutral haplotype will show the selection signature for the shared region, despite no target of selection being located on that specific haplotype (S1 Fig). Thus, if only a subset of the founder haplotypes is sequenced, following the

haplotype specific markers of only a subset of haplotypes could identify rising blocks that are due to hitchhikers with a selected site on an unknown haplotype. While the complete sequencing of all founder chromosomes would avoid such errors, the sequencing costs could become prohibitive, in particular when recommendations to increase the number of founder haplotypes (Kofler and Schlötterer 2013; Baldwin-Brown et al. 2014) are being followed. Moreover, if founder populations are established from natural populations or only recently established isofemale lines, that are not completely inbred, sequencing of all segregating founder chromosomes is technically impossible.

Previous approaches to infer haplotype information from Pool-Seq data in experimental evolution studies estimate local frequencies of known founder haplotypes using regression (Long et al. 2011), a hidden Markov model (Cubillos et al. 2013), maximum likelihood estimation (Kessner et al. 2013) and a system of linear equations (Cao and Sun 2015). These methods, however, rely on the complete knowledge of experimental founder haplotypes (Cubillos et al. 2013) and / or are limited to haplotype frequency estimation for a restricted window size as error rates increase with within-window recombination (Long et al. 2011; Kessner et al. 2013; Cao and Sun 2015).

Here, we propose a new strategy to reconstruct selected haplotype-blocks of founder haplotypes without the necessity for sequencing founder chromosomes individually. Taking advantage of correlated allele frequency trajectories of hitchhiking SNPs that start from low frequencies, we identify sets of markers that are linked in the founding population. We show that such clusters are detected for large genomic regions spanning several Mb. We test our method on simulated data and in an E&R experiment with *Drosophila melanogaster*. Both computer simulations and experimental data indicate that the new approach of haplotype reconstruction has minimal error rates and provides a valuable tool for analyzing Pool-Seq data in experimental evolution studies.

132

133 New Approaches

134

135 Concept of haplotype reconstruction

136 Experimental evolution studies using outcrossing organisms typically have moderate
137 population sizes, start from a limited set of different founder haplotypes, frequently
138 present in multiple copies, and are exposed to selection for a limited number of
139 generations, usually less than 100 (reviewed in Schlötterer et al. 2015). This setup limits
140 the number of recombination events during the experiment, and causes pronounced
141 haplotype structures in evolving populations. In absence of recombination, the
142 frequencies of all SNPs specific to a given founder haplotype are correlated throughout
143 the experiment. Thus, within the limits of accuracy of the allele frequency estimate, it is
144 possible to identify blocks of founder haplotypes, i.e. haplotype-blocks, in replicated time
145 series data through correlated allele frequency changes. The accuracy of the haplotype
146 reconstruction increases with the magnitude of frequency changes throughout the
147 experiment, as this improves the signal to noise ratio. Haplotypes carrying a selected
148 allele are expected to change more in frequency during the experiment than the
149 remaining ones. Thus, they are not only more interesting, but also easier to reconstruct.
150 Typical evolve and resequence studies uncover more than 10^6 SNPs, which makes it
151 challenging to determine the correlation of all SNP pairs in the data set. Hence, we
152 prioritize the SNPs by 1) focusing only on initially low frequency SNPs, since haplotype-
153 specific SNPs are most informative, and 2) initially analyzing only SNPs located in a
154 window. The choice of the window size depends on the number of recombination events
155 during the experiment, with the preferred window size and recombination frequency
156 being negatively correlated. To extend a haplotype beyond the initial window, we
157 perform a sliding window analysis and link haplotypes detected in two overlapping
158 windows if they share SNPs. Our method is implemented in the R package
159 *haploReconstruct* (Materials & Methods).

160

161

162 ## Results

163

164 ### Proof of principle

165 We determined the validity of our approach by simulating two independently selected
166 sites in five replicates, each with 400 homozygous individuals founded by 200 different
167 haplotypes, i.e. 200 different founder haplotypes, each present in 4 copies. Both selected
168 alleles were unique to one of the 200 founder haplotypes and ca. 1 Mb apart from each
169 other. The second selected allele was included to account for the possibility that more
170 than one target of selection is segregating, which complicates the reconstruction of
171 selected haplotypes. For simplicity, we concentrated on the performance of our approach
172 to reconstruct of one of the selected haplotypes.

173

174 In all five replicates the focal selected allele continuously increased in frequency until
175 generation 60 (Fig 1). As expected from the stochasticity of recombination, the genomic
176 region affected by the spread of the beneficial allele varies in width and results in varying
177 trajectories of the founder haplotype carrying the selected allele along the chromosome
178 as well as among replicates (Fig 1 A). Applying our algorithm for the reconstruction of the
179 focal selected haplotype, we successfully reconstructed a 7.57 Mb region consisting of
180 1,347 markers (Fig 1 B). In this region, 100% of the block markers were correct alleles of
181 the corresponding founder haplotype and the inferred haplotype trajectory matched the
182 simulated data (Fig 1 A vs. B), indicating that a selected haplotype can be identified with
183 high confidence.

184

185 ### Validation based on simulated data

186 The accuracy of the haplotype-block reconstruction depends on many specifics of the
187 experiment, such as strength of selection, population size and duration of the experiment.

