# New criteria for sympatric speciation in the genomic era

Emilie J Richards[1], Maria R. Servedio[1], Christopher H Martin[1]*

[1] Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill NC

*corresponding author: chmartin@unc.edu

Key words: sympatric speciation, genomics, gene flow, introgression, selection, cichlid

Main text: 3,139 words, 1 Figure, 3 Boxes

**Abstract**

Sympatric speciation illustrates how natural and sexual selection may create new species in isolation without geographic barriers. However, so far, all genomic reanalyses of classic examples of sympatric speciation indicate secondary gene flow occurred. Thus, there is a need to revisit criteria for demonstrating sympatric speciation in the face of widespread gene flow. We summarize theoretical differences between sympatric speciation and speciation-with-gene-flow models and propose genomic criteria for sympatric speciation: 1) timing of fine-scale introgression; 2) timing of selective sweeps and 3) functional annotation of this introgressed variation; and 4) the absence of similar sweeps in outgroups. Monophyly is an insufficient criterion for sympatric speciation; we must take a locus-specific approach to investigate whether any introgression contributed to reproductive isolation.

## What is sympatric speciation?

Sympatric speciation is the evolution of reproductive isolation within a single panmictic population without the aid of any geographic isolation [1]. It represents the most extreme and controversial endpoint on the divergence with gene flow continuum: panmictic gene flow and no initial divergence at the start of speciation [2–5]. In the context of theoretical speciation models, sympatric speciation is the most difficult process because the starting conditions involve no pre-existing divergence, potentially tied with physical linkage, among loci involved in reproductive isolation (i.e. 'barrier' loci [6,7]); instead, **linkage disequilibrium** (see Glossary) must build up through time within a population through the action of disruptive natural selection and strong assortative mating by ecotype, despite the countervailing eroding force of recombination [8–11].

We can thus distinguish different types of scenarios that will result in two sister species being found in sympatry based on whether **secondary gene flow** aided population divergence: 1) classic sympatric speciation without gene flow; 2) sympatric speciation in the presence of a) neutral secondary gene flow or b) after differential sorting of an ancestral hybrid swarm. In the latter case, it is important to distinguish whether the ancestral hybrid swarm population achieved panmixia before later divergence (i.e. sympatric divergence); otherwise, differential sorting of haplotypes within the hybrid swarm is better described by secondary contact speciation with gene flow models. 3) Speciation may be aided by secondary gene flow that a) triggers initial sympatric divergence or b) increases divergence after initial divergence in sympatry becomes stalled, an outcome of many sympatric speciation models without sufficiently strong disruptive selection [12–15]. Finally, 4) secondary contact after a period of allopatry between two populations can result in coexistence or reinforcement, if there is not collapse into a single

70 admixed population [16–19]. We consider scenarios 1 and 2 to be examples of sympatric
71 speciation, whereas scenarios 3 and 4 would be examples of speciation aided by secondary gene
72 flow. Interestingly, hybrid swarm scenarios exist in a gray area, since substantial initial gene
73 flow from multiple sources may increase ecological or preference variation within a population
74 that is sufficient to trigger later sympatric divergence, even without segregating inversions or
75 genetic incompatibilities [20–22]. So far, we know of no examples of scenario 1 within any case
76 study of sympatric sister species examined using genomic tools; even long diverged species
77 show some evidence of **introgression** in their past (e.g. [23]). In contrast, sympatric speciation
78 with neutral gene flow (Scenario 2) and speciation aided by gene flow (Scenarios 3 and 4)
79 frequently appear to operate concurrently even within a single sympatric adaptive radiation (e.g.
80 [24–26]).

81 It is important to distinguish these scenarios because theoretical models predict that
82 sympatric divergence unaided by any form of secondary gene flow is substantially more difficult
83 than other speciation with gene flow scenarios (Box 1). Gene flow throughout the speciation
84 process allows recombination to break down linkage disequilibrium among alleles associated
85 with ecological divergence and assortative mating. There are actually three different classes of
86 sympatric speciation models to consider: the most difficult process involves independently
87 segregating loci for ecotype, female preferences, and male traits within the population, whereas
88 sympatric divergence is much easier if any of these three types of traits are combined, such as
89 assortative mating based on phenotype matching instead of separate loci for preference and traits
90 [27,28] or "magic" traits (such as assortative mating based on microhabitat preference;
91 [9,10,29]). Sympatric speciation by sexual selection alone is also theoretically possible if there is

substantial preference variation either initially within the population or through secondary gene flow [20,30,31].

Any form of linkage disequilibrium among ecological and mate choice loci formed in allopatry, whether due to physical linkage, selection, or drift, can tend to shift the initial starting conditions of panmixia in favor of sympatric divergence [17]. However, linkage disequilibrium without physical linkage subsides within a few generations after secondary sympatry and thus may not allow sufficient time for the evolution of assortative mating within the population. In contrast, pre-existing physical linkage among ecological loci has been shown to increase the probability of divergence, especially when it captures already divergent alleles as is more likely after allopatric divergence [32,33]. Similarly, physical linkage can cause preference and trait alleles to mimic phenotype matching, although even tight linkage can break down over long timescales (shown in a model with population structure: [34]). Segregating inversions in the ancestral population are now well-known empirical examples of physical linkage promoting divergence in sympatry [35–37]. Sympatric divergence is also limited by many other restrictive conditions regarding the costs of female choosiness and strengths of disruptive selection and assortative mating (Box 1).

Despite extensive searches for examples of sympatric speciation in the wild, there are few convincing case studies due to the difficulty of ruling out historical allopatric scenarios (see below) and the new difficulty of ruling out a role of introgression in speciation. Furthermore, the role of magic traits or matching vs. preference/trait mechanisms is not fully understood in any existing case study. Thus, we still have very limited empirical tests of an extensive theoretical

113 literature and diverse competing models of the notoriously difficult process of sympatric
114 speciation [24,27,28,38–42].
115
116 **The classic problem of sympatric speciation**
117 There are four traditional criteria for demonstrating classic sympatric speciation (Scenario 1): 1)
118 sister species which are reproductively isolated, 2) form a **monophyletic** group, 3) largely
119 overlap in ranges, and 4) have biogeographic and evolutionary histories that make periods of
120 allopatric divergence highly unlikely [1]. Very few case studies have been able to meet these
121 rigorous criteria despite intense searches [1,5]. This has led to the prominent status of crater lake
122 cichlid radiations as some of the best examples of sympatric speciation in the wild due to the
123 uniform shape of isolated volcanic lakes which convincingly rule out phases of allopatry due to
124 water level changes (Box 2; [43–45]).
125 The monophyly criterion assumes that monophyly arises only when a single ancestral
126 population underlies the present-day daughter species. This is typically met by inferring a single
127 phylogeny from one or more loci. This single point-estimate view of evolutionary history is
128 problematic because it obscures the presence of non-bifurcating relationships among organisms
129 (e.g. sister species which derived ancestry from multiple source populations due to extensive
130 gene flow or hybrid speciation) and the real variation in evolutionary histories among genes
131 across the genome itself (e.g. [46]). Few regions of the genome may initially contribute to
132 reproductive isolation resulting in a heterogeneous genomic landscape of differentiation among
133 incipient species [47], a pattern now extensively supported across case studies [48–51].
134 Therefore, monophyletic relationships are consistent with, but not exclusive to a scenario of

135 sympatric speciation. Examining heterogeneous evolutionary histories across regions relevant to
136 speciation is thus crucial for understanding the processes and conditions under which sympatric
137 divergence can occur.
138
139 **The 'new' problem of sympatric speciation**
140 While genomics has increased our ability to resolve evolutionary relationships among organisms,
141 it has also revealed more complex evolutionary histories of multiple colonization and extensive
142 secondary gene flow in all examples of sympatric speciation that have been examined with
143 genomic data so far [52–68]. Indeed, only a handful of genes may directly contribute to the
144 speciation process whereas the rest of the genome is porous to gene flow while reproductive
145 isolation is incomplete [47,69]. Examples of 'classic' sympatric speciation without secondary
146 gene flow (Scenario 1) are now unknown after applying modern genomic tools to search for
147 introgression (and paleogenomics is therefore unnecessary to provide historical point estimates
148 of spatial isolation as recently suggested [70]). Instead, it is still possible that sympatric
149 speciation occurs in the face of secondary gene flow in nearly all these examples (Scenario 2;
150 [67]). Importantly, most evidence of secondary gene flow comes from genome-wide tests of
151 introgression from outgroup lineages, not gene flow between diverging populations in sympatry
152 (e.g. [64,65]). Therefore, introgression detected at the genome-wide level from lineages outside
153 the speciation event tells us little about the divergence process among incipient sympatric species
154 and how gene flow shaped the process of speciation.
155 The challenge of sympatric speciation in the genomic era is establishing or rejecting a
156 functional role for the ubiquitous secondary gene flow present during the speciation process, in

