# Discovery of ongoing selective sweeps within *Anopheles* mosquito populations using deep learning

Short title: Classifying partial sweeps in *Anopheles* populations via deep learning

Alexander T. Xue[1,2,*], Daniel R. Schrider[3], Andrew D. Kern[1,4], Ag1000G Consortium

[1] Human Genetics Institute of New Jersey and Department of Genetics, Rutgers University, 604 Allison Road, Piscataway, NJ 08854

[2] Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, NY 11724

[3] Department of Genetics, University of North Carolina, 120 Mason Farm Road, Chapel Hill, NC 27514

[4] Institute of Ecology and Evolution, 5289 University of Oregon, Onyx Bridge 272, Eugene, OR 97403

[*] Corresponding Author: xanderxue@gmail.com

## Abstract

Identification of partial sweeps, which include both hard and soft sweeps that have not currently reached fixation, provides crucial information about ongoing evolutionary responses. To this end, we introduce a deep learning approach that uses a convolutional neural network for image processing, which is trained with coalescent simulations incorporating population-specific history, to discover selective sweeps from population genomic data. This approach distinguishes between completed versus partial sweeps, hard versus soft sweeps, and regions directly affected by selection versus those merely linked to nearby selective sweeps. We perform several simulation experiments under various demographic scenarios to demonstrate the performance of our deep learning classifier *partialS/HIC*, which exhibits unprecedented resolution for detecting partial sweeps. We also apply our method to whole genomes from eight mosquito populations sampled across sub-Saharan Africa by the *Anopheles gambiae* 1000 Genomes Consortium, elucidating both continent-wide patterns as well as sweeps unique to specific geographic regions. These populations have experienced intense insecticide exposure over the past two decades, and we observe a strong overrepresentation of sweeps at insecticide resistance loci. Our analysis thus provides a catalog of candidate adaptive loci that may aid mosquito control efforts. More broadly, the success of our supervised machine learning approach introduces a powerful method to distinguish between completed and partial sweeps, as well as between hard and soft sweeps, under a variety of demographic scenarios. As whole-genome data rapidly accumulate for a greater diversity of organisms, *partialS/HIC* addresses an increasing demand for useful selection scan tools that can track in-progress evolutionary dynamics.

## Author Summary

Recent successful efforts to reduce malaria transmission are in danger of collapse due to evolving insecticide resistance in the mosquito vector *Anopheles gambiae*. We aim to understand the genetic basis of current adaptation to vector control efforts by deploying a novel method that can classify multiple categories of selective sweeps from population genomic data. In recent years, there has been great progress made in the identification of completed selective sweeps through the use of supervised machine learning (SML), but SML techniques have rarely been applied to partial or ongoing selective sweeps. Partial sweeps represent an important facet of evolution as they reflect present-day selection and thus may give insight into future dynamics. However, the genomic impact left by partial sweeps is more subtle than that left by completed sweeps, making such signatures more difficult to detect. To this end, we extend a recent SML method to partial sweep inference and apply it to elucidate ongoing selective sweeps from *Anopheles* population genomic samples.

## Introduction

Malaria represents an enormous burden on human health, with an estimated 214 million cases and 438,000 deaths in 2015 [1]. As mosquitos of the *Anopheles gambiae* species complex are the major vector for *Plasmodium* parasites, roughly 70% of global malaria relief budgets have been focused on mosquito control, including insecticide treated bed-nets, indoor residual spraying, and larva control through the direct modification of habitats as well as the application of larvicide. While these vector control efforts have successfully produced major reductions of malaria transmission rates over the past 15 years [1], there has been an alarming increase in mosquitos resistant to insecticides, specifically pyrethroids, observed across nearly all areas of the world covered by anti-malarial efforts [2]. Pyrethroids are the only class of insecticide used in long-lasting insecticidal nets and are applied in many indoor spraying programs, thus the evolutionary innovation of resistance is a well-recognized Achilles heel of anti-malarial efforts.

The increase in insecticide resistant mosquitoes is to be expected from an evolutionary standpoint: anti-malaria control efforts exert a strong selective pressure to which mosquito populations will respond through the differential survivorship and reproduction of those individuals that can best cope with the applied insecticides. Pyrethroid resistance was reported within African malaria vectors first in Sudan during the 1970s, then later in West Africa during the 1990s, most likely stemming from accidental exposure of mosquitos to crop applications of pyrethroids [3,4]. Subsequent analysis showed this earliest resistance to be a result of mutations in the knockdown resistance locus *kdr*, which is known to contribute to pyrethroid resistance in other insect species [5]. Mutations conferring resistance at *kdr* as well as other loci have since spread throughout Africa, and threaten to nullify the gains in malaria control achieved over the past decade [6]. While control efforts are now looking toward non-pyrethroid insecticides [2,7] as well as gene drive technologies [8], it is anticipated that resistance to these control modalities will eventually evolve as well [9,10]. Hence, an important goal in the continued fight against malaria is to identify genomic targets of resistance in *Anopheles*, especially in such a way that might inform vector managers in the field.

88 Alleles that confer resistance to control efforts should rapidly increase in frequency
89 within *Anopheles* populations in a manner consistent with selective sweeps. When an allele
90 increases in frequency under selection, its linked genetic background comes with it in a
91 process known as genetic hitchhiking. Selective sweeps, through this hitchhiking effect,
92 lead to decreased levels of polymorphism [11–13], skewed allele frequency spectra [14,15],
93 and increases in linkage disequilibrium surrounding the site under selection [16].
94 Classically, methods for finding sweeps have focused on a particular aspect of genetic
95 variation, for instance observing the site frequency spectrum (SFS) at a locus and
96 comparing it to expectations under neutrality and selective sweeps [19]. More recently, the
97 field has made excellent progress in combining signals across multiple features of genetic
98 variation through supervised machine learning (SML) [20–27], which has substantially
99 improved power, accuracy, and robustness in what have been stubbornly difficult inference
100 problems within population genetics [28]. While much attention has been paid to applying
101 SML for the identification and classification of completed selective sweeps in the genome
102 [24,26], less effort has been made for using SML to identify sweeps that are incomplete
103 within a population, sometimes called partial sweeps (although see Sugden *et al.* [27] for a
104 recent example). In these cases, the beneficial allele is not currently fixed within the
105 population, thereby creating a weaker hitchhiking effect in comparison to a completed
106 sweep, and accordingly a more subtle perturbation of patterns of genetic variation [29].
107 Partial sweeps may perhaps be implicated in cases where recently initiated selective forces
108 cause presently ongoing adaptation, directional selection ceases prior to fixation, or an
109 intermediate allele frequency is favored by balancing, polygenic, and/or pleiotropic
110 selection.

111 Here we introduce *partialS/HIC*, an extension of *S/HIC* [24] and *diploS/HIC* [26] that
112 includes both hard and soft partial sweeps along with their associated linked classes (*i.e.*
113 regions adjacent to either a partial hard or soft sweep) as selection states for which a
114 genomic window can be classified. We apply our extended method to data from phase I of
115 the *Anopheles gambiae* 1000 genomes project (Ag1000G) [6], with particular emphasis on
116 discovering sweeps currently in progress that might be the result of vector control efforts
117 such as insecticide spraying. We find that our method is considerably more powerful for

finding partial sweeps than *iHS* [30], even in the face of complex population size histories such as those found among the Ag1000G samples. Application of our method to the Ag1000G data reveals a large number of partial sweeps as well as completed sweeps from standing genetic variation. Moreover, we find that our sweep candidates are highly enriched for loci that have been previously identified as contributing to insecticide resistance.

## Results

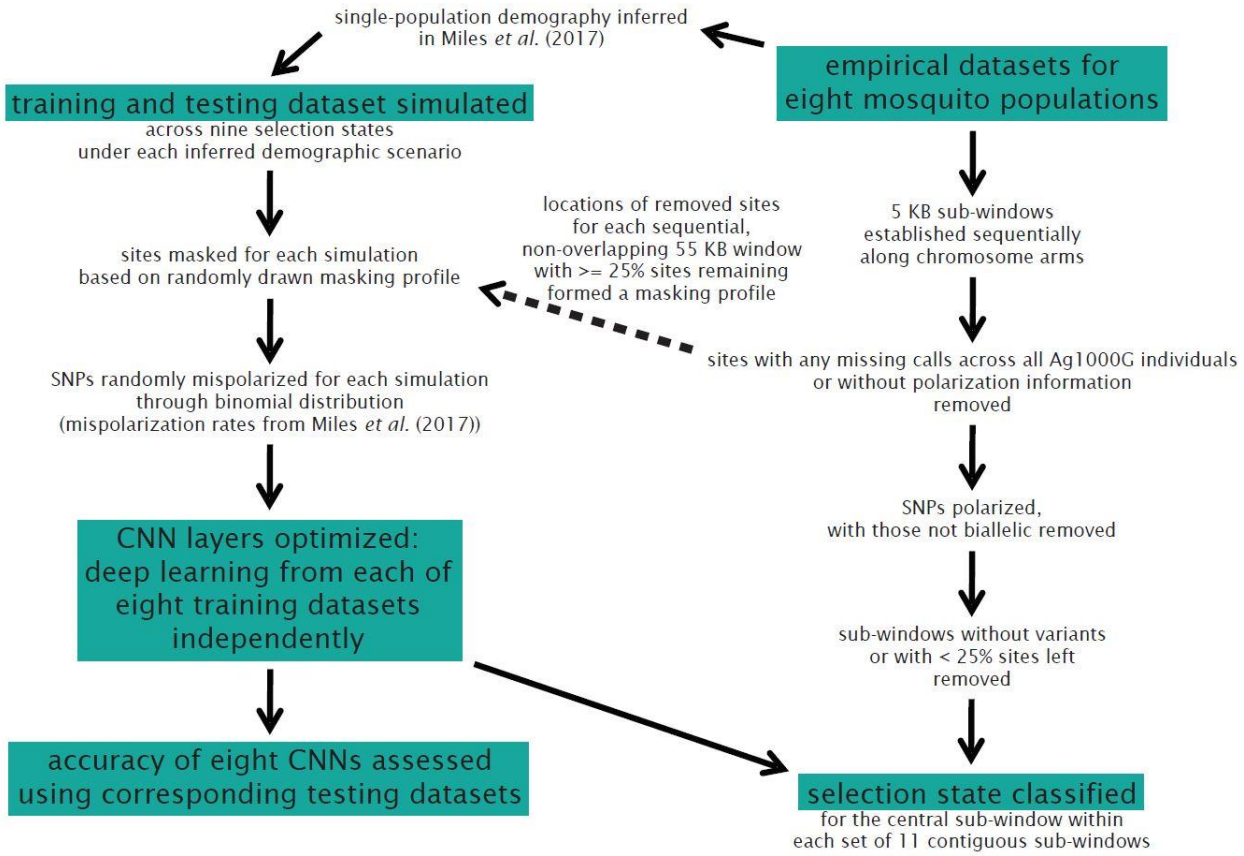



***Figure 1. Flowchart of methodological procedure.*** *Major steps highlighted in green. Note that processing of simulated data incorporated uncertainties within the empirical data, specifically with respect to masking sites from calculation of summary statistics.*

### Coalescent simulations of feature vector images for *partialS/HIC* training

131 We aimed to classify genomic segments into one of nine states: unaffected by selection (*i.e.*
132 neutral); containing a completed hard, completed soft, partial hard, or partial soft sweep,
133 respectively; or linked to a completed hard, completed soft, partial hard, or partial soft
134 sweep, respectively (Figure 1). To this end, we developed *partialS/HIC*, a SML classifier that
135 uses a deep convolutional neural network (CNN) to classify a genomic window, which is
136 represented by a two-dimensional matrix constructed from a large collection of summary
137 statistics. To train *partialS/HIC*, we deployed the program *discoal* [31] to perform
138 coalescent simulations of completed and partial as well as hard and soft selective sweeps,
139 along with simulations without sweeps, in a manner analogous to Schrider and Kern [24].
140 This was conducted for each of eight population size histories corresponding to the
141 empirical Ag1000G population datasets, which were previously inferred as part of the
142 initial data release [6]. These *Anopheles* population datasets from Miles *et al.* [6] are labeled
143 here as AOM (*A. coluzzii* from Angola), BFM (*A. coluzzii* from Burkina Faso), BFS (*A.*
144 *gambiae* from Burkina Faso), CMS (*A. gambiae* from Cameroon), GAS (*A. gambiae* from
145 Gabon), GNS (*A. gambiae* from Guinea), GWA (*Anopheles* of uncertain species from Guinea-
146 Bissau), and UGS (*A. gambiae* from Uganda). Individual simulations were converted into
147 two-dimensional matrices, or feature vector images, built from 89 rows corresponding to
148 different summary statistics, and 11 columns corresponding to adjacent sub-windows. The
149 89 statistics include 17 that are implemented in *diploS/HIC* along with 72 derivatives of the
150 recently developed SNP-specific SAFE statistic [32]. We defined the four completed
151 hard/soft and partial hard/soft selective sweep states as containing a sweep within the
152 central, focal sub-window. In contrast, the four linked selection states were defined as
153 having a sweep of the given type within one of the remaining ten sub-windows. Heatmaps
154 constructed from median values across simulations reveal expected spatial patterns, such
155 that values immediately flanking a sweep are substantially different than those further
156 from the focal sub-window, while neutral regions display no discernible pattern among
157 sub-windows (Figure S1). Additionally, spatial patterns of statistics differ qualitatively
158 between selection states. These observations are consistent regardless of mosquito
159 population history, suggesting that there is signal within this collection of summary
160 statistics to isolate the location of a sweep to a specific sub-window as well as distinguish
161 among neutral regions and types of selective sweeps.

## Deep learning excels in detecting selective sweeps, including partial hard sweeps

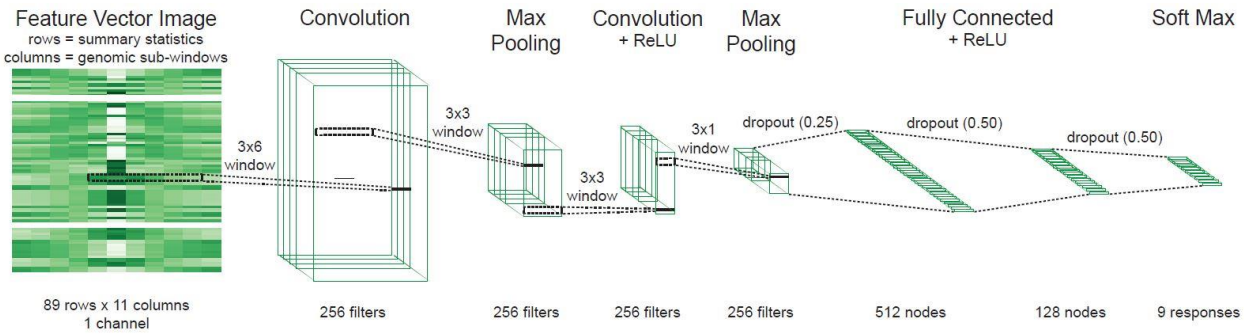



***Figure 2. CNN architecture of neural network layers.*** *The input feature vector image is passed to a 2D convolutional layer with 256 filters using a 3×6 receptive field. Next is a 2D max pooling layer given a 3×3 receptive field. Then there is a second 2D convolutional layer of 256 filters based on a 3×3 receptive field and ReLU activation. Afterward is a second 2D max pooling layer, also with a 3×3 receptive field. The tensor is then flattened and passed to two fully-connected layers with ReLU activation, resulting in 512 and 128 nodes respectively. Finally, a softmax activation layer classifies the 9 responses.*

We utilized *partialS/HIC* to train a separate CNN for nine-state classification on each of the eight demographic histories associated with the Ag1000G population samples (Figures 1–2). In order to assess accuracy, each CNN was subsequently tested against another set of simulated data that was generated under the same specifications as the training dataset. Among the eight test sets, there was moderate overall accuracy for this simulation experiment (median accuracy = 66.4%; Table S1). However, confusion matrix heatmaps provide a more informative view of our classifier's performance, which was generally sufficient for identifying neutral regions, completed hard and soft sweeps, partial hard sweeps, and regions linked to completed hard/completed soft/partial hard sweeps (Figures 3, S2). Assignment accuracy was highest for completed hard sweeps in all demographic scenarios save for AOM (median accuracy = 96.0%). We also had excellent accuracy for identification of linked completed hard regions, demonstrating a strong ability to localize completed hard sweeps. Accuracy for completed soft sweeps was also quite good (median accuracy = 84.2%); when misclassification did occur, it was generally to either the

187 neutral or partial soft sweep state, thus completed soft sweeps were rarely incorrectly
188 classified as hard sweeps. Moreover, sub-windows linked to completed soft sweeps beyond
189 one sub-window away had low levels of misclassification to one of the non-linked sweep
190 states, again allowing for excellent localization of the sweep.

191 Importantly, the purpose of *partialS/HIC* is to extend our state space to identify
192 ongoing selective sweeps while distinguishing these from completed sweeps. We find that
193 our ability to identify partial hard sweeps was generally strong across population histories
194 (median accuracy = 74.6%) and is often comparable to that of completed soft sweeps.
195 However, localization of partial hard sweeps along the chromosome was more difficult
196 than for completed sweeps, as can be seen from the moderate levels of confusion between
197 partial hard sweep and linked partial hard sweep sub-windows. Undoubtedly, this is due to
198 the limited amount of time recombination has had to whittle down the haplotype carrying
199 the beneficial mutation.

200 Identifying partial soft sweeps was a much more challenging task (median accuracy
201 = 45.3%), with a high false negative rate (median rate of misclassification as neutral =
202 27.6%) as well as a substantial probability of misclassification as a completed soft sweep
203 (median rate = 14.4%). It is encouraging though that partial soft sweeps were almost never
204 misclassified as a completed nor partial hard sweep. Additionally, while our accuracy in
205 classifying partial soft sweeps was poor, false positives were not a substantial concern
206 (median rate of misclassifying neutral regions as partial soft sweep = 3.2%). Therefore,
207 *partialS/HIC* should underestimate the true number of partial soft sweeps when applied to
208 a given dataset. Furthermore, linked partial soft sweeps beyond one sub-window away
209 from the focal sub-window were rarely mistaken for a sweep state, and likewise partial soft
210 sweeps were seldom confused for being linked (median rate of misclassification as linked
211 partial soft sweep = 4.3%), thus demonstrating that localization of partial soft sweeps may
212 be possible.

213 In summary, *partialS/HIC* has excellent ability to distinguish partial from completed
214 sweeps for *de novo* mutations, and lesser yet still substantial power for sweeps from
215 standing variation. Moreover, we demonstrated very strong performance in differentiating

between hard and soft sweeps, regardless of whether a sweep was completed or incomplete. Importantly, this is all while maintaining an acceptable false positive rate across each of the population histories tested (median accuracy for neutral regions = 85.1%; median rate of misclassifying neutral regions as any one of the four non-linked sweep states = 4.6%).

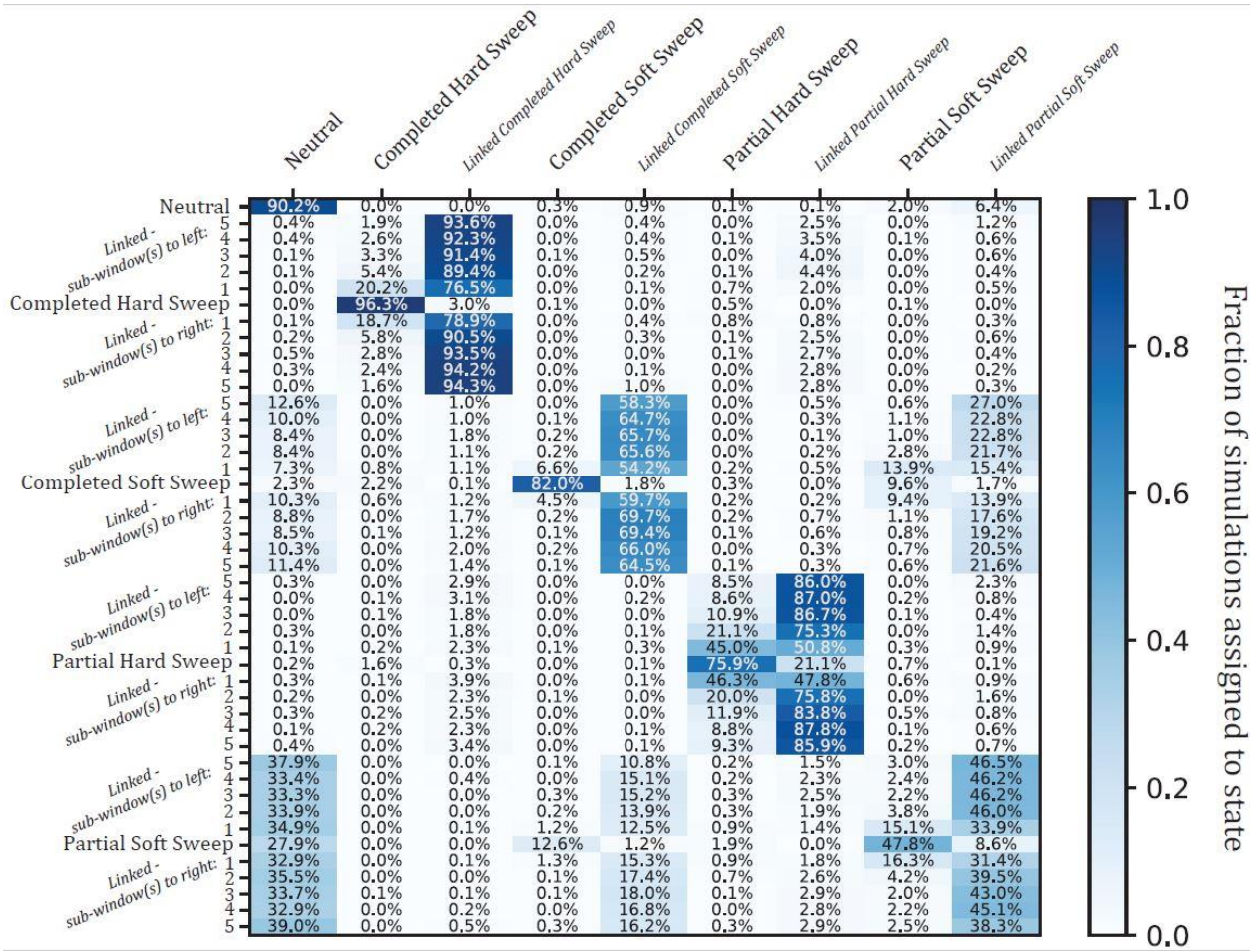



***Figure 3. Confusion matrix heatmap of simulation experiment including partial sweep training.*** *Given the BFS population history.*

## Robustness to demographic model misspecification

To assess robustness to demographic misspecification, we applied a CNN trained on simulations from one population sample to data generated from an alternate demographic

228 history (Figure 1). Specifically, we used training data from the GAS population size history,
229 which was fairly stable over time, and leveraged it against the CMS test dataset, which
230 experienced a dramatic population expansion (overall accuracy = 55.0%; rate of
231 misclassifying neutral regions as any one of the four non-linked sweep states = 1.7%).
232 Despite this misspecification, the confusion matrix (Figure S3) strongly resembles the
233 corresponding matrix that is correctly specified for demography (*i.e.* for CMS in Figure S2).
234 In particular, accuracies for finding neutral regions, completed hard sweeps, and partial
235 hard sweeps are roughly equivalent between the correctly specified model and
236 misspecified model (Figure S3). For soft sweeps, while confusion between completed and
237 partial sweeps is increased for the misspecified model, the overall ability to distinguish
238 sweeps from neutrality is largely preserved. Moreover, the rates at which examples from
239 the linked classes were mistaken for sweeps are seemingly unaffected. Together, these
240 results indicate that sensitivity for sweep discovery and localization was not strongly
241 impacted by the demographic model misspecification during training.

242

## 243 Partial sweeps are unpredictably misclassified as either completed 244 sweep or neutral when not explicitly considered

245 Since the previous versions of *partialS/HIC* (*S/HIC* and *diploS/HIC*) did not allow for partial
246 sweep selection states, we were interested in how such five-state classifiers would behave
247 when confronted with partial sweeps. To explore this, we conducted a simulation
248 experiment that first removed partial hard and soft sweeps as well as their associated
249 linked classes from the CNN training process, thus training on only five states rather than
250 all nine (Figure 1). Next, in an effort to examine the classification behavior for these five-
251 state CNNs, we applied the full test set that included the partial sweep classes.
252 Unsurprisingly, the trend was for partial sweeps to be most often confused for linked
253 selection (Figures 4, S4). Perhaps more concerning is the false negative rate (*i.e.* rate at
254 which partial sweeps were misclassified as neutral), which was substantial in partial hard
255 sweeps for several populations (median = 8.8%; max = 32.4%; >1% in all populations) and
256 extreme in partial soft sweeps (>50% in three populations, >40% in three more

populations, and >24% in all populations). Partial hard sweeps that were discovered were also often misclassified as a completed soft sweep (median rate of misclassification as completed soft sweep = 5.5%). However, when training included partial sweeps, there is universal and dramatic improvement in both finding sweeps and correctly identifying the model of selection (Figures 3, S2). Meanwhile, overall accuracy remains similar among the five-state and nine-state classifiers with respect to simulations of neutral, completed sweep, and linked classes exclusively (Figures 4, S4; Table S2). As a result, accuracy only stands to benefit from incorporating partial sweeps into training since ignoring such information leads to unacceptably high false negative rates of partial sweeps being called neutral or linked.