188 To account for the impact of these (typically unknown) parameters on the haplotype
189 reconstruction procedure, we simulated a range of parameter combinations. Each
190 produced time series datasets with two independently selected alleles. The selected
191 alleles were separated by about 1Mb and each was located on a single, but different
192 founder haplotype (see Materials & Methods). Haplotype-blocks were then reconstructed
193 from these simulated data, using different values for eight reconstruction parameters
194 (Table 1) and different sequencing coverages. The accuracy of haplotype reconstruction
195 was most strongly affected by the reconstruction parameters: the minimum frequency
196 change (minfreqchange) in a minimum number of replicates (minrepl) and the minimum
197 correlation (min.cl.cor) between markers in a local cluster (Fig 2). The percentage of
198 reconstructions with negligible error rates (correct allele fraction >0.99) increased with
199 the minimum frequency change in more replicates and with higher correlation within
200 local clusters. A frequency increase of at least 0.15 in 3 replicates and cluster correlation
201 of at least 0.7 resulted in more than 90% high quality reconstructions. To obtain a similar
202 reconstruction quality with only two replicates, a frequency increase of at least 0.2 and a
203 cluster correlation of 0.8 was required. Importantly, low quality reconstructions were
204 typically occurring in only a few of the 82 simulation scenarios evaluated (S2 Fig). The
205 remaining reconstruction parameters had only subtle effects on the accuracy of the
206 reconstructed haplotype-blocks.
207 Generally, error rates of the inferred haplotype-blocks decreased with lower starting
208 frequencies, smaller window size, higher sequencing coverage, more intersecting
209 markers for cluster elongation, more time points and the minimum cluster size (S3 A-E
210 Fig). It is important to note that the influence of these parameters is particularly
211 pronounced for 2 replicates and smaller minimum cluster correlations (S3 Fig). In
212 reconstructions with three increasing replicates and minimum cluster correlations of at
213 least 0.7 the remaining parameters had almost no effect on accuracy (S3 Fig). For
214 subsequent analyses, we focused on parameter combinations with a correlation of at

215 least 0.7 and at least 20% frequency change since these result in very reliable
216 reconstructions (Fig 2).
217
218 Alternative approaches to evaluate the quality of reconstructed blocks are the block
219 length and the accuracy of the inferred haplotype frequencies throughout the experiment.
220 The length of the selected haplotype block that can be reconstructed is important to
221 determine the entire region affected by hitchhiking, and to locate targets of selection
222 based on haplotype trajectories. We found a tradeoff between haplotype-block length and
223 block accuracy. Parameter values for the minimum starting frequency, window size and
224 sequencing coverage that yield a higher fraction of correctly called haplotype-block
225 alleles generally resulted in slightly shorter blocks, while the influence of the minimum
226 cluster size and intersection was negligible (S4 A-E Fig). Increasing the number of time
227 points did not only result in higher quality reconstructions but also slightly longer blocks
228 (S4 F Fig). While more relaxed parameter settings often increased the reconstructed
229 block length up to several Mb, the loss in accuracy was typically negligible (S4 and S5
230 Figs, S1 Table).
231 One important application of haplotype-block reconstruction is the inference of
232 haplotype trajectories along the chromosome. Thus, we also determined how well
233 haplotype frequencies were inferred by the mean squared error of true versus estimated
234 frequencies across windows (see Materials & Methods). We noted that a large
235 contribution to the mean squared error came from alleles that were incorrectly classified
236 as haplotype specific. Reconstructions with minimum cluster correlation of at least 0.7 in
237 2 or 3 replicates with minimum allele frequency changes of at least 0.2 almost exclusively
238 contained correctly called alleles (≥ 82.2% or ≥ 91%, respectively). Nevertheless, a small
239 subset of them had high mean squared errors in haplotype frequency estimates for at
240 least one replicate (≥ 0.1, S6 Fig). In most cases such inaccurate reconstructions could be
241 largely avoided when the minimum cluster correlation was increased to 0.8 (S6 Fig, e.g.
242 sim 19, 39, 51). A closer inspection showed that errors in calling haplotye-specific alleles

243 can be due to 1) higher frequency markers shared with another haplotype, and 2)
244 incorrectly linking the two independently selected haplotypes. A particularly challenging
245 case arose for simulation 51, where in one replicate both selected alleles recombined into
246 a single haplotype at an early time point (S7 A Fig). Nevertheless, in this case the
247 inaccurate reconstructions were almost entirely eliminated by increasing the minimum
248 cluster correlation (S6 Fig, S7 B-C Fig).
249

250 **Application to experimental data**

251 We applied our haplotype-block reconstruction method to time series data for one
252 chromosome arm, from an ongoing laboratory natural selection experiment with *D.*
253 *melanogaster* (Orozco-terWengel et al. 2012; Tobler et al. 2014; Franssen et al. 2015).
254 The analysis of real data is more complex than our computer simulations, since multiple
255 SNPs may be selected and the targets of selection are not known. Thus, we compared the
256 haplotype reconstruction for several reliable parameter combinations (see above) and
257 focused on those haplotypes carrying at least one outlier (i.e: candidate SNP) identified
258 by the Cochran-Mantel-Haenszel (CMH) test (see Materials & Methods, Franssen et al.
259 2015). Genome-wide block reconstructions were quite consistent among the different
260 parameter settings (S8 Fig). All six parameter combinations identified the same major
261 block, which differed only slightly in length (6.30 – 6.93 Mb) and harbored 88-99%
262 identical marker SNPs for reconstructions with the same maximum starting frequency
263 (Fig 3, Table 2).
264