157 effect ruling out scenarios 3 and 4 in favor of scenario 2 (Fig. 1). Even if signatures of secondary
158 gene flow are detected, speciation could still have occurred solely via mechanisms of sympatric
159 speciation if that secondary gene flow did not play a causal role in divergence. Secondary gene
160 flow could play a causal role if it introduced novel genetic variation or physically linked alleles
161 (e.g. a segregating inversion) that promote speciation before the start of divergence, such as
162 **hybrid swarm (**Scenario 2b**;** ([21,22,25,71]), adaptive introgression (Scenario 3; [72–77]),
163 **transgressive segregation** (Scenarios 2-3; [78,79]), or hybrid speciation (Scenario 4; [80]). We
164 note a distinction between ancestral introgression of segregating haplotypes promoting speciation
165 (speciation with gene flow models apply) versus sufficient time for recombination to break down
166 these haplotypes after a hybrid swarm and create panmictic conditions before the start of
167 divergence (sympatric speciation models apply) versus simply inflated ecological and preference
168 variation within a population due to hybrid swarm (gray area; sympatric speciation more likely
169 due to this initial gene flow). Here we propose and discuss new genomic criteria to help establish
170 or reject a functional role of secondary gene flow in the speciation process (Fig. 1). This is
171 necessary to identify putative cases of the sympatric speciation process when gene flow appears
172 to be nearly universal in the wild, particularly among sympatric diverging populations.
173

174 **New criteria for sympatric speciation in the genomic era**

175 Although genome-wide analyses of introgression provide a starting point, ultimately
176 consideration of the time of arrival and functional role of each introgressed region within extant
177 sympatric sister species pairs will be necessary to distinguish between sympatric speciation with
178 gene flow (Scenario 2) or speciation aided by secondary gene flow (Scenario 3; e.g. segregating

179 inversions [35,81] or ancient balancing selection on regions containing multiple barrier loci
180 [82,83]). We suggest four major types of genomic analyses as new criteria to help identify
181 sympatric speciation with gene flow: 1) estimate the timing of introgression into sympatric sister
182 species relative to their divergence time, 2) infer the presence and timing of selective sweeps
183 within sympatric sister species, 3) annotate candidate adaptive introgression regions for
184 functional elements or trait associations that may be relevant to speciation, and 4) if closely
185 related non-speciating outgroups are available, confirm the lack of selective sweeps of these
186 regions in outgroups. Combining these statistics will aid in distinguishing where case studies fall
187 along the speciation with gene flow continuum and whether the starting conditions of panmixia
188 in sympatric speciation models will apply (Fig. 1).
189
190 *1) Secondary gene flow is constant across the speciation process or not concurrent with*
191 *divergence times*
192 Estimating the duration of gene flow and the timing of introgression relative to the timing of
193 divergence between sympatric sister species will help distinguish between scenarios of sympatric
194 speciation, speciation with gene flow, and secondary contact. If populations diverged in
195 sympatry independent of any concurrent secondary gene flow (Scenario 2), we might expect to
196 see weak concordance of the timing of gene flow with divergence times among species. This
197 discordance could be in the form of discrete gene flow events that date well before or after
198 divergence times among species (Fig 1A). In the case of continuous gene flow from the time of
199 colonization to the present, more information about the functional role and selection on
200 introgressed regions will be needed.

201 The timing of introgression is also useful in ruling out other evolutionary phenomena that
202 can leave similar genomic signatures. The random or biased assortment of ancestral variation
203 among lineages during the speciation process can create similar phylogenetic patterns to
204 introgression resulting from secondary gene flow [82,84]. Timing is important for differentiating
205 introgression from the sorting of ancient ancestral polymorphisms due to processes such as
206 balancing selection. For example, if genetic divergence in an introgressed region shared between
207 sister species is greater (e.g. elevated $\boldsymbol{D_{xy}}$) than expected given divergence time between the
208 sister species, this pattern suggests differential sorting of ancestral variation and doesn't rule out
209 a scenario of sympatric speciation (Scenarios 1 & 2). If introgression after secondary contact did
210 occur, genetic divergence in these regions between recipient sister species should be lower than
211 expected given their divergence time. Increasingly sophisticated approaches for detecting fine-
212 scale patterns of introgression are available to estimate the timing and duration of gene flow
213 from genomic data (Box 3).
214
215 *2) Lack of selective sweeps or non-concurrent timing of sweeps in regions that have experience*
216 *gene flow*
217 We can use information about selective sweeps of introgressed variation to further characterize
218 the role of gene flow in sympatric divergence. When an allele is selectively favored in a
219 population, positive selection may cause it to increase in frequency and form a localized
220 selective sweep of reduced genetic variation surrounding the adaptive variant [85]. Such regions
221 of high differentiation in recently diverged species are often targeted as candidates for speciation
222 genes, although other processes not directly associated with speciation can lead to similar

223 patterns of high heterogeneity in differentiation across a genome (reviewed in [86–88]); indeed,
224 there is still no evidence that these regions are associated with reproductive isolation or reduced
225 gene flow and can also result from adaptive introgression [89,90]). If speciation was recent or
226 ongoing, there may be strong signatures of a selective sweep for particular haplotypes in at least
227 one of the sister species for regions involved in the divergence process (Fig. 1B). If secondary
228 gene flow was neutral with respect to speciation, we may find no signatures of selective sweeps
229 in those introgressed regions.

230 Importantly, a sweep of the same introgressed region in both sympatric sister species may
231 be interpreted as adaptation to the same new environment, which may not contribute to
232 reproductive isolation between the pair (dependent on their respective genetic backgrounds; e.g.
233 [90,91]). However, this pattern is also consistent with the sweep of a region contributing to a
234 'one-allele' mechanism of mate choice [27,28,92], such as increased female choosiness in both
235 sympatric sister species (e.g. [93]), which *would* contribute to reproductive isolation. Thus,
236 selective sweeps of an introgressed region in both sympatric sister species do not rule out its role
237 in aiding the speciation process.

238 Alternatively, if selective sweeps are detected, the timing of selective sweeps in the
239 regions affected by this gene flow can give indirect evidence about the selective pressure
240 underlying the sweep. It is challenging to infer something about the importance of an
241 introgressed region if the timing of introgression predates the timing of the selective sweep
242 because linkage disequilibrium among loci relevant to speciation may take time to build up, a
243 process involved in most speciation models [9,10]. However, the absence of selective sweeps or
244 introgression until long after population divergence would suggest that introgression was not

245 relevant to the speciation process (Scenario 2a). Introgressed regions that have undergone soft
246 selective sweeps and were important for divergence may easily be missed, but increasingly
247 sensitive methods [94,95] are making it easier to detect them.
248
249 *3) Weak support for casual role of secondary gene flow based on functional genetic analyses of*
250 *variants in the region*
251 Another potential source of evidence for the functional importance of gene flow can come from
252 associations between variants in introgressed regions with traits involved in ecological and
253 sexual isolation between sister species from genome-wide association studies (GWAS).
254 However, many complex traits are driven by a large number of variants of small effect and ruling
255 out a functional role for gene flow from any annotations is difficult (e.g. omnigenic model; [96]).
256 The conservation of sequences within introgressed regions across taxa may also provide strong
257 evidence of a functional role (e.g. PhastCons [97]). Finally, and most powerfully, genome editing
258 and gene expression reporter systems are increasingly tractable in non-model systems (e.g. [98–
259 101]. This is ultimately an asymmetric problem: finding evidence that an introgressed region
260 may have contributed to reproductive isolation is far easier than demonstrating that no
261 introgressed regions contributed to reproductive isolation in any way [67]. Finding evidence for
262 sympatric speciation in the wild is now the difficult problem of functional genetic analyses of
263 introgressed regions.
264
265 *4) Similar patterns of selection or divergence in the introgessed regions in closely related*
266 *outgroup populations*

267 Thorough investigations of these same regions in outgroups to the sympatric species gives added
268 power to distinguish whether secondary gene flow aided sympatric divergence. If non-
269 diversifying, closely related species exist in similar environments and haven't diversified in a
270 similar manner but share signatures of selective sweeps in the same regions, then the observed
271 introgression may have been neutral relative to speciation, e.g. due to adaptations to shared
272 changes in climate or pathogens or shared regions of reduced recombination or increased
273 background selection. For example, several studies comparing genomic landscapes of
274 differentiation across closely related taxa have found that high differentiation observed in the
275 same genomic regions across taxa reflects the action of linked selection across low-
276 recombination regions rather than selection against gene flow at barrier loci [102–110].
277

278 **Concluding Remarks**

279 Sympatric speciation remains among the most controversial evolutionary processes, beloved by
280 theorists and long sought after by empiricists. While evidence for this process appeared to be
281 mounting using traditional criteria [5], genomic data has now cast doubt on all these putative
282 examples due to the ubiquity of secondary gene flow. Furthermore, nearly all our existing case
283 studies involve some form of automatic magic trait, such as assortative mating by habitat
284 [35,111,112], along a depth gradient [63], or environment-induced phenology shifts [113]. We
285 think an outstanding remaining question is whether sympatric divergence can occur in nature
286 without the aid of some form of magic trait, as originally demonstrated to be possible in theory
287 [9,114].