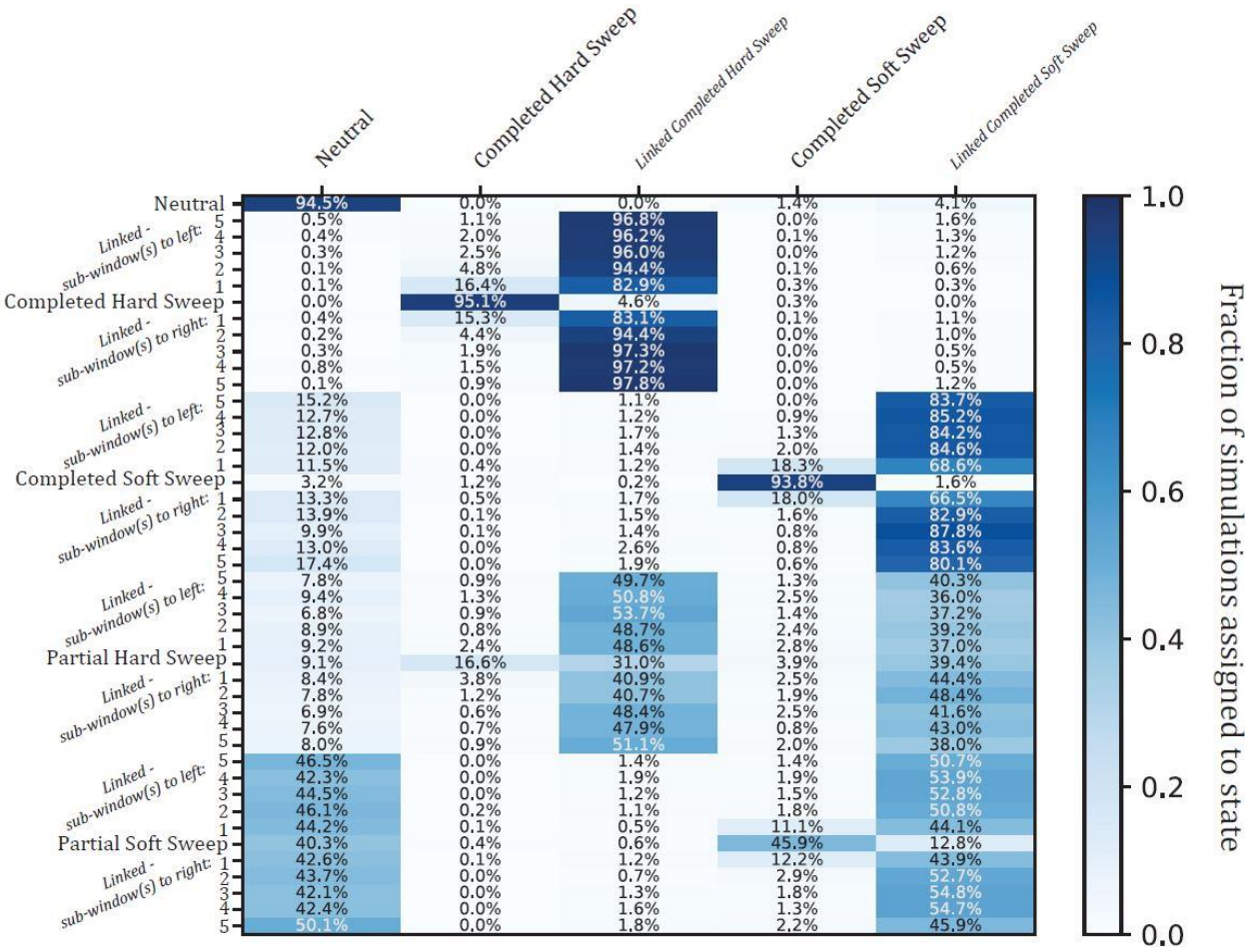



***Figure 4. Confusion matrix heatmap of simulation experiment with partial sweeps ignored during training.*** *Given the BFS population history.*

## *partialS/HIC* binary classification outperforms a competing individual summary statistic approach

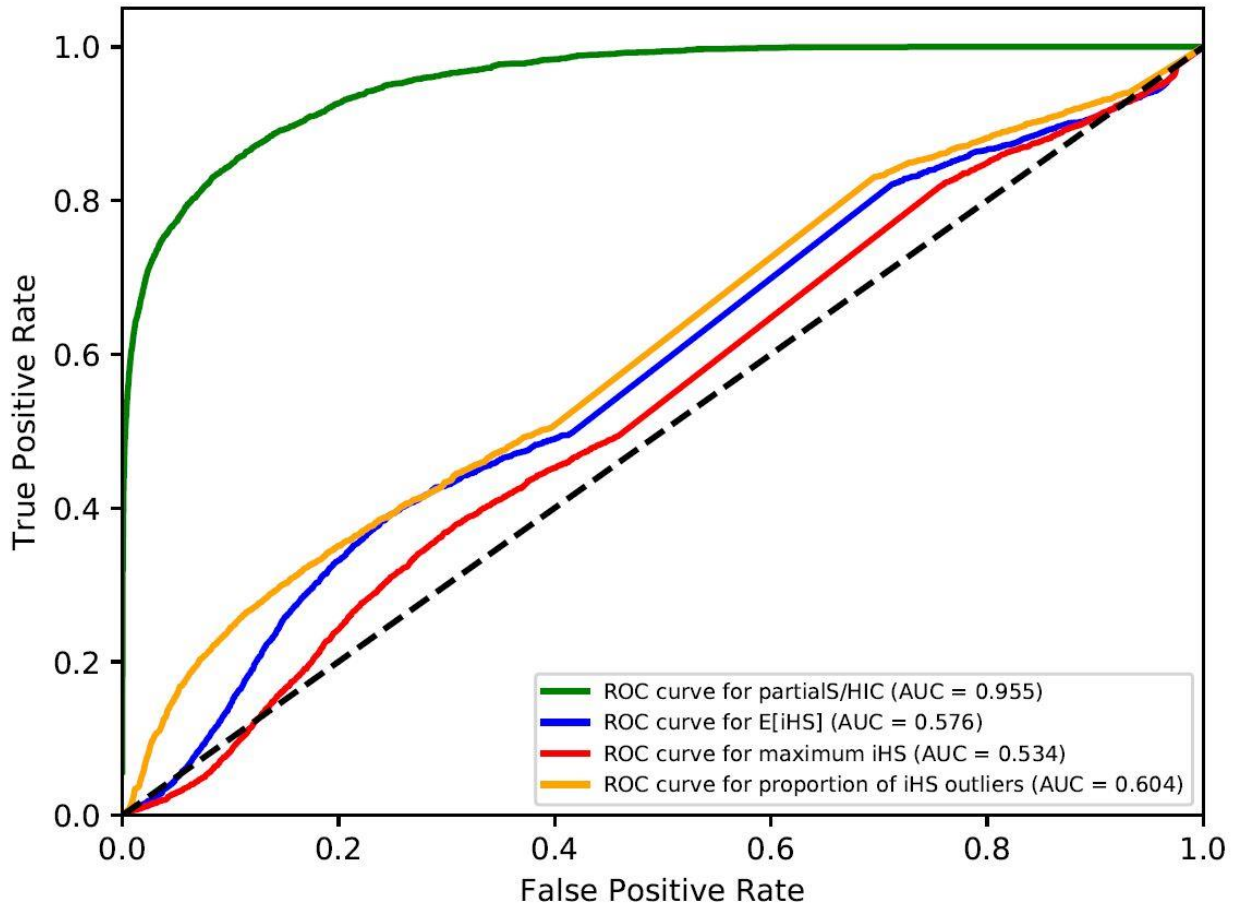


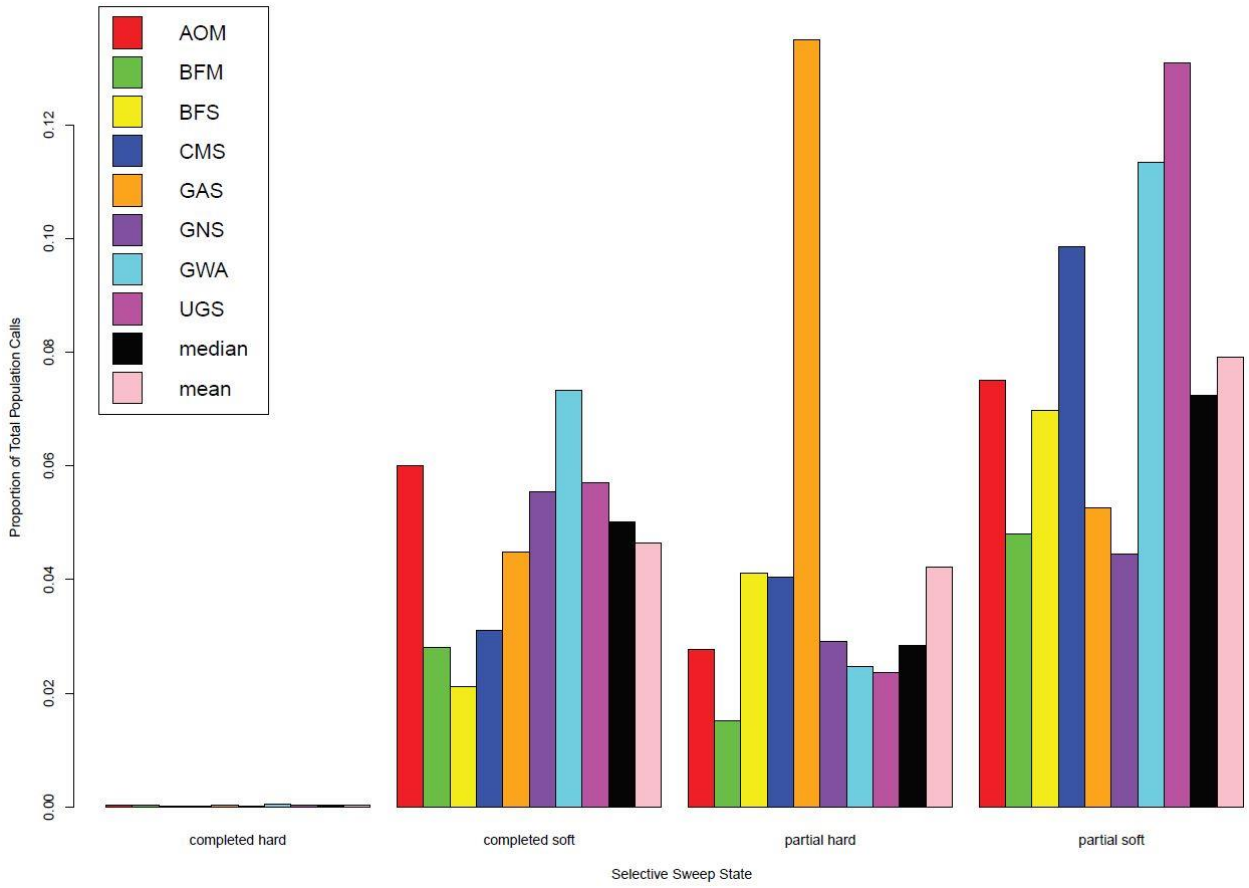

***Figure 5. ROC curves for binary classification of selective sweeps versus regions neutrally evolving or under linked selection.*** *Given the BFS population history.*

To assess whether our deep learning method extends inferential resolution beyond the signal conferred by *iHS*, a statistic explicitly designed for detecting partial sweeps [30], we compared receiver operating characteristic (ROC) curves for the binary classification task of broadly detecting selective sweeps (*i.e.* any of the four selection states involving a sweep within the focal sub-window) vs. neutral regions or linked sweeps (Figure 1). Producing a ROC curve, which plots true positive against false positive rates given varying thresholds, for *partialS/HIC* per population required optimizing a separate CNN, keeping the same whole training dataset as well as architecture of network layers (Figure 2) except with two

284 final output responses (*i.e.* sweep in central sub-window vs. no sweep in central sub-
285 window) instead of nine. Mean, maximum, and proportion of outlier *iHS* values across SNPs
286 within the central sub-window of the training simulations were used to obtain three more
287 respective ROC curves for each population history. All sub-window variants of the *iHS*
288 statistic performed quite poorly (median AUC for: $E[iHS]$ = 0.555; maximum *iHS* = 0.505),
289 especially in contrast to the *partialS/HIC* binary classifier (median AUC = 0.939), in
290 identifying selective sweeps to the focal sub-window (Figures 5, S5).

291

## 292 Soft and partial sweeps are commonplace among *A. gambiae* populations




293

294 ***Figure 6. Genome-wide sweep calls across empirical mosquito population datasets.***

295 Turning our attention to the Ag1000G phase I data, we applied our nine-state CNNs to the
296 corresponding *A. gambiae* population datasets, classifying 5 KB segments using a 55 KB full

297 sliding window throughout the whole genome (Figure 1). Each of the eight mosquito
298 populations contains a large number of sub-windows identified as completed soft sweeps
299 (median fraction of total calls genome-wide = 5.01%) as well as partial sweeps (median
300 fraction of total calls genome-wide for partial hard sweep = 2.84%; median fraction of total
301 calls genome-wide for partial soft sweep = 7.24%), coupled with only a handful of
302 completed hard sweep predictions (median fraction of total calls genome-wide = 0.03%)
303 (Figure 6; Table S3). Partial soft sweeps were typically discovered the most often (median
304 proportion of sweep calls = 52.59%), with completed soft sweeps often following (median
305 proportion of sweep calls = 28.80%) and partial hard sweeps usually being the third most
306 numerous class of detected sweep (median proportion of sweep calls = 19.75%). Notably,
307 our estimated false discovery rates (FDRs) are higher for soft sweeps (median FDR for
308 completed soft sweeps = 11.09%; median FDR for partial soft sweeps = 12.20%) compared
309 to hard sweeps (median FDR for completed hard sweeps = 0.00%; median FDR for partial
310 hard sweeps = 0.39%); this implies that individual soft sweep candidates should be viewed
311 with more caution, though we should be able to estimate the genome-wide proportion of
312 these classes well. After false discovery correction, classifications for partial soft sweeps
313 still outnumber those for completed soft sweeps in AOM, BFS, CMS, and UGS, as well as for
314 partial hard sweeps in all populations but GNS. Importantly, had partial sweeps not been
315 accounted for in the training process, our results suggest that we would have both
316 underestimated the total number of sweeps and incorrectly labeled many of our partial
317 sweeps (Figures 4, S4). This would have led to the conclusion that adaptation from
318 standing variation rather than *de novo* mutations dominate selective sweep dynamics in
319 these *A. gambiae* populations. While it is clear that soft sweeps are indeed more common in
320 these data, our results suggest that hard sweeps often occur as well, though with few
321 reaching fixation. Furthermore, the *partialS/HIC* classifications indicate that most selective
322 sweeps in these population samples are incomplete, suggesting that we are capturing a
323 view of selection in progress.

324

325 **Selective sweeps are significantly enriched in functional regions of the *A.***
326 ***gambiae* genome**

327 To elucidate broad characteristics underlying the genomic targets of selection, we used
328 permutation tests of sweep call locations to discover enrichment patterns in the following
329 DNA regions of interest: gene, mRNA, exon, CDS, five-prime UTR, and three-prime UTR
330 (Figure 1). Permutation tests were based on the total number of calls for the four selection
331 states with sweeps occurring within the focal sub-window, as well as the individual
332 number of calls for each of these states respectively. Across all eight population datasets
333 and for all six DNA regions under investigation, there is a statistically significant
334 enrichment of total sweep calls along with completed soft sweeps calls in particular,
335 whereas completed hard sweep calls are not significantly enriched in any single case
336 (Figures 7, S6; Table S4). Conversely, partial sweep enrichment varies among populations
337 as well as individual DNA regions. Specifically, partial hard sweeps are significantly
338 enriched in five of the six DNA regions for BFS and CMS, and four of the DNA regions for
339 UGS; while partial soft sweeps are significantly enriched in all six DNA regions for UGS, five
340 of the DNA regions for BFM, and four of the DNA regions for BFS.

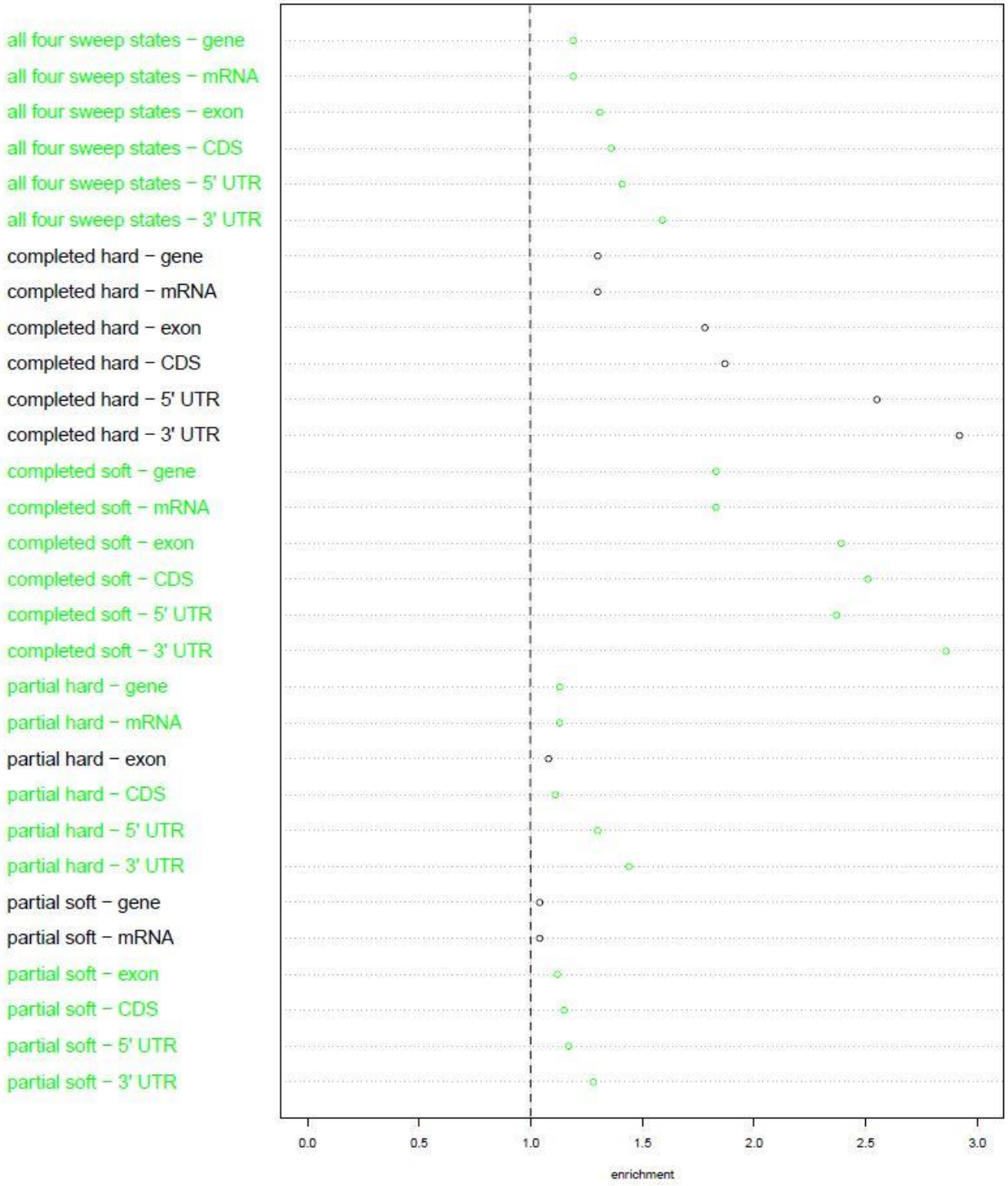



***Figure 7. Enrichments of selective sweeps within DNA regions.*** *Statistically significant (p < 0.05) enrichments in green.*

## Insecticide resistance loci, especially related to metabolism, are significantly enriched for selective sweeps

We performed a similar permutation analysis for four sets of genes known to confer insecticide resistance (IR) (Figure 1), finding at least one set of IR genes to be statistically significant for every population dataset in enrichment of total sweep calls (*i.e.* aggregate of all four sweep classes) and completed soft sweeps, respectively (Figures 8, S7; Table S5). In particular, metabolism-related IR genes are significantly enriched for each of these cases. Furthermore, IR genes corresponding to well-characterized resistance loci (*i.e.* target sites) are significantly enriched for AOM and BFM total sweep calls as well as completed soft sweeps in AOM, BFM, and GNS; IR genes associated with behavior are significantly enriched for BFM and GNS total sweep calls as well as completed soft sweeps in AOM, CMS, and GAS; and IR genes affiliated with cuticular activity are significantly enriched for completed soft sweeps in GWA and UGS. In contrast, completed hard sweeps are only significantly enriched in BFS, GNS, GWA, and UGS for IR genes connected to behavior (as well as metabolism for BFS). For partial sweeps, significant enrichment only occurs within BFM (partial soft sweeps in metabolism as well as behavior IR genes), CMS (partial hard sweeps in metabolism as well as target site IR genes), and UGS (partial hard sweeps in metabolism IR genes).

## Completed soft sweeps are significantly enriched within the same GO term annotations across populations

To uncover further functional traits targeted by selection, we used our permuted datasets to ask for which individual gene ontology (GO) terms our sweep candidates are enriched (Figure 1). Only completed soft sweeps, as well as the total of all four sweep states in combination, are significantly enriched for any GO terms, except for two single cellular components GO terms in UGS ("membrane" is enriched for completed hard sweeps and "nuclear cohesion complex" is enriched for partial hard sweeps); partial soft sweeps are therefore not significantly enriched for any GO terms among populations (Table S6). For

373 the completed soft sweep significant enrichments, we found six cases of the same GO term
374 in all eight populations. Three of these belong to the cellular components GO domain
375 ("nucleus", "membrane" and "integral component of membrane"), and the other three are
376 connected to molecular function, specifically binding ("nucleic acid binding", "protein
377 binding" and "ATP binding"). All six of these terms, especially those conferring binding
378 function, are also enriched for total sweep calls across multiple datasets: "nucleic acid
379 binding" and "ATP binding" in seven populations; "protein binding" and "nucleus" in six
380 populations; "integral component of membrane" in five populations; and "membrane" in
381 three populations, one of which is significantly enriched for this GO term of completed hard
382 sweeps as well, the single example of such among all populations. Other cases involving the
383 same GO term significantly enriched for completed soft sweeps in over half of the
384 populations include: "binding", "cytoplasm", and "zinc ion binding" in seven populations;
385 "RNA binding" in six populations; and "mRNA splicing, via spliceosome" and "ATPase
386 activity" in five populations.

***Figure 8. Enrichments of selective sweeps within groupings of insecticide resistance genes.*** *Statistically significant (p < 0.05) enrichments in green.*

391 # Discussion

392 ## *partialS/HIC* elucidates both species-wide and population-specific sweep
393 dynamics within *A. gambiae*

394 The Ag1000G data provided the opportunity to investigate selection at both the continental
395 scale, where wide-reaching impact across the whole species complex could be uncovered,
396 and the regional level, revealing population-specific sweep dynamics. For the former, we
397 observed that *A. gambiae* populations consistently experienced very few completed hard
398 sweeps, with nearly all sweeps being partial and/or soft. In fact, the impact of completed
399 hard sweeps on the adaptive process within mosquitos appears to be even more limited
400 than what was observed previously in humans [33]. This is likely a result of the much
401 larger population sizes and concordant levels of genetic variation that are maintained
402 within *Anopheles* populations. Importantly, we find a large number of ongoing selective
403 sweeps within these populations, particularly in comparison to the number of completed
404 sweeps. There are multiple reasons why this might be the case. A trivial explanation may
405 simply be that we only have power to detect sweeps that have completed in the past few
406 hundred generations, though this seems unlikely. More plausibly, a large number of
407 ongoing sweeps might be expected given the recent change in environment induced by
408 vector control efforts. Another possible explanation is that the frequency dynamics of
409 beneficial alleles within a population is often more complex than assumed and may indeed
410 contain an overdominant component [34]. This would mean that some portion of the
411 partial sweeps that we are observing in *Anopheles* is actually balanced, or transiently
412 balanced, polymorphisms. A fourth class of explanation is that beneficial mutations may
413 not be able to fix in populations due to competition with beneficial mutations on other
414 genes that have originated in different parts of the species range [35]. Indeed, each of these
415 factors may play some role in our reported abundance of partial sweeps.

416 Although such genome-wide sweep patterns occur species-wide, enrichment
417 behavior seems much more population-specific. For instance, while every population
418 possesses significant enrichment of completed soft sweeps coupled with no completed
419 hard sweep significant enrichments for the six functional DNA regions studied here, partial

420 sweep enrichments vary widely among datasets. Sweep behavior is even more
421 idiosyncratic for insecticide resistance genes, as the only constant between populations is
422 that metabolism is a recurring target of selection, especially for completed soft sweeps.

423 These findings from the Ag1000G data provide important genomic resources that
424 will inform continental-wide strategies that apply to the entire *A. gambiae* species complex,
425 as well as aid management efforts in specializing to certain populations and localities,
426 which can improve malaria control efficacy. Such insight into mosquito vector evolution
427 may also help curb future insecticide resistance adaptation, and in turn prevent impending
428 crises of vector control failure. However, it is important to consider that our partial sweep
429 calls could be capturing more complex selective dynamics at play, for example polygenic
430 and quantitative trait adaptation [36,37], balancing selection [38], and introgression of
431 beneficial alleles. These could lead to different modes of adaptation for the same genomic
432 region across populations, for instance a favorable SNP undergoing a soft sweep at its
433 origin and then carried to neighboring populations (as was suggested in Miles *et al.* [6])
434 may appear to be experiencing a partial hard sweep in those recipient populations. Such
435 complicated interactions merit further investigation on the Ag1000G data, which would not
436 only continue advancing methodological development for population genetics, but also
437 address interesting questions for a widespread and ecologically important organism that
438 has crucial ramifications on wildlife management and public health.

439

## 440 *partialS/HIC* offers powerful and unprecedented detection of partial 441 sweeps

442 Supervised machine learning approaches are rapidly gaining traction among population
443 geneticists, with deep learning in particular beginning to experience increased attention
444 and methodological development due to its exciting potential to unlock classic population
445 genetics problems. Examples of successful SML implementation in population genomics
446 include demographic model choice [39], demographic parameter inference [40],
447 comparative analysis of independent single-population size changes [41], identification of

448 introgressed regions [42], recombination rate estimation [43–45], and genomic scans of
449 selective sweeps [24]; deep learning specifically has been employed for joint inference of
450 demography and selection [25], discovery of recombination hotspots [46], estimation of
451 demographic and recombination parameters [47], discovery of functional variants [48],
452 and differentiating between hard and soft sweeps from neutral regions [26]. These
453 applications especially benefit from the ability to handle high dimensional input data and
454 bypassing the need of a likelihood function, which is due to SML uncovering data patterns
455 from leveraging *a priori* information through a training algorithm [25,28]. CNNs expand
456 this utility to image processing, which has been demonstrated with *diploS/HIC* to be a
457 powerful tool for exploiting the genomic spatial distribution of multiple population-level
458 summary statistics to detect selective sweeps [26].

459 Here, we demonstrated with *partialS/HIC* that deep learning can be extended to
460 partial sweeps, especially partial hard sweeps, yielding greater accuracy and robustness
461 than has been previously attained. We also showcased consistent performance in the face
462 of several underlying demographic backgrounds. Specifically, *partialS/HIC* achieves
463 dependable discovery of selective sweeps, excellent simultaneous disambiguation between
464 partial and completed sweeps as well as between hard and soft sweeps, and reliable spatial
465 localization of selection targets in the genome. Moreover, we have shown that partial
466 sweeps remain mostly undetected if ignored from the training process, even though such
467 selection may be commonplace throughout a genome as with the Ag1000G data. As a result,
468 many previous studies scanning for either complete or ongoing selective sweeps solely (*i.e.*
469 not jointly inferring both types of selection) may have overlooked an important subset of
470 evolutionary events [35]. Researchers may then be interested in reexamining datasets with
471 *partialS/HIC* to elucidate the relative contributions of fixed versus incomplete sweeps to
472 adaptive evolution.

473 Importantly, the efficacy of *partialS/HIC* relies on several factors that are
474 unexplored here, including simulation prior specifications, CNN architecture with respect
475 to construction and parameterization of neural network layers, and data structure. Hence,
476 it is prudent for future implementations to validate performance by testing a range of
477 configurations, given a project's individual intricacies, to assess robustness and inherent

assumptions. In particular, future exploration of alternate image constructions could be potentially of great methodological benefit. Such images could be derived from different ordering schemes and/or suites of summary statistics, as well as without summary statistics entirely, instead directly exploiting sequence alignments [46,47] or even raw reads. More broadly, CNNs can be further extended to address other long-standing efforts in evolutionary biology, such as parameter inference under complex isolation-migration models or phylogenetic reconstruction.