265 We used 24 full genome haplotype sequences from generation F67 of replicate 2
266 (Franssen et al. 2015) to validate the reconstructed haplotype-block. While evolved
267 haplotypes are recombinants of original founder haplotypes, long consecutive stretches
268 of marker alleles spanning an entire reconstructed block within an evolved haplotype
269 imply correct inference of a block from a founder haplotype (Fig 4). Following this
270 rationale, we compared the reconstructed haplotype-block to the evolved haplotypes that

271 shared the longest stretches of identical markers and obtained an empirical error rate for
272 the inferred haplotype. Depending on the reconstruction parameter, the error rate
273 ranged from 0.46% to 0.87% (Table 2). This result clearly indicates that we correctly
274 identified the haplotype in the founder population, which carried at least one beneficial
275 allele and increased in frequency during the experiment.
276

277 **Avoiding false signals of selection due to IBD segments**

278 We compared the haplotype-block reconstructed from correlated allele frequency
279 changes to the one obtained by mapping candidate SNPs onto a subset of founder
280 haplotypes (Franssen et al. 2015) (Fig 5). Blocks identified in both analyses (new block
281 based on parameter set P6, S1 Table) correspond to a region in founder haplotype B.19
282 and start at the same genomic position (2R, 22,338), but the reconstructed block is 2.84
283 Mb longer (2R, 4,116,872 vs. 6,952,744). In the overlapping region both blocks are
284 identical (error rate 1/139), but the remaining 2.84 Mb of the reconstructed block are
285 clearly different from the homologous region of haplotype B.19 from the founder
286 population (Fig 5). This raises the question which of the two backgrounds – the newly
287 reconstructed block or the background of founder haplotype B.19 – is associated with the
288 selected site.
289

290 For validation of the reconstructed founder block we only sequenced haplotypes from
291 one evolved replicate (Fig 4, Table 2). Hence, it may be possible that the reconstructed
292 haplotype does not represent a founder haplotype, but one generated by recombination
293 during the experiment. Thus, we plotted the trajectory of the entire haplotype in all five
294 replicates and found that the entire reconstructed haplotype increases in frequency (Fig
295 3), indicating that the reconstructed haplotype was present in the founder population
296 and carried the selected allele. In contrast, the frequency increase of the founder
297 haplotype B.19 is restricted to the IBD region (Franssen et al. 2015). If haplotype B.19
298 also carried the beneficial allele, then the frequency increase would not be restricted to

299 the IBD region, but would also extend to the flanking sequence (see also S1 Fig).
300 Therefore, we conclude that the target of selection is located outside of the IBD region on
301 the reconstructed haplotype-block.
302
303
304 # Discussion
305
306 **Incorporating linkage information to map targets of selection in E&R data**
307 While most E&R studies assume independence among SNPs (e.g. Turner et al. 2011;
308 Burke et al. 2010; Orozco-terWengel et al. 2012; Topa et al. 2015), it has been recognized
309 that the inclusion of linkage information could improve the mapping of selection targets.
310 Kessner and Novembre (2015) used linkage information from the founder haplotypes to
311 determine haplotype frequencies in moderately sized (200 kb) windows for the evolved
312 populations. The increase in accuracy of their approach stems from a more accurate
313 frequency estimate compared to single marker analyses. In another recent approach
314 Terhorst et al. (2015) describe a multi-locus model of selection for replicated time-series
315 data. Here, a small number of SNPs adjacent to each focal SNP are used to increase the
316 information content of the data by taking into account the local haplotype structure and
317 recombination. While this improves the inference of the selected SNP, the actual
318 haplotype structure in the evolved populations is a nuisance parameter and remains
319 ultimately unknown.
320 Our approach differs from these two methods by primarily focusing on the
321 reconstruction of selected haplotypes. Furthermore, we validate for the first time
322 reconstructed haplotypes experimentally by sequencing evolved flies. We anticipate that
323 time series trajectories of selected haplotypes will not only facilitate the mapping of
324 targets of selection, but also provide the unique opportunity to match the observed
325 patterns against the expectations of classic population genetics models.
326

327 **Impact of IBD regions**

328 Natural *D. melanogaster* populations have low levels of linkage disequilibrium (Mackay et
329 al. 2012; Langley et al. 2012; Franssen et al. 2015). Nevertheless, it is becoming
330 increasingly clear that large genomic regions can be shared among individuals from the
331 same population. Such regions of identity by descent (IBD) are due to sampling of related
332 individuals, which is to be expected from local population structure. In the context of
333 adaptation, however, they have a severe impact on the mapping of selected alleles. The
334 comparison of the reconstructed haplotype to the haplotype-block identified from a
335 subset of the founder chromosomes (Franssen et al. 2015) is a particularly good
336 demonstration of how IBD blocks could result in wrong conclusions. Since the selected
337 chromosome shared a large IBD region with another chromosome, which does not
338 include the target of selection, the IBD region may be incorrectly identified as the target
339 of selection if the selected haplotype is not known. While sequencing of all founder
340 haplotypes would avoid this problem, previous suggestions to increase the number of
341 founder haplotypes for a reliable mapping of selection targets (Kofler and Schlötterer
342 2013; Baldwin-Brown et al. 2014) argue against this strategy. Furthermore, experiments
343 starting from freshly established isofemale lines contain multiple haplotypes (at least 4
344 with their resulting recombinants), which further complicates the inference of founder
345 haplotypes. The haplotype reconstruction method introduced here seems a more
346 promising and resource efficient approach.