288 Future fine-scale investigations of introgression will likely continue to paint a complex
289 picture of the role of secondary gene flow in nearly all speciation events. The highly polygenic

290 and multi-dimensional nature of adaptation and mate choice suggests that an 'all-of-the-above'
291 speciation scenario containing a mix of preference/trait, magic trait, and phenotype matching,
292 each spread across a wide distribution of allelic effect sizes with varying times of arrival, will be
293 the norm in nature. In contrast, although numerous and diverse, most speciation models continue
294 to address these mechanisms in a piecemeal fashion with an assumption of large effect alleles. It
295 remains unclear how different mechanisms, effect sizes, and times of arrival will interact and
296 compete within a single model.
297 Interestingly, strict isolation of sympatric environments such as crater lakes may
298 ultimately become less important, since even these isolated environments are not isolated from
299 secondary gene flow [115,116]. Instead, recent sympatric divergence combined with well
300 characterized introgression and functional annotations may be the new limiting factor for
301 convincing case studies of sympatric divergence in the genomic era.
302
303
304
305
306
307
308
309
310
311
312

313 ***Box 1*. Why do we care whether speciation is sympatric?**

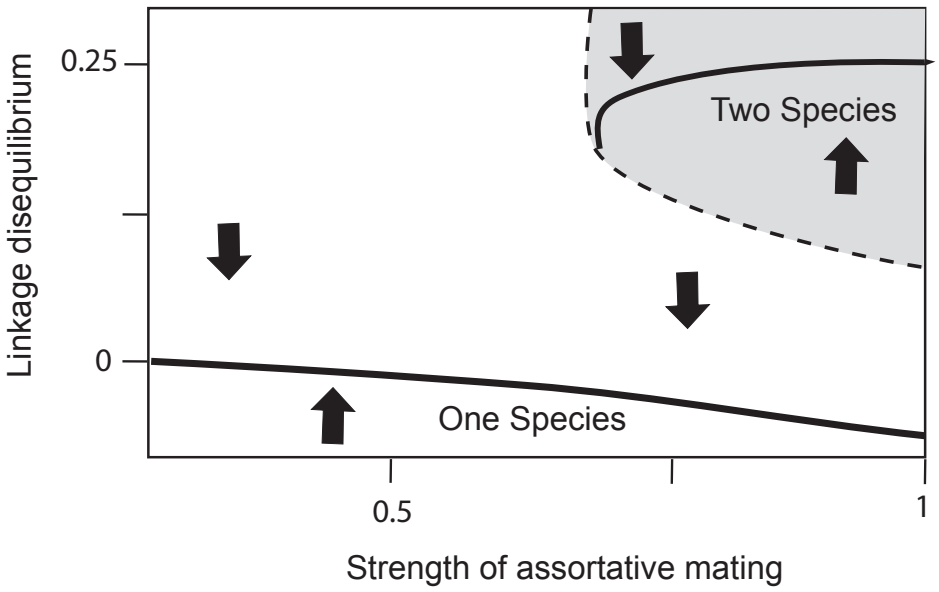
314 An increasingly common claim in the empirical speciation literature is that there is no difference
315 between 'speciation with gene flow' and 'sympatric speciation' (e.g. [70]). This contrasts sharply
316 with the rich theoretical literature differentiating models of sympatric speciation with models of
317 speciation with gene flow. Indeed, theory teaches us that we should care about the real
318 differences between the process of sympatric speciation (i.e. population divergence in sympatry
319 without the aid of introgression contributing to reproductive isolation) and other models of
320 speciation with gene flow. Sympatric speciation is uniquely and notoriously difficult [117], in
321 part because quite specific conditions of resource availability (e.g., [9,118]), mating traits and
322 preferences (e.g., [31,119]), and search costs (e.g., [120]) must be met for it to occur.

323 Inferences from theoretical models predict that, under a scenario of speciation with gene
324 flow (Scenario 3), introgression can make the process of speciation much easier in three ways.
325 First, by introducing additional variation in ecological traits into the population, introgression
326 could potentially facilitate a branching process due to competition for resources (although we are
327 not aware of a model that assesses this precise situation, it can be inferred from the dynamics of
328 [9]). Second, introgression of novel alleles for mating preferences may provide a boost in
329 preference variation that could be an important trigger to aid the evolution of assortative mating
330 under a preference/trait mechanism, which requires preference variation to be large ([20,31]).
331 Indeed, we see exactly this pattern of secondary gene flow of olfactory alleles shortly before the
332 rapid divergence of a Cameroon cichlid radiation in Lake Ejagham [66]. Third, secondary
333 sympatry may lead to increased linkage disequilibrium between assortative mating and
334 ecological loci or among ecological loci. It seems logical that this might facilitate sympatric

335 speciation as this metric is often described as progress along the speciation continuum. However,
336 initial linkage disequilibrium has been shown not to matter much in at least some scenarios [8],
337 because without physical linkage, linkage disequilibrium will break down quickly. However,
338 physical linkage may enable these alleles to remain in association for a sufficient time for
339 assortative mating to evolve within the population (e.g., [34]). Initial linkage disequilibrium may
340 also increase the probability of allelic capture by an inversion or for selection for new mutations
341 within the inversion that may affect both ecology and assortment [32]. Increased linkage
342 disequilibrium among ecological loci may also increase the probability of sympatric divergence,
343 but this is in effect similar to varying effect sizes of alleles at ecological loci (e.g. many small
344 effect alleles within a region resemble a large-effect locus [121–124]).

345 The fundamental difference between sympatric speciation and speciation with gene flow,
346 including secondary contact scenarios, lies in the fact that very often multiple equilibrium states
347 exist in speciation models, such that loss of divergence and maintenance of divergence in the
348 presence of gene flow are both possible outcomes, depending on the starting conditions of a
349 population (this is nicely illustrated for one measure of divergence by [17], Fig. I). In such
350 cases, speciation is much more easily reached from starting conditions that match those of two
351 populations that have diverged largely in allopatry due to the large amount of allelic variation or
352 pre-existing phenotypic bimodality and assortative mating. Even for scenarios of speciation with
353 gene flow that are much easier, such as geographic separation between two incipient species that
354 are undergoing gene flow, differentiation is much more difficult to reach or maintain from an
355 initially homogeneous population than from an initially differentiated one [125,126].

356 **Fig. I**




357

358

359 ***Fig I.*** *Two equilibrium cases exist for the linkage disequilibrium (LD), a proxy for*

360 *differentiation into two distinct "species" in this model, that can be maintained between two loci*

361 *that are under disruptive selection and determine assortative mating. With little initial LD, the*

362 *one-species equilibrium is likely to be reached even when the intensity of assortment is high.*

363 *When LD in the traits is initially large, as can be the case if there is initially divergence in*

364 *allopatry, the two-species equilibrium can be reached instead. Modified from* [17].