# Methods

## Simulations for training and testing CNN classifier

We used *discoal* [31] to simulate training and test datasets corresponding to each *A. gambiae* population under nine different selection states: neutrally evolving, completed hard sweep, completed soft sweep, partial hard sweep, partial soft sweep, and linked region for every one of the four sweep classes (Figure 1; Table S7). For the four sweep types, the target SNP was located in the exact middle position within the central, or sixth in sequence, sub-window of 11 in total; the selected SNP was placed in the middle within one of the other ten sub-windows for linked sweeps. There were 2,000 training examples per selection state (with selected sub-window randomized for linked sweeps) and 1,000 test examples per class (including for each of the ten linked sweep locations), resulting in a training dataset of 18,000 simulations and a test dataset of 45,000 simulations given each demographic history, thus totaling 144,000 training and 360,000 test simulations. To conduct single-population simulations with *discoal*, we used the *stairway plot* [53] point estimates from Miles *et al.* [6] for size change parameters as well as $N_0$ (present-day effective population size), assumed a mutation rate ($\mu$) of $3.5\times10^{-9}$ mutations per base pair per generation, and performed random draws for locus-wide mutation and recombination rates from the following distributions for each independent replicate: $\theta \sim U(\frac{2\times E[\theta]}{11.0}, \frac{2\times E[\theta]}{1.1})$, where $E[\theta] = 4N_0\mu L$ and $L$ is the length of the simulated sequence with $L$ = 55,000 base pairs; $\rho \sim TEXP(2\times E[\theta], 6\times E[\theta])$, where $\rho = 4N_0rL$, $r$ is the recombination rate per base

506 pair, and $TEXP(\beta$, maximum value) is a truncated exponential distribution with mean $\beta$; $s \sim$
507 $U(1.0\times10^{-4}, 1.0\times10^{-2})$; end time of sweep $\sim U(0, 2{,}000)$ generations ago, which represents
508 fixation for completed sweeps and the transition back to neutral evolution for partial
509 sweeps; selected SNP allele frequency at onset of soft sweep $\sim U(\frac{1}{N_0}, 0.2)$; and selected SNP
510 allele frequency at end of partial sweep $\sim U(0.20, 0.99)$.

511

## 512 Constructing two-dimensional feature vector images of summary 513 statistics

514 The eight training and test datasets, as well as empirical datasets, were converted into two-
515 dimensional feature vector matrices for downstream deep learning (Figure 1); this was
516 performed within the Python environment and required usage of the module *numpy*. Prior
517 to this two-dimensional transformation, the simulated data were modified to better
518 account for uncertainty within the empirical data, specifically: 1) sites that were missing
519 any individual calls or could not be polarized against the outgroup were excluded; and 2)
520 incorrect identification of the derived allele. For the former, each simulation randomly
521 drew from a distribution of 1,552 masking profiles (with test simulations drawing without
522 replacement per selection class of 1,000 simulations), which determined the exact sites to
523 be omitted from further analysis; a masking profile consisted of the site positions within a
524 single full 55 KB window on the *A. gambiae* genome that had absent at least one sample
525 throughout the entirety of the Ag1000G data and/or ancestral state information, and the
526 total set represented all 1,552 sequential, non-overlapping windows (*e.g.* 2L: 1–55,000; 2L:
527 55,001–110,000; 2L: 110,001–165,000; etc.) where the proportion of masked sites did not
528 exceed 75% in any of the constituent sub-windows (*i.e.* 1,250 sites). To account for
529 mispolarization, estimated rates were obtained from Miles *et al.* [6] and exploited via a
530 binomial distribution to mispolarize a random subset of SNPs to the other allele per
531 simulation.

532 The empirical data similarly underwent processing for compatibility with the
533 simulated data. First, chromosomes were delineated into sequential 5 KB sub-windows

534 (*e.g.* positions 1-5,000 formed the first sub-window, positions 5,001-10,000 formed the
535 second sub-window, etc.), with the aforementioned masking criteria applied across sites
536 and remaining SNPs polarized. Within each population dataset, all polymorphic positions
537 containing more than two alleles were further removed from analysis, such that only
538 polarized monomorphic and biallelic sites comprising a full data matrix of no missing data
539 were left. Sub-windows containing no SNPs or less than 25% of the original sites were
540 subsequently discarded, and every configuration of 11 contiguous 5 KB sub-windows of
541 those remaining formed a single full window, which would be classified into one of the nine
542 selection states based upon its central sub-window while using spatial information from
543 the neighboring five sub-windows on either side. To clarify, this eliminated any window
544 that did not contain a consecutive sequence of 11 sub-windows that survived data filtering,
545 and resulted in a sliding window that progressed a single sub-window at a time, such that
546 succeeding full windows could be overlapping by up to ten sub-windows.

547 Every independent simulation totaling 55 KB in length from 11 sub-windows of 5
548 KB, as well as empirical sequence of 11 adjacent sub-windows per population, was then
549 transformed into 89 separate summary statistic vectors, each with 11 elements that
550 captured population-level variation across the sampled individuals per sub-window. The
551 first 17 summary statistics, which were $\pi$ [54], $\theta_W$ [55], Tajima's $D$ [14], $\theta_H$ [15], Fay-Wu's $H$
552 [15], number of haplotypes, $H_1$ [52], $H_{12}$ [52], $H_2/H_1$ [52], $Z_{nS}$ [56], $\omega$ [16], $E[iHS]$ [30],
553 maximum *iHS* [30], proportion of outlier *iHS* values [30], variance of pairwise genotype
554 distances [26], skewness of pairwise genotype distances [26], and kurtosis of pairwise
555 genotype distances [26], were calculated with the Python package *scikit-allel* as done with
556 *diploS/HIC* [26]. Values for *iHS* were standardized within 50 derived allele frequency bins,
557 following mispolarization in the case of simulated data. Outlier *iHS* values were defined as
558 within either 2.0% tail of the distribution obtained from simulations of neutral evolution
559 under the appropriate demographic history.

560 The remaining 72 summary statistics originated from SAFE, a recently developed
561 statistic [32], and its various components; these SAFE-derived statistics included
562 summaries for the distribution of values for: haplotype allele frequency (HAF), which is the
563 sum of derived allele counts across all the derived alleles present within a sequence;

unique HAF score (*i.e.* each unique HAF value is counted only once, even if representing multiple individuals); $\varphi$, which is the sum of HAF scores for sequences harboring the derived allele, divided by the total sum of HAF scores across all sequences; κ, which is the proportion of distinct HAF scores that carry the derived allele; derived allele frequency; and SAFE itself, which is the difference between $\varphi$ and κ normalized against the derived allele frequency. Notably, HAF is calculated per sequence, whereas $\varphi$, κ, derived allele frequency, and SAFE are calculated per polymorphism. The following distribution summaries were utilized to construct individual values spanning a sub-window: mean, median, mode; 2.5%, 25%, 75%, and 97.5% quartiles; maximum, variance, standard deviation, skewness, and kurtosis. Importantly, each summary statistic vector was normalized, in the same manner as the preceding versions to *partialS/HIC* [24,26], to capture signal solely from the relative spatial distribution of the summary statistics across the 11 sub-windows, rather than allowing influence from absolute values. Subsequently, the 89 vectors were vertically collated to form a two-dimensional matrix that could then be exploited for image processing. The arrangement of these vectors were such that the 11 columns corresponded to the series of sub-windows from left to right, and the 89 rows of summary statistics were in the order presented here (with the distribution summaries iterating first for every SAFE component, *e.g.* row 52: skewness of $\varphi$ values; row 53: kurtosis of $\varphi$ values; row 54: $E$[κ]). Importantly, column and row order affects deep learning optimization, which may have consequences on overall efficacy, due to the convolutional and pooling windows employed by the CNN architecture, hence related summary statistics were grouped together (*e.g.* alternative distribution summaries of a SNP-based statistic, various SAFE derivatives). Heatmap images, based on median values per statistic and sub-window, were generated in R for the neutral case and each of the four sweep states under every population history from the training simulations.

## Training and testing CNNs for deep learning implementation

The architecture of our CNN was composed of the following sequential layers: 1) 2D convolutional layer with 256 filters using 3×6 windows and "same" padding; 2) 2D max

593 pooling layer given a 3×3 window; 3) a second 2D convolutional layer of 256 filters based
594 on 3×3 windows, "same" padding, and ReLU activation; 4) a second 2D max pooling layer
595 also with a 3×3 window; 5) dropout layer with p=0.25; 6) flattening layer; 7) fully-
596 connected layer with ReLU activation to 512 responses; 8) a second dropout layer with
597 p=0.50; 9) a second fully-connected layer with ReLU activation to 128 elements; 10)
598 another dropout layer with p=0.50; and 11) softmax activation layer to 9 states (Figures 1–
599 2). This architecture was trained using the Python module *Keras* [57] given the *Adam*
600 optimizer [58], with 20 epochs, batch size of 32 simulations per step within an epoch, and
601 10% of the training data (*e.g.* 1800 simulations from the total nine-state training dataset)
602 randomly removed as a validation set during optimization. Training was performed for
603 every population demography under three experimental settings: 1) given the full set of
604 training data distributed across nine selection states; 2) exploiting a subset of the training
605 data from only five of the selection states, specifically those involving neutral regions or
606 completed sweeps; 3) deploying the entire training data, but with binary classification
607 between selective sweeps in the focal sub-window and all unselected classes (*i.e.* neutral
608 class together with every linked class). The same test datasets were used for the first two
609 simulation experiments, producing overall accuracy measures as well as confusion matrix
610 heatmaps to assess misclassification bias for each of 45 classes, in this case treating
611 different sub-window placements of linked sweeps discretely. Moreover, to explore the
612 effect of demographic misspecification, we conducted a single additional test under the first
613 experimental set-up whereby the CNN trained on the GAS simulations was applied to the
614 CMS test simulations. Regarding the binary classification experiment, the Python module
615 *sklearn*, which is available at http://scikit-learn.org/stable/ [59], was used for building
616 ROC curves to evaluate accuracy and sensitivity. For comparison, ROC curves were also
617 constructed from the training simulations based on the focal sub-window mean *iHS*,
618 maximum *iHS*, and proportion of *iHS* values that were outliers, respectively.

619

620 **Detecting selective sweeps for *A. gambiae* population datasets**

To scan the genome for signatures of selective sweeps, the nine-state trained CNNs were applied to the eight empirical mosquito datasets, with the underlying simulated demography matched to the sampled population (Figure 1). Calls were corrected for false discovery by exploiting the accuracy and error rates for neutral regions from the nine-state simulation experiment, such that the amount of neutral calls was assumed to be underestimated while the amount of calls for the remaining eight selection states were assumed to be inflated. Subsequently, we produced sets of 10,000 randomly permuted calls across the genome to derive null expectations of sweep enrichment, following Schrider and Kern [33]. Using the gene annotation file "Anopheles-gambiae-PEST_BASEFEATURES_AgamP4.7.gff3.gz" from VectorBase, we exploited these permuted datasets to assess statistically significant enrichment within certain DNA regions, groupings of known IR genes (N. Harding, *pers. comm.*), and all basic GO term definitions from http://www.geneontology.org (last accessed February 18, 2015). The DNA regions of interest included gene, mRNA, exon, CDS, five-prime UTR, and three-prime UTR; IR genes were assigned to four functional categories: metabolism, target sites, behavior, and cuticular. To determine significant enrichment, the number of inferred calls for a particular DNA region or IR gene category had to have a $p$-value $< 0.05$ based on the respective distribution of 10,000 permutations; for the GO terms, we deployed a corrected $q$-value $< 0.05$ due to concerns of false discovery stemming from the large number of terms tested for enrichment.

## Acknowledgments

We thank Jeff Adrion for comments on the manuscript. This work was supported by NIH award no. R01GM117241 to ADK and NIH award no. K99HG008696 to DRS.

645 # References


646 1. WHO. World Malaria Report 2015 [Internet]. 2015. Available:
647 http://www.who.int/malaria/publications/world-malaria-report-2015/en/

648 2. Hemingway J, Ranson H, Magill A, Kolaczinski J, Fornadel C, Gimnig J, et al. Averting a
649 malaria disaster: Will insecticide resistance derail malaria control? Lancet. 2016;387:
650 1785–1788. doi:10.1016/S0140-6736(15)00417-1

651 3. Brown AW. Insecticide resistance in mosquitoes: a pragmatic review. J Am Mosq
652 Control Assoc. 1986;2: 123—140.

653 4. Elissa N, Mouchet J, Rivière F, Meunier J-Y, Yao K. Resistance of Anopheles gambiae
654 s.s. to pyrethroids in Côte d'Ivoire. Ann Soc Belg Med Trop (1920). 1993;73: 291–
655 294.

656 5. Martinez-Torres D, Chandre F, Williamson MS, Darriet F, Bergé JB, Devonshire AL, et
657 al. Molecular characterization of pyrethroid knockdown resistance (kdr) in the major
658 malaria vector Anopheles gambiae s.s. Insect Mol Biol. 1998;7: 179–184.
659 doi:10.1046/j.1365-2583.1998.72062.x

660 6. Miles A, Harding NJ, Bottà G, Clarkson CS, Antão T, Kozak K, et al. Genetic diversity of
661 the African malaria vector anopheles gambiae. Nature. 2017;552: 96–100.
662 doi:10.1038/nature24995

663 7. Oxborough RM, N'Guessan R, Jones R, Kitau J, Ngufor C, Malone D, et al. The activity of
664 the pyrrole insecticide chlorfenapyr in mosquito bioassay: Towards a more rational
665 testing and screening of non-neurotoxic insecticides for malaria vector control. Malar
666 J. 2015;14: 124. doi:10.1186/s12936-015-0639-x

667 8. Hammond A, Galizi R, Kyrou K, Simoni A, Siniscalchi C, Katsanos D, et al. A CRISPR-
668 Cas9 gene drive system targeting female reproduction in the malaria mosquito
669 vector Anopheles gambiae. Nat Biotechnol. 2016;34: 78–83. doi:10.1038/nbt.3439.A

670 9. Unckless RL, Messer PW, Connallon T, Clark AG. Modeling the manipulation of
671 natural populations by the mutagenic Chain reaction. Genetics. 2015;201: 425–431.
672 doi:10.1534/genetics.115.177592

673 10. Unckless RL, Clark AG, Messer PW. Evolution of resistance against CRISPR/Cas9 gene
674 drive. Genetics. 2017;205: 827–841. doi:10.1534/genetics.116.197285

675 11. Maynard Smith J, Haigh J. The hitch-hiking effect of a favourable gene. Genet Res
676 (Camb). Rutgers University; 1974;23: 23–35. doi:10.1017/S0016672308009579

677 12. Kaplan NL, Hudson RR, Langley CH. The "hitchhiking effect" revisited. Genetics.
678 1989;123: 887–899.

679 13. Stephan W, Wiehe THE, Lenz MW. The effect of strongly selected substitutions on
680 neutral polymorphism: analytical results based on diffusion theory. Theor Popul Biol.

681 1992;41: 237–254. doi:10.1016/0040-5809(92)90045-U

682 14. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA
683 polymorphism. Genetics. 1989;123: 585–595.

684 15. Fay JC, Wu C-I. Hitchhiking under positive Darwinian selection. Genetics. 2000;155:
685 1405–1413.

686 16. Kim Y, Neilsen R. Linkage disequilibrium as a signature of selective sweeps. Genetics.
687 2004;167: 1513–1524. doi:10.1534/genetics.103.025387

688 17. Kern AD, Jones CD, Begun DJ. Genomic effects of nucleotide substitutions in
689 Drosophila simulans. Genetics. 2002;162: 1753–1761.

690 18. Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, et al. Detecting
691 recent positive selection in the human genome from haplotype structure. Nature.
692 2002;419: 832–837. doi:10.1038/nature01027.1.

693 19. Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, Bustamante C. Genomic scans
694 for selective sweeps using SNP data. Genome Res. 2005;15: 1566–75.
695 doi:10.1101/gr.4252305

696 20. Pavlidis P, Jensen JD, Stephan W. Searching for footprints of positive selection in
697 whole-genome SNP data from nonequilibrium populations. Genetics. 2010;185: 907–
698 922. doi:10.1534/genetics.110.116459

699 21. Lin K, Li H, Schlötterer C, Futschik A. Distinguishing positive selection from neutral
700 evolution: Boosting the performance of summary statistics. Genetics. 2011;187: 229–
701 244. doi:10.1534/genetics.110.122614

702 22. Ronen R, Udpa N, Halperin E, Bafna V. Learning natural selection from the site
703 frequency spectrum. Genetics. 2013;195: 181–193.
704 doi:10.1534/genetics.113.152587

705 23. Pybus M, Luisi P, Dall'Olio GM, Uzkudun M, Laayouni H, Bertranpetit J, et al.
706 Hierarchical boosting: A machine-learning framework to detect and classify hard
707 selective sweeps in human populations. Bioinformatics. 2015;31: 3946–3952.
708 doi:10.1093/bioinformatics/btv493

709 24. Schrider DR, Kern AD. S/HIC: Robust Identification of Soft and Hard Sweeps Using
710 Machine Learning. PLoS Genet. 2016;12: e1005928.
711 doi:10.1371/journal.pgen.1005928

712 25. Sheehan S, Song YS. Deep Learning for Population Genetic Inference. PLoS Comput
713 Biol. 2016;12: e1004845. doi:10.1371/journal.pcbi.1004845

714 26. Kern AD, Schrider DR. diploS/HIC: An Updated Approach to Classifying Selective
715 Sweeps. G3 Genes, Genomes, Genet. 2018;8: 1959–1970. doi:10.1534/g3.118.200262

716 27. Sugden LA, Atkinson EG, Fischer AP, Rong S, Henn BM, Ramachandran S. Localization

717 of adaptive variants in human genomes using averaged one-dependence estimation.
718 Nat Commun. Springer US; 2018;9: 703. doi:10.1038/s41467-018-03100-7

719 28. Schrider DR, Kern AD. Supervised Machine Learning for Population Genetics: A New
720 Paradigm. Trends Genet. Elsevier Ltd; 2018;34: 301–312.
721 doi:10.1016/j.tig.2017.12.005

722 29. Coop G, Ralph P. Patterns of neutral diversity under general models of selective
723 sweeps. Genetics. 2012;192: 205–224. doi:10.1534/genetics.112.141861

724 30. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in
725 the human genome. PLoS Biol. 2006;4: e72. doi:10.1371/journal.pbio.0040072

726 31. Kern AD, Schrider DR. Discoal: Flexible Coalescent Simulations With Selection.
727 Bioinformatics. 2016;32: 3839–3841. doi:10.1093/bioinformatics/btw556

728 32. Akbari A, Vitti JJ, Iranmehr A, Bakhtiari M, Sabeti PC, Mirarab S, et al. Identifying the
729 favored mutation in a positive selective sweep. Nat Methods. Nature Publishing
730 Group; 2018;15: 279–282. doi:10.1038/nmeth.4606

731 33. Schrider DR, Kern AD. Soft sweeps are the dominant mode of adaptation in the
732 human genome. Mol Biol Evol. 2017;34: 1863–1877. doi:10.1093/molbev/msx154

733 34. Sellis D, Callahan BJ, Petrov DA, Messer PW. Heterozygote advantage as a natural
734 consequence of adaptation in diploids. Proc Natl Acad Sci. 2011;108: 20666–20671.
735 doi:10.1073/pnas.1114573108

736 35. Ralph PL, Coop G. Parallel Adaptation: One or Many Waves of Advance of an
737 Advantageous Allele? Genetics. 2010;186: 647–668.
738 doi:10.1534/genetics.110.119594

739 36. Pritchard JK, Pickrell JK, Coop G. The Genetics of Human Adaptation: Hard Sweeps,
740 Soft Sweeps, and Polygenic Adaptation. Curr Biol. Elsevier Ltd; 2010;20: R208–R215.
741 doi:10.1016/j.cub.2009.11.055

742 37. Booker TR, Jackson BC, Keightley PD. Detecting positive selection in the genome.
743 BMC Biol. BMC Biology; 2017;15: 98. doi:10.1186/s12915-017-0434-y

744 38. Connallon T, Clark AG. Antagonistic versus nonantagonistic models of balancing
745 selection: Characterizing the relative timescales and hitchhiking effects of partial
746 selective sweeps. Evolution (N Y). 2013;67: 908–917. doi:10.1111/j.1558-
747 5646.2012.01800.x

748 39. Smith ML, Ruffley M, Espíndola A, Tank DC, Sullivan J, Carstens BC. Demographic
749 Model Selection using Random Forests and the Site Frequency Spectrum. Mol Ecol.
750 2017; doi:10.1111/mec.14223

751 40. Pudlo P, Marin J-M, Estoup A, Cornuet J-M, Gauthier M, Robert CP. Reliable ABC
752 model choice via random forests. Bioinformatics. 2016;32: 859–866.

753 doi:10.1093/bioinformatics/btv684

754 41. Xue AT, Hickerson MJ. Multi-DICE : R package for comparative population genomic
755 inference under hierarchical co-demographic models of independent single-
756 population size changes. Mol Ecol Resour. 2017; doi:10.1111/1755-0998.12686

757 42. Schrider DR, Ayroles J, Matute DR, Kern AD. Supervised machine learning reveals
758 introgressed loci in the genomes of Drosophila simulans and D. sechellia. PLOS Genet.
759 2018;14: e1007341. doi:10.1371/journal.pgen.1007341

760 43. Lin K, Futschik A, Li H. A fast estimate for the population recombination rate based
761 on regression. Genetics. 2013;194: 473–484. doi:10.1534/genetics.113.150201

762 44. Adrian AB, Corchado JC, Comeron JM. Predictive models of recombination rate
763 variation across the drosophila melanogaster genome. Genome Biol Evol. 2016;8:
764 2597–2612. doi:10.1093/gbe/evw181

765 45. Gao F, Ming C, Hu W, Li H. New Software for the Fast Estimation of Population
766 Recombination Rates (FastEPRR) in the Genomic Era. G3 Genes, Genomes, Genet.
767 2016;6: 1563–1571. doi:10.1534/g3.116.028233

768 46. Chan J, Perrone V, Spence JP, Jenkins PA, Mathieson S, Song YS. A Likelihood-Free
769 Inference Framework for Population Genetic Data using Exchangeable Neural
770 Networks. bioRxiv. 2018; 1802.06153. doi:10.1101/267211

771 47. Flagel L, Brandvain Y, Schrider DR. The Unreasonable Effectiveness of Convolutional
772 Neural Networks in Population Genetic Inference. bioRxiv. 2018; 336073.
773 doi:10.1101/336073

774 48. Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning-
775 based sequence model. Nat Methods. 2015;12: 931–934. doi:10.1038/nmeth.3547

776 49. Mu J, Myers RA, Jiang H, Liu S, Ricklefs S, Waisberg M, et al. Plasmodium falciparum
777 genome-wide scans for positive selection, recombination hot spots and resistance to
778 antimalarial drugs. Nat Genet. 2010;42: 268–271. doi:10.1038/ng.528

779 50. Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, et al. Genome-wide
780 comparative diversity uncovers multiple targets of selection for improvement in
781 hexaploid wheat landraces and cultivars. Proc Natl Acad Sci. 2013;110: 8057–8062.
782 doi:10.1073/pnas.1217133110

783 51. Andersen EC, Gerke JP, Shapiro JA, Crissman JR, Ghosh R, Bloom JS, et al.
784 Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic
785 diversity. Nat Genet. 2012;44: 285–290. doi:10.1038/ng.1050

786 52. Garud NR, Messer PW, Buzbas EO, Petrov DA. Recent Selective Sweeps in North
787 American Drosophila melanogaster Show Signatures of Soft Sweeps. PLoS Genet.
788 2015;11: e1005004. doi:10.1371/journal.pgen.1005004

53. Liu X, Fu Y-X. Exploring population size changes using SNP frequency spectra. Nat Genet. 2015;47: 555–559. doi:10.1038/ng.3254

54. Tajima F. Evolutionary relationship of DNA sequences in finite populations. Genetics. 1983;105: 437–460.

55. Watterson GA. On the number of segregating sites in genetical models without recombination. Theor Popul Biol. 1975;7: 256–276. doi:10.1016/0040-5809(75)90020-9

56. Kelly JK. A test of neutrality based on interlocus associations. Genetics. 1997;146: 1197–1206.

57. Chollet F. Keras [Internet]. 2015. Available: https://keras.io

58. Kingma DP, Ba JL. Adam: A Method for Stochastic Optimization. arXiv. 2014; 1412.6980. doi:http://doi.acm.org.ezproxy.lib.ucf.edu/10.1145/1830483.1830503

59. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011;12: 2825–2830. doi:10.1007/s13398-014-0173-7.2

805 # Figure Captions

806 **Figure 1. Flowchart of methodological procedure.** Major steps highlighted in green.
807 Note that processing of simulated data incorporated uncertainties within the empirical
808 data, specifically with respect to masking sites from calculation of summary statistics.

809 **Figure 2. CNN architecture of neural network layers.** The input feature vector image is
810 passed to a 2D convolutional layer with 256 filters using a 3×6 receptive field. Next is a 2D
811 max pooling layer given a 3×3 receptive field. Then there is a second 2D convolutional layer
812 of 256 filters based on a 3×3 receptive field and ReLU activation. Afterward is a second 2D
813 max pooling layer, also with a 3×3 receptive field. The tensor is then flattened and passed
814 to two fully-connected layers with ReLU activation, resulting in 512 and 128 nodes
815 respectively. Finally, a softmax activation layer classifies the 9 responses.

816 **Figure 3. Confusion matrix heatmap of simulation experiment including partial**
817 **sweep training.** Given the BFS population history.

818 **Figure 4. Confusion matrix heatmap of simulation experiment with partial sweeps**
819 **ignored during training.** Given the BFS population history.

820 **Figure 5. ROC curves for binary classification of selective sweeps versus regions**
821 **neutrally evolving or under linked selection.** Given the BFS population history.

822 **Figure 6. Genome-wide sweep calls across empirical mosquito population datasets.**

823 **Figure 7. Enrichments of selective sweeps within DNA regions.** Statistically significant
824 ($p < 0.05$) enrichments in green.

825 **Figure 8. Enrichments of selective sweeps within groupings of insecticide resistance**
826 **genes.** Statistically significant ($p < 0.05$) enrichments in green.

827 **Figure S1. Feature vector heatmaps for neutral and four sweep states under each**
828 **population history.** Heatmaps derived from median values of a single training dataset,
829 each assuming a different population history. Note focal sub-window for four sweep states
830 versus neutral state across populations.

**Figure S2. Confusion matrix heatmaps of simulation experiment including partial sweep training under each population history.**

**Figure S3. Confusion matrix heatmap of demographic misspecification test.** Training on simulations (including partial sweeps) given the GAS population history while testing on simulations given the CMS population history.

**Figure S4. Confusion matrix heatmaps of simulation experiment with partial sweeps ignored during training under each population history.**

**Figure S5. ROC curves for binary classification of selective sweeps versus regions neutrally evolving or under linked selection for each population history.**

**Figure S6. Enrichments of selective sweeps within DNA regions for each empirical mosquito population dataset.** Statistically significant ($p < 0.05$) enrichments in green.

**Figure S7. Enrichments of selective sweeps within groupings of insecticide resistance genes for each empirical mosquito population dataset.** Statistically significant ($p < 0.05$) enrichments in green.