347 One further challenge of IBD regions for the reconstruction of selected haplotypes arises
348 when they are shared between the selected chromosome and multiple non-selected
349 haplotypes resulting in an intermediate frequency of the IBD region. In this setting, the
350 selected haplotype will not carry any haplotype specific markers in the IBD region and
351 thus prevent the extension of the haplotype-block across the IBD region. Therefore, it is
352 possible that the length of the selected haplotype-block is severely underestimated.

353

354 **Limitations of the haplotype reconstruction**

355 Our approach to reconstruct haplotypes is targeted at selected alleles that are present at
356 low frequencies in the founder population. While this seems to be the predominant
357 genomic response in *D. melanogaster* populations adapting to hot environments (Tobler
358 et al. 2014; Franssen et al. 2015), other studies have found that selected alleles are at
359 intermediate frequencies (e.g. Turner et al. 2011; Turner and Miller 2012). Such common
360 alleles are expected to occur in multiple chromosomal backgrounds with few sites being
361 in high LD, which results in fewer hitchhikers obscuring signal from the selection targets
362 (e.g. S33 Fig in Kofler and Schlötterer 2013). The reconstruction of high frequency
363 selected clusters may therefore be more limited. On the other hand, if only a small
364 number of high frequency candidate SNPs emerge from the analysis, the identification of
365 the actual targets of selection is substantially simplified relative to cases of low starting
366 frequencies.
367
368
369 ## Materials and Methods
370
371 **Haplotype reconstruction based on correlated marker dynamics**
372 The reconstruction of founder haplotypes rising in frequency during experimental
373 evolution with the *haploReconstruct* software tool (XX will be available at CRAN) consists
374 of four main steps: 1) data upload and allele polarization, 2) data filtering, 3) marker
375 clustering in sliding-windows and 4) cluster extension to generate haplotype-blocks
376 across multiple windows (Table 1).
377 The underlying principle of the method is the clustering of linked alleles in the founder
378 population, based on frequency trajectories across multiple time points and replicates.
379 When reconstructing founder haplotypes, haplotype specific markers (i.e. singleton
380 markers of the founder population) carry the strongest signal, as their trajectories mirror
381 exclusively the trajectory of their respective haplotype. Therefore replicated, genome-
382 wide time series data are initially polarized for the minor allele using the estimated allele

383 frequencies in the founder population (step 1). Next, SNPs are filtered for informative
384 alleles defining a maximum frequency of the minor allele ("max.minor.freq") in the
385 starting population and a minimum frequency change ("minfreqchange") in a minimum
386 number of replicates ("minrepl") (step 2). The maximum frequency ensures that only
387 alleles with high information content with respect to the true haplotype frequency are
388 included, while the minimum frequency changes ensure that trajectories are specific to a
389 rising haplotype. For each overlapping sliding-window (parameter "winsize") pair-wise
390 correlations between SNPs are estimated. This is achieved by 1) square root
391 transformation of the frequency data, 2) scaling transformed frequencies for each SNP to
392 a mean of zero and a variance of one. These two transformations are performed for data
393 normalization prior to estimating the pair-wise correlations between SNPs with
394 Pearson's correlation coefficient. Correlations are subsequently transformed to into
395 distances (1 – correlation). Clusters are determined based on the distance matrix with the
396 average linkage clustering algorithm using the minimum average correlation
397 ("min.cl.cor") parameter as cutoff to build clusters (R package stats, hclust(… ,
398 method="average"), cutree(…)). Resulting clusters with fewer than the minimum number
399 of markers ("min.cl.size") are discarded (step 3). Finally, clusters are elongated across
400 overlapping windows (overlap = window size / 2) based on a minimum number of
401 identical markers ("min.inter"): all clusters identified for two overlapping windows are
402 tested pair-wise and a local cluster is combined across both windows if at least
403 "min.inter" markers are identical. If allele frequencies are available for several time
404 points, subsets can be chosen with the "use.libs" option. We generated a pseudo-code of
405 the reconstruction algorithm for more clarity of the reonstruction procedure (Text S1,
406 supplemental material). The source code of the R package *haploReconstruct* is available at
407 XXCRAN addressXX.
408