365

366

367

368

369

370 ***Box 2.*** **Evidence for sympatric speciation from crater lake cichlid radiations**

371 There are relatively few volcanic chains of crater lakes containing fishes in the tropics, notably
372 found only in Cameroon, Nicaragua, Tanzania, Uganda, Madagascar, and Papua New Guinea
373 [63,127,128]. Although sympatric radiations of endemic fishes are known from other isolated
374 saline, alkali, and ancient lakes [129–132], only three lineages of cichlids have radiated in the
375 world's crater lakes (Fig. II). The most diverse is Barombi Mbo, Cameroon with eleven endemic
376 species, followed by Lake Bermin, Cameroon with nine [43]. Nicaraguan crater lakes reach up to
377 five species [45,55,133,134], the East African craters never exceed two sympatric species
378 [63,68], and Madagascar's several crater lakes each contain only a single endemic cichlid [128].
379 It remains unknown why regional and lineage diversity varies so greatly because there appears to
380 be no relationship with crater lake size or age (up to approximately 5 km diameter and 2 million
381 years old) until reaching the much larger sizes of the East African rift lakes ([135], but also see
382 [134]).

383 In contrast to claims in a recent review [70], the evidence for sympatric speciation with
384 secondary gene flow is rather consistently in favor and remarkably similar across all crater lake
385 cichlid radiations. In all cases examined with genomic data so far, secondary gene flow was
386 detected, but there was little evidence it came from substantial divergence in allopatry followed
387 by secondary contact. Instead, nearly all studies have concluded sympatric divergence with
388 periodic or continuous gene flow, frequently from an initial hybrid swarm population (i.e.
389 introgression from multiple outgroup populations; [54,63–65,68]).

390 We think that the best evidence for secondary gene flow as a trigger of sympatric
391 divergence in cichlids comes from a radiation of three *Coptodon* species in Lake Ejagham:

392 demographic analyses of whole genomes suggest that this population did not diversify for 8,000
393 years in the face of frequent gene flow until an influx of olfactory receptor alleles coinciding
394 with the first sympatric divergence event in the lake [66]. Similarly in Lake Victoria, segregating
395 opsin alleles in riverine cichlid populations were differentially sorted among Lake Victorian
396 cichlids and may have triggered their diversification [25].

397 Evidence for sympatric divergence in crater lake cichlids without the aid of secondary
398 gene flow remains elusive. Malinsky et al. [63] showed that introgression and hybrid swarm
399 predated the divergence of a shallow/deep-water pair of cichlids in Lake Massoko, Tanzania;
400 however, these ancestral segregating haplotypes may have later aided sympatric divergence
401 (which admittedly is very difficult to rule out – see text). Very recent sympatric divergence in
402 some crater lakes on the order of thousands of years may also suggest that ongoing divergence
403 occurred in sympatry without the aid of gene flow [133,136]; however, it remains unclear if this
404 incipient divergence will become stalled as in other sympatric radiations [24]. Very rare
405 secondary gene flow without a clear functional role into the Barombi Mbo cichlid radiation ( <
406 1% introgressed regions in nearly all species) provides weak evidence of sympatric divergence,
407 but more functional characterization and timing of introgression is needed [67]. The recent
408 advent of transgenic reporters, CRISPR-Cas9, and *in situ* hybridization genetic tools within
409 Nicaraguan crater lake cichlids provides much promise for future investigations of the role of
410 introgression in sympatric divergence [99,137].

411

412

413

**Fig. II**

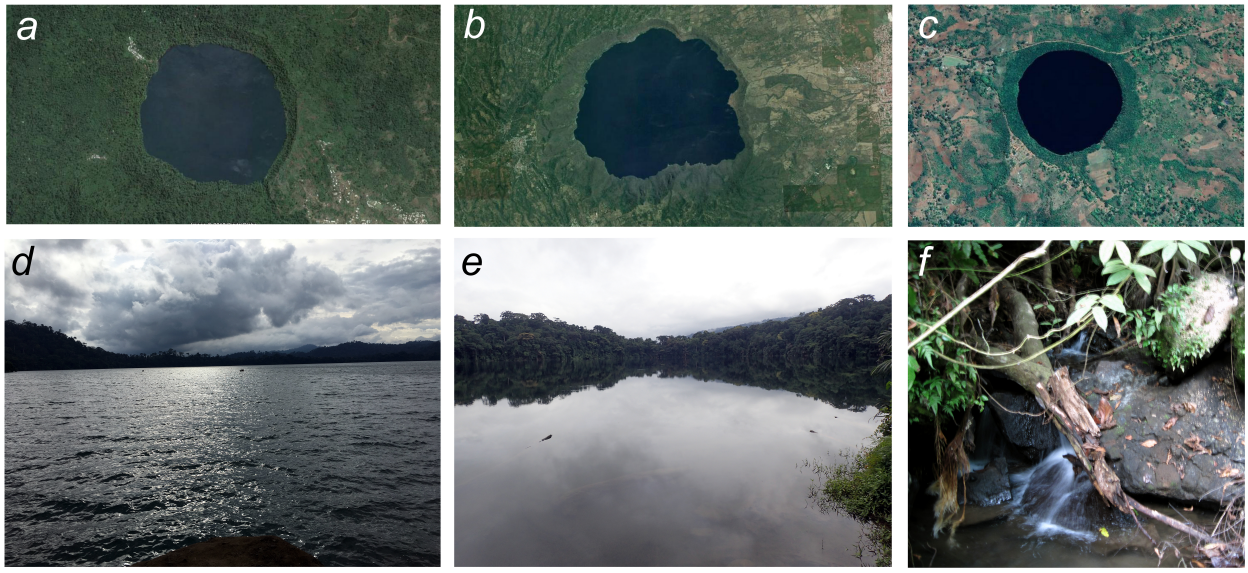



**Fig. II** Examples of volcanic crater lakes containing endemic cichlid radiations around the globe: *a,d,f*) Barombi Mbo, Cameroon and its only outlet stream; *b*) Lake Apoyo, Nicaragua, *c*) Lake Massoko, Tanzania, *e*) Lake Bermin, Cameroon. Satellite images (a-c) from Google Earth; other images by CHM.

429 ***Box 3.*** **Tools for detecting and dating local gene flow and selective sweeps**

430 *1) Detecting and dating local secondary gene flow*

431 While there are a variety of tests to detect gene flow on a local scale or within sliding genomic
432 windows [23,138–142], dating gene flow events is still accomplished mostly at the level of the
433 entire genome (although see [141] for dating introgression relative to divergence). Currently,
434 three types of demographic **coalescent** modeling approaches can infer local gene flow timing
435 based on different information from genomic data: 1) the distribution of allele frequencies from
436 genotype data (site frequency spectrum: [143,144]), 2) the distribution of haplotype block
437 lengths from phased genomes [145–148], and 3) variation in coalescent patterns among gene
438 trees [149]. Alternative approaches using **conditional random fields** [150] and **hidden Markov**
439 **models** [151,152] have also been used to detect and date gene flow events.

440

441 *2) Dating selective sweeps and ages of beneficial alleles:*

442 Methods for estimating the age of a sweep of a beneficial allele exploit several different aspects
443 about the pattern of variation surrounding the allele on its haplotypic background. These include
444 heuristic approaches that use point estimates of mean haplotype length or the number of derived
445 mutations within a chosen distance of the site [153–156] and model-based approaches that use
446 demographic information and summary statistics of allele frequencies and linkage disequilibrium
447 to model a distribution of ages that fit the observed data [157–160]. Alternatively, full sequence
448 data about haplotype structure on chromosomes and models that leverage the length of ancestral
449 haplotypes surrounding the beneficial allele and the accumulation of derived mutations can be
450 used to estimate the age of beneficial alleles [161–163].

3) *Functional genetic analyses of introgressed variants*

Functional assessments of introgressed regions minimally involve searching an annotated reference genome for genes with relevant functions known for model organisms (or a pipeline for assembling and annotating the organism of interest if an annotated reference is not already available (reviewed in [164,165]). Introgressed regions that are unannotated can be searched for evidence of potential functional importance based on strong sequence conservation across taxa [97] or potential regulatory elements (reviewed in [166]). Additionally, genome wide association studies (GWAS) can highlight variants in introgressed regions that may underlie complex traits of interest, including novel variants previously unknown in model organisms (GWAS reviewed in [167–169]). Functional validation of gene and regulatory element variants through genome-editing experiments is also becoming increasingly tractable for non-model organisms (e.g. [98,99,101,170]).