Neutral Region
Completed Hard Sweep
Completed Soft Sweep
Partial Hard Sweep
Partial Soft Sweep
AOM
BFM
BFS
CMS
GAS
GNS
GWA
UGS

## AOM

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep | Partial Hard Sweep | Linked Partial Hard Sweep | Partial Soft Sweep | Linked Partial Soft Sweep |
|---|---|---|---|---|---|---|---|---|---|
| Neutral | 91.9% | 0.0% | 0.0% | 1.8% | 1.9% | 0.1% | 0.4% | 1.4% | 2.5% |
| Linked - sub-window(s) to left: 5 | 0.5% | 5.1% | 90.6% | 0.3% | 1.7% | 0.0% | 1.5% | 0.0% | 0.3% |
| 4 | 0.6% | 5.6% | 90.6% | 0.5% | 1.8% | 0.1% | 0.6% | 0.0% | 0.2% |
| 3 | 0.7% | 9.1% | 87.6% | 0.7% | 1.2% | 0.0% | 0.5% | 0.0% | 0.2% |
| 2 | 0.3% | 13.9% | 80.3% | 1.0% | 3.0% | 0.3% | 0.9% | 0.1% | 0.2% |
| 1 | 1.4% | 31.3% | 60.5% | 2.7% | 2.7% | 0.7% | 0.2% | 0.4% | 0.1% |
| Completed Hard Sweep | 0.0% | 90.7% | 3.6% | 5.6% | 0.0% | 0.1% | 0.0% | 0.0% | 0.0% |
| Linked - sub-window(s) to right: 1 | 0.5% | 33.9% | 62.5% | 1.4% | 0.9% | 0.4% | 0.3% | 0.1% | 0.0% |
| 2 | 0.7% | 14.2% | 82.2% | 1.0% | 1.1% | 0.1% | 0.1% | 0.2% | 0.4% |
| 3 | 0.7% | 10.3% | 86.1% | 0.4% | 1.2% | 0.0% | 0.5% | 0.1% | 0.7% |
| 4 | 0.4% | 5.9% | 91.2% | 0.2% | 2.1% | 0.0% | 0.2% | 0.0% | 0.0% |
| 5 | 0.5% | 5.6% | 89.4% | 0.4% | 2.7% | 0.1% | 0.7% | 0.0% | 0.6% |
| Linked - sub-window(s) to left: 5 | 13.2% | 1.0% | 14.1% | 2.6% | 62.4% | 0.1% | 1.0% | 0.6% | 5.0% |
| 4 | 8.4% | 0.9% | 14.8% | 3.0% | 65.7% | 0.2% | 0.7% | 0.7% | 5.6% |
| 3 | 8.7% | 1.1% | 11.5% | 2.5% | 69.9% | 0.0% | 0.4% | 1.6% | 4.3% |
| 2 | 8.6% | 1.4% | 8.0% | 6.5% | 69.6% | 0.2% | 0.4% | 1.3% | 4.0% |
| 1 | 7.0% | 2.6% | 6.3% | 28.8% | 48.4% | 0.1% | 0.4% | 4.5% | 1.9% |
| Completed Soft Sweep | 2.3% | 8.7% | 0.0% | 85.4% | 0.9% | 0.0% | 0.0% | 2.5% | 0.2% |
| Linked - sub-window(s) to right: 1 | 6.1% | 3.9% | 7.0% | 36.7% | 40.9% | 0.4% | 0.1% | 3.3% | 1.6% |
| 2 | 10.4% | 1.0% | 13.1% | 8.0% | 61.4% | 0.0% | 0.2% | 1.3% | 4.6% |
| 3 | 7.6% | 0.7% | 11.7% | 3.0% | 70.6% | 0.1% | 0.2% | 1.9% | 4.2% |
| 4 | 6.3% | 0.8% | 14.3% | 2.7% | 70.7% | 0.1% | 0.2% | 0.6% | 4.3% |
| 5 | 11.0% | 0.8% | 13.8% | 1.9% | 66.3% | 0.0% | 0.4% | 0.4% | 5.4% |
| Linked - sub-window(s) to left: 5 | 2.9% | 0.0% | 5.5% | 0.3% | 1.5% | 10.4% | 72.7% | 0.5% | 6.2% |
| 4 | 2.3% | 0.0% | 5.0% | 0.1% | 0.6% | 11.8% | 75.3% | 0.7% | 4.2% |
| 3 | 2.8% | 0.4% | 4.0% | 0.2% | 1.1% | 13.8% | 71.4% | 0.8% | 5.5% |
| 2 | 2.2% | 0.1% | 4.6% | 0.4% | 0.6% | 24.7% | 61.8% | 0.7% | 4.9% |
| 1 | 2.0% | 0.8% | 3.3% | 0.4% | 0.4% | 42.5% | 45.0% | 2.7% | 2.9% |
| Partial Hard Sweep | 1.8% | 2.4% | 0.6% | 0.4% | 0.1% | 56.0% | 32.5% | 5.0% | 1.2% |
| Linked - sub-window(s) to right: 1 | 1.4% | 1.1% | 4.2% | 0.5% | 0.2% | 40.5% | 46.7% | 2.9% | 2.5% |
| 2 | 2.4% | 0.2% | 6.9% | 0.3% | 0.5% | 23.1% | 61.9% | 0.3% | 4.4% |
| 3 | 1.9% | 0.3% | 5.2% | 0.1% | 0.8% | 12.1% | 75.4% | 0.6% | 3.6% |
| 4 | 2.2% | 0.1% | 8.7% | 0.3% | 0.6% | 8.7% | 74.9% | 0.7% | 3.8% |
| 5 | 2.3% | 0.2% | 7.7% | 0.3% | 0.6% | 7.4% | 75.7% | 0.3% | 5.5% |
| Linked - sub-window(s) to left: 5 | 46.9% | 0.2% | 1.0% | 1.7% | 13.4% | 1.1% | 6.1% | 1.8% | 27.8% |
| 4 | 41.7% | 0.2% | 1.9% | 1.5% | 14.4% | 1.3% | 5.9% | 1.5% | 31.6% |
| 3 | 38.2% | 0.1% | 0.5% | 2.1% | 13.9% | 3.4% | 9.0% | 2.3% | 30.5% |
| 2 | 38.9% | 0.1% | 0.7% | 3.6% | 12.2% | 3.8% | 5.7% | 5.9% | 29.1% |
| 1 | 41.5% | 0.1% | 0.7% | 8.3% | 9.0% | 5.3% | 3.4% | 18.3% | 13.4% |
| Partial Soft Sweep | 28.0% | 0.3% | 0.2% | 8.3% | 1.5% | 8.2% | 2.0% | 49.1% | 2.4% |
| Linked - sub-window(s) to right: 1 | 40.9% | 0.4% | 1.2% | 8.2% | 7.2% | 5.3% | 3.8% | 17.4% | 15.6% |
| 2 | 41.9% | 0.1% | 1.9% | 3.3% | 12.5% | 3.3% | 5.2% | 4.8% | 27.0% |
| 3 | 39.0% | 0.0% | 2.4% | 2.4% | 14.2% | 1.6% | 8.9% | 2.4% | 29.1% |
| 4 | 43.3% | 0.1% | 1.1% | 1.5% | 12.8% | 1.1% | 8.1% | 1.2% | 30.8% |
| 5 | 44.3% | 0.2% | 0.9% | 2.5% | 12.3% | 1.2% | 8.4% | 1.3% | 28.9% |

Fraction of simulations assigned to state (0.0 – 1.0)

## GAS

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep | Partial Hard Sweep | Linked Partial Hard Sweep | Partial Soft Sweep | Linked Partial Soft Sweep |
|---|---|---|---|---|---|---|---|---|---|
| Neutral | 85.2% | 0.0% | 0.0% | 1.3% | 2.0% | 0.3% | 0.0% | 1.7% | 9.5% |
| Linked - sub-window(s) to left: 5 | 0.4% | 5.1% | 87.1% | 0.1% | 1.6% | 0.2% | 5.0% | 0.1% | 0.4% |
| 4 | 0.1% | 5.4% | 88.3% | 0.0% | 1.8% | 0.0% | 3.5% | 0.0% | 0.9% |
| 3 | 0.1% | 7.0% | 88.7% | 0.2% | 2.2% | 0.1% | 1.4% | 0.0% | 0.3% |
| 2 | 0.4% | 14.5% | 81.7% | 0.3% | 1.1% | 0.1% | 1.2% | 0.0% | 0.7% |
| 1 | 0.5% | 31.6% | 64.7% | 0.4% | 1.0% | 0.7% | 0.7% | 0.0% | 0.4% |
| Completed Hard Sweep | 0.0% | 95.2% | 2.2% | 1.9% | 0.1% | 0.6% | 0.0% | 0.0% | 0.0% |
| Linked - sub-window(s) to right: 1 | 0.2% | 28.0% | 67.3% | 0.8% | 2.2% | 0.9% | 0.2% | 0.2% | 0.2% |
| 2 | 0.1% | 13.0% | 83.1% | 0.3% | 1.8% | 0.3% | 0.9% | 0.0% | 0.5% |
| 3 | 0.1% | 7.5% | 88.5% | 0.4% | 1.9% | 0.0% | 1.0% | 0.0% | 0.6% |
| 4 | 0.2% | 5.9% | 91.2% | 0.0% | 1.3% | 0.0% | 1.1% | 0.0% | 0.3% |
| 5 | 0.5% | 3.7% | 92.1% | 0.2% | 1.2% | 0.0% | 2.2% | 0.0% | 0.1% |
| Linked - sub-window(s) to left: 5 | 8.2% | 0.0% | 5.5% | 0.6% | 69.7% | 0.1% | 0.7% | 0.5% | 14.7% |
| 4 | 6.4% | 0.1% | 6.2% | 1.0% | 73.4% | 0.0% | 0.7% | 0.4% | 11.8% |
| 3 | 5.2% | 0.4% | 6.3% | 0.6% | 76.3% | 0.0% | 0.9% | 0.4% | 9.9% |
| 2 | 6.0% | 1.1% | 5.2% | 3.6% | 76.2% | 0.0% | 0.1% | 0.7% | 7.1% |
| 1 | 5.8% | 2.4% | 4.6% | 19.3% | 58.9% | 0.1% | 0.1% | 3.4% | 5.4% |
| Completed Soft Sweep | 2.6% | 7.5% | 0.3% | 83.2% | 3.0% | 0.3% | 0.0% | 2.7% | 0.4% |
| Linked - sub-window(s) to right: 1 | 7.3% | 2.3% | 6.6% | 16.0% | 60.1% | 0.2% | 0.2% | 2.9% | 4.4% |
| 2 | 6.0% | 0.5% | 4.7% | 1.5% | 79.1% | 0.2% | 0.2% | 0.9% | 6.9% |
| 3 | 5.8% | 0.3% | 5.1% | 1.2% | 78.9% | 0.3% | 0.3% | 0.3% | 7.8% |
| 4 | 5.0% | 0.2% | 8.3% | 0.8% | 78.1% | 0.1% | 0.5% | 0.4% | 6.6% |
| 5 | 7.7% | 0.5% | 10.3% | 0.4% | 70.3% | 0.0% | 0.8% | 0.4% | 9.6% |
| Linked - sub-window(s) to left: 5 | 1.2% | 0.1% | 1.7% | 0.2% | 0.3% | 17.5% | 74.8% | 0.2% | 4.0% |
| 4 | 1.4% | 0.1% | 1.4% | 0.0% | 0.3% | 18.7% | 75.1% | 0.1% | 2.9% |
| 3 | 0.3% | 0.1% | 2.9% | 0.0% | 0.4% | 21.3% | 71.0% | 0.0% | 4.0% |
| 2 | 0.9% | 0.1% | 3.5% | 0.2% | 0.3% | 29.8% | 61.1% | 0.3% | 3.8% |
| 1 | 0.6% | 1.0% | 3.0% | 0.0% | 0.2% | 47.7% | 44.0% | 0.9% | 2.6% |
| Partial Hard Sweep | 0.7% | 2.5% | 0.7% | 0.1% | 0.0% | 71.9% | 21.3% | 1.7% | 1.1% |
| Linked - sub-window(s) to right: 1 | 0.9% | 0.1% | 3.3% | 0.1% | 0.3% | 49.9% | 42.5% | 0.7% | 2.2% |
| 2 | 0.8% | 0.3% | 4.2% | 0.0% | 0.8% | 29.6% | 60.3% | 0.6% | 3.4% |
| 3 | 1.1% | 0.2% | 3.3% | 0.0% | 0.3% | 19.3% | 73.3% | 0.0% | 2.5% |
| 4 | 0.8% | 0.2% | 3.8% | 0.0% | 0.4% | 14.0% | 77.3% | 0.1% | 3.4% |
| 5 | 1.1% | 0.1% | 2.3% | 0.0% | 0.2% | 13.3% | 78.1% | 0.2% | 4.7% |
| Linked - sub-window(s) to left: 5 | 33.5% | 0.0% | 0.3% | 0.7% | 8.1% | 0.9% | 11.1% | 1.2% | 44.2% |
| 4 | 32.5% | 0.0% | 0.3% | 0.7% | 9.9% | 1.2% | 10.3% | 1.3% | 43.8% |
| 3 | 27.5% | 0.0% | 0.4% | 0.9% | 9.1% | 1.5% | 9.7% | 2.2% | 48.7% |
| 2 | 30.7% | 0.2% | 0.2% | 1.9% | 10.0% | 1.7% | 8.1% | 4.8% | 42.4% |
| 1 | 31.3% | 0.1% | 0.1% | 4.5% | 10.2% | 5.8% | 4.4% | 9.4% | 34.2% |
| Partial Soft Sweep | 24.5% | 0.3% | 0.2% | 11.3% | 2.0% | 10.5% | 1.3% | 42.4% | 7.5% |
| Linked - sub-window(s) to right: 1 | 32.3% | 0.3% | 0.2% | 5.9% | 9.0% | 4.4% | 5.4% | 10.2% | 32.3% |
| 2 | 33.2% | 0.1% | 0.6% | 1.5% | 12.4% | 2.0% | 7.0% | 3.3% | 39.9% |
| 3 | 29.3% | 0.0% | 1.6% | 1.0% | 11.7% | 1.7% | 8.3% | 1.2% | 45.2% |
| 4 | 30.1% | 0.1% | 0.7% | 0.7% | 12.1% | 1.1% | 10.6% | 1.2% | 43.4% |
| 5 | 37.7% | 0.0% | 0.4% | 0.9% | 10.6% | 1.2% | 8.7% | 1.6% | 38.9% |

Fraction of simulations assigned to state (0.0 – 1.0)

## BFM

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep | Partial Hard Sweep | Linked Partial Hard Sweep | Partial Soft Sweep | Linked Partial Soft Sweep |
|---|---|---|---|---|---|---|---|---|---|
| Neutral | 92.7% | 0.0% | 0.2% | 0.7% | 0.7% | 0.0% | 0.0% | 2.8% | 2.9% |
| Linked - sub-window(s) to left: 5 | 0.3% | 2.9% | 94.0% | 0.0% | 1.8% | 0.0% | 0.7% | 0.0% | 0.3% |
| 4 | 0.6% | 3.1% | 94.0% | 0.0% | 1.6% | 0.0% | 0.5% | 0.0% | 0.2% |
| 3 | 0.0% | 4.2% | 94.2% | 0.1% | 0.4% | 0.0% | 0.3% | 0.1% | 0.7% |
| 2 | 0.2% | 8.7% | 89.4% | 0.1% | 0.4% | 0.0% | 0.2% | 0.0% | 1.0% |
| 1 | 0.3% | 20.3% | 78.0% | 0.1% | 0.4% | 0.3% | 0.2% | 0.3% | 0.1% |
| Completed Hard Sweep | 0.0% | 95.7% | 2.4% | 1.3% | 0.0% | 0.3% | 0.1% | 0.2% | 0.0% |
| Linked - sub-window(s) to right: 1 | 0.1% | 24.9% | 72.1% | 0.0% | 0.3% | 0.3% | 1.6% | 0.0% | 0.7% |
| 2 | 0.3% | 8.1% | 89.8% | 0.0% | 0.6% | 0.0% | 0.8% | 0.0% | 0.4% |
| 3 | 0.2% | 4.9% | 93.5% | 0.0% | 0.5% | 0.0% | 0.5% | 0.0% | 0.4% |
| 4 | 0.2% | 2.5% | 95.3% | 0.0% | 0.2% | 0.0% | 1.2% | 0.0% | 0.6% |
| 5 | 0.0% | 2.5% | 94.7% | 0.0% | 0.9% | 0.0% | 1.1% | 0.1% | 0.7% |
| Linked - sub-window(s) to left: 5 | 25.5% | 0.1% | 1.7% | 0.2% | 62.4% | 0.1% | 0.0% | 0.6% | 9.4% |
| 4 | 18.3% | 0.1% | 1.5% | 0.0% | 68.9% | 0.0% | 0.1% | 0.2% | 10.9% |
| 3 | 13.9% | 0.0% | 2.4% | 0.2% | 74.0% | 0.0% | 0.1% | 0.4% | 9.0% |
| 2 | 17.7% | 0.0% | 2.6% | 1.1% | 67.3% | 0.1% | 0.1% | 0.7% | 10.4% |
| 1 | 14.3% | 0.8% | 1.5% | 12.1% | 54.7% | 0.2% | 0.1% | 8.7% | 7.6% |
| Completed Soft Sweep | 3.4% | 1.6% | 0.2% | 85.2% | 0.4% | 0.1% | 0.1% | 8.5% | 0.5% |
| Linked - sub-window(s) to right: 1 | 15.7% | 1.0% | 1.0% | 12.5% | 51.9% | 0.1% | 0.0% | 8.5% | 9.3% |
| 2 | 16.3% | 0.0% | 2.4% | 2.1% | 62.0% | 0.0% | 0.3% | 1.8% | 15.1% |
| 3 | 14.0% | 0.1% | 1.9% | 0.9% | 69.2% | 0.1% | 0.3% | 1.2% | 12.3% |
| 4 | 18.2% | 0.2% | 3.2% | 0.2% | 62.3% | 0.3% | 0.3% | 0.7% | 14.6% |
| 5 | 20.6% | 0.2% | 2.2% | 0.3% | 61.2% | 0.0% | 0.1% | 0.9% | 14.5% |
| Linked - sub-window(s) to left: 5 | 0.6% | 0.1% | 7.7% | 0.1% | 0.4% | 3.3% | 84.1% | 0.2% | 3.5% |
| 4 | 0.3% | 0.0% | 6.7% | 0.0% | 0.3% | 4.8% | 86.1% | 0.0% | 1.8% |
| 3 | 0.6% | 0.0% | 6.8% | 0.0% | 0.3% | 7.2% | 81.8% | 0.0% | 3.3% |
| 2 | 0.3% | 0.2% | 6.1% | 0.1% | 0.0% | 13.5% | 77.3% | 0.2% | 2.3% |
| 1 | 0.2% | 0.3% | 3.9% | 0.0% | 0.2% | 35.8% | 58.2% | 0.7% | 0.7% |
| Partial Hard Sweep | 0.1% | 2.3% | 0.3% | 0.1% | 0.0% | 66.1% | 29.8% | 1.2% | 0.1% |
| Linked - sub-window(s) to right: 1 | 0.6% | 0.3% | 2.9% | 0.0% | 0.0% | 35.8% | 59.0% | 0.1% | 1.3% |
| 2 | 0.3% | 0.0% | 3.7% | 0.0% | 0.1% | 14.9% | 80.0% | 0.0% | 1.0% |
| 3 | 0.4% | 0.0% | 4.4% | 0.0% | 0.1% | 5.5% | 88.1% | 0.1% | 1.4% |
| 4 | 0.3% | 0.0% | 5.2% | 0.0% | 0.1% | 4.5% | 88.9% | 0.1% | 0.9% |
| 5 | 0.4% | 0.0% | 4.5% | 0.0% | 0.1% | 4.5% | 89.3% | 0.1% | 1.1% |
| Linked - sub-window(s) to left: 5 | 57.2% | 0.0% | 0.2% | 0.5% | 15.9% | 0.0% | 1.5% | 2.3% | 22.4% |
| 4 | 51.0% | 0.0% | 0.3% | 0.3% | 16.5% | 0.2% | 2.2% | 1.2% | 28.3% |
| 3 | 48.9% | 0.0% | 0.2% | 0.5% | 17.6% | 0.2% | 2.2% | 1.8% | 28.6% |
| 2 | 50.7% | 0.0% | 0.4% | 0.9% | 15.2% | 0.5% | 1.6% | 2.5% | 28.2% |
| 1 | 47.0% | 0.0% | 0.1% | 4.2% | 14.0% | 1.6% | 1.4% | 13.5% | 18.2% |
| Partial Soft Sweep | 34.7% | 0.0% | 0.0% | 15.1% | 0.1% | 1.7% | 0.3% | 45.6% | 2.5% |
| Linked - sub-window(s) to right: 1 | 48.3% | 0.0% | 0.1% | 3.0% | 9.6% | 1.3% | 0.9% | 16.1% | 20.7% |
| 2 | 50.5% | 0.0% | 0.1% | 0.8% | 14.4% | 0.4% | 2.4% | 2.1% | 29.3% |
| 3 | 47.5% | 0.0% | 0.6% | 0.5% | 15.6% | 0.3% | 2.9% | 2.0% | 30.6% |
| 4 | 49.7% | 0.1% | 0.2% | 0.3% | 14.8% | 0.3% | 3.0% | 1.0% | 30.6% |
| 5 | 52.0% | 0.0% | 0.3% | 0.2% | 13.5% | 0.4% | 3.3% | 1.4% | 28.9% |

Fraction of simulations assigned to state (0.0 – 1.0)

## GNS

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep | Partial Hard Sweep | Linked Partial Hard Sweep | Partial Soft Sweep | Linked Partial Soft Sweep |
|---|---|---|---|---|---|---|---|---|---|
| Neutral | 84.9% | 0.0% | 0.0% | 2.9% | 3.4% | 0.0% | 0.4% | 3.6% | 4.8% |
| Linked - sub-window(s) to left: 5 | 0.1% | 4.3% | 89.4% | 0.0% | 0.9% | 0.5% | 4.5% | 0.0% | 0.3% |
| 4 | 0.2% | 7.4% | 88.2% | 0.0% | 1.5% | 0.2% | 2.5% | 0.0% | 0.0% |
| 3 | 0.1% | 8.9% | 86.1% | 0.0% | 0.9% | 0.2% | 3.5% | 0.0% | 0.3% |
| 2 | 0.0% | 17.0% | 77.2% | 0.1% | 2.0% | 0.6% | 3.0% | 0.0% | 0.1% |
| 1 | 0.0% | 33.9% | 60.6% | 0.0% | 1.1% | 2.6% | 1.7% | 0.0% | 0.1% |
| Completed Hard Sweep | 0.0% | 90.7% | 4.4% | 1.1% | 0.0% | 2.7% | 0.3% | 0.8% | 0.0% |
| Linked - sub-window(s) to right: 1 | 0.2% | 33.1% | 59.5% | 0.2% | 1.8% | 2.7% | 2.1% | 0.2% | 0.2% |
| 2 | 0.0% | 15.0% | 78.6% | 0.0% | 0.9% | 0.6% | 4.4% | 0.0% | 0.5% |
| 3 | 0.1% | 8.3% | 82.6% | 0.0% | 1.1% | 0.4% | 6.7% | 0.0% | 0.8% |
| 4 | 0.3% | 4.1% | 90.6% | 0.0% | 1.1% | 0.1% | 3.5% | 0.0% | 0.3% |
| 5 | 0.2% | 4.4% | 88.3% | 0.0% | 1.9% | 0.6% | 4.3% | 0.0% | 0.3% |
| Linked - sub-window(s) to left: 5 | 36.0% | 0.0% | 2.0% | 0.9% | 47.9% | 0.2% | 1.2% | 2.7% | 9.1% |
| 4 | 28.3% | 0.0% | 2.2% | 1.0% | 58.5% | 0.0% | 0.6% | 1.5% | 7.9% |
| 3 | 25.6% | 0.0% | 2.5% | 1.1% | 61.9% | 0.1% | 0.9% | 1.0% | 6.9% |
| 2 | 25.1% | 0.2% | 1.8% | 2.0% | 57.9% | 0.2% | 0.7% | 3.0% | 9.1% |
| 1 | 22.7% | 0.5% | 1.1% | 16.3% | 45.2% | 0.3% | 0.2% | 7.9% | 5.8% |
| Completed Soft Sweep | 8.8% | 0.9% | 0.1% | 78.3% | 3.2% | 0.0% | 0.1% | 7.3% | 1.3% |
| Linked - sub-window(s) to right: 1 | 24.3% | 0.5% | 0.6% | 18.0% | 39.4% | 0.5% | 0.2% | 9.2% | 7.3% |
| 2 | 23.4% | 0.4% | 1.8% | 2.5% | 59.4% | 0.1% | 0.6% | 2.2% | 9.6% |
| 3 | 23.0% | 0.0% | 2.3% | 0.7% | 63.6% | 0.3% | 1.0% | 1.1% | 8.0% |
| 4 | 24.2% | 0.1% | 1.3% | 0.4% | 63.2% | 0.0% | 0.5% | 1.7% | 8.6% |
| 5 | 33.7% | 0.1% | 1.7% | 1.4% | 51.2% | 0.1% | 0.4% | 1.9% | 9.5% |
| Linked - sub-window(s) to left: 5 | 0.3% | 0.1% | 4.5% | 0.1% | 0.7% | 17.7% | 75.7% | 0.0% | 0.9% |
| 4 | 0.4% | 0.5% | 5.1% | 0.0% | 0.7% | 21.1% | 70.8% | 0.4% | 1.0% |
| 3 | 0.2% | 0.1% | 4.6% | 0.2% | 0.9% | 25.2% | 66.7% | 0.6% | 1.5% |
| 2 | 0.2% | 0.2% | 4.0% | 0.0% | 0.4% | 37.8% | 56.0% | 0.1% | 1.3% |
| 1 | 0.2% | 0.9% | 3.5% | 0.2% | 0.6% | 59.9% | 33.3% | 0.6% | 0.8% |
| Partial Hard Sweep | 0.0% | 3.3% | 0.2% | 0.3% | 0.0% | 80.1% | 13.9% | 1.7% | 0.5% |
| Linked - sub-window(s) to right: 1 | 0.0% | 0.9% | 1.8% | 0.1% | 0.8% | 61.0% | 33.8% | 0.3% | 1.3% |
| 2 | 0.5% | 0.2% | 2.8% | 0.0% | 0.3% | 41.4% | 53.7% | 0.2% | 0.9% |
| 3 | 0.0% | 0.1% | 3.0% | 0.0% | 0.1% | 25.3% | 69.6% | 0.0% | 1.9% |
| 4 | 0.4% | 0.2% | 4.4% | 0.1% | 0.5% | 19.5% | 73.6% | 0.2% | 1.1% |
| 5 | 0.5% | 0.2% | 3.4% | 0.0% | 0.5% | 16.9% | 76.3% | 0.3% | 1.9% |
| Linked - sub-window(s) to left: 5 | 57.4% | 0.0% | 0.2% | 2.2% | 19.8% | 0.2% | 3.7% | 3.2% | 13.3% |
| 4 | 52.4% | 0.0% | 0.5% | 1.2% | 24.3% | 0.4% | 3.2% | 2.9% | 15.1% |
| 3 | 51.7% | 0.0% | 0.8% | 1.2% | 27.5% | 0.7% | 2.5% | 1.9% | 13.7% |
| 2 | 52.5% | 0.0% | 0.5% | 2.2% | 24.6% | 0.7% | 2.1% | 5.4% | 12.0% |
| 1 | 48.8% | 0.0% | 0.1% | 6.7% | 17.2% | 1.4% | 1.5% | 13.3% | 11.0% |
| Partial Soft Sweep | 37.5% | 0.1% | 0.1% | 22.4% | 3.3% | 1.9% | 0.3% | 30.2% | 4.2% |
| Linked - sub-window(s) to right: 1 | 49.7% | 0.0% | 0.0% | 7.7% | 18.5% | 1.5% | 1.1% | 10.6% | 10.9% |
| 2 | 52.5% | 0.0% | 0.2% | 2.7% | 23.2% | 0.5% | 2.8% | 4.6% | 13.5% |
| 3 | 51.5% | 0.0% | 0.3% | 1.7% | 26.2% | 0.4% | 3.3% | 1.8% | 14.8% |
| 4 | 52.1% | 0.0% | 0.4% | 1.1% | 24.7% | 0.7% | 3.1% | 2.0% | 15.9% |
| 5 | 56.1% | 0.0% | 0.1% | 1.6% | 22.3% | 0.4% | 3.5% | 3.2% | 12.8% |

Fraction of simulations assigned to state (0.0 – 1.0)