409 **Validation of haploReconstruct on simulated data**

410 **Simulation of time series data.** We explored the properties of our haplotype-block
411 reconstruction using computer simulations and different settings for the main
412 reconstruction parameters: max.minor.freq, minfreqchange, minrepl, winsize, min.cl.cor,
413 min.cl.size, min.inter, use.libs (specification of time points and replicates to be used for
414 reconstruction) and different sequencing coverages. Since for real data many important
415 parameters are not known, we simulated 100 time series data sets based on a range of
416 parameters and used these data to determine the influence of the reconstruction
417 parameters on the accuracy of reconstruction. Specifically, we used mimicrEE (Kofler and
418 Schlötterer 2013) to perform forward simulations based on a founder population
419 generated from 200 different *D. melanogaster* haplotypes of chromosomal arm 2R from
420 the DGRP lines (Mackay et al. 2012). Simulations included two selected sites in the middle
421 of the chromosomal arm that are separated by approximately 1Mb. Each selected allele
422 was unique to one of the 200 different founder haplotypes. Two independently selected
423 sites in close proximity on different chromosomal backgrounds were simulated to create
424 similar trajectories that have to be differentiated by the clustering method. For the
425 validation of this method, however, we concentrated on the reconstruction performance
426 of only one focal selected haplotype. In total we simulated 100 time series data sets for 80
427 generations and five replicates each. Thus, the simulations matched closely the
428 experimental data from Orozco-terWengel et al. (2012); Tobler et al. (2014); Franssen et
429 al. (2015). For each data set (five replicates over 80 generations) we randomly picked
430 one effective population size N_e {200 or 400} and a selection coefficient {0.1 - 0.2} with
431 complete dominance (heterozygous effect, h=1). For both selected sites, the selection
432 coefficients were independently drawn. Simulation parameters were chosen to match
433 previous estimates of N_e (Tobler et al. 2014; Franssen et al. 2015) and result in long range
434 hitchhiking (S9 Fig).
435
436 **Haplotype reconstruction on simulated data.** For each of the 100 simulations we first
437 tested if the selected site had survived the initial drift phase at the low initial population

438 frequency. The selected site was still present in at least 2 or 3 replicates (frequency ≥0.1
439 until generation 60) in 82 and 57 simulation sets, respectively. For each of the 82
440 simulation sets, in which the selected site had survived the drift phase, 800 haplotype
441 reconstruction runs were performed, 200 for each of 4 different coverages, i.e. 30, 60, 90
442 and infinite (i.e. true frequency without sampling error). We modeled the variation of
443 Pool-Seq at a given coverage by binomial sampling using all chromosomes in the
444 population. For each reconstruction, parameter values were randomly chosen from the
445 following parameter ranges: winsize ∈ {0.5, 1, 2} Mb, min.cl.cor ∈ {0.5, 0.6, 0.7, 0.8},
446 min.cl.size ∈ {4, 8}, min.inter ∈ {2, 4}, use.libs ∈ {20-40-60, 10-20-30-40-50-60, 20-40-60-
447 80, 60} and max.minor.freq ∈ {0.005, 0.01, 0.015, 0.02, 0.025, 0.03} (equivalent to the
448 presence in 1, 2, 3, 4, 5, 6 different founder haplotypes, respectively). The minimum
449 number of replicates (minrepl) was set to 2 when the selected site remained in 2
450 replicates only and was randomly chosen from {2, 3} when it survived initial drift in at
451 least 3 replicates. Minimum frequency changes (minfreqchange) were randomly chosen
452 from {0.1, 0.15, 0.2, 0.25, 0.3} until the requirement was fulfilled for the selected site in
453 the corresponding minimum number of replicates. This procedure ensured exclusion of
454 parameter combinations for which no haplotype-block reconstruction is possible since
455 the extent of hitchhiking is small / negligible.
456
457 **Performance measures for reconstructed haplotype-blocks.** We measured the
458 performance of haplotype-block reconstruction as the fraction of alleles in the
459 reconstructed haplotype-block (consisting only of markers, which are at a low frequency
460 in the founder population) that had the same character state in the selected haplotype. As
461 in a reconstruction run multiple blocks – including neutral ones – may be reconstructed,
462 we compared the reconstructed haplotype that shared most of 14 singleton markers
463 flanking the selected allele with the selected haplotype to evaluate the quality of the
464 haplotype reconstruction. Note that this accuracy measure does not relate to the full
465 sequence, but only to those SNPs, which were assigned to the haplotype-block.

466 In addition to the total number of markers shared with the selected haplotype, we also
467 measured the accuracy of the estimated haplotype-block frequency since alleles assigned
468 to the selected haplotype may additionally occur on other haplotypes, which adds noise
469 to the estimated frequencies. We determined the mean squared error between the true
470 haplotype frequency and the frequency of the corresponding reconstructed haplotype.
471 True haplotype frequencies were determined by the mean frequency of unique haplotype
472 markers for non-overlapping windows (window size = “winsize”/2) at generation F60. To
473 account for sampling variation, we calculated the true mean for each coverage separately.
474 Mean deviations in frequency across windows were calculated for each replicate by the
475 mean squared error. We denoted the largest deviation across replicates as “max(sqDiff)”.
476

477 **Application of *halpoReconstruct* to experimental data**

478 The accuracy of the haplotype-block reconstruction was also tested on replicated time
479 series data from an experimental evolution study with *D. melanogaster* populations
480 adapting to a new high temperature regime using Pool-Seq data. In total, 14 samples from
481 5 experimental replicates were used: time points F0, F15, F37 and F59 for the first three
482 and F15 from additional two replicates (data taken from: Orozco-terWengel et al. 2012;
483 Tobler et al. 2014; Franssen et al. 2015). A synchronized file containing genome-wide
484 allele counts of all 14 population samples (sync format see Kofler et al. 2011) was
485 obtained as described in Franssen et al. (2015). Haplotype-reconstruction was performed
486 with *haploReconstruct* in R (R Development Core Team 2008) for chromosomal arm 2R
487 and several parameters settings (S8 Fig). Reconstructed haplotype-blocks were filtered to
488 contain at least one putatively selected site identified by an independent single-locus test
489 for selection (see below for details). Haplotype-blocks were validated using 24
490 experimentally phased haplotypes from generation F67, replicate 2 (Franssen et al.
491 2015).
492