**Fig. 1**

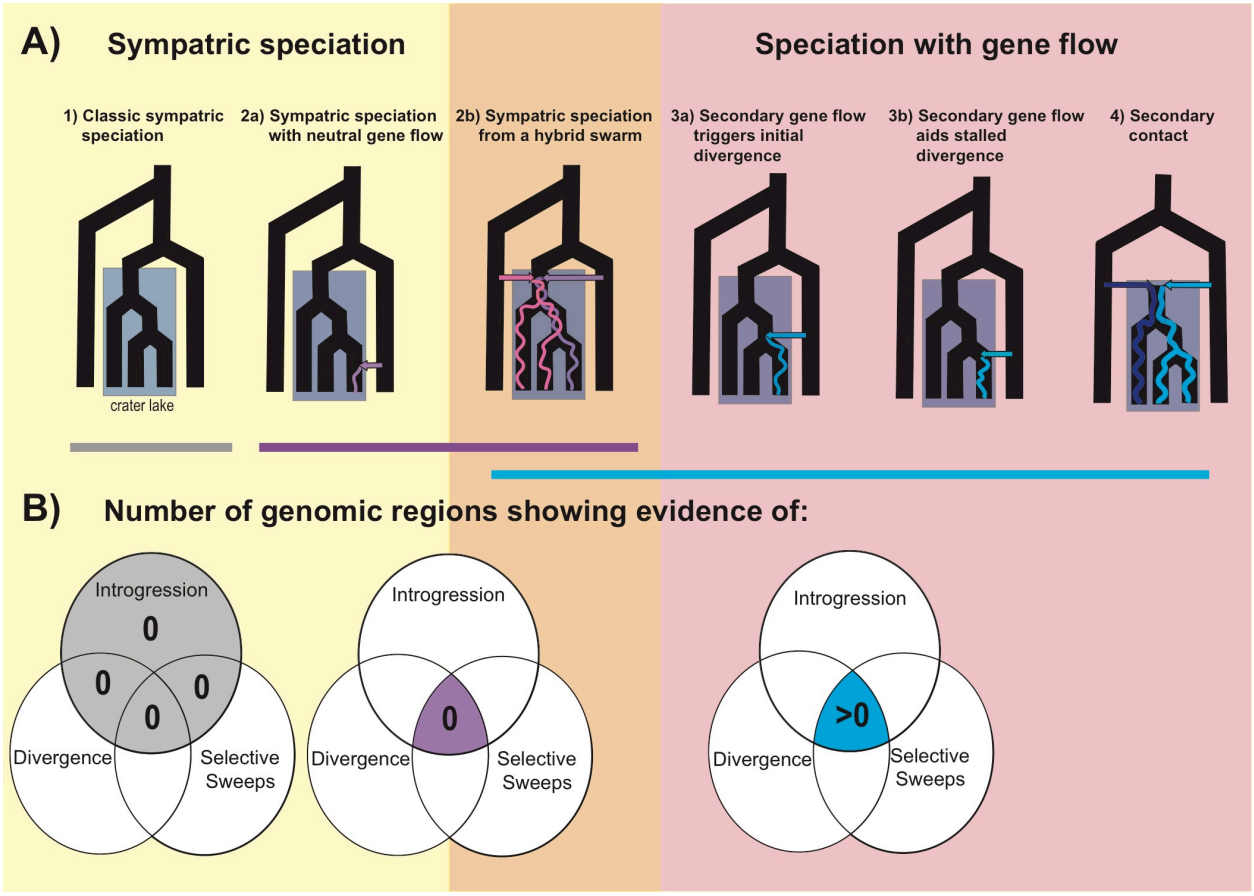



**Figure 1: Genomic signatures of sympatric speciation and speciation with gene flow.** Speciation scenarios are grouped into sympatric speciation scenarios (yellow box) and speciation with gene flow scenarios (red box). Speciation from a hybrid swarm (orange box) can fall under either category and additional information is necessary to determine what category of speciation models best describe this process. **A)** The timing of gene flow relative to divergence can be used to distinguish between speciation scenarios. The colored arrows represent gene flow events and the colored lines within the tree represent a signature of introgression from that gene flow event into the sympatric species. **B)** Venn diagrams illustrating the number of genomic windows across the entire genome expected to have overlapping signatures of introgression (e.g. $f_d$ outliers), genetic divergence (e.g. $F_{st}$ and $D_{xy}$ outliers), and selective sweeps (e.g. SweeD) for each

speciation scenario (e.g. see [77]). The highlighted sections of the Venn diagrams indicate the key signature that can be used to distinguish between the scenarios. The scenarios that are expected to leave very similar signatures of overlap are grouped by the bars colored with their respective Venn diagram.

**Glossary Box**

**Coalescence**: The event of two sampled lineages from different populations merging back in time in a shared ancestral lineage.

$\boldsymbol{D_{xy}}$: A measure of absolute genetic divergence between populations calculated as the average number of pairwise differences between sequences from two populations, excluding all the comparisons between sequences within populations.

**Hybrid swarm**: A genetically diverse population with unique allele combinations derived from the hybridization of multiple distinct taxa and subsequent backcrossing with hybrids and crossing between hybrids themselves.

**Hidden Markov model**: A statistical modeling approach used to infer hidden states from observed data along a sequence, where each hidden variable is independent of all others and conditional only on the state of the previous hidden variable.

**Conditional random field**: A statistical modeling approach similar to hidden Markov models except that each hidden variable can be conditional on regional hidden variables, not just the immediately previous one.

**Incomplete lineage sorting**: The imperfect sorting of ancestral alleles between diverging lineages that creates variable signatures about the evolutionary relationships among organisms.

**Introgression**: The movement and incorporation of genetic material from one distinct lineage into another upon hybridization between the two and subsequent backcrossing with one of the parent species.

**Linkage disequilibrium**: A non-random association of alleles at two or more loci.

**Monophyletic**: A group of lineages where the most recent ancestor of the group is not an ancestor of any lineages outside the group.

**Secondary gene flow**: Any gene flow event from non-sympatric populations after the initial colonization of the area that sympatric sister species diverged in. Introgression into the diverging sister species following such events potentially brings in variation that has evolved in allopatry that can aid the speciation process.

**Transgressive segregation**: The formation of extreme phenotypes in a segregating hybrid population that are outside the range of phenotypes observed in parental species.

**Acknowledgements**

We thank J. McGirr, M. St. John, J. Poelstra, B. Reinhard, and J. Hermisson for valuable discussion of content in this manuscript. This work was supported by the National Science Foundation DEB CAREER Grant #1749764 and by the University of North Carolina to CHM.

## References

1 Coyne, J.A. and Orr, H.A. (2004) *Speciation*, Sinauer Associates.
2 Mallet, J. *et al.* (2009) Space, sympatry and speciation. *J. Evol. Biol.* 22, 2332–41
3 Fitzpatrick, B.M. *et al.* (2009) Pattern, process and geographic modes of speciation. *J. Evol. Biol.* 22, 2342–2347
4 Fitzpatrick, B.M. *et al.* (2008) What, if anything, is sympatric speciation? *J. Evol. Biol.* 21, 1452–1459
5 Bolnick, D.I. and Fitzpatrick, B.M. (2007) Sympatric Speciation: Models and Empirical Evidence. *Annu. Rev. Ecol. Evol. Syst.* 38, 459–487
6 Nosil, P. and Schluter, D. (2011) The genes underlying the process of speciation. *Trends Ecol. Evol.* 26, 160–167
7 Ravinet, M. *et al.* (2017) Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *J. Evol. Biol.* 30, 1450–1477
8 Felsenstein, J. (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution (N. Y).* 35, 124–138
9 Dieckmann, U. and Doebeli, M. (1999) On the origin of species by sympatric speciation. *Nature* 400, 354
10 Gavrilets, S. (2004) *Fitness landscape and origin of species*, Princeton University Press.
11 Doebeli, M. *et al.* (2005) What we have also learned: adaptive speciation is theoretically plausible. *Evolution (N. Y).* 59, 691–699
12 Matessi, C. *et al.* (2002) Long-term Buildup of Reproductive Isolation Promoted by Disruptive Selection: How Far Does it Go? *Selection* 2, 41–64
13 Bolnick, D.I. and Doebeli, M. (2003) Sexual dimorphism: two sides of the same ecological coin. *Evolution (N. Y).* 57, 2433–2449
14 Bolnick, D.I. (2006) Multi-species outcomes in a common model of sympatric speciation. *J. Theor. Biol.* 241, 734–744
15 Bürger, R. and Schneider, K.A. (2006) Intraspecific Competitive Divergence and Convergence under Assortative Mating. *Am. Nat.* 167, 190–205
16 Otto, S.P. *et al.* (2008) Frequency-dependent selection and the evolution of assortative mating. *Genetics* 179, 2091–2112
17 Kirkpatrick, M. and Ravigné, V. (2002) Speciation by Natural and Sexual Selection: Models and Experiments. *Am. Nat.* 159, S22–S35
18 Turelli, M. *et al.* (2001) Theory and speciation. *Trends Ecol. Evol.* 16, 330–343
19 Garner, A.G. *et al.* (2018) Genomic signatures of reinforcement. *Genes (Basel).* 9,
20 van Doorn, G.S. and Weissing, F.J. (2002) Ecological versus Sexual Selection Models of Sympatric Speciation: A Synthesis. *Selection* 2, 17–40