## BFS

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep | Partial Hard Sweep | Linked Partial Hard Sweep | Partial Soft Sweep | Linked Partial Soft Sweep |
|---|---|---|---|---|---|---|---|---|---|
| Neutral | 90.2% | 0.0% | 0.0% | 0.3% | 0.9% | 0.1% | 0.1% | 2.0% | 6.4% |
| Linked - sub-window(s) to left: 5 | 0.4% | 1.9% | 93.6% | 0.0% | 0.4% | 0.0% | 2.5% | 0.0% | 1.2% |
| 4 | 0.4% | 2.6% | 92.3% | 0.0% | 0.4% | 0.1% | 3.5% | 0.1% | 0.6% |
| 3 | 0.1% | 3.3% | 91.4% | 0.1% | 0.5% | 0.0% | 4.0% | 0.0% | 0.6% |
| 2 | 0.1% | 5.4% | 89.4% | 0.0% | 0.2% | 0.1% | 4.4% | 0.0% | 0.4% |
| 1 | 0.0% | 20.2% | 76.5% | 0.0% | 0.1% | 0.7% | 2.0% | 0.0% | 0.5% |
| Completed Hard Sweep | 0.0% | 96.3% | 3.0% | 0.1% | 0.0% | 0.5% | 0.0% | 0.1% | 0.0% |
| Linked - sub-window(s) to right: 1 | 0.1% | 18.7% | 78.9% | 0.0% | 0.4% | 0.8% | 0.8% | 0.0% | 0.3% |
| 2 | 0.2% | 5.8% | 90.5% | 0.0% | 0.3% | 0.1% | 2.5% | 0.0% | 0.6% |
| 3 | 0.5% | 2.8% | 93.5% | 0.0% | 0.0% | 0.1% | 2.7% | 0.0% | 0.4% |
| 4 | 0.3% | 2.4% | 94.2% | 0.0% | 0.1% | 0.0% | 2.8% | 0.0% | 0.2% |
| 5 | 0.0% | 1.6% | 94.3% | 0.0% | 1.0% | 0.0% | 2.8% | 0.0% | 0.3% |
| Linked - sub-window(s) to left: 5 | 12.6% | 0.0% | 1.0% | 0.0% | 58.3% | 0.0% | 0.5% | 0.6% | 27.0% |
| 4 | 10.0% | 0.0% | 1.0% | 0.1% | 64.7% | 0.0% | 0.3% | 1.1% | 22.8% |
| 3 | 8.4% | 0.0% | 1.8% | 0.2% | 65.7% | 0.0% | 0.1% | 1.0% | 22.8% |
| 2 | 8.4% | 0.0% | 1.1% | 0.2% | 65.6% | 0.0% | 0.2% | 2.8% | 21.7% |
| 1 | 7.3% | 0.8% | 1.1% | 6.6% | 54.2% | 0.2% | 0.5% | 13.9% | 15.4% |
| Completed Soft Sweep | 2.3% | 2.2% | 0.1% | 82.0% | 1.8% | 0.3% | 0.0% | 9.6% | 1.7% |
| Linked - sub-window(s) to right: 1 | 10.3% | 0.6% | 1.2% | 4.5% | 59.7% | 0.2% | 0.2% | 9.4% | 13.9% |
| 2 | 8.8% | 0.0% | 1.7% | 0.2% | 69.7% | 0.2% | 0.7% | 1.1% | 17.6% |
| 3 | 8.5% | 0.1% | 1.2% | 0.1% | 69.4% | 0.1% | 0.6% | 0.8% | 19.2% |
| 4 | 10.3% | 0.0% | 2.0% | 0.2% | 66.0% | 0.0% | 0.3% | 0.7% | 20.5% |
| 5 | 11.4% | 0.0% | 1.4% | 0.1% | 64.5% | 0.1% | 0.3% | 0.6% | 21.6% |
| Linked - sub-window(s) to left: 5 | 0.3% | 0.0% | 2.9% | 0.0% | 0.0% | 8.5% | 86.0% | 0.0% | 2.3% |
| 4 | 0.0% | 0.1% | 3.1% | 0.0% | 0.2% | 8.6% | 87.0% | 0.2% | 0.8% |
| 3 | 0.0% | 0.1% | 1.8% | 0.0% | 0.0% | 10.9% | 86.7% | 0.1% | 0.4% |
| 2 | 0.3% | 0.0% | 1.8% | 0.0% | 0.1% | 21.1% | 75.3% | 0.0% | 1.4% |
| 1 | 0.1% | 0.2% | 2.3% | 0.1% | 0.3% | 45.0% | 50.8% | 0.3% | 0.9% |
| Partial Hard Sweep | 0.2% | 1.6% | 0.3% | 0.0% | 0.1% | 75.9% | 21.1% | 0.7% | 0.1% |
| Linked - sub-window(s) to right: 1 | 0.3% | 0.1% | 3.9% | 0.0% | 0.1% | 46.3% | 47.8% | 0.6% | 0.9% |
| 2 | 0.2% | 0.0% | 2.3% | 0.1% | 0.0% | 20.0% | 75.8% | 0.0% | 1.6% |
| 3 | 0.3% | 0.2% | 2.5% | 0.0% | 0.0% | 11.9% | 83.8% | 0.5% | 0.8% |
| 4 | 0.1% | 0.2% | 2.3% | 0.0% | 0.1% | 8.8% | 87.8% | 0.1% | 0.6% |
| 5 | 0.4% | 0.0% | 3.4% | 0.0% | 0.1% | 9.3% | 85.9% | 0.2% | 0.7% |
| Linked - sub-window(s) to left: 5 | 37.9% | 0.0% | 0.0% | 0.1% | 10.8% | 0.2% | 1.5% | 3.0% | 46.5% |
| 4 | 33.4% | 0.0% | 0.4% | 0.0% | 15.1% | 0.2% | 2.3% | 2.4% | 46.2% |
| 3 | 33.3% | 0.0% | 0.0% | 0.3% | 15.2% | 0.3% | 2.5% | 2.2% | 46.2% |
| 2 | 33.9% | 0.0% | 0.0% | 0.2% | 13.9% | 0.3% | 1.9% | 3.8% | 46.0% |
| 1 | 34.9% | 0.0% | 0.1% | 1.2% | 12.5% | 0.9% | 1.4% | 15.1% | 33.9% |
| Partial Soft Sweep | 27.9% | 0.0% | 0.0% | 12.6% | 1.2% | 1.9% | 0.0% | 47.8% | 8.6% |
| Linked - sub-window(s) to right: 1 | 32.9% | 0.0% | 0.1% | 1.3% | 15.3% | 0.9% | 1.8% | 16.3% | 31.4% |
| 2 | 35.5% | 0.0% | 0.0% | 0.1% | 17.4% | 0.7% | 2.6% | 4.2% | 39.5% |
| 3 | 33.7% | 0.1% | 0.1% | 0.1% | 18.0% | 0.1% | 2.9% | 2.0% | 43.0% |
| 4 | 32.9% | 0.0% | 0.2% | 0.0% | 16.8% | 0.0% | 2.8% | 2.2% | 45.1% |
| 5 | 39.0% | 0.0% | 0.5% | 0.3% | 16.2% | 0.3% | 2.9% | 2.5% | 38.3% |

Fraction of simulations assigned to state (0.0 – 1.0)

## GWA

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep | Partial Hard Sweep | Linked Partial Hard Sweep | Partial Soft Sweep | Linked Partial Soft Sweep |
|---|---|---|---|---|---|---|---|---|---|
| Neutral | 78.6% | 0.0% | 0.0% | 2.7% | 3.3% | 0.0% | 0.0% | 8.6% | 6.8% |
| Linked - sub-window(s) to left: 5 | 0.2% | 4.4% | 91.2% | 0.0% | 1.4% | 0.2% | 2.4% | 0.0% | 0.2% |
| 4 | 0.5% | 5.1% | 90.5% | 0.1% | 1.5% | 0.1% | 1.9% | 0.0% | 0.3% |
| 3 | 0.7% | 8.1% | 88.7% | 0.0% | 1.1% | 0.1% | 0.7% | 0.0% | 0.6% |
| 2 | 0.5% | 15.4% | 80.7% | 0.0% | 0.4% | 0.1% | 2.1% | 0.1% | 0.7% |
| 1 | 0.4% | 30.1% | 65.8% | 0.1% | 0.3% | 1.7% | 1.3% | 0.0% | 0.3% |
| Completed Hard Sweep | 0.0% | 96.4% | 2.5% | 0.4% | 0.1% | 0.5% | 0.1% | 0.0% | 0.0% |
| Linked - sub-window(s) to right: 1 | 0.2% | 28.7% | 65.3% | 0.5% | 0.6% | 1.6% | 3.0% | 0.0% | 0.1% |
| 2 | 0.4% | 12.5% | 80.2% | 0.0% | 0.6% | 0.4% | 5.3% | 0.1% | 0.5% |
| 3 | 0.1% | 5.8% | 88.4% | 0.1% | 0.9% | 0.4% | 3.8% | 0.0% | 0.5% |
| 4 | 0.4% | 4.2% | 91.3% | 0.1% | 0.7% | 0.0% | 2.9% | 0.0% | 0.4% |
| 5 | 0.1% | 4.3% | 88.2% | 0.0% | 0.8% | 0.0% | 5.5% | 0.1% | 1.0% |
| Linked - sub-window(s) to left: 5 | 14.7% | 0.1% | 2.1% | 0.9% | 63.4% | 0.1% | 0.0% | 2.9% | 15.8% |
| 4 | 11.3% | 0.2% | 2.5% | 0.9% | 

| | | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep | Partial Hard Sweep | Linked Partial Hard Sweep | Partial Soft Sweep | Linked Partial Soft Sweep |
|---|---|---|---|---|---|---|---|---|---|---|
| Neutral | | 96.3% | 0.0% | 0.0% | 0.3% | 0.1% | 0.7% | 0.3% | 0.7% | 1.6% |
| Linked - sub-window(s) to left: | 5 | 0.5% | 2.4% | 87.7% | 0.1% | 0.5% | 0.0% | 8.6% | 0.0% | 0.2% |
| | 4 | 0.4% | 3.9% | 87.8% | 0.0% | 0.5% | 0.1% | 6.8% | 0.0% | 0.5% |
| | 3 | 0.7% | 3.3% | 91.7% | 0.1% | 0.2% | 0.0% | 3.0% | 0.0% | 1.0% |
| | 2 | 0.5% | 7.0% | 89.5% | 0.2% | 0.5% | 0.0% | 2.0% | 0.0% | 0.3% |
| | 1 | 0.6% | 20.9% | 75.0% | 0.1% | 0.7% | 0.6% | 1.8% | 0.1% | 0.2% |
| Completed Hard Sweep | | 0.1% | 97.3% | 1.2% | 0.6% | 0.0% | 0.5% | 0.1% | 0.1% | 0.1% |
| Linked - sub-window(s) to right: | 1 | 0.3% | 18.9% | 78.1% | 0.0% | 0.5% | 1.1% | 0.8% | 0.0% | 0.3% |
| | 2 | 0.3% | 6.5% | 90.3% | 0.0% | 0.9% | 0.0% | 1.2% | 0.0% | 0.8% |
| | 3 | 0.5% | 3.2% | 93.2% | 0.0% | 1.0% | 0.0% | 1.2% | 0.0% | 0.9% |
| | 4 | 0.1% | 2.4% | 95.3% | 0.0% | 0.6% | 0.0% | 1.2% | 0.0% | 0.4% |
| | 5 | 1.0% | 1.2% | 94.3% | 0.0% | 0.3% | 0.2% | 2.6% | 0.0% | 0.4% |
| Linked - sub-window(s) to left: | 5 | 45.8% | 0.0% | 3.0% | 0.2% | 17.7% | 1.3% | 10.3% | 0.2% | 21.5% |
| | 4 | 34.5% | 0.1% | 0.7% | 0.0% | 29.4% | 1.6% | 6.7% | 0.6% | 26.4% |
| | 3 | 30.4% | 0.1% | 2.3% | 0.1% | 31.5% | 2.1% | 7.7% | 0.4% | 25.4% |
| | 2 | 38.5% | 0.0% | 1.8% | 0.3% | 28.1% | 2.5% | 4.3% | 1.6% | 22.9% |
| | 1 | 35.4% | 0.4% | 1.5% | 8.5% | 24.6% | 5.9% | 2.0% | 7.5% | 14.2% |
| Completed Soft Sweep | | 21.1% | 1.6% | 0.1% | 63.3% | 0.5% | 2.5% | 0.2% | 9.1% | 1.6% |
| Linked - sub-window(s) to right: | 1 | 36.3% | 0.1% | 1.8% | 7.2% | 24.2% | 4.6% | 1.5% | 7.2% | 17.1% |
| | 2 | 35.8% | 0.1% | 1.6% | 0.2% | 35.8% | 2.4% | 2.9% | 1.4% | 19.8% |
| | 3 | 32.6% | 0.0% | 1.6% | 0.1% | 39.8% | 1.2% | 2.9% | 0.8% | 21.0% |
| | 4 | 32.7% | 0.0% | 1.6% | 0.1% | 37.6% | 1.2% | 4.2% | 0.9% | 21.7% |
| | 5 | 38.5% | 0.0% | 1.7% | 0.4% | 34.3% | 2.0% | 5.3% | 0.5% | 17.3% |
| Linked - sub-window(s) to left: | 5 | 0.5% | 0.0% | 1.5% | 0.0% | 0.0% | 11.2% | 85.0% | 0.0% | 1.8% |
| | 4 | 0.7% | 0.2% | 1.7% | 0.1% | 0.1% | 13.0% | 82.4% | 0.0% | 1.8% |
| | 3 | 0.5% | 0.0% | 2.1% | 0.0% | 0.0% | 15.5% | 79.4% | 0.1% | 2.4% |
| | 2 | 0.8% | 0.0% | 2.3% | 0.0% | 0.0% | 24.7% | 69.6% | 0.3% | 2.3% |
| | 1 | 0.6% | 0.3% | 2.5% | 0.0% | 0.1% | 50.4% | 44.8% | 0.2% | 1.1% |
| Partial Hard Sweep | | 0.5% | 1.9% | 0.3% | 0.0% | 0.1% | 78.8% | 17.9% | 0.5% | 0.0% |
| Linked - sub-window(s) to right: | 1 | 0.6% | 0.1% | 4.3% | 0.1% | 0.0% | 49.1% | 44.7% | 0.1% | 1.0% |
| | 2 | 0.9% | 0.0% | 4.9% | 0.0% | 0.0% | 26.2% | 66.5% | 0.0% | 1.5% |
| | 3 | 0.3% | 0.1% | 3.9% | 0.0% | 0.4% | 17.9% | 75.7% | 0.0% | 1.7% |
| | 4 | 0.7% | 0.1% | 4.8% | 0.0% | 0.0% | 11.6% | 80.3% | 0.1% | 2.4% |
| | 5 | 0.3% | 0.0% | 3.6% | 0.0% | 0.2% | 9.7% | 83.2% | 0.1% | 2.9% |
| Linked - sub-window(s) to left: | 5 | 66.6% | 0.0% | 0.1% | 0.0% | 1.3% | 2.8% | 9.8% | 0.5% | 18.9% |
| | 4 | 60.5% | 0.0% | 0.0% | 0.2% | 2.9% | 2.1% | 11.3% | 0.1% | 22.9% |
| | 3 | 57.9% | 0.1% | 0.0% | 0.0% | 3.2% | 2.7% | 8.6% | 0.5% | 27.0% |
| | 2 | 66.4% | 0.0% | 0.0% | 0.2% | 3.3% | 3.1% | 3.2% | 1.5% | 22.3% |
| | 1 | 62.5% | 0.0% | 0.0% | 1.9% | 2.6% | 7.7% | 3.5% | 6.7% | 15.1% |
| Partial Soft Sweep | | 49.0% | 0.2% | 0.0% | 10.9% | 0.2% | 8.6% | 0.7% | 28.5% | 1.9% |
| Linked - sub-window(s) to right: | 1 | 62.1% | 0.0% | 0.2% | 1.7% | 2.7% | 6.1% | 2.9% | 6.9% | 17.4% |
| | 2 | 66.3% | 0.0% | 0.2% | 0.3% | 4.6% | 2.3% | 5.3% | 1.3% | 19.7% |
| | 3 | 60.8% | 0.0% | 0.3% | 0.0% | 5.5% | 1.4% | 5.5% | 0.3% | 26.2% |
| | 4 | 61.8% | 0.0% | 0.3% | 0.0% | 6.9% | 1.8% | 5.9% | 0.5% | 22.8% |
| | 5 | 65.6% | 0.1% | 0.1% | 0.1% | 4.8% | 1.7% | 6.2% | 0.1% | 21.3% |

Fraction of simulations assigned to state: 1.0, 0.8, 0.6, 0.4, 0.2, 0.0

## AOM

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep |
|---|---|---|---|---|---|
| Neutral | 94.1% | 0.0% | 0.0% | 1.2% | 4.7% |
| Linked - sub-window(s) to left: 5 | 0.4% | 4.1% | 87.8% | 0.2% | 7.5% |
| 4 | 0.7% | 3.6% | 90.1% | 0.2% | 5.4% |
| 3 | 0.7% | 6.6% | 87.7% | 0.2% | 4.8% |
| 2 | 0.6% | 12.1% | 81.8% | 0.1% | 5.4% |
| 1 | 0.9% | 38.9% | 54.1% | 1.0% | 5.1% |
| Completed Hard Sweep | 0.0% | 91.0% | 6.2% | 2.6% | 0.2% |
| Linked - sub-window(s) to right: 1 | 0.4% | 28.9% | 67.7% | 0.1% | 2.9% |
| 2 | 0.8% | 10.8% | 85.5% | 0.0% | 2.9% |
| 3 | 0.7% | 7.3% | 88.6% | 0.0% | 3.4% |
| 4 | 0.2% | 4.8% | 89.5% | 0.0% | 5.5% |
| 5 | 0.4% | 4.8% | 89.5% | 0.0% | 5.3% |
| Linked - sub-window(s) to left: 5 | 10.4% | 0.5% | 8.1% | 0.5% | 80.5% |
| 4 | 8.2% | 0.5% | 9.4% | 0.8% | 81.1% |
| 3 | 7.4% | 1.0% | 8.2% | 0.5% | 82.9% |
| 2 | 7.5% | 0.9% | 6.4% | 1.7% | 83.5% |
| 1 | 6.5% | 3.5% | 5.4% | 10.6% | 74.0% |
| Completed Soft Sweep | 2.3% | 10.0% | 1.1% | 82.8% | 3.8% |
| Linked - sub-window(s) to right: 1 | 4.4% | 2.7% | 7.5% | 14.2% | 71.2% |
| 2 | 6.3% | 1.0% | 11.1% | 1.2% | 80.4% |
| 3 | 6.0% | 0.8% | 9.2% | 0.4% | 83.6% |
| 4 | 5.3% | 0.3% | 10.0% | 0.4% | 84.0% |
| 5 | 10.8% | 0.5% | 9.0% | 0.4% | 79.3% |
| Linked - sub-window(s) to left: 5 | 26.1% | 1.5% | 31.8% | 0.4% | 40.2% |
| 4 | 25.2% | 1.4% | 34.8% | 1.1% | 37.5% |
| 3 | 22.1% | 2.3% | 35.0% | 1.0% | 39.6% |
| 2 | 25.1% | 3.1% | 34.3% | 1.0% | 36.5% |
| 1 | 28.3% | 5.4% | 29.9% | 2.6% | 33.8% |
| Partial Hard Sweep | 27.8% | 12.9% | 23.6% | 5.8% | 29.9% |
| Linked - sub-window(s) to right: 1 | 24.9% | 5.1% | 33.1% | 2.6% | 34.3% |
| 2 | 27.0% | 2.4% | 33.9% | 2.1% | 34.6% |
| 3 | 22.6% | 1.8% | 36.2% | 1.1% | 38.3% |
| 4 | 24.6% | 0.9% | 38.6% | 1.1% | 34.8% |
| 5 | 25.5% | 0.8% | 32.3% | 0.6% | 40.8% |
| Linked - sub-window(s) to left: 5 | 51.2% | 0.1% | 3.0% | 0.5% | 45.2% |
| 4 | 49.5% | 0.3% | 3.9% | 0.6% | 45.7% |
| 3 | 45.2% | 0.1% | 3.4% | 1.0% | 50.3% |
| 2 | 49.2% | 0.2% | 4.1% | 2.4% | 44.1% |
| 1 | 52.2% | 1.2% | 3.1% | 7.5% | 36.0% |
| Partial Soft Sweep | 51.5% | 1.6% | 2.0% | 22.1% | 22.8% |
| Linked - sub-window(s) to right: 1 | 50.8% | 0.7% | 3.1% | 7.8% | 37.6% |
| 2 | 49.0% | 0.1% | 3.1% | 1.7% | 46.1% |
| 3 | 48.1% | 0.1% | 4.2% | 1.2% | 46.4% |
| 4 | 52.7% | 0.2% | 2.5% | 0.9% | 43.7% |
| 5 | 53.8% | 0.1% | 2.3% | 0.9% | 42.9% |

## GAS

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep |
|---|---|---|---|---|---|
| Neutral | 97.7% | 0.0% | 0.0% | 0.2% | 2.1% |
| Linked - sub-window(s) to left: 5 | 0.6% | 4.6% | 92.8% | 0.0% | 2.0% |
| 4 | 0.4% | 4.6% | 92.3% | 0.1% | 2.6% |
| 3 | 0.2% | 6.5% | 90.2% | 0.1% | 3.0% |
| 2 | 0.5% | 11.8% | 84.2% | 0.1% | 3.4% |
| 1 | 0.6% | 26.0% | 71.2% | 0.2% | 2.0% |
| Completed Hard Sweep | 0.2% | 93.0% | 4.4% | 2.3% | 0.1% |
| Linked - sub-window(s) to right: 1 | 0.6% | 28.3% | 68.0% | 0.4% | 2.7% |
| 2 | 0.4% | 12.5% | 84.1% | 0.2% | 2.8% |
| 3 | 0.4% | 6.7% | 89.4% | 0.0% | 3.5% |
| 4 | 0.3% | 4.6% | 91.5% | 0.0% | 3.6% |
| 5 | 0.7% | 2.9% | 93.3% | 0.0% | 3.1% |
| Linked - sub-window(s) to left: 5 | 13.1% | 0.2% | 7.2% | 0.1% | 79.4% |
| 4 | 10.1% | 0.2% | 7.6% | 0.4% | 81.7% |
| 3 | 9.9% | 0.3% | 6.3% | 0.4% | 83.1% |
| 2 | 10.8% | 1.0% | 5.7% | 1.0% | 81.5% |
| 1 | 12.4% | 1.8% | 5.5% | 10.2% | 70.1% |
| Completed Soft Sweep | 7.1% | 6.7% | 0.2% | 81.6% | 4.4% |
| Linked - sub-window(s) to right: 1 | 14.5% | 2.4% | 6.2% | 12.7% | 64.2% |
| 2 | 12.3% | 0.4% | 6.6% | 1.0% | 79.7% |
| 3 | 11.5% | 0.4% | 4.5% | 0.8% | 82.8% |
| 4 | 10.7% | 0.0% | 7.8% | 0.5% | 81.0% |
| 5 | 14.2% | 0.3% | 9.4% | 0.4% | 75.7% |
| Linked - sub-window(s) to left: 5 | 26.1% | 0.3% | 33.6% | 0.5% | 39.5% |
| 4 | 25.1% | 0.3% | 35.6% | 0.6% | 38.4% |
| 3 | 27.7% | 0.7% | 37.3% | 0.6% | 33.7% |
| 2 | 28.0% | 1.0% | 33.1% | 0.3% | 37.6% |
| 1 | 27.6% | 3.7% | 35.4% | 1.3% | 32.0% |
| Partial Hard Sweep | 32.4% | 13.4% | 27.1% | 5.2% | 21.9% |
| Linked - sub-window(s) to right: 1 | 29.8% | 3.0% | 38.9% | 1.7% | 26.6% |
| 2 | 27.1% | 1.7% | 41.6% | 1.3% | 28.3% |
| 3 | 24.2% | 0.6% | 44.3% | 0.3% | 30.6% |
| 4 | 26.0% | 0.6% | 42.2% | 0.5% | 30.7% |
| 5 | 24.1% | 0.3% | 39.0% | 0.5% | 36.1% |
| Linked - sub-window(s) to left: 5 | 53.3% | 0.0% | 3.4% | 0.5% | 42.8% |
| 4 | 52.5% | 0.0% | 4.7% | 0.3% | 42.5% |
| 3 | 52.8% | 0.2% | 2.7% | 0.4% | 43.9% |
| 2 | 54.4% | 0.1% | 2.9% | 0.7% | 41.9% |
| 1 | 54.9% | 0.2% | 2.0% | 3.0% | 39.9% |
| Partial Soft Sweep | 50.6% | 2.5% | 2.6% | 25.4% | 18.9% |
| Linked - sub-window(s) to right: 1 | 60.7% | 0.3% | 3.5% | 6.3% | 29.2% |
| 2 | 60.8% | 0.1% | 3.2% | 2.1% | 33.8% |
| 3 | 57.3% | 0.0% | 4.3% | 0.7% | 37.7% |
| 4 | 52.2% | 0.2% | 3.2% | 0.3% | 44.1% |
| 5 | 58.8% | 0.0% | 2.8% | 0.4% | 38.0% |

## BFM

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep |
|---|---|---|---|---|---|
| Neutral | 92.1% | 0.0% | 0.1% | 0.9% | 6.9% |
| Linked - sub-window(s) to left: 5 | 0.3% | 2.6% | 95.7% | 0.0% | 1.4% |
| 4 | 0.3% | 2.9% | 96.2% | 0.0% | 0.6% |
| 3 | 0.1% | 4.1% | 95.5% | 0.0% | 0.3% |
| 2 | 0.2% | 8.5% | 90.6% | 0.2% | 0.5% |
| 1 | 0.5% | 18.4% | 80.3% | 0.2% | 0.6% |
| Completed Hard Sweep | 0.0% | 95.2% | 3.5% | 1.1% | 0.2% |
| Linked - sub-window(s) to right: 1 | 0.1% | 23.5% | 75.2% | 0.0% | 1.2% |
| 2 | 0.2% | 7.1% | 91.3% | 0.0% | 1.4% |
| 3 | 0.2% | 4.9% | 93.3% | 0.0% | 1.6% |
| 4 | 0.2% | 2.7% | 95.8% | 0.0% | 1.3% |
| 5 | 0.2% | 2.6% | 95.0% | 0.0% | 2.2% |
| Linked - sub-window(s) to left: 5 | 20.2% | 0.1% | 2.2% | 0.2% | 77.3% |
| 4 | 15.3% | 0.0% | 2.8% | 0.4% | 81.5% |
| 3 | 12.9% | 0.0% | 2.6% | 0.1% | 84.4% |
| 2 | 13.8% | 0.1% | 2.9% | 0.9% | 82.3% |
| 1 | 12.5% | 0.7% | 1.2% | 10.8% | 74.8% |
| Completed Soft Sweep | 4.8% | 2.3% | 0.1% | 88.9% | 3.9% |
| Linked - sub-window(s) to right: 1 | 12.0% | 0.7% | 1.3% | 8.2% | 77.8% |
| 2 | 10.8% | 0.0% | 2.4% | 0.8% | 86.0% |
| 3 | 9.3% | 0.1% | 1.4% | 0.1% | 89.1% |
| 4 | 14.4% | 0.1% | 3.1% | 0.1% | 82.3% |
| 5 | 13.7% | 0.1% | 1.8% | 0.3% | 84.1% |
| Linked - sub-window(s) to left: 5 | 4.6% | 0.5% | 45.8% | 0.7% | 48.4% |
| 4 | 4.6% | 0.2% | 49.9% | 0.9% | 44.4% |
| 3 | 4.6% | 0.4% | 50.4% | 1.2% | 43.4% |
| 2 | 5.1% | 0.9% | 51.1% | 0.9% | 42.0% |
| 1 | 4.3% | 1.3% | 46.8% | 1.1% | 46.5% |
| Partial Hard Sweep | 4.2% | 11.3% | 31.4% | 3.6% | 49.5% |
| Linked - sub-window(s) to right: 1 | 4.2% | 1.8% | 43.7% | 1.6% | 48.7% |
| 2 | 5.3% | 0.4% | 44.2% | 0.9% | 49.2% |
| 3 | 5.0% | 0.1% | 46.2% | 0.3% | 48.4% |
| 4 | 4.4% | 0.2% | 45.9% | 0.3% | 49.2% |
| 5 | 6.6% | 0.1% | 40.8% | 0.7% | 51.8% |
| Linked - sub-window(s) to left: 5 | 55.0% | 0.0% | 1.1% | 0.9% | 43.0% |
| 4 | 49.2% | 0.0% | 2.0% | 0.8% | 48.0% |
| 3 | 49.7% | 0.0% | 2.6% | 1.0% | 46.7% |
| 2 | 50.8% | 0.0% | 1.8% | 1.0% | 46.4% |
| 1 | 48.9% | 0.2% | 1.1% | 5.6% | 44.2% |
| Partial Soft Sweep | 43.2% | 0.8% | 1.0% | 39.6% | 15.4% |
| Linked - sub-window(s) to right: 1 | 52.2% | 0.0% | 0.5% | 5.0% | 42.3% |
| 2 | 48.5% | 0.0% | 0.8% | 0.9% | 49.8% |
| 3 | 48.3% | 0.0% | 0.7% | 0.6% | 50.4% |
| 4 | 49.2% | 0.0% | 0.6% | 0.4% | 49.8% |
| 5 | 52.2% | 0.0% | 0.8% | 0.4% | 46.6% |