493 **Independent single-locus test for selection.** CMH tests were performed as an outlier
494 test for pair-wise comparisons between the experimental starting population and each of
495 the evolved time points at generation F15, F37 and F59 (Franssen et al. 2015). Tests
496 always included replicates 1-3 for euchromatic SNPs. Putative targets of selection were
497 among the top 2,000 ranked SNPs in either of the three comparisons (Orozco-terWengel
498 et al. 2012; Tobler et al. 2014). Note that neutral simulations consistently suggested that
499 2,000 candidate SNPs are a conservative estimate for the number of loci affected by
500 selection (Orozco-terWengel et al. 2012; Tobler et al. 2014; Franssen et al. 2015).
501
502

Acknowledgments

The authors thank all members of the Institute of Population Genetics for discussion and support on the project and particularly N. Barghi for helpful comments on earlier versions of the manuscript. This work was supported by the European Research Council (ERC) grants "ArchAdapt" and "250152".

Author contributions

SUF, NHB and CS conceived the study. SUF performed the data analysis. SUF and CS wrote the manuscript.

Fig 1. Visualization and reconstruction of a selected haplotype. The genome-wide frequency of SNPs specific for a selected haplotype is shown for chromosomal arm 2R in *D. melanogaster* in 5 different replicates (R1-R5) for simulated E&R data: A) The true frequency of the selected haplotype using only unique markers of the single founder haplotype that carries the selected allele (s_{focal}=0.1915) for frequency estimation at a sequencing coverage of 90. B) Frequencies of the identical selected haplotype as in A) but estimated by the markers of the corresponding reconstructed haplotype-block (parameters: minfreqchange=0.2, minrepl=3, max.minor.freq=4/200, winsize=0.5Mb, use.libs=10-20-30-40-50-60, min.cl.cor=0.7, min.cl.size=4, min.inter=2) using the same sequencing coverage of 90. The position of the selected site is indicated by a vertical black line. Single dots represent mean frequencies for overlapping windows of 5 markers. Simulation parameters, parameters for haplotype-block reconstruction and reconstruction performance are summarized in S1 Table.

528 **Fig 2. Fraction of correct alleles in reconstructed haplotype-blocks for different**
529 **reconstruction parameters.** The performance of haplotype reconstruction is
530 determined for 200 randomly chosen reconstruction parameter combinations for each of
531 four different sequencing coverages (30, 60, 90 and infinite) and 82 simulated E&R
532 datasets, where a single founder haplotype is selected (see Materials & Methods). The
533 percentage of reconstructions of the focal selected haplotype with a fraction >0.99 of
534 correctly reconstructed alleles is shown in response to the three parameters with the
535 largest influence on reconstruction performance: the minimum correlation within a
536 cluster and the minimum frequency change in a minimum number of replicates. The data
537 are summarized across the remaining parameters: max.minor.freq, use.libs, winsize,
538 min.cl.size, min.inter and different sequencing coverages. Smaller frequency changes in
539 fewer replicates and usage of low cluster correlations increase error rates in
540 reconstructed blocks. Importantly, if a minimum frequency change of 0.2 in at least 3
541 replicates is required and minimum cluster correlations of 0.7 are used, errors rates are
542 negligible, regardless of all other parameters tested.
543
544 **Fig 3. Trajectory of a putatively selected haplotype-block, reconstructed from**
545 **experimental data in *D. melanogaster*.** The frequencies of SNPs specific for this
546 haplotype are shown for chromosomal arm 2R in 5 different replicates (R1-R5). Single
547 dots represent mean frequencies for overlapping windows of 5 markers. Parameters for
548 haplotype-block reconstruction and reconstruction performance are summarized in table
549 S1, parameter set 6. Despite only a single time point was available for replicate number 4
550 and 5, it is apparent that at generation 15 in these replicates the same haplotype
551 increased in frequency.
552
553 **Fig 4. Validation of a reconstructed haplotype-block using 24 experimentally**
554 **determined haplotypes from a replicate population evolved for 67 generations.**
555 Rows are labeled with the respective haplotype ID or indicated as the reconstructed

556 block (hbr, highlighted by a black frame, reconstruction parameters P6). Each column
557 indicates a single marker; alleles shared with the reconstructed haplotype are colored in
558 red, the alternative alleles in cyan. Missing information is shown in gray. Haplotype R2.2,
559 R2.5 and R2.15 cover the full length of the reconstructed haplotype indicating that the
560 6.93 Mb haplotype block was correctly reconstructed.
561

562 **Fig 5. Comparison of a reconstructed haplotype-block to an experimentally**
563 **determined subset of founder haplotypes suggests the presence of IBD segments.**
564 Reconstructed haplotype-block 1 of parameter set 6 is visualized together with 24 of a
565 total of ~113 founder haplotypes. Rows are labeled with the respective haplotype ID or
566 indicated as the reconstructed block (hbr, highlighted by a black frame). Coloring
567 indicates the haplotype-block allele (red), the alternate alleles (green, blue) or an
568 unknown nucleotide (gray). Columns correspond to positions of hbr marker alleles. In the
569 first half of the region (indicated by an arrow on the x axis), where a rising haplotype-
570 block was identified for B.19 in a previous approach (Franssen et al. 2015), the newly
571 reconstructed block and the previous block are identical, while they are clearly different
572 in the second half.
573

574 **Table 1. Parameters for haplotype-block reconstruction.** Note: parameter base.pops
575 is not variable for a given experiment as SNP frequency estimates of the founder
576 population are mandatory.