21 Seehausen, O. (2004) Hybridization and adaptive radiation. *Trends Ecol. Evol.* 19, 198–207
22 Seehausen, O. (2013) Conditions when hybridization might predispose populations for adaptive radiation. *J. Evol. Biol.* 26, 279–281
23 Turissini, D.A. and Matute, D.R. (2017) Fine scale mapping of genomic introgressions within the Drosophila yakuba clade. *PLOS Genet.* 13, e1006971
24 Martin, C.H. (2012) Weak disruptive selection and incomplete phenotypic divergence in two classic examples of sympatric speciation: Cameroon crater lake cichlids. *Am. Nat.* 180, E90–E109
25 Meier, J.I. *et al.* (2017) Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun.* 8, 1–11
26 Wagner, C.E. and McCune, A.R. (2009) Contrasting patterns of spatial genetic structure in sympatric rock-dwelling cichlid fishes. *Evolution (N. Y).* 63, 1312–1326
27 Servedio, M.R. and Boughman, J.W. (2017) The Role of Sexual Selection in Local Adaptation and Speciation. *Annu. Rev. Ecol. Evol. Syst.* 48, annurev-ecolsys-110316-022905
28 Kopp, M. *et al.* (2018) Mechanisms of Assortative Mating in Speciation with Gene Flow: Connecting Theory and Empirical Research. *Am. Nat.* 191, 1–20
29 Servedio, M.R. *et al.* (2011) Magic traits in speciation: 'magic' but not rare? *Trends Ecol. Evol.* 26, 389–397
30 Turner, G.F. and Burrows, M.T. (1995) A model of sympatric speciation by sexual selection. *Proc. R. Soc. London. Ser. B Biol. Sci.* 260, 287–292
31 Weissing, F.J. *et al.* (2011) Adaptive speciation theory: A conceptual review. *Behav. Ecol. Sociobiol.* 65, 461–480
32 Kirkpatrick, M. and Barton, N.H. (2006) Chromosome inversions, local adaptation and speciation. *Genetics* 173, 419–434
33 Feder, J.L. and Nosil, P. (2009) Chromosomal inversions and species differences: When are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions? *Evolution (N. Y).* 63, 3061–3075
34 Servedio, M. and Bürger, R. The Effects on Parapatric Divergence of Linkage between Preference and Trait Loci versus Pleiotropy. , *Genes* , 9. (2018)
35 Feder, J.L. *et al.* (2003) Allopatric genetic origins for sympatric host-plant shifts and race formation in Rhagoletis. *Proc. Natl. Acad. Sci.* 100, 10314–10319
36 Fuller, Z. *et al.* The role of chromosomal inversions in speciation. *biorxiv.org* at <https://www.biorxiv.org/content/early/2017/11/24/211771.abstract>
37 Kirkpatrick, M. *et al.* Chromosome inversions, local adaptation and speciation. *Genet. Soc Am.* at <http://www.genetics.org/content/173/1/419.short>
38 Martin, C.H. (2013) Strong assortative mating by diet, color, size, and morphology but limited progress toward sympatric speciation in a classic example: Cameroon crater lake cichlids. *Evolution (N. Y).* 67, 2114–23
39 Bolnick, D.I. (2011) Sympatric Speciation in Threespine Stickleback: Why Not? *Int. J. Ecol.* 2011, 1–15
40 Gavrilets, S. *et al.* (2007) Case studies and mathematical models of ecological speciation. 1. Cichlids in a crater lake. *Mol. Ecol.* 16, 2893–909
41 Gavrilets, S. and Vose, A. (2007) Case studies and mathematical models of ecological speciation. 2. Palms on an oceanic island. *Mol. Ecol.* 16, 2910–21
42 Gavrilets, S. (2014) Models of speciation: Where are we now? *J. Hered.* 105, 743–755
43 Schliewen, U.K. *et al.* (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature* 368, 629–631
44 Schliewen, U.K. and Klee, B. (2004) Reticulate sympatric speciation in Cameroonian crater lake cichlids. *Front. Zool.* 1, 9–14
45 Barluenga, M. *et al.* (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439, 719–723
46 Hahn, M.W. and Nakhleh, L. (2015) Irrational exuberance for resolved species trees. *Evolution (N.*

*Y).* DOI: 10.1111/evo.12832
47 Wu, C.-I. (2001) The genic view of the process of speciation. *J. Evol. Biol.* 14, 851–865
48 McGirr, J.A. and Martin, C.H. (2016) Novel candidate genes underlying extreme trophic specialization in Caribbean pupfishes. *Mol. Biol. Evol.* 34, 873–888
49 Martin, S.H. *et al.* (2013) Genome-wide evidence for speciation with gene flow in Heliconius butterflies. *Genome Res.* 23, 1817–1828
50 Fontaine, M.C. *et al.* (2015) Extensive introgression in a malaria vector species complex revealed by phylogenomics. *Science (80-. ).* 347, 1258522–1258522
51 Campbell, C.R. *et al.* (2018) What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biol. J. Linn. Soc.*
52 Papadopolus *et al.* (2011) Speciation with gene flow on Lord Howe Island. *Proc. Natl. Acad. Sci. U. S. A.* 108, 13188–93
53 Geiger, M.F. *et al.* (2013) Crater Lake Apoyo Revisited - Population Genetics of an Emerging Species Flock. *PLoS One* 8, 1–17
54 Elmer, K.R. *et al.* (2014) Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nat. Commun.* 5, 1–8
55 Barluenga, M. and Meyer, A. (2010) Phylogeography, colonization and population history of the Midas cichlid species complex (Amphilophus spp.) in the Nicaraguan crater lakes. *BMC Evol. Biol.* 10, 326
56 Kusche, H. *et al.* (2015) Sympatric ecological divergence associated with a color polymorphism. *BMC Biol.* 13,
57 Pfaender, J. *et al.* (2016) Rugged adaptive landscapes shape a complex, sympatric radiation. *Proc. R. Soc. B Biol. Sci.* 283, 11–13
58 Hadid, Y. *et al.* (2013) Possible incipient sympatric ecological speciation in blind mole rats (Spalax). *Proc. Natl. Acad. Sci.* 110, 2587–2592
59 Elmer, K.R. *et al.* (2010) Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. *Mol. Ecol.* 19, 197–211
60 Foote, A. *et al.* Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. *nature.com* at <https://www.nature.com/articles/ncomms11693>
61 Alcaide, M. *et al.* (2014) Genomic divergence in a ring species complex. *Nature* 511, 83–85
62 Igea, J. *et al.* (2015) A comparative analysis of island floras challenges taxonomy-based biogeographical models of speciation. *Evolution (N. Y).* 69, 482–491
63 Malinsky, M. *et al.* (2015) Genomic islands of speciation separate Cichlid ectomorphs in an East African crater lake. *Science (80-. ).* 350, 1493–1498
64 Martin, C.H. *et al.* (2015) Complex histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on one of the clearest examples of sympatric speciation. *Evolution (N. Y).* 69, 1406–1422
65 Kautt, A.F. *et al.* (2016) Multispecies Outcomes of Sympatric Speciation after Admixture with the Source Population in Two Radiations of Nicaraguan Crater Lake Cichlids. *PLoS Genet.* 12, 1–33
66 Poelstra, J. *et al.* (2018) Speciation in sympatry with ongoing secondary gene flow and an olfactory trigger in a radiation of Cameroon cichlids. *bioRxiv* 229864,
67 Richards, E. *et al.* (2018) Don't throw out the sympatric speciation with the crater lake water: fine-scale investigation of introgression provides equivocal support for causal role of secondary gene flow in one of the clearest examples of sympatric speciation. *bioRxiv*
68 Machado-Schiaffino, G. *et al.* (2015) Parallel evolution in Ugandan crater lakes: Repeated evolution of limnetic body shapes in haplochromine cichlid fish. *BMC Evol. Biol.* 15,
69 Wu, C.-I. and Ting, C.-T. (2004) Genes and speciation. *Nat. Rev. Genet.* 5, 114–122
70 Foote, A.D. (2017) Sympatric Speciation in the Genomic Era. *Trends Ecol. Evol.* 33, 85–95
71 Martin, C.H. (2016) The cryptic origins of evolutionary novelty : 1000-fold faster trophic diversification rates without increased ecological opportunity or hybrid swarm. DOI: 10.1111/evo.13046