## GNS

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep |
|---|---|---|---|---|---|
| Neutral | 73.6% | 0.0% | 0.1% | 11.3% | 15.0% |
| Linked - sub-window(s) to left: 5 | 0.3% | 3.9% | 94.0% | 0.0% | 1.8% |
| 4 | 0.2% | 6.9% | 91.2% | 0.1% | 1.6% |
| 3 | 0.3% | 6.9% | 91.3% | 0.0% | 1.5% |
| 2 | 0.0% | 16.2% | 81.8% | 0.2% | 1.8% |
| 1 | 0.1% | 34.8% | 63.8% | 0.2% | 1.1% |
| Completed Hard Sweep | 0.0% | 95.1% | 3.5% | 1.1% | 0.3% |
| Linked - sub-window(s) to right: 1 | 0.5% | 35.3% | 62.1% | 0.2% | 1.9% |
| 2 | 0.0% | 16.3% | 81.8% | 0.0% | 1.9% |
| 3 | 0.0% | 9.3% | 88.7% | 0.2% | 1.8% |
| 4 | 0.0% | 4.5% | 93.0% | 0.0% | 2.5% |
| 5 | 0.4% | 6.0% | 90.9% | 0.1% | 2.6% |
| Linked - sub-window(s) to left: 5 | 30.0% | 0.0% | 2.9% | 5.3% | 61.8% |
| 4 | 22.2% | 0.0% | 2.7% | 2.7% | 72.4% |
| 3 | 19.5% | 0.0% | 2.3% | 2.4% | 75.8% |
| 2 | 15.7% | 0.1% | 2.0% | 6.1% | 76.1% |
| 1 | 13.3% | 0.6% | 0.9% | 18.6% | 66.6% |
| Completed Soft Sweep | 5.1% | 1.2% | 0.1% | 89.1% | 4.5% |
| Linked - sub-window(s) to right: 1 | 15.4% | 0.3% | 0.2% | 22.4% | 61.7% |
| 2 | 14.4% | 0.5% | 1.9% | 4.9% | 78.3% |
| 3 | 16.0% | 0.0% | 1.9% | 2.0% | 80.1% |
| 4 | 16.1% | 0.0% | 1.0% | 2.3% | 80.6% |
| 5 | 22.4% | 0.0% | 1.3% | 3.4% | 72.9% |
| Linked - sub-window(s) to left: 5 | 2.2% | 0.6% | 40.4% | 3.4% | 53.4% |
| 4 | 1.7% | 1.5% | 40.8% | 2.2% | 53.8% |
| 3 | 1.3% | 0.8% | 45.1% | 2.8% | 50.0% |
| 2 | 1.0% | 1.4% | 43.8% | 3.9% | 49.9% |
| 1 | 1.9% | 4.8% | 41.6% | 5.1% | 46.6% |
| Partial Hard Sweep | 1.3% | 20.4% | 26.1% | 6.7% | 45.5% |
| Linked - sub-window(s) to right: 1 | 1.5% | 7.1% | 36.1% | 6.2% | 49.1% |
| 2 | 1.2% | 1.7% | 41.4% | 3.7% | 52.0% |
| 3 | 1.2% | 1.4% | 46.2% | 3.2% | 48.0% |
| 4 | 1.0% | 0.6% | 49.8% | 2.6% | 46.0% |
| 5 | 1.3% | 0.6% | 46.3% | 3.3% | 48.5% |
| Linked - sub-window(s) to left: 5 | 50.1% | 0.0% | 0.7% | 8.2% | 41.0% |
| 4 | 46.0% | 0.0% | 0.4% | 5.9% | 47.7% |
| 3 | 43.3% | 0.0% | 1.5% | 6.3% | 48.9% |
| 2 | 43.3% | 0.0% | 1.7% | 10.3% | 44.7% |
| 1 | 39.4% | 0.4% | 1.2% | 18.4% | 40.6% |
| Partial Soft Sweep | 31.4% | 0.7% | 0.8% | 48.7% | 18.4% |
| Linked - sub-window(s) to right: 1 | 39.8% | 0.2% | 0.7% | 19.9% | 39.4% |
| 2 | 43.0% | 0.0% | 0.9% | 8.9% | 47.2% |
| 3 | 40.1% | 0.0% | 1.7% | 6.4% | 51.8% |
| 4 | 42.2% | 0.0% | 1.1% | 6.9% | 49.8% |
| 5 | 48.5% | 0.1% | 1.1% | 6.6% | 43.7% |

## BFS

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep |
|---|---|---|---|---|---|
| Neutral | 94.5% | 0.0% | 0.0% | 1.4% | 4.1% |
| Linked - sub-window(s) to left: 5 | 0.5% | 1.1% | 96.8% | 0.0% | 1.6% |
| 4 | 0.4% | 2.0% | 96.2% | 0.1% | 1.3% |
| 3 | 0.3% | 2.5% | 96.0% | 0.0% | 1.2% |
| 2 | 0.1% | 4.8% | 94.4% | 0.1% | 0.6% |
| 1 | 0.1% | 16.4% | 82.9% | 0.3% | 0.3% |
| Completed Hard Sweep | 0.0% | 95.1% | 4.6% | 0.3% | 0.0% |
| Linked - sub-window(s) to right: 1 | 0.4% | 15.3% | 83.1% | 0.1% | 1.1% |
| 2 | 0.2% | 4.4% | 94.4% | 0.0% | 1.0% |
| 3 | 0.3% | 1.9% | 97.3% | 0.0% | 0.5% |
| 4 | 0.8% | 1.5% | 97.2% | 0.0% | 0.5% |
| 5 | 0.1% | 0.9% | 97.8% | 0.0% | 1.2% |
| Linked - sub-window(s) to left: 5 | 15.2% | 0.0% | 1.1% | 0.0% | 83.7% |
| 4 | 12.7% | 0.0% | 1.2% | 0.9% | 85.2% |
| 3 | 12.8% | 0.0% | 1.7% | 1.3% | 84.2% |
| 2 | 12.0% | 0.0% | 1.4% | 2.0% | 84.6% |
| 1 | 11.5% | 0.4% | 1.2% | 18.3% | 68.6% |
| Completed Soft Sweep | 3.2% | 1.2% | 0.2% | 93.8% | 1.6% |
| Linked - sub-window(s) to right: 1 | 13.3% | 0.5% | 1.7% | 18.0% | 66.5% |
| 2 | 13.9% | 0.1% | 1.5% | 1.6% | 82.9% |
| 3 | 9.9% | 0.1% | 1.4% | 0.8% | 87.8% |
| 4 | 13.0% | 0.0% | 2.6% | 0.8% | 83.6% |
| 5 | 17.4% | 0.0% | 1.9% | 0.6% | 80.1% |
| Linked - sub-window(s) to left: 5 | 7.8% | 0.9% | 49.7% | 1.3% | 40.3% |
| 4 | 9.4% | 1.3% | 50.8% | 2.5% | 36.0% |
| 3 | 6.8% | 0.9% | 53.7% | 1.4% | 37.2% |
| 2 | 8.9% | 0.8% | 48.7% | 2.4% | 39.2% |
| 1 | 9.2% | 2.4% | 48.6% | 2.8% | 37.0% |
| Partial Hard Sweep | 9.1% | 16.6% | 31.0% | 3.9% | 39.4% |
| Linked - sub-window(s) to right: 1 | 8.4% | 3.8% | 40.9% | 2.5% | 44.4% |
| 2 | 7.8% | 1.2% | 40.7% | 1.9% | 48.4% |
| 3 | 6.9% | 0.6% | 48.4% | 2.5% | 41.6% |
| 4 | 7.6% | 0.7% | 47.9% | 0.8% | 43.0% |
| 5 | 8.0% | 0.9% | 51.1% | 2.0% | 38.0% |
| Linked - sub-window(s) to left: 5 | 46.5% | 0.0% | 1.4% | 1.4% | 50.7% |
| 4 | 42.3% | 0.0% | 1.9% | 1.9% | 53.9% |
| 3 | 44.5% | 0.0% | 1.2% | 1.5% | 52.8% |
| 2 | 46.1% | 0.2% | 1.1% | 1.8% | 50.8% |
| 1 | 44.2% | 0.1% | 0.5% | 11.1% | 44.1% |
| Partial Soft Sweep | 40.3% | 0.4% | 0.6% | 45.9% | 12.8% |
| Linked - sub-window(s) to right: 1 | 42.6% | 0.1% | 1.2% | 12.2% | 43.9% |
| 2 | 43.7% | 0.0% | 0.7% | 2.9% | 52.7% |
| 3 | 42.1% | 0.0% | 1.3% | 1.8% | 54.8% |
| 4 | 42.4% | 0.0% | 1.6% | 1.3% | 54.7% |
| 5 | 50.1% | 0.0% | 1.8% | 2.2% | 45.9% |

## GWA

| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep |
|---|---|---|---|---|---|
| Neutral | 96.0% | 0.0% | 0.2% | 2.1% | 1.7% |
| Linked - sub-window(s) to left: 5 | 0.6% | 4.0% | 94.4% | 0.1% | 0.9% |
| 4 | 0.5% | 6.4% | 92.0% | 0.1% | 1.0% |
| 3 | 1.1% | 9.6% | 88.5% | 0.0% | 0.8% |
| 2 | 0.8% | 18.2% | 80.3% | 0.1% | 0.6% |
| 1 | 0.4% | 41.1% | 57.9% | 0.4% | 0.2% |
| Completed Hard Sweep | 0.0% | 98.3% | 1.1% | 0.6% | 0.0% |
| Linked - sub-window(s) to right: 1 | 0.5% | 42.6% | 56.1% | 0.4% | 0.4% |
| 2 | 0.4% | 15.8% | 83.1% | 0.2% | 0.5% |
| 3 | 0.4% | 6.9% | 92.2% | 0.0% | 0.5% |
| 4 | 0.7% | 6.1% | 91.7% | 0.0% | 1.5% |
| 5 | 0.6% | 5.4% | 92.5% | 0.2% | 1.3% |
| Linked - sub-window(s) to left: 5 | 27.6% | 0.2% | 2.8% | 0.6% | 68.8% |
| 4 | 21.4% | 0.2% | 3.6% | 1.0% | 73.8% |
| 3 | 20.5% | 0.2% | 3.3% | 0.6% | 75.4% |
| 2 | 19.5% | 0.5% | 2.1% | 4.0% | 73.9% |
| 1 | 16.8% | 1.4% | 1.8% | 29.3% | 50.7% |
| Completed Soft Sweep | 4.5% | 2.1% | 0.0% | 92.6% | 0.8% |
| Linked - sub-window(s) to right: 1 | 18.4% | 1.5% | 1.4% | 26.1% | 52.6% |
| 2 | 23.5% | 0.2% | 2.0% | 3.7% | 70.6% |
| 3 | 25.2% | 0.2% | 2.0% | 1.0% | 71.6% |
| 4 | 22.9% | 0.1% | 2.2% | 1.1% | 73.7% |
| 5 | 28.8% | 0.1% | 1.8% | 1.6% | 67.7% |
| Linked - sub-window(s) to left: 5 | 9.2% | 2.1% | 40.7% | 3.1% | 44.9% |
| 4 | 7.9% | 2.2% | 42.6% | 4.4% | 42.9% |
| 3 | 10.4% | 2.1% | 41.2% | 3.0% | 43.3% |
| 2 | 10.0% | 5.3% | 37.8% | 6.1% | 40.8% |
| 1 | 9.6% | 15.0% | 34.4% | 7.6% | 33.4% |
| Partial Hard Sweep | 8.5% | 32.9% | 19.2% | 9.9% | 29.5% |
| Linked - sub-window(s) to right: 1 | 9.1% | 12.6% | 37.1% | 8.2% | 33.0% |
| 2 | 8.0% | 5.2% | 47.0% | 4.7% | 35.1% |
| 3 | 6.8% | 3.4% | 47.2% | 3.4% | 39.2% |
| 4 | 9.6% | 3.1% | 44.5% | 2.7% | 40.1% |
| 5 | 8.1% | 2.6% | 40.4% | 3.1% | 45.8% |
| Linked - sub-window(s) to left: 5 | 66.9% | 0.1% | 1.2% | 1.7% | 30.1% |
| 4 | 60.6% | 0.0% | 2.2% | 1.6% | 35.6% |
| 3 | 59.2% | 0.1% | 1.1% | 1.5% | 38.1% |
| 2 | 61.2% | 0.1% | 1.3% | 5.4% | 32.0% |
| 1 | 57.5% | 0.7% | 1.1% | 17.8% | 22.9% |
| Partial Soft Sweep | 48.3% | 1.0% | 0.2% | 45.4% | 5.1% |
| Linked - sub-window(s) to right: 1 | 58.5% | 0.2% | 1.6% | 16.5% | 23.2% |
| 2 | 61.5% | 0.2% | 1.7% | 3.7% | 32.9% |
| 3 | 64.6% | 0.3% | 1.4% | 1.4% | 32.3% |
| 4 | 63.6% | 0.1% | 1.2% | 1.7% | 33.4% |
| 5 | 65.9% | 0.0% | 0.6% | 1.9% | 31.6% |

## CMS

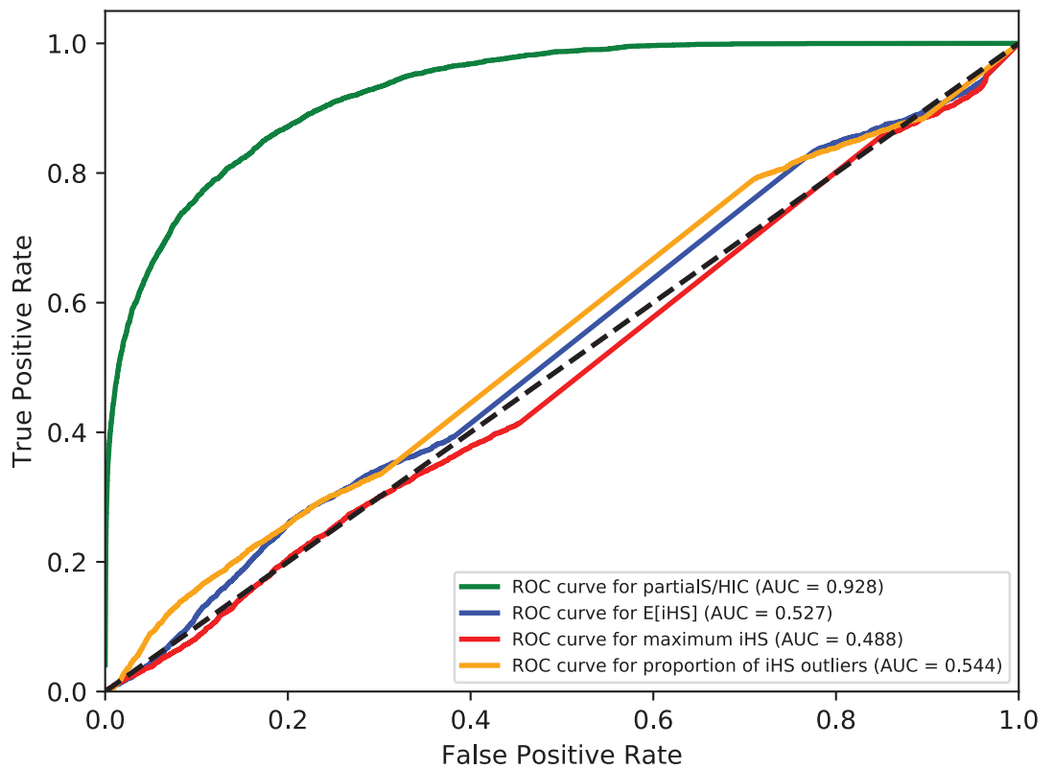
| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep |
|---|---|---|---|---|---|
| Neutral | 78.7% | 0.1% | 0.0% | 6.2% | 15.0% |
| Linked - sub-window(s) to left: 5 | 0.4% | 1.0% | 97.6% | 0.0% | 1.0% |
| 4 | 0.6% | 1.3% | 97.2% | 0.1% | 0.8% |
| 3 | 0.6% | 1.4% | 97.7% | 0.0% | 0.3% |
| 2 | 0.4% | 4.1% | 95.1% | 0.2% | 0.2% |
| 1 | 0.5% | 14.5% | 84.1% | 0.1% | 0.8% |
| Completed Hard Sweep | 0.0% | 96.1% | 2.8% | 1.0% | 0.1% |
| Linked - sub-window(s) to right: 1 | 0.3% | 21.2% | 77.3% | 0.5% | 0.7% |
| 2 | 0.5% | 7.7% | 90.0% | 0.1% | 1.7% |
| 3 | 0.9% | 2.9% | 95.2% | 0.0% | 1.0% |
| 4 | 0.2% | 2.3% | 96.7% | 0.0% | 0.8% |
| 5 | 1.2% | 0.9% | 96.3% | 0.1% | 1.5% |
| Linked - sub-window(s) to left: 5 | 9.8% | 0.1% | 3.0% | 2.1% | 85.0% |
| 4 | 6.5% | 0.0% | 1.3% | 0.7% | 91.5% |
| 3 | 3.0% | 0.1% | 2.3% | 1.2% | 93.4% |
| 2 | 6.5% | 0.0% | 1.3% | 2.1% | 90.1% |
| 1 | 4.0% | 0.4% | 0.8% | 19.9% | 74.9% |
| Completed Soft Sweep | 6.9% | 1.1% | 0.2% | 89.9% | 1.9% |
| Linked - sub-window(s) to right: 1 | 2.7% | 0.4% | 0.8% | 28.2% | 67.9% |
| 2 | 2.8% | 0.2% | 1.5% | 4.9% | 90.6% |
| 3 | 2.7% | 0.2% | 1.2% | 0.8% | 95.1% |
| 4 | 6.4% | 0.0% | 1.3% | 1.9% | 90.4% |
| 5 | 4.2% | 0.1% | 0.9% | 2.8% | 92.0% |
| Linked - sub-window(s) to left: 5 | 6.2% | 0.4% | 42.0% | 2.6% | 48.8% |
| 4 | 4.8% | 0.2% | 47.0% | 2.7% | 45.3% |
| 3 | 5.8% | 0.2% | 48.1% | 2.4% | 43.5% |
| 2 | 6.8% | 0.3% | 42.5% | 3.1% | 47.3% |
| 1 | 5.8% | 2.5% | 40.0% | 6.0% | 45.7% |
| Partial Hard Sweep | 7.7% | 15.0% | 26.0% | 7.7% | 43.6% |
| Linked - sub-window(s) to right: 1 | 6.3% | 2.2% | 42.8% | 4.9% | 43.8% |
| 2 | 6.8% | 0.8% | 42.6% | 3.6% | 46.2% |
| 3 | 5.8% | 0.5% | 45.8% | 3.0% | 44.9% |
| 4 | 6.2% | 0.4% | 48.9% | 3.1% | 41.4% |
| 5 | 7.3% | 0.3% | 46.2% | 3.2% | 43.0% |
| Linked - sub-window(s) to left: 5 | 34.1% | 0.0% | 0.4% | 5.6% | 59.9% |
| 4 | 26.3% | 0.0% | 1.0% | 4.6% | 68.1% |
| 3 | 28.1% | 0.1% | 1.5% | 4.0% | 66.3% |
| 2 | 30.0% | 0.1% | 0.6% | 4.6% | 64.7% |
| 1 | 28.3% | 0.0% | 0.8% | 16.9% | 54.0% |
| Partial Soft Sweep | 24.7% | 0.5% | 0.4% | 54.3% | 20.1% |
| Linked - sub-window(s) to right: 1 | 21.0% | 0.1% | 1.4% | 22.0% | 55.5% |
| 2 | 28.1% | 0.1% | 1.0% | 6.3% | 64.5% |
| 3 | 25.4% | 0.0% | 1.2% | 4.3% | 69.1% |
| 4 | 31.8% | 0.1% | 1.7% | 3.6% | 62.8% |
| 5 | 36.8% | 0.0% | 0.7% | 3.3% | 59.2% |

## UGS

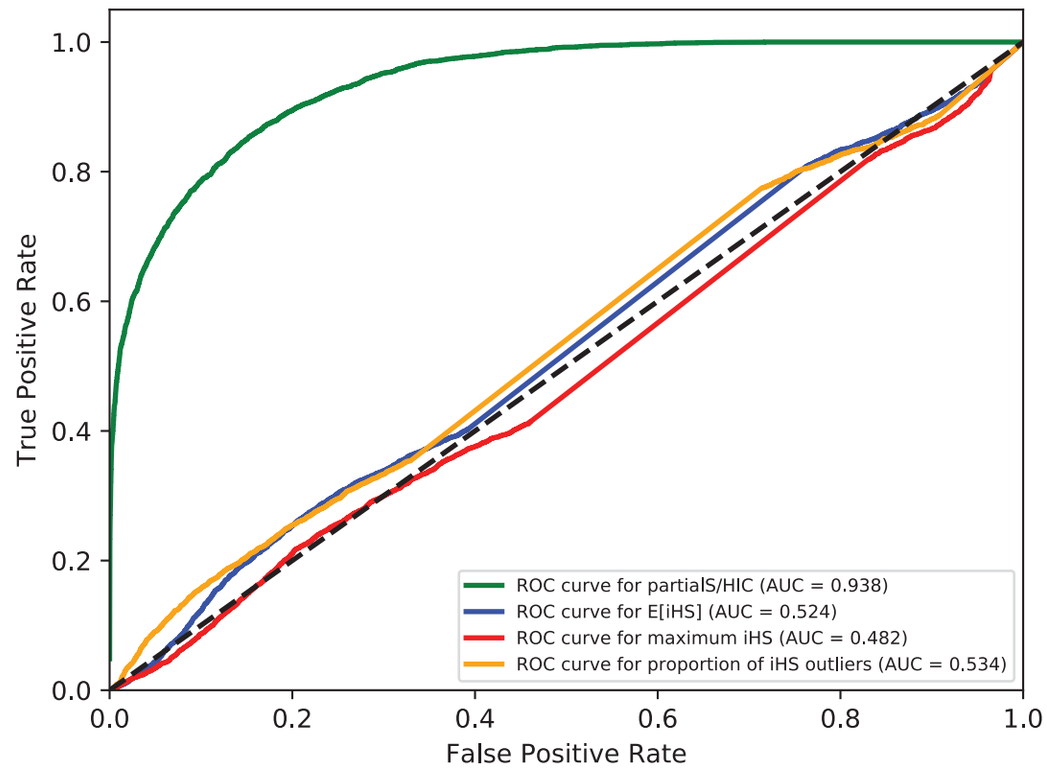
| | Neutral | Completed Hard Sweep | Linked Completed Hard Sweep | Completed Soft Sweep | Linked Completed Soft Sweep |
|---|---|---|---|---|---|
| Neutral | 96.0% | 0.0% | 0.0% | 0.4% | 3.6% |
| Linked - sub-window(s) to left: 5 | 0.3% | 1.5% | 96.6% | 0.1% | 1.5% |
| 4 | 0.2% | 1.5% | 96.7% | 0.0% | 1.6% |
| 3 | 0.5% | 3.4% | 94.6% | 0.0% | 1.5% |
| 2 | 0.9% | 6.0% | 92.3% | 0.0% | 0.8% |
| 1 | 0.5% | 13.4% | 85.8% | 0.0% | 0.3% |
| Completed Hard Sweep | 0.1% | 96.3% | 3.0% | 0.6% | 0.0% |
| Linked - sub-window(s) to right: 1 | 0.1% | 15.2% | 84.0% | 0.0% | 0.7% |
| 2 | 0.2% | 7.3% | 91.3% | 0.0% | 1.2% |
| 3 | 0.3% | 2.4% | 96.3% | 0.0% | 1.0% |
| 4 | 0.4% | 1.0% | 97.6% | 0.0% | 1.0% |
| 5 | 0.4% | 1.6% | 96.2% | 0.0% | 1.8% |
| Linked - sub-window(s) to left: 5 | 11.2% | 0.0% | 1.9% | 0.1% | 86.8% |
| 4 | 5.8% | 0.0% | 1.4% | 0.0% | 92.8% |
| 3 | 7.6% | 0.1% | 2.0% | 0.1% | 90.2% |
| 2 | 8.6% | 0.2% | 2.1% | 0.6% | 88.5% |
| 1 | 7.4% | 0.5% | 1.9% | 9.3% | 80.9% |
| Completed Soft Sweep | 6.0% | 1.5% | 0.5% | 88.6% | 3.4% |
| Linked - sub-window(s) to right: 1 | 9.6% | 0.0% | 1.6% | 9.8% | 79.0% |
| 2 | 7.2% | 0.6% | 1.4% | 0.1% | 90.7% |
| 3 | 9.0% | 0.2% | 1.8% | 0.0% | 89.0% |
| 4 | 8.9% | 0.3% | 2.4% | 0.1% | 88.3% |
| 5 | 12.9% | 0.0% | 1.5% | 0.0% | 85.6% |
| Linked - sub-window(s) to left: 5 | 15.3% | 0.1% | 45.0% | 0.2% | 39.4% |
| 4 | 15.4% | 0.1% | 46.2% | 0.0% | 38.3% |
| 3 | 18.2% | 0.1% | 47.4% | 0.0% | 34.3% |
| 2 | 15.0% | 0.3% | 48.2% | 0.1% | 36.4% |
| 1 | 16.3% | 0.3% | 48.8% | 0.5% | 34.1% |
| Partial Hard Sweep | 17.6% | 9.0% | 40.4% | 0.8% | 32.2% |
| Linked - sub-window(s) to right: 1 | 17.5% | 1.0% | 51.4% | 0.8% | 29.3% |
| 2 | 16.0% | 0.3% | 50.3% | 0.1% | 33.3% |
| 3 | 13.8% | 0.3% | 55.9% | 0.4% | 29.6% |
| 4 | 13.7% | 0.0% | 52.5% | 0.4% | 33.4% |
| 5 | 14.5% | 0.1% | 52.1% | 0.6% | 32.7% |
| Linked - sub-window(s) to left: 5 | 54.2% | 0.0% | 1.5% | 0.0% | 44.3% |
| 4 | 51.4% | 0.1% | 0.7% | 0.0% | 47.8% |
| 3 | 48.4% | 0.0% | 1.5% | 0.2% | 49.9% |
| 2 | 51.7% | 0.0% | 1.3% | 0.7% | 46.3% |
| 1 | 52.7% | 0.1% | 1.3% | 2.6% | 43.3% |
| Partial Soft Sweep | 51.8% | 0.2% | 1.8% | 24.5% | 21.7% |
| Linked - sub-window(s) to right: 1 | 47.9% | 0.0% | 2.8% | 2.4% | 46.9% |
| 2 | 39.8% | 0.1% | 1.4% | 0.7% | 58.0% |
| 3 | 48.2% | 0.0% | 2.0% | 0.3% | 49.5% |
| 4 | 46.5% | 0.0% | 3.3% | 0.1% | 50.1% |
| 5 | 52.5% | 0.0% | 3.1% | 0.6% | 43.8% |