Reconstruction parameters	
1) Allele polarization	
base.pops	Specification of the sequencing libraries of the experimental founder population. They are used for allele polarization to the minor allele in the founder population.
2) SNP Filtering	
max.minor.freq	Maximum minor allele frequency in the founder population. Only SNPs with a starting frequency smaller or equal this value will be kept.

minfreqchange	Minimum frequency change of an allele between the starting frequency and the frequency at any evolved time point in one replicate for a SNP to be kept.
minrepl	The minimum number of replicate populations with the specified minimum frequency change.
use.libs	Specification of the libraries (replicate and time point), for which minimum frequency changes are tested and that are used for subsequent marker clustering. Note: If libraries of the experimental founder population are not specified here they will not be used for marker clustering in the next step. However, starting frequencies are always required and will be used for 1) allele polarization and 2) estimation of minimum frequency changes.
3) Marker clustering in sliding windows	
winsize	Window size of the local marker clustering in Mb.
min.cl.cor	Minimum pair-wise correlation between SNPs in allele frequency trajectories, required to be grouped into one window-based cluster.
min.cl.size	Minimum number of markers required for one window based cluster in order to be reported.
4) Cluster extension across windows	
min.inter	Minimum number of markers that have to be shared between clusters of overlapping windows in order to be extended into a larger haplotype-block.

577

578

579 **Table 2. Summary statistics for reconstructed haplotype-blocks for parameter sets**
580 **that resulted in disjunct selected blocks.** (see also S8 Fig)

Parameter set	Chr	haplotype-block ID	# marker	Start position	End position	Length [bp]	# marker / Mb	# occurrence in F67 R2 haplotypes	Mean experimental error rate [%]
P1	2R	1	168	197613	6493901	6.30	26.7	3 - 8	0.64
P2	2R	1	184	22338	6862392	6.84	26.9	3 - 8	0.69
P3	2R	1	182	22338	6862392	6.84	26.6	3 - 8	0.61
P4	2R	1	383	22338	6952744	6.93	55.3	3 - 8	0.79
P5	2R	1	434	22338	6952744	6.93	62.6	3 - 8	0.87
P6	2R	1	322	22338	6952744	6.93	46.5	3 - 8	0.46

581

582 ## References

583

584 Baldwin-Brown JG, Long AD, Thornton KR. 2014. The Power to Detect Quantitative Trait
585 Loci Using Resequenced, Experimentally Evolved Populations of Diploid, Sexual
586 Organisms. Mol. Biol. Evol. 31:1040–1055.

587 Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD. 2010. Genome-
588 wide analysis of a long-term evolution experiment with *Drosophila*. Nature 467:587–
589 590.

590 Burke MK, Liti G, Long AD. 2014. Standing genetic variation drives repeatable experimental
591 evolution in outcrossing populations of *Saccharomyces cerevisiae*. Mol. Biol. Evol.
592 31:3228–3239.

593 Cao C-C, Sun X. 2015. Accurate estimation of haplotype frequency from pooled sequencing
594 data and cost-effective identification of rare haplotype carriers by overlapping pool
595 sequencing. Bioinformatics 31:515–522.

596 Cubillos FA, Parts L, Salinas F, Bergström A, Scovacricchi E, Zia A, Illingworth CJR,
597 Mustonen V, Ibstedt S, Warringer J, et al. 2013. High-Resolution Mapping of Complex
598 Traits with a Four-Parent Advanced Intercross Yeast Population. Genetics 195:1141–
599 1155.

600 Franssen SU, Nolte V, Tobler R, Schlötterer C. 2015. Patterns of Linkage Disequilibrium and
601 Long Range Hitchhiking in Evolving Experimental *Drosophila melanogaster*
602 Populations. Mol. Biol. Evol. 32:495–509.

603 Jha AR, Miles CM, Lippert NR, Brown CD, White KP, Kreitman M. 2015. Whole genome
604 resequencing of experimental populations reveals polygenic basis of egg size variation
605 in Drosophila melanogaster. Mol. Biol. Evol.:msv136.

606 Kessner D, Novembre J. 2015. Power Analysis of Artificial Selection Experiments Using
607 Efficient Whole Genome Simulation of Quantitative Traits. Genetics 199:991–1005.

608 Kessner D, Turner TL, Novembre J. 2013. Maximum Likelihood Estimation of Frequencies of
609 Known Haplotypes from Pooled Sequence Data. Mol. Biol. Evol. 30:1145–1158.

610 Kofler R, Pandey RV, Schlötterer C. 2011. PoPoolation2: identifying differentiation between
611 populations using sequencing of pooled DNA samples (Pool-Seq). Bioinformatics
612 27:3435–3436.

613 Kofler R, Schlötterer C. 2013. A guide for the design of evolve and resequencing studies. Mol.
614 Biol. Evol. 31:474–483.