72 Dasmahapatra, K.K. *et al.* (2012) Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487, 94–98
73 Hedrick, P.W. (2013) Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol. Ecol.* 22, 4606–4618
74 Huerta-Sanchez, E. *et al.* (2014) Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512, 194–197
75 Stankowski, S. and Streisfeld, M.A. (2015) Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. *Proc. R. Soc. B Biol. Sci.* 282,
76 Juan, E. *et al.* (2017) Evolution of novel mimicry rings facilitated by adaptive introgression in tropical butterflies. *Mol. Ecol.* 26, 5160–5172
77 Richards, E.J. and Martin, C.H. (2017) Adaptive introgression from distant Caribbean islands contributed to the diversification of a microendemic adaptive radiation of trophic specialist pupfishes. *PLOS Genet.* 13, e1006919
78 Rieseberg, L.H. *et al.* (1999) Transgressive segregation, adaptation and speciation. *Heredity (Edinb).* 83, 363–372
79 Kagawa, K. and Takimoto, G. (2017) Hybridization can promote adaptive radiation by means of transgressive segregation. *Ecol. Lett.* 21, 264–274
80 Schumer, M. *et al.* (2014) How common is homoploid hybrid speciation? *Evolution (N. Y).* 68, 1553–1560
81 Fishman, L. *et al.* (2013) Chromosomal rearrangements and the genetics of reproductive barriers in mimulus (monkey flowers). *Evolution (N. Y).* 67, 2547–2560
82 Guerrero, R.F. and Hahn, M.W. (2017) Speciation as a sieve for ancestral polymorphism. *Mol. Ecol.* DOI: 10.1111/mec.14290
83 Nelson, T.C. and Cresko, W.A. (2018) Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evol. Lett.* 2, 9–21
84 Durand, E.Y. *et al.* (2011) Testing for ancient admixture between closely related populations. *Mol. Biol. Evol.* 28, 2239–2252
85 Smith, J.M. and Haigh, J. (1974) The hitch-hiking effect of a favourable gene. *Genet. Res.* 23, 23–35
86 Ravinet, M. *et al.* (2017) Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *J. Evol. Biol.* 30, 1450–1477
87 Nachman, M.W. and Payseur, B.A. (2012) Recombination rate variation and speciation : theoretical predictions and empirical results from rabbits and mice. DOI: 10.1098/rstb.2011.0249
88 Cruickshank, T.E. and Hahn, M.W. (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* 23, 3133–3157
89 Bay, R.A. and Ruegg, K. (2017) Genomic islands of divergence or opportunities for introgression? *Proc. R. Soc. B Biol. Sci.* 284,
90 Clarkson, C.S. *et al.* (2014) Adaptive introgression between Anopheles sibling species eliminates a major genomic island but not reproductive isolation. *Nat. Commun.* 5, 4248
91 Stölting, K.N. *et al.* (2015) Genome-wide patterns of differentiation and spatially varying selection between postglacial recolonization lineages of *Populus alba* (Salicaceae), a widespread forest tree. *New Phytol.* 207, 723–734
92 Felsenstein, J. (1981) Skepticism towards Santa-Rosalia or why are there so few kinds of animals. *Evolution (N. Y).* 35, 124–138
93 Ortíz-Barrientos, D. and Noor, M.A.F. (2005) Evidence for a one-allele assortative mating locus. *Science* 310, 1467
94 Schrider, D.R. and Kern, A.D. (2016) S/HIC: Robust Identification of Soft and Hard Sweeps Using Machine Learning. *PLOS Genet.* 12, e1005928
95 Kern, A.D. and Schrider, D.R. (2018) diploS/HIC: An Updated Approach to Classifying Selective Sweeps. *G3 Genes|Genomes|Genetics*
96 Boyle, E.A. *et al.* (2017) An Expanded View of Complex Traits: From Polygenic to Omnigenic.

*Cell* 169, 1177–1186
97 Siepel, A. *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* 15, 1034–1050
98 Erickson, P.A. *et al.* (2016) Microinjection for transgenesis and genome editing in threespine sticklebacks. *J. Vis. Exp.* DOI: 10.3791/54055
99 Kratochwil, C.F. *et al.* (2017) Tol2 transposon-mediated transgenesis in the Midas cichlid (Amphilophus citrinellus) — towards understanding gene function and regulatory evolution in an ecological model system for rapid phenotypic diversification. *BMC Dev. Biol.* 17, 15
100 Hart, J.C. and Miller, C.T. (2017) Sequence-Based Mapping and Genome Editing Reveal Mutations in Stickleback <em>Hps5</em> Cause Oculocutaneous Albinism and the <em>casper</em> Phenotype. *G3 Genes|Genomes|Genetics* 7, 3123 LP-3131
101 Cleves, P.A. *et al.* (2018) An intronic enhancer of Bmp6 underlies evolved tooth gain in sticklebacks. *PLOS Genet.* 14, e1007449
102 Kronforst, M.R. *et al.* (2013) Hybridization Reveals the Evolving Genomic Architecture of Speciation. *Cell Rep.* 5, 666–677
103 Renaut, S. *et al.* (2013) Genomic islands of divergence are not affected by geography of speciation in sunflowers. *Nat. Commun.* 4, 1–8
104 Burri, R. *et al.* (2015) Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Res.* 25, 1656–1665
105 Irwin, D.E. *et al.* (2016) Recurrent selection explains parallel evolution of genomic regions of high relative but low absolute differentiation in a ring species. *Mol. Ecol.* 25, 4488–4507
106 Burri, R. (2017) Interpreting differentiation landscapes in the light of long-term linked selection. *Evol. Lett.* 1, 118–131
107 Vijay, N. *et al.* (2017) Genomewide patterns of variation in genetic diversity are shared among populations, species and higher-order taxa. *Mol. Ecol.* 26, 4284–4295
108 Van Doren, B.M. *et al.* (2017) Correlated patterns of genetic diversity and differentiation across an avian family. *Mol. Ecol.* 26, 3982–3997
109 Ma, T. *et al.* (2018) Ancient polymorphisms and divergence hitchhiking contribute to genomic islands of divergence within a poplar species complex. *Proc. Natl. Acad. Sci.* 115, E236 LP-E243
110 Delmore, K.E. *et al.* (2018) Comparative analysis examining patterns of genomic differentiation across multiple episodes of population divergence in birds. *Evol. Lett.* 2, 76–87
111 Bush, G.L. (1975) Modes of Animal Speciation. *Annu. Rev. Ecol. Syst.* 6, 339–364
112 Sorenson, M.D. *et al.* (2003) Speciation by host switch in brood parasitic indigobirds. 424, 31–32
113 Savolainen, V. *et al.* (2006) Sympatric speciation in palms on an oceanic island. *Nature* 441, 210–213
114 Kondrashov, A.S. and Kondrashov, F.A. (1999) Interactions among quantitative traits in the course of sympatric speciation. *Nature* 400, 351–354
115 Martin, C.H. and Turner, B.J. (2018) Long-distance dispersal over land by fishes: extremely rare ecological events become probable over millennial timescales. *Proc. R. Soc. B Biol. Sci.* 285,
116 Martin, C.H. *et al.* (2016) Diabolical survival in Death Valley: recent pupfish colonization, gene flow, and genetic assimilation in the smallest species range on earth. *Proc. R. Soc. B Biol. Sci.* 283, 23–34
117 A., C.J. and Allen, O.H. (1989) Patterns of speciation in Drosophila. *Evolution (N. Y).* 43, 362–381
118 Polechová, J. *et al.* (2005) SPECIATION THROUGH COMPETITION: A CRITICAL REVIEW. *Evolution (N. Y).* 59, 1194–1210
119 Norvaišas, P. and Kisdi, E. (2012) Revisiting Santa Rosalia to Unfold a Degeneracy of Classic Models of Speciation. *Am. Nat.* 180, 388–393
120 Kopp, M. and Hermisson, J. (2008) Competitive speciation and costs of choosiness. *J. Evol. Biol.* 21, 1005–1023