Fraction of simulations assigned to state

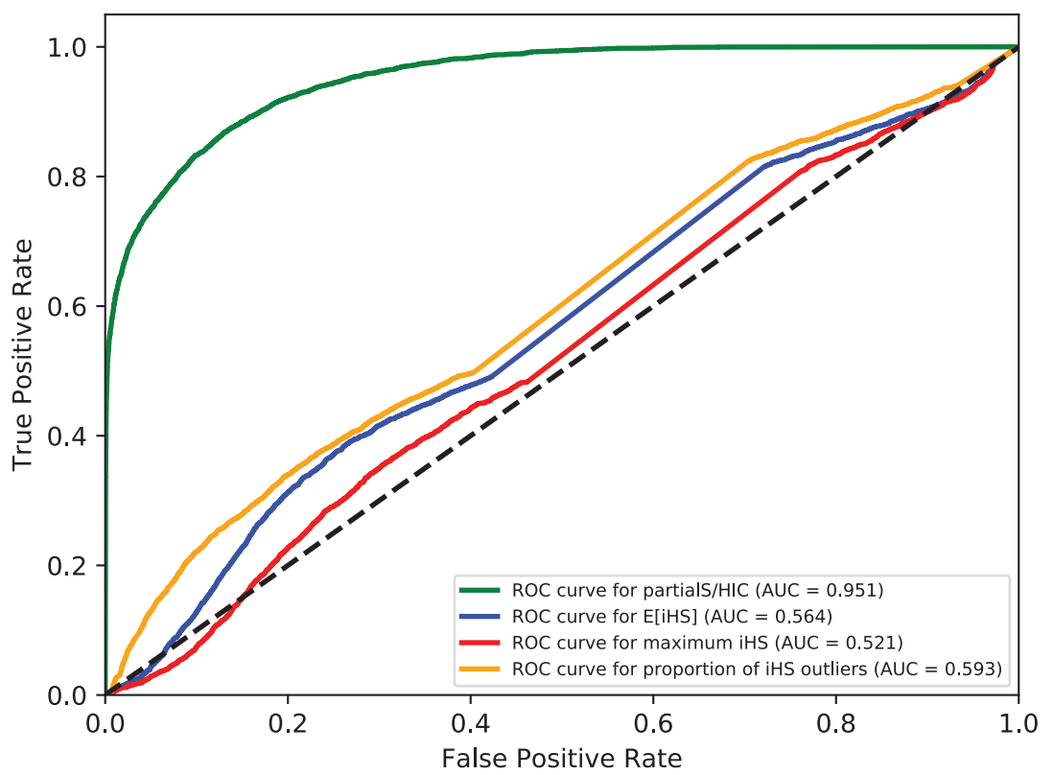
AOM



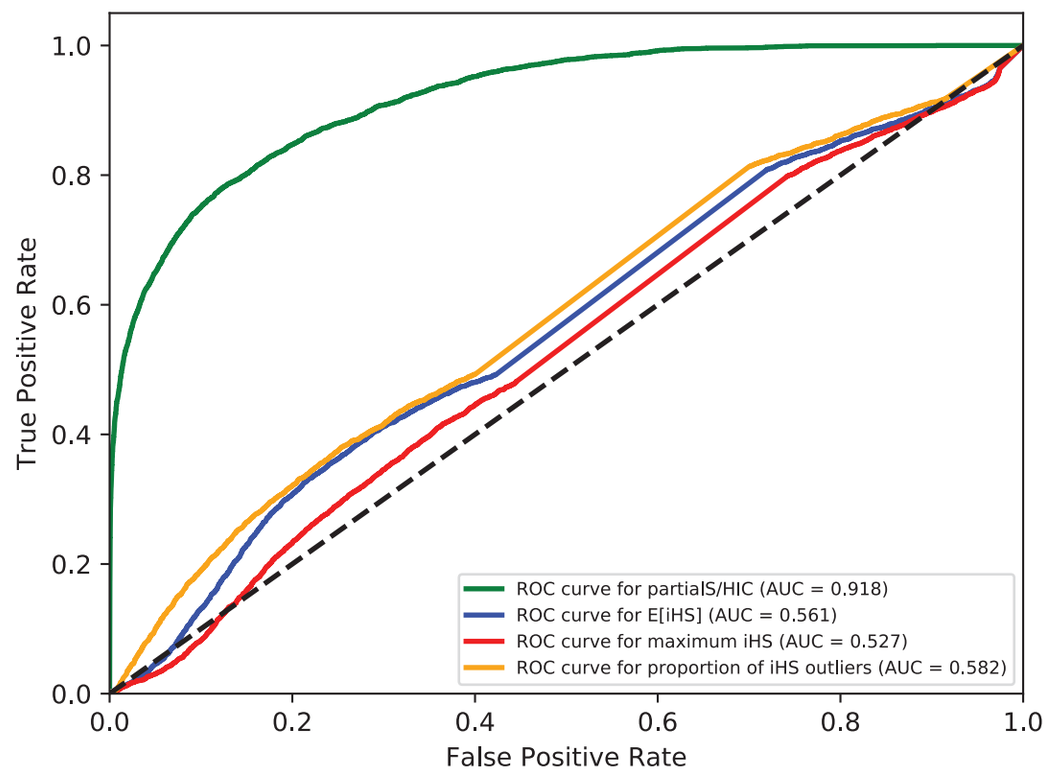

GAS



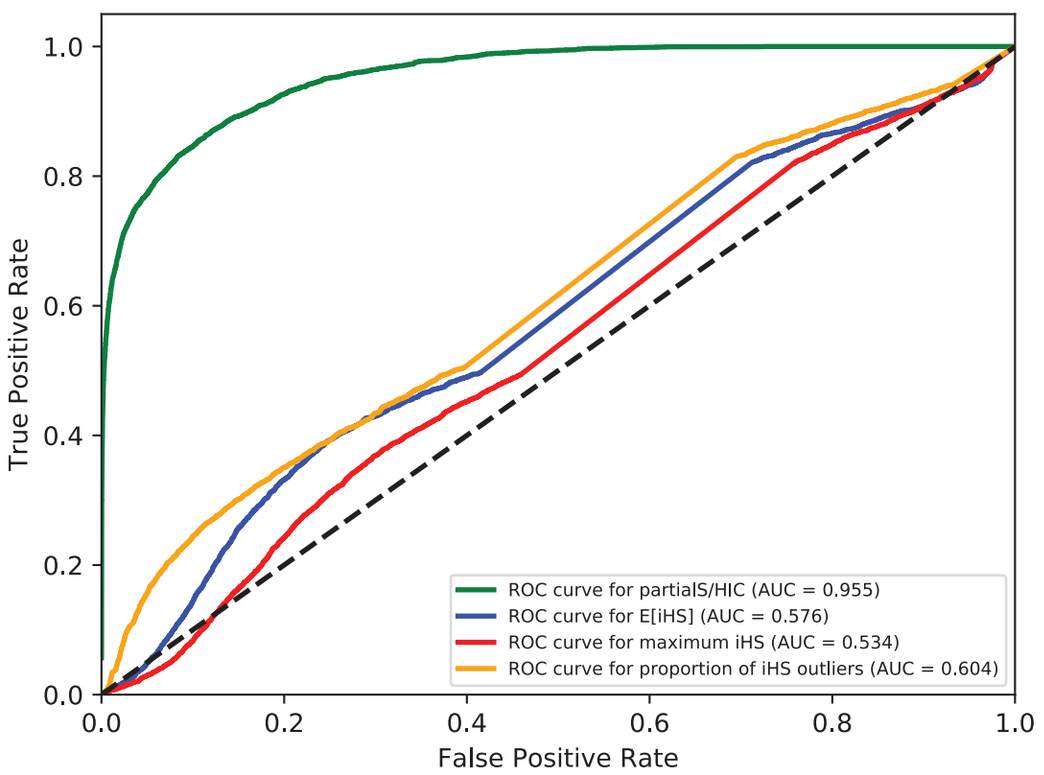

BFM



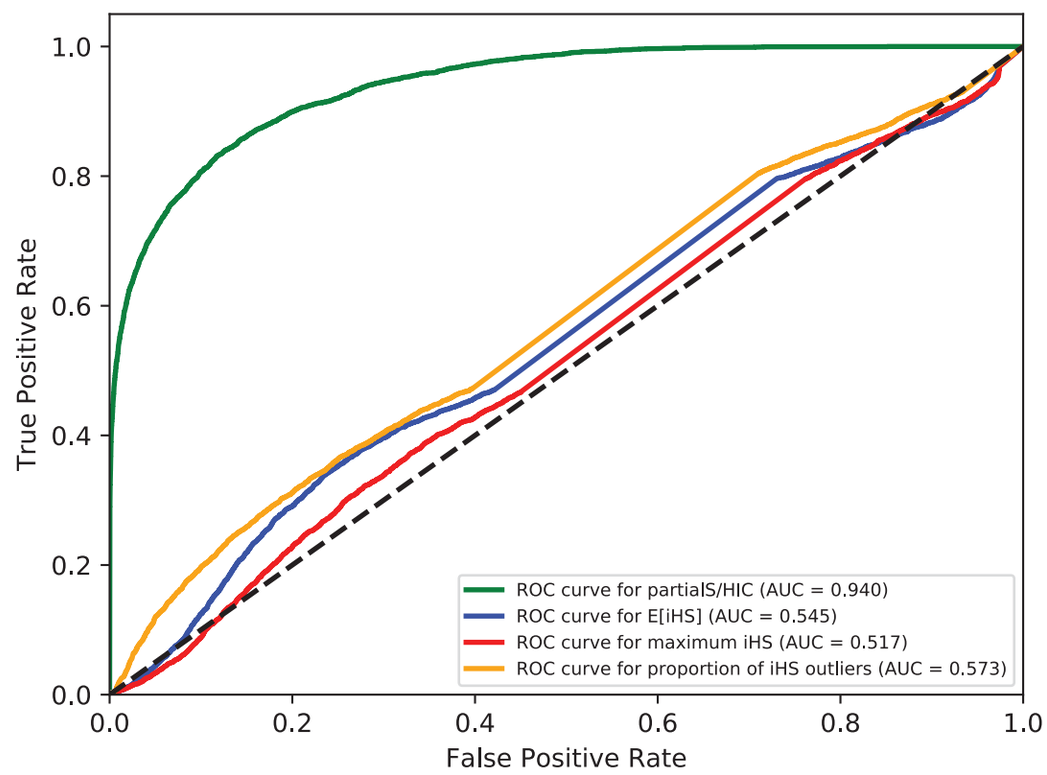

GNS



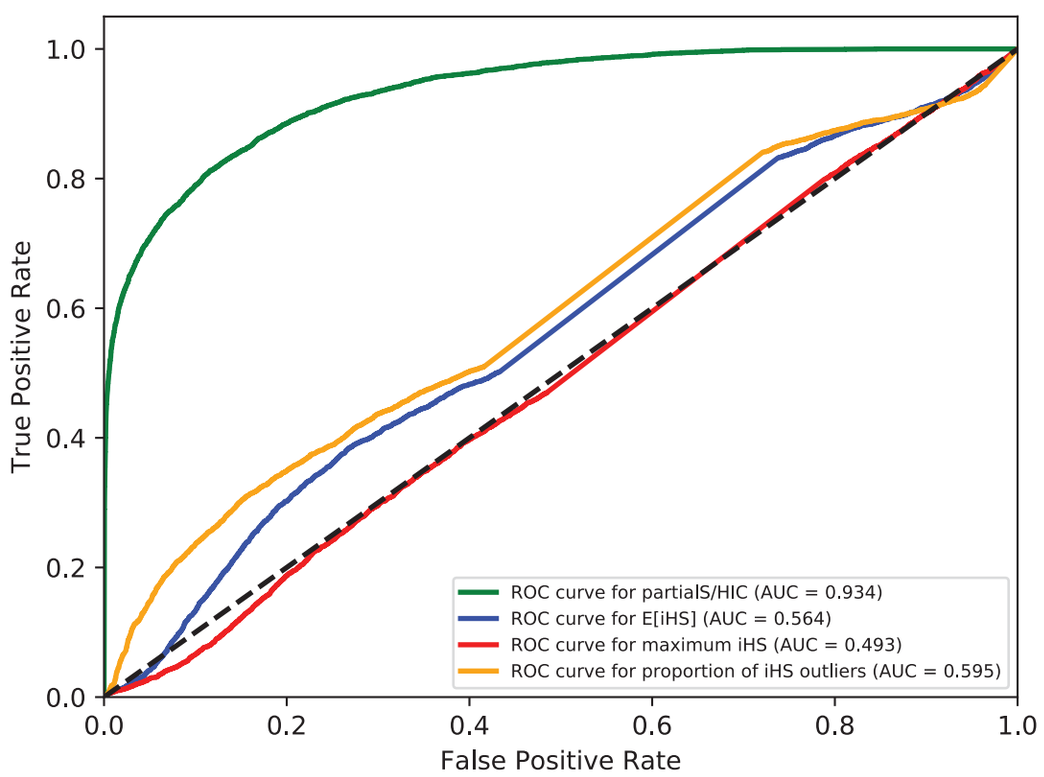

BFS



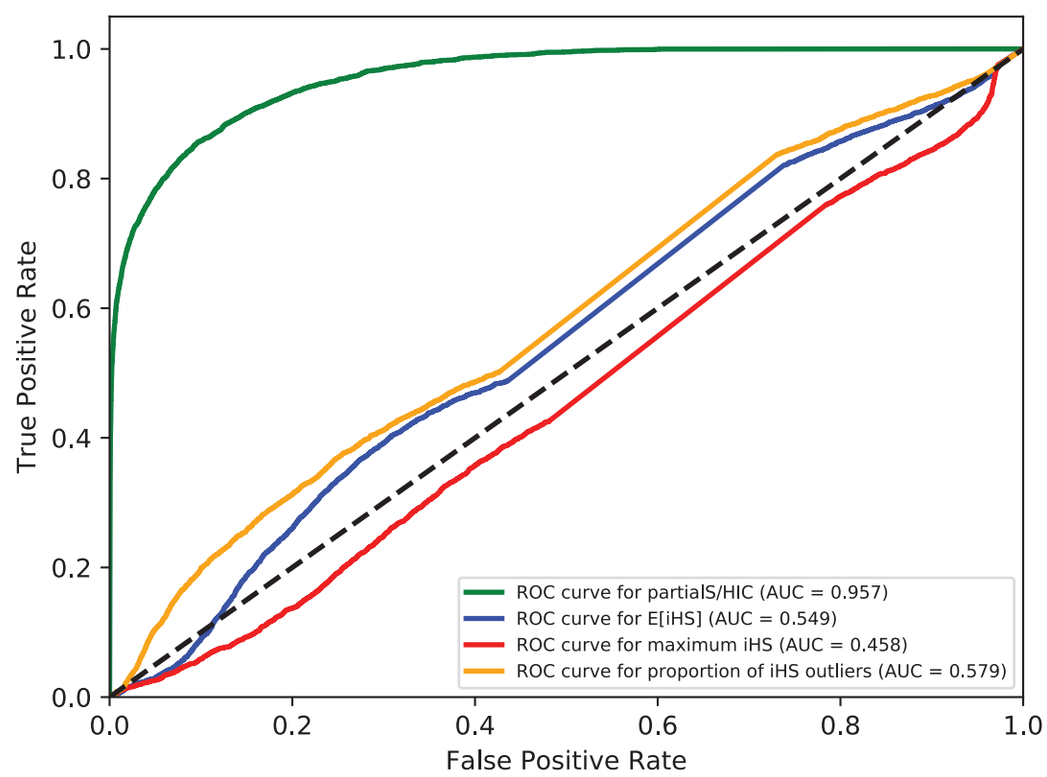

GWA



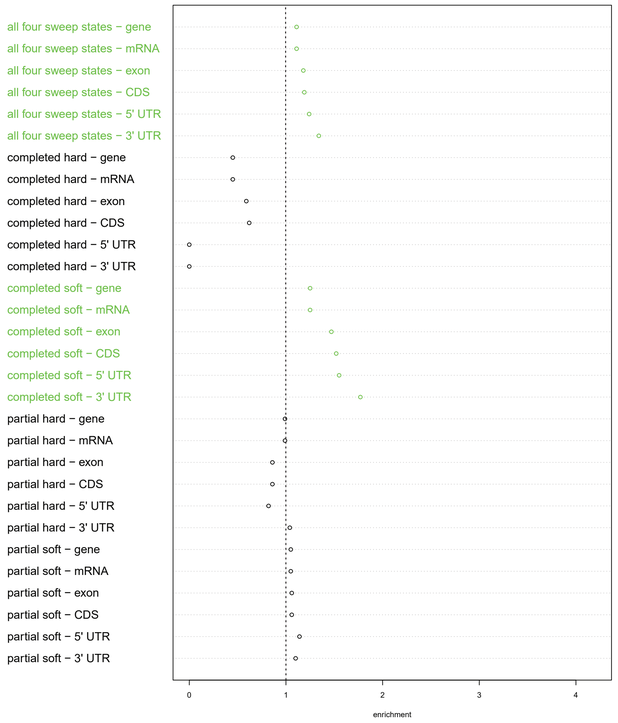

CMS



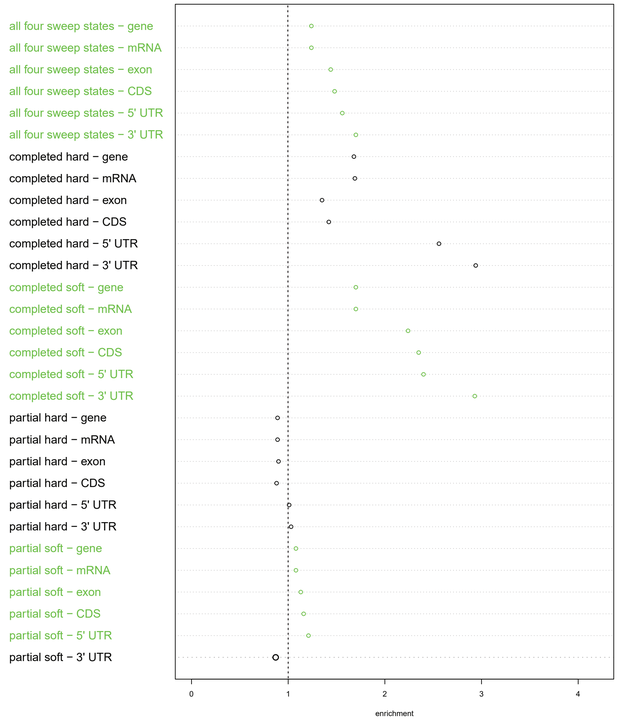

UGS



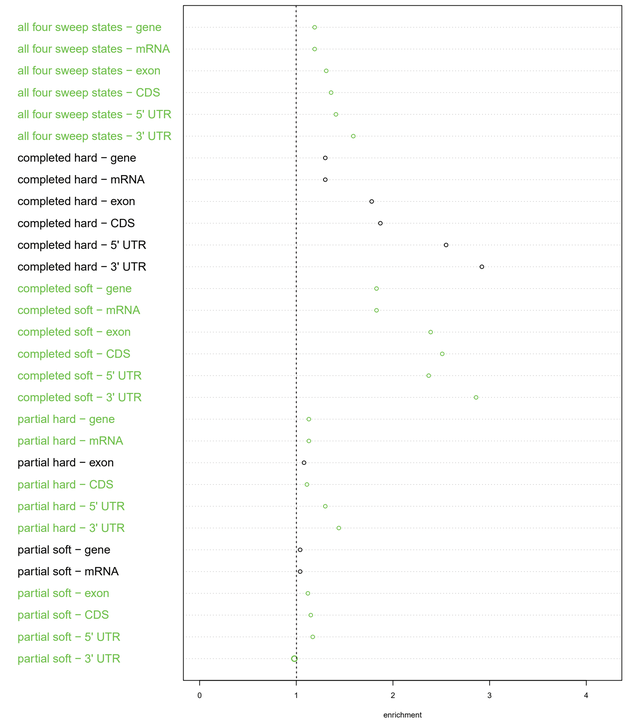
AOM



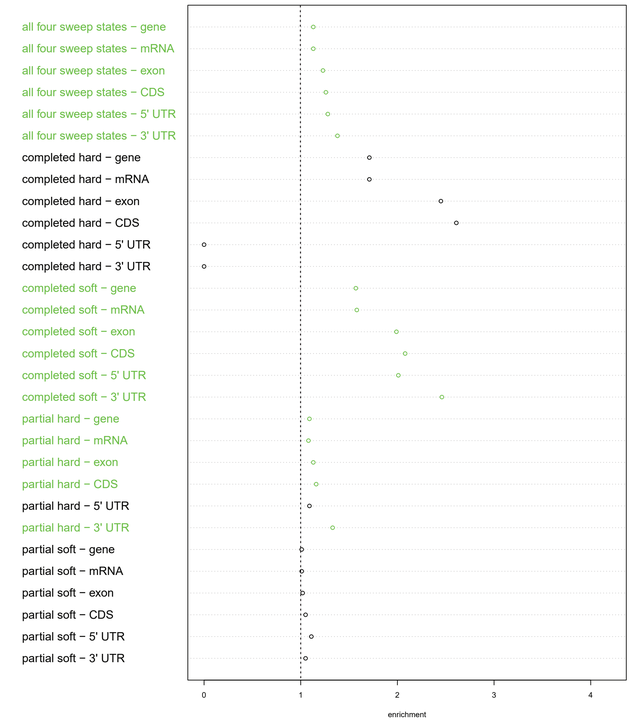

BFM



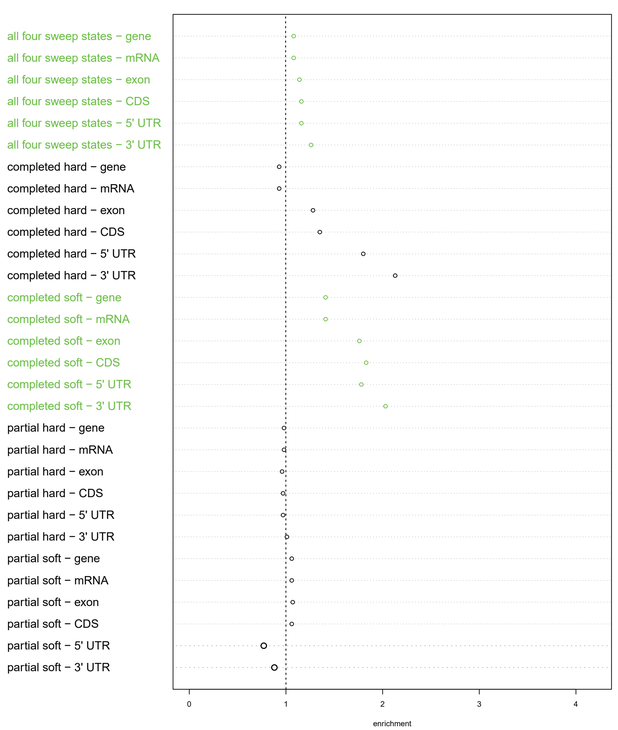

BFS



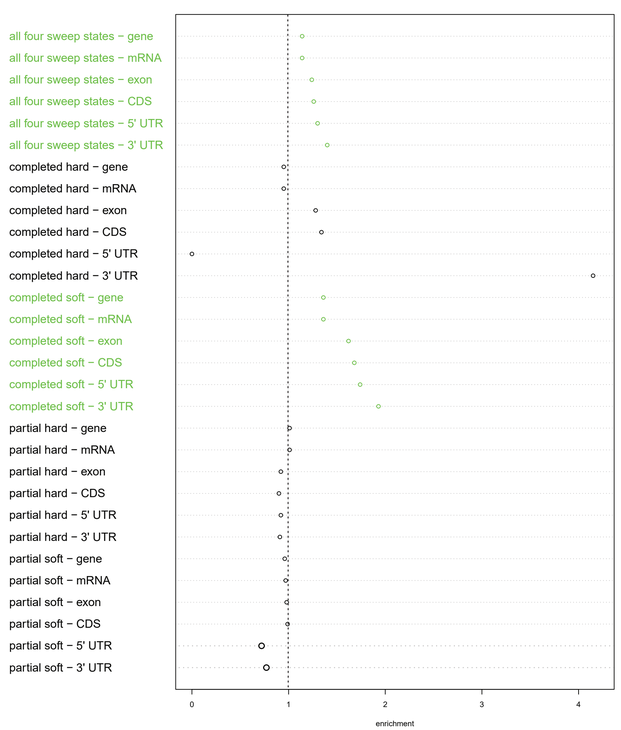

CMS



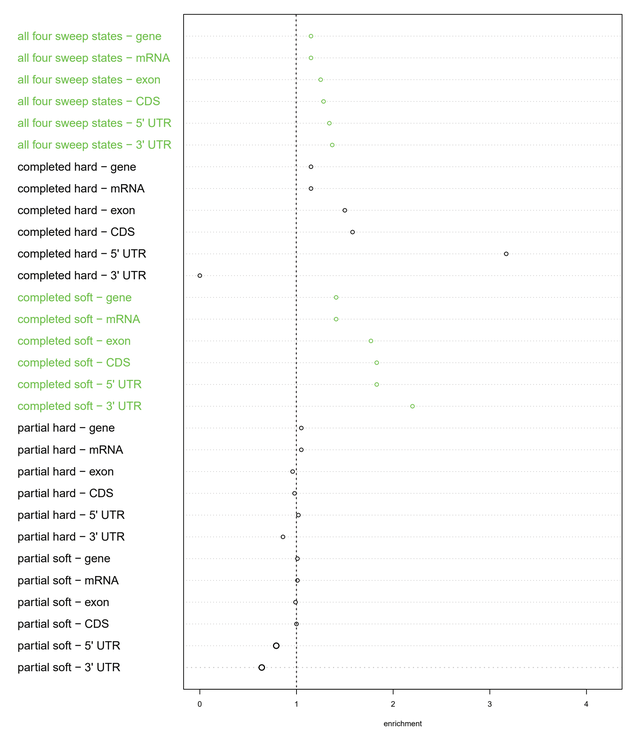

GAS



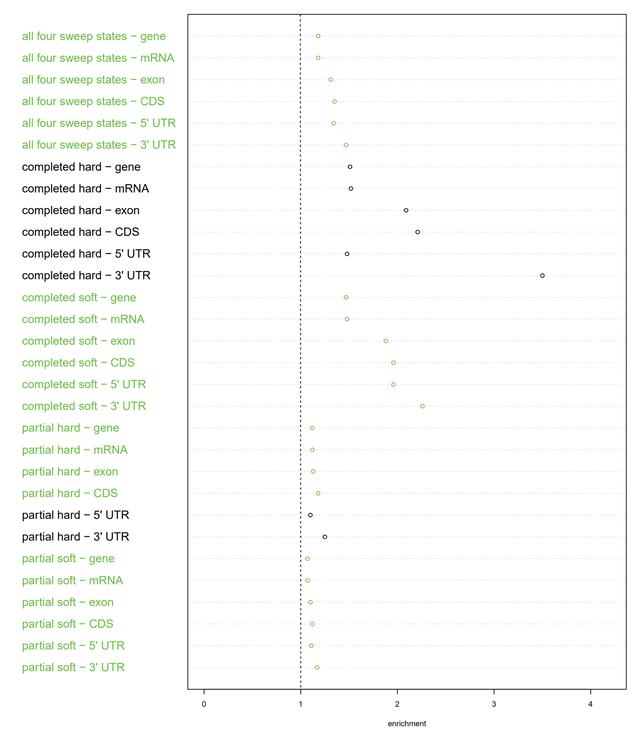

GNS




GWA




UGS



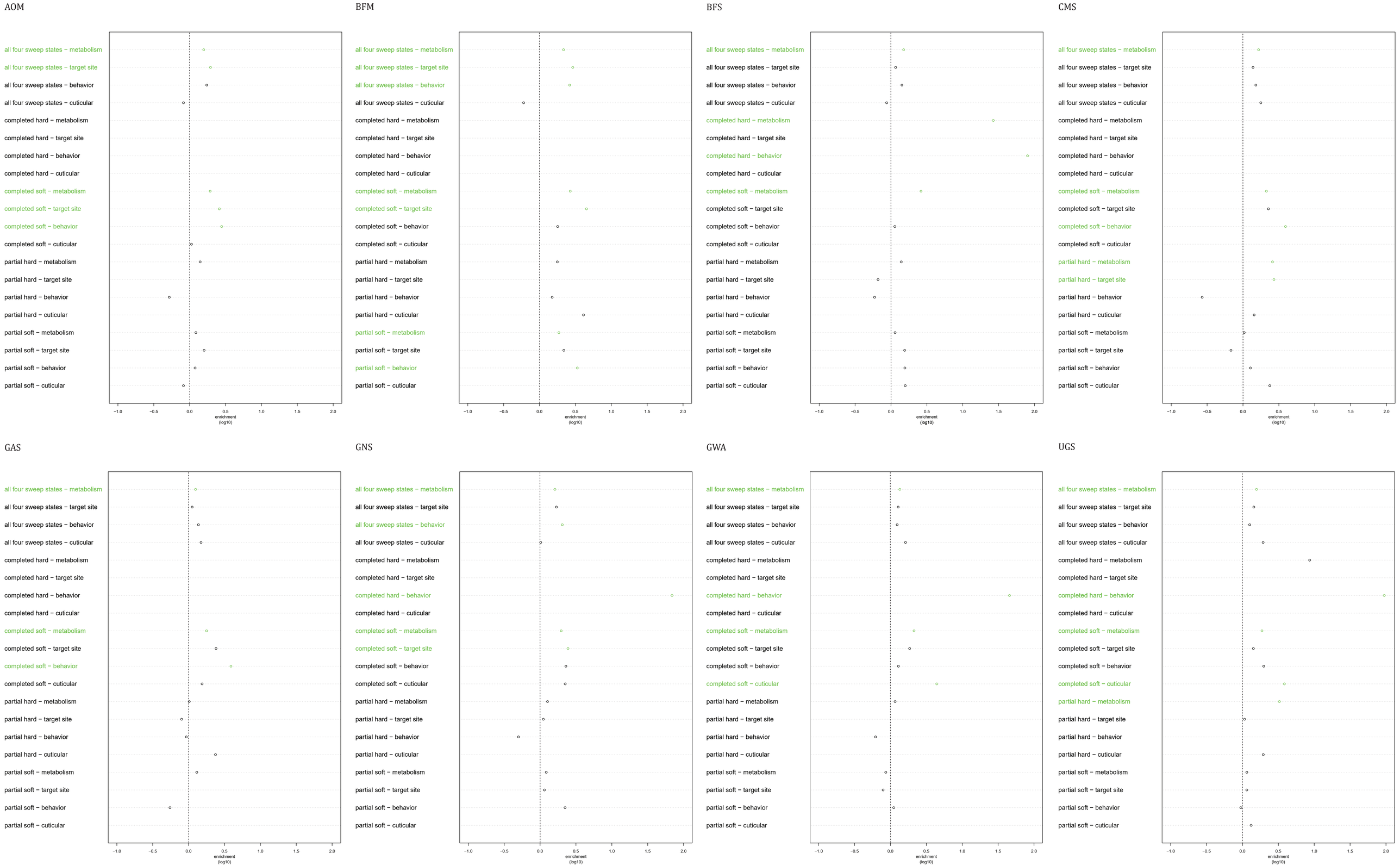

AOM
BFM
BFS
CMS
GAS
GNS
GWA
UGS
all four sweep states − metabolism
all four sweep states − target site
all four sweep states − behavior
all four sweep states − cuticular
completed hard − metabolism
completed hard − target site
completed hard − behavior
completed hard − cuticular
completed soft − metabolism
completed soft − target site
completed soft − behavior
completed soft − cuticular
partial hard − metabolism
partial hard − target site
partial hard − behavior
partial hard − cuticular
partial soft − metabolism
partial soft − target site
partial soft − behavior
partial soft − cuticular
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
enrichment
(log10)

**Table S1.** Accuracy rate for deep learning *in silico* test per simulated population dataset.

| **Population** | **Accuracy** |
|---|---|
| AOM | 0.610 |
| BFM | 0.661 |
| BFS | 0.691 |
| CMS | 0.711 |
| GAS | 0.667 |
| GNS | 0.546 |
| GWA | 0.625 |
| UGS | 0.678 |
| *median* | 0.664 |
| *mean* | 0.648 |

**Table S2.** Accuracy rate for five-state deep learning *in silico* test per simulated population dataset (neutral regions and completed sweeps only).

| Population | Accuracy |
|---|---|
| AOM | 0.822 |
| BFM | 0.872 |
| BFS | 0.881 |
| CMS | 0.897 |
| GAS | 0.830 |
| GNS | 0.792 |
| GWA | 0.780 |
| UGS | 0.906 |
| *median* | 0.851 |
| *mean* | 0.848 |

**Table S3.** Nine-state calls for individual chromosome arms and genome-wide for each empirical mosquito population dataset.

| | AOM | BFM | BFS | CMS | GAS | GNS | GWA | UGS | *median* | *mean* |
|---|---|---|---|---|---|---|---|---|---|---|
| | ***neutral*** *(proportion)* | | | | | | | | | |
| **2L** | 0.4524 | 0.6649 | 0.2929 | 0.0167 | 0.3153 | 0.2960 | 0.0395 | 0.1990 | 0.2944 | 0.2846 |
| **2R** | 0.4217 | 0.3822 | 0.2899 | 0.0046 | 0.2227 | 0.3465 | 0.0335 | 0.2231 | 0.2565 | 0.2405 |
| **3L** | 0.5452 | 0.6728 | 0.4012 | 0.0067 | 0.2709 | 0.5540 | 0.3315 | 0.3663 | 0.3838 | 0.3936 |
| **3R** | 0.5181 | 0.6202 | 0.3123 | 0.0069 | 0.2676 | 0.4124 | 0.2433 | 0.3327 | 0.3225 | 0.3392 |
| **genome-wide** | 0.4767 | 0.5608 | 0.3172 | 0.0081 | 0.2626 | 0.3912 | 0.1483 | 0.2761 | 0.2967 | 0.3051 |
| | ***completed hard*** *(proportion)* | | | | | | | | | |
| **2L** | 0.0000 | 0.0000 | 0.0000 | 0.0006 | 0.0004 | 0.0000 | 0.0000 | 0.0004 | 0.0000 | 0.0002 |
| **2R** | 0.0010 | 0.0004 | 0.0002 | 0.0000 | 0.0002 | 0.0002 | 0.0009 | 0.0002 | 0.0002 | 0.0004 |
| **3L** | 0.0000 | 0.0004 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0008 | 0.0000 | 0.0000 | 0.0002 |
| **3R** | 0.0000 | 0.0008 | 0.0005 | 0.0000 | 0.0005 | 0.0003 | 0.0003 | 0.0007 | 0.0004 | 0.0004 |
| **genome-wide** | 0.0003 | 0.0004 | 0.0002 | 0.0001 | 0.0003 | 0.0002 | 0.0005 | 0.0004 | 0.0003 | 0.0003 |
| | ***completed hard – linked*** *(proportion)* | | | | | | | | | |
| **2L** | 0.0004 | 0.0004 | 0.0017 | 0.0035 | 0.0000 | 0.0019 | 0.0000 | 0.0007 | 0.0006 | 0.0011 |
| **2R** | 0.0045 | 0.0018 | 0.0013 | 0.0034 | 0.0016 | 0.0009 | 0.0018 | 0.0024 | 0.0018 | 0.0022 |
| **3L** | 0.0010 | 0.0020 | 0.0004 | 0.0027 | 0.0005 | 0.0013 | 0.0004 | 0.0008 | 0.0009 | 0.0011 |
| **3R** | 0.0028 | 0.0025 | 0.0020 | 0.0007 | 0.0003 | 0.0000 | 0.0000 | 0.0012 | 0.0009 | 0.0012 |
| **genome-wide** | 0.0026 | 0.0017 | 0.0014 | 0.0025 | 0.0007 | 0.0009 | 0.0007 | 0.0014 | 0.0014 | 0.0015 |
| | ***completed soft*** *(proportion)* | | | | | | | | | |
| **2L** | 0.0587 | 0.0309 | 0.0187 | 0.0273 | 0.0393 | 0.0596 | 0.0665 | 0.0540 | 0.0466 | 0.0444 |
| **2R** | 0.0763 | 0.0307 | 0.0254 | 0.0259 | 0.0664 | 0.0580 | 0.0817 | 0.0647 | 0.0614 | 0.0536 |
| **3L** | 0.0558 | 0.0242 | 0.0234 | 0.0403 | 0.0332 | 0.0511 | 0.0676 | 0.0547 | 0.0457 | 0.0438 |
| **3R** | 0.0446 | 0.0257 | 0.0168 | 0.0329 | 0.0304 | 0.0520 | 0.0716 | 0.0522 | 0.0388 | 0.0408 |
| **genome-wide** | 0.0601 | 0.0281 | 0.0212 | 0.0310 | 0.0449 | 0.0554 | 0.0732 | 0.0571 | 0.0501 | 0.0464 |
| | ***completed soft – linked*** *(proportion)* | | | | | | | | | |
| **2L** | 0.1106 | 0.0872 | 0.1333 | 0.2428 | 0.1537 | 0.3515 | 0.3744 | 0.2050 | 0.1793 | 0.2073 |
| **2R** | 0.1267 | 0.1635 | 0.1310 | 0.2699 | 0.2286 | 0.2983 | 0.4573 | 0.1666 | 0.1976 | 0.2302 |
| **3L** | 0.0837 | 0.0890 | 0.1257 | 0.3544 | 0.1034 | 0.2003 | 0.2353 | 0.1480 | 0.1368 | 0.1675 |
| **3R** | 0.0823 | 0.0920 | 0.1310 | 0.3211 | 0.0783 | 0.2407 | 0.2631 | 0.1572 | 0.1441 | 0.1707 |
| **genome-wide** | 0.1033 | 0.1139 | 0.1305 | 0.2955 | 0.1487 | 0.2755 | 0.3448 | 0.1682 | 0.1584 | 0.1975 |
| | ***partial hard*** *(proportion)* | | | | | | | | | |
| **2L** | 0.0281 | 0.0103 | 0.0463 | 0.0331 | 0.1380 | 0.0331 | 0.0395 | 0.0427 | 0.0363 | 0.0464 |
| **2R** | 0.0363 | 0.0241 | 0.0554 | 0.0836 | 0.0917 | 0.0359 | 0.0276 | 0.0293 | 0.0361 | 0.0480 |
| **3L** | 0.0168 | 0.0071 | 0.0138 | 0.0113 | 0.1624 | 0.0074 | 0.0118 | 0.0119 | 0.0118 | 0.0303 |
| **3R** | 0.0237 | 0.0135 | 0.0388 | 0.0176 | 0.1676 | 0.0315 | 0.0193 | 0.0113 | 0.0215 | 0.0404 |
| **genome-wide** | 0.0277 | 0.0151 | 0.0412 | 0.0405 | 0.1350 | 0.0291 | 0.0248 | 0.0236 | 0.0284 | 0.0422 |
| | ***partial hard – linked*** *(proportion)* | | | | | | | | | |
| **2L** | 0.0768 | 0.0223 | 0.1382 | 0.0864 | 0.1325 | 0.0581 | 0.0512 | 0.0823 | 0.0796 | 0.0810 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **2R** | 0.1011 | 0.0767 | 0.1338 | 0.1535 | 0.1334 | 0.0882 | 0.0372 | 0.0533 | 0.0947 | 0.0971 |
| **3L** | 0.0635 | 0.0115 | 0.0461 | 0.0329 | 0.1867 | 0.0234 | 0.0248 | 0.0184 | 0.0289 | 0.0509 |
| **3R** | 0.0811 | 0.0353 | 0.1112 | 0.0652 | 0.1821 | 0.0746 | 0.0341 | 0.0204 | 0.0699 | 0.0755 |
| **genome-wide** | 0.0841 | 0.0419 | 0.1122 | 0.0914 | 0.1564 | 0.0670 | 0.0370 | 0.0433 | 0.0755 | 0.0791 |
| | ***partial soft*** *(proportion)* | | | | | | | | | |
| **2L** | 0.0885 | 0.0475 | 0.0598 | 0.1040 | 0.0531 | 0.0510 | 0.1170 | 0.1146 | 0.0741 | 0.0794 |
| **2R** | 0.0708 | 0.0504 | 0.0737 | 0.0535 | 0.0499 | 0.0426 | 0.1024 | 0.1292 | 0.0621 | 0.0716 |
| **3L** | 0.0659 | 0.0461 | 0.0734 | 0.0839 | 0.0538 | 0.0438 | 0.1105 | 0.1296 | 0.0696 | 0.0759 |
| **3R** | 0.0761 | 0.0473 | 0.0702 | 0.1547 | 0.0545 | 0.0426 | 0.1249 | 0.1447 | 0.0732 | 0.0894 |
| **genome-wide** | 0.0751 | 0.0481 | 0.0698 | 0.0986 | 0.0526 | 0.0445 | 0.1133 | 0.1308 | 0.0724 | 0.0791 |
| | ***partial soft – linked*** *(proportion)* | | | | | | | | | |
| **2L** | 0.1846 | 0.1365 | 0.3092 | 0.4855 | 0.1678 | 0.1489 | 0.3119 | 0.3013 | 0.2429 | 0.2557 |
| **2R** | 0.1616 | 0.2704 | 0.2893 | 0.4055 | 0.2054 | 0.1293 | 0.2575 | 0.3312 | 0.2640 | 0.2563 |
| **3L** | 0.1683 | 0.1469 | 0.3159 | 0.4679 | 0.1891 | 0.1188 | 0.2172 | 0.2704 | 0.2031 | 0.2368 |
| **3R** | 0.1712 | 0.1629 | 0.3173 | 0.4010 | 0.2186 | 0.1459 | 0.2436 | 0.2795 | 0.2311 | 0.2425 |
| **genome-wide** | 0.1702 | 0.1898 | 0.3062 | 0.4322 | 0.1989 | 0.1362 | 0.2575 | 0.2991 | 0.2282 | 0.2487 |
| | ***neutral*** *(count)* | | | | | | | | | |
| **2L** | 1,125 | 1,875 | 848 | 52 | 802 | 795 | 108 | 568 | 798.5 | 771.6 |