615 Lang GI, Desai MM. 2014. The spectrum of adaptive mutations in experimental evolution.
616 Genomics 104:412–416.

617 Langley CH, Stevens K, Cardeno C, Lee YCG, Schrider DR, Pool JE, Langley SA, Suarez C,
618 Corbett-Detig RB, Kolaczkowski B, et al. 2012. Genomic Variation in Natural
619 Populations of *Drosophila melanogaster*. Genetics 192:533–598.

620 Logeswaran S, Barton NH. 2011. Mapping Mendelian traits in asexual progeny using changes
621 in marker allele frequency. Genet. Res. 93:221–232.

622 Long A, Liti G, Luptak A, Tenaillon O. 2015. Elucidating the molecular architecture of
623 adaptation via evolve and resequence experiments. Nat. Rev. Genet. 16:567–582.

624 Long Q, Jeffares DC, Zhang Q, Ye K, Nizhynska V, Ning Z, Tyler-Smith C, Nordborg M.
625 2011. PoolHap: Inferring Haplotype Frequencies from Pooled Samples by Next
626 Generation Sequencing. PLoS ONE 6:e15292.

627 Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, Casillas S, Han Y,
628 Magwire MM, Cridland JM, et al. 2012. The *Drosophila melanogaster* Genetic
629 Reference Panel. Nature 482:173–178.

630 Martins NE, Faria VG, Nolte V, Schlötterer C, Teixeira L, Sucena É, Magalhães S. 2014. Host
631 adaptation to viruses relies on few genes with different cross-resistance properties.
632 Proc. Natl. Acad. Sci. 111:5938–5943.

633 Masri L, Branca A, Sheppard AE, Papkou A, Laehnemann D, Guenther PS, Prahl S, Saebelfeld
634 M, Hollensteiner J, Liesegang H, et al. 2015. Host–Pathogen Coevolution: The
635 Selective Advantage of *Bacillus thuringiensis* Virulence and Its Cry Toxin Genes.
636 PLoS Biol 13:e1002169.

637 Nuzhdin SV, Turner TL. 2013. Promises and limitations of hitchhiking mapping. Curr. Opin.
638 Genet. Dev. 23:694–699.

639 Orozco-terWengel P, Kapun M, Nolte V, Kofler R, Flatt T, Schlötterer C. 2012. Adaptation of
640 *Drosophila* to a novel laboratory environment reveals temporally heterogeneous
641 trajectories of selected alleles. Mol. Ecol. 21:4931–4941.

642 Pool JE, Corbett-Detig RB, Sugino RP, Stevens KA, Cardeno CM, Crepeau MW, Duchen P,
643 Emerson JJ, Saelao P, Begun DJ, et al. 2012. Population Genomics of Sub-Saharan
644 Drosophila melanogaster: African Diversity and Non-African Admixture. PLoS Genet
645 8:e1003080.

646 Reed LK, Lee K, Zhang Z, Rashid L, Poe A, Hsieh B, Deighton N, Glassbrook N, Bodmer R,
647 Gibson G. 2014. Systems Genomics of Metabolic Phenotypes in Wild-Type *Drosophila*
648 *melanogaster*. Genetics 197:781–793.

649 Remolina SC, Chang PL, Leips J, Nuzhdin SV, Hughes KA. 2012. Genomic Basis of Aging
650 and Life-History Evolution in *Drosophila Melanogaster*. Evolution 66:3390–3403.

651 Sanjuán R, Cuevas JM, Moya A, Elena SF. 2005. Epistasis and the Adaptability of an RNA
652 Virus. Genetics 170:1001–1008.

653 Schlötterer C, Kofler R, Versace E, Tobler R, Franssen SU. 2015. Combining experimental
654 evolution with next-generation sequencing: a powerful tool to study adaptation from
655 standing genetic variation. Heredity 114:431–440.

656 Sheng Z, Pettersson ME, Honaker CF, Siegel PB, Carlborg Ö. 2015. Standing genetic variation
657 as a major contributor to adaptation in the Virginia chicken lines selection experiment.
658 Genome Biol. 16:219.

659 Teotónio H, Chelo IM, Bradić M, Rose MR, Long AD. 2009. Experimental evolution reveals
660 natural selection on standing genetic variation. Nat. Genet. 41:251–257.

661 Terhorst J, Schlötterer C, Song YS. 2015. Multi-locus Analysis of Genomic Time Series Data
662 from Experimental Evolution. PLoS Genet 11:e1005069.

663 Tobler R, Franssen SU, Kofler R, Orozco-terWengel P, Nolte V, Hermisson J, Schlötterer C.
664 2014. Massive Habitat-Specific Genomic Response in *D. melanogaster* Populations
665 during Experimental Evolution in Hot and Cold Environments. Mol. Biol. Evol.
666 31:364–375.

667 Topa H, Jónás Á, Kofler R, Kosiol C, Honkela A. 2015. Gaussian process test for high-
668 throughput sequencing time series: application to experimental evolution.
669 Bioinformatics 31:1762–1770.

670 Turner TL, Miller PM. 2012. Investigating Natural Variation in *Drosophila* Courtship Song by
671 the Evolve and Resequence Approach. Genetics 191:633–642.

672 Turner TL, Stewart AD, Fields AT, Rice WR, Tarone AM. 2011. Population-Based
673 Resequencing of Experimentally Evolved Populations Reveals the Genetic Basis of
674 Body Size Variation in *Drosophila melanogaster*. PLoS Genet 7:e1001336.

675

676