121 Yeaman, S. and Whitlock, M.C. (2011) The genetic architecture of adaptation under migration-selection balance. *Evolution (N. Y).* 65, 1897–1911
122 Bürger, R. and Akerman, A. (2011) The effects of linkage and gene flow on local adaptation: A two-locus continent–island model. *Theor. Popul. Biol.* 80, 272–288
123 Akerman, A. and Bürger, R. (2014) The consequences of gene flow for local adaptation and differentiation: a two-locus two-deme model. *J. Math. Biol.* 68, 1135–1198
124 Yeaman, S. *et al.* (2016) The evolution of genomic islands by increased establishment probability of linked alleles. *Mol. Ecol.* 25, 2542–2558
125 Cotto, O. and Servedio, M.R. (2017) The Roles of Sexual and Viability Selection in the Evolution of Incomplete Reproductive Isolation: From Allopatry to Sympatry. *Am. Nat.* 190, 680–693
126 Sachdeva, H. and Barton, N.H. (2017) Divergence and evolution of assortative mating in a polygenic trait model of speciation with gene flow. *Evolution (N. Y).* 71, 1478–1493
127 Seehausen, O. (2006) African cichlid fish: A model system in adaptive radiation research. *Proc. R. Soc. B Biol. Sci.* 273, 1987–1998
128 Sparks, J.S. (2004) Molecular phylogeny and biogeography of the Malagasy and South Asian cichlids (Teleostei: Perciformes: Cichlidae). *Mol. Phylogenet. Evol.* 30, 599–614
129 Martin, C.H. and Wainwright, P.C. (2013) A Remarkable Species Flock of Cyprinodon Pupfishes Endemic to San Salvador Island, Bahamas. *Bull. Peabody Museum Nat. Hist.* 54, 231–241
130 Ford, A.G.P. *et al.* (2015) High levels of interspecific gene flow in an endemic cichlid fish adaptive radiation from an extreme lake environment. *Mol. Ecol.* 24, 3421–3440
131 Herder, F. *et al.* (2008) Adaptive sympatric speciation of polychromatic “roundfin” sailfin silverside fish in Lake Matano (Sulawesi). *Evolution (N. Y).* 62, 2178–2195
132 Sibbing, F.A. *et al.* (1998) Speciation of endemic Lake Tana barbs (Cyprinidae, Ethiopia) driven by trophic resource partitioning; a molecular and ecomorphological approach. *Aquat. Ecol.* 32, 217–227
133 Elmer, K.R. *et al.* (2010) Rapid sympatric ecological differentiation of crater lake cichlid fishes within historic times. *BMC Biol.* 8,
134 Kautt, A.F. *et al.* (2018) Lessons from a natural experiment : Allopatric morphological divergence and sympatric diversification in the Midas cichlid species complex are largely influenced by ecology in a deterministic way. *Evol. Lett.* DOI: 10.1002/evl3.64
135 Wagner, C.E. *et al.* (2014) Cichlid species-area relationships are shaped by adaptive radiations that scale with area. *Ecol. Lett.* 17, 583–592
136 Kautt, A.F. *et al.* (2016) Incipient sympatric speciation in Midas cichlid fish from the youngest and one of the smallest crater lakes in Nicaragua due to differential use of the benthic and limnetic habitats? *Ecol. Evol.* 6, 5342–5357
137 Kratochwil, C.F. *et al.* (2015) Embryonic and larval development in the Midas cichlid fish species flock (Amphilophus spp.): A new evo-devo model for the investigation of adaptive novelties and species differences Evolutionary developmental biology. *BMC Dev. Biol.* 15, 1–15
138 Martin, S.H. *et al.* (2015) Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol. Biol. Evol.* 32, 244–257
139 Rosenzweig, B.K. *et al.* (2016) Powerful methods for detecting introgressed regions from population genomic data. *Mol. Ecol.* 25, 2387–2397
140 Schrider, D.R. *et al.* (2018) Supervised machine learning reveals introgressed loci in the genomes of Drosophila simulans and D. sechellia. *PLOS Genet.* 14, e1007341
141 Hibbins, M.S. and Hahn, M.W. (2018) Population genetic tests for the direction and relative timing of introgression. *bioRxiv*
142 Elworth, R. *et al.* (2018) DGEN: A Test Statistic for Detection of General Introgression Scenarios. *bioRxiv*
143 Excoffier, L. *et al.* (2013) Robust Demographic Inference from Genomic and SNP Data. *PLoS Genet.* 9,
144 Gutenkunst, R.N. *et al.* (2009) Inferring the Joint Demographic History of Multiple Populations

from Multidimensional SNP Frequency Data. *PLOS Genet.* 5, e1000695
145 Pool, J.E. and Nielsen, R. (2009) Inference of Historical Changes in Migration Rate From the Lengths of Migrant Tracts. *Genetics* 181, 711 LP-719
146 Gravel, S. (2012) Population Genetics Models of Local Ancestry. *Genetics* 191, 607 LP-619
147 Loh, P.-R. *et al.* (2013) Inferring Admixture Histories of Human Populations Using Linkage Disequilibrium. *Genetics* 193, 1233 LP-1254
148 Vernot, B. and Akey, J.M. (2014) Resurrecting Surviving Neandertal Lineages from Modern Human Genomes. *Science (80-. ).*
149 Gronau, I. *et al.* (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* 43, 1031
150 Patterson, N. *et al.* (2014) The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* 507, 354–7
151 Prüfer, K. *et al.* (2013) The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505, 43
152 Seguin-Orlando, A. *et al.* (2014) Genomic structure in Europeans dating back at least 36,200 years. *Science (80-. ).* 346, 1113 LP-1118
153 Thomson, R. *et al.* (2000) Recent common ancestry of human Y chromosomes: Evidence from DNA sequence data. *Proc. Natl. Acad. Sci.* 97, 7360 LP-7365
154 Tang, H. *et al.* (2002) Frequentist Estimation of Coalescence Times From Nucleotide Sequence Data Using a Tree-Based Partition. *Genetics* 161, 447 LP-459
155 Hudson, R.R. (2007) The Variance of Coalescent Time Estimates from DNA Sequences. *J. Mol. Evol.* 64, 702–705
156 Coop, G. *et al.* (2008) The Timing of Selection at the Human FOXP2 Gene. *Mol. Biol. Evol.* 25, 1257–1259
157 Peter, B.M. *et al.* (2012) Distinguishing between Selective Sweeps from Standing Variation and from a De Novo Mutation. *PLOS Genet.* 8, e1003011
158 Beleza, S. *et al.* (2013) The Timing of Pigmentation Lightening in Europeans. *Mol. Biol. Evol.* 30, 24–35
159 Nakagome, S. *et al.* (2016) Estimating the Ages of Selection Signals from Different Epochs in Human History. *Mol. Biol. Evol.* 33, 657–669
160 Ormond, L. *et al.* (2016) Inferring the age of a fixed beneficial allele. *Mol. Ecol.* 25, 157–169
161 Chen, H. and Slatkin, M. (2013) Inferring Selection Intensity and Allele Age from Multilocus Haplotype Structure. *G3 Genes,Genomes,Genetics* 3, 1429–1442
162 Chen, H. *et al.* (2015) A hidden Markov model for investigating recent positive selection through haplotype structure. *Theor. Popul. Biol.* 99, 18–30
163 Smith, J. *et al.* (2018) Estimating time to the common ancestor for a beneficial allele. *Mol. Biol. Evol.* 35, 1003–1017
164 Ekblom, R. and Wolf, J.B.W. (2014) A field guide to whole-genome sequencing, assembly and annotation. *Evol. Appl.* 7, 1026–1042
165 Matz, M. V (2018) Fantastic Beasts and How To Sequence Them: Ecological Genomics for Obscure Model Organisms. *Trends Genet.* 34, 121–132
166 Chatterjee, S. and Ahituv, N. (2017) Gene Regulatory Elements, Major Drivers of Human Disease. *Annu. Rev. Genomics Hum. Genet.* 18, 45–63
167 Witte, J.S. (2010) Genome-Wide Association Studies and Beyond. *Annu. Rev. Public Health* 31, 9–20
168 Price, A.L. *et al.* (2015) Progress and promise in understanding the genetic basis of common diseases. *Proc. R. Soc. B Biol. Sci.* 282, 20151684
169 Gibson, G. (2018) Population genetics and GWAS: A primer. *PLoS Biol.* 16, 1–6
170 Hart, J.C. and Miller, C.T. (2017) Sequence-based mapping and genome editing reveal mutations in stickleback Hps5 cause oculocutaneous albinism and the Casper phenotype. *G3 Genes|Genomes|Genetics* 7, 3123 LP-3131

881