| **2R** | 1,764 | 1,730 | 1,361 | 23 | 950 | 1,552 | 147 | 1,034 | 1,197.5 | 1,070.1 |
| **3L** | 1,134 | 1,694 | 1,044 | 20 | 579 | 1,278 | 789 | 958 | 1,001.0 | 937.0 |
| **3R** | 1,858 | 2,441 | 1,264 | 31 | 977 | 1,569 | 935 | 1,352 | 1,308.0 | 1,303.4 |
| **genome-wide** | 5,881 | 7,740 | 4,517 | 126 | 3,308 | 5,194 | 1,979 | 3,912 | 4,214.5 | 4,082.1 |
| | ***completed hard*** *(count)* | | | | | | | | | |
| **2L** | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 1 | 0.0 | 0.5 |
| **2R** | 4 | 2 | 1 | 0 | 1 | 1 | 4 | 1 | 1.0 | 1.8 |
| **3L** | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 0.0 | 0.4 |
| **3R** | 0 | 3 | 2 | 0 | 2 | 1 | 1 | 3 | 1.5 | 1.5 |
| **genome-wide** | 4 | 6 | 3 | 2 | 4 | 2 | 7 | 5 | 4.0 | 4.1 |
| | ***completed hard – linked*** *(count)* | | | | | | | | | |
| **2L** | 1 | 1 | 5 | 11 | 0 | 5 | 0 | 2 | 1.5 | 3.1 |
| **2R** | 19 | 8 | 6 | 17 | 7 | 4 | 8 | 11 | 8.0 | 10.0 |
| **3L** | 2 | 5 | 1 | 8 | 1 | 3 | 1 | 2 | 2.0 | 2.9 |
| **3R** | 10 | 10 | 8 | 3 | 1 | 0 | 0 | 5 | 4.0 | 4.6 |
| **genome-wide** | 32 | 24 | 20 | 39 | 9 | 12 | 9 | 20 | 20.0 | 20.6 |
| | ***completed soft*** *(count)* | | | | | | | | | |
| **2L** | 146 | 87 | 54 | 85 | 100 | 160 | 182 | 154 | 123.0 | 121.0 |
| **2R** | 319 | 139 | 119 | 129 | 283 | 260 | 358 | 300 | 271.5 | 238.4 |
| **3L** | 116 | 61 | 61 | 121 | 71 | 118 | 161 | 143 | 117.0 | 106.5 |
| **3R** | 160 | 101 | 68 | 148 | 111 | 198 | 275 | 212 | 154.0 | 159.1 |
| **genome-wide** | 741 | 388 | 302 | 483 | 565 | 736 | 976 | 809 | 650.5 | 625.0 |
| | ***completed soft – linked*** *(count)* | | | | | | | | | |
| **2L** | 275 | 246 | 386 | 756 | 391 | 944 | 1,024 | 585 | 488.0 | 575.9 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **2R** | 530 | 740 | 615 | 1,343 | 975 | 1,336 | 2,005 | 772 | 873.5 | 1,039.5 |
| **3L** | 174 | 224 | 327 | 1,065 | 221 | 462 | 560 | 387 | 357.0 | 427.5 |
| **3R** | 295 | 362 | 530 | 1,443 | 286 | 916 | 1,011 | 639 | 584.5 | 685.3 |
| **genome-wide** | 1,274 | 1,572 | 1,858 | 4,607 | 1,873 | 3,658 | 4,600 | 2,383 | 2,128.0 | 2,728.1 |
| | ***partial hard*** *(count)* | | | | | | | | | |
| **2L** | 70 | 29 | 134 | 103 | 351 | 89 | 108 | 122 | 105.5 | 125.8 |
| **2R** | 152 | 109 | 260 | 416 | 391 | 161 | 121 | 136 | 156.5 | 218.3 |
| **3L** | 35 | 18 | 36 | 34 | 347 | 17 | 28 | 31 | 32.5 | 68.3 |
| **3R** | 85 | 53 | 157 | 79 | 612 | 120 | 74 | 46 | 82.0 | 153.3 |
| **genome-wide** | 342 | 209 | 587 | 632 | 1,701 | 387 | 331 | 335 | 364.5 | 565.5 |
| | ***partial hard – linked*** *(count)* | | | | | | | | | |
| **2L** | 191 | 63 | 400 | 269 | 337 | 156 | 140 | 235 | 213.0 | 223.9 |
| **2R** | 423 | 347 | 628 | 764 | 569 | 395 | 163 | 247 | 409.0 | 442.0 |
| **3L** | 132 | 29 | 120 | 99 | 399 | 54 | 59 | 48 | 79.0 | 117.5 |
| **3R** | 291 | 139 | 450 | 293 | 665 | 284 | 131 | 83 | 287.5 | 292.0 |
| **genome-wide** | 1,037 | 578 | 1,598 | 1,425 | 1,970 | 889 | 493 | 613 | 963.0 | 1,075.4 |
| | ***partial soft*** *(count)* | | | | | | | | | |
| **2L** | 220 | 134 | 173 | 324 | 135 | 137 | 320 | 327 | 196.5 | 221.3 |
| **2R** | 296 | 228 | 346 | 266 | 213 | 191 | 449 | 599 | 281.0 | 323.5 |
| **3L** | 137 | 116 | 191 | 252 | 115 | 101 | 263 | 339 | 164.0 | 189.3 |
| **3R** | 273 | 186 | 284 | 695 | 199 | 162 | 480 | 588 | 278.5 | 358.4 |
| **genome-wide** | 926 | 664 | 994 | 1,537 | 662 | 591 | 1,512 | 1,853 | 960.0 | 1,092.4 |
| | ***partial soft – linked*** *(count)* | | | | | | | | | |
| **2L** | 459 | 385 | 895 | 1,512 | 427 | 400 | 853 | 860 | 656.0 | 723.9 |
| **2R** | 676 | 1,224 | 1,358 | 2,018 | 876 | 579 | 1,129 | 1,535 | 1,176.5 | 1,174.4 |
| **3L** | 350 | 370 | 822 | 1,406 | 404 | 274 | 517 | 707 | 460.5 | 606.3 |
| **3R** | 614 | 641 | 1,284 | 1,802 | 798 | 555 | 936 | 1,136 | 867.0 | 970.8 |
| **genome-wide** | 2,099 | 2,620 | 4,359 | 6,738 | 2,505 | 1,808 | 3,435 | 4,238 | 3,027.5 | 3,475.3 |
| | ***neutral*** *(proportion)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.4922 | 0.7173 | 0.3247 | 0.0387 | 0.3700 | 0.3486 | 0.0502 | 0.2369 | 0.3367 | 0.3223 |
| **2R** | 0.4589 | 0.4122 | 0.3214 | 0.0107 | 0.2614 | 0.4081 | 0.0427 | 0.2656 | 0.2935 | 0.2726 |
| **3L** | 0.5932 | 0.7257 | 0.4448 | 0.0154 | 0.3180 | 0.6525 | 0.4218 | 0.4361 | 0.4405 | 0.4510 |
| **3R** | 0.5638 | 0.6690 | 0.3463 | 0.0160 | 0.3141 | 0.4857 | 0.3095 | 0.3960 | 0.3712 | 0.3876 |
| **genome-wide** | 0.5188 | 0.6050 | 0.3517 | 0.0188 | 0.3082 | 0.4608 | 0.1887 | 0.3287 | 0.3402 | 0.3476 |
| | ***completed hard*** *(proportion)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.0000 | 0.0000 | 0.0000 | 0.0006 | 0.0004 | 0.0000 | 0.0000 | 0.0004 | 0.0000 | 0.0002 |
| **2R** | 0.0010 | 0.0004 | 0.0002 | 0.0000 | 0.0002 | 0.0002 | 0.0009 | 0.0002 | 0.0002 | 0.0004 |
| **3L** | 0.0000 | 0.0004 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0008 | 0.0000 | 0.0000 | 0.0002 |
| **3R** | 0.0000 | 0.0008 | 0.0005 | 0.0000 | 0.0005 | 0.0003 | 0.0003 | 0.0007 | 0.0004 | 0.0004 |
| **genome-wide** | 0.0003 | 0.0004 | 0.0002 | 0.0001 | 0.0003 | 0.0002 | 0.0005 | 0.0004 | 0.0003 | 0.0003 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | ***completed hard – linked*** *(proportion)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.0004 | -0.0011 | 0.0017 | 0.0035 | 0.0000 | 0.0019 | 0.0000 | 0.0007 | 0.0006 | 0.0009 |
| **2R** | 0.0045 | 0.0009 | 0.0013 | 0.0034 | 0.0016 | 0.0009 | 0.0018 | 0.0024 | 0.0017 | 0.0021 |
| **3L** | 0.0010 | 0.0005 | 0.0004 | 0.0027 | 0.0005 | 0.0013 | 0.0004 | 0.0008 | 0.0006 | 0.0009 |
| **3R** | 0.0028 | 0.0012 | 0.0020 | 0.0007 | 0.0003 | 0.0000 | 0.0000 | 0.0012 | 0.0009 | 0.0010 |
| **genome-wide** | 0.0026 | 0.0005 | 0.0014 | 0.0025 | 0.0007 | 0.0009 | 0.0007 | 0.0014 | 0.0012 | 0.0013 |
| | ***completed soft*** *(proportion)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.0498 | 0.0258 | 0.0177 | 0.0267 | 0.0345 | 0.0495 | 0.0652 | 0.0485 | 0.0415 | 0.0397 |
| **2R** | 0.0680 | 0.0278 | 0.0244 | 0.0258 | 0.0630 | 0.0462 | 0.0805 | 0.0586 | 0.0524 | 0.0493 |
| **3L** | 0.0451 | 0.0191 | 0.0221 | 0.0400 | 0.0291 | 0.0322 | 0.0563 | 0.0447 | 0.0361 | 0.0361 |
| **3R** | 0.0345 | 0.0210 | 0.0158 | 0.0327 | 0.0263 | 0.0380 | 0.0632 | 0.0431 | 0.0336 | 0.0343 |
| **genome-wide** | 0.0507 | 0.0239 | 0.0202 | 0.0307 | 0.0408 | 0.0421 | 0.0681 | 0.0495 | 0.0415 | 0.0407 |
| | ***completed soft – linked*** *(proportion)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.1012 | 0.0822 | 0.1304 | 0.2419 | 0.1463 | 0.3396 | 0.3727 | 0.2036 | 0.1749 | 0.2022 |
| **2R** | 0.1180 | 0.1606 | 0.1281 | 0.2697 | 0.2234 | 0.2844 | 0.4559 | 0.1650 | 0.1942 | 0.2256 |
| **3L** | 0.0724 | 0.0839 | 0.1217 | 0.3541 | 0.0971 | 0.1781 | 0.2214 | 0.1454 | 0.1335 | 0.1592 |
| **3R** | 0.0716 | 0.0873 | 0.1278 | 0.3207 | 0.0721 | 0.2242 | 0.2529 | 0.1549 | 0.1414 | 0.1639 |
| **genome-wide** | 0.0934 | 0.1097 | 0.1273 | 0.2951 | 0.1425 | 0.2598 | 0.3385 | 0.1662 | 0.1544 | 0.1916 |
| | ***partial hard*** *(proportion)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.0277 | 0.0103 | 0.0460 | 0.0330 | 0.1369 | 0.0331 | 0.0395 | 0.0425 | 0.0363 | 0.0461 |
| **2R** | 0.0359 | 0.0241 | 0.0551 | 0.0836 | 0.0909 | 0.0359 | 0.0276 | 0.0291 | 0.0359 | 0.0478 |
| **3L** | 0.0162 | 0.0071 | 0.0134 | 0.0113 | 0.1614 | 0.0074 | 0.0118 | 0.0114 | 0.0116 | 0.0300 |
| **3R** | 0.0231 | 0.0135 | 0.0384 | 0.0175 | 0.1667 | 0.0315 | 0.0193 | 0.0109 | 0.0212 | 0.0401 |
| **genome-wide** | 0.0272 | 0.0151 | 0.0409 | 0.0405 | 0.1341 | 0.0291 | 0.0248 | 0.0233 | 0.0282 | 0.0419 |
| | ***partial hard – linked*** *(proportion)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.0748 | 0.0223 | 0.1378 | 0.0863 | 0.1325 | 0.0567 | 0.0512 | 0.0821 | 0.0785 | 0.0805 |
| **2R** | 0.0993 | 0.0767 | 0.1335 | 0.1535 | 0.1334 | 0.0866 | 0.0372 | 0.0530 | 0.0929 | 0.0966 |
| **3L** | 0.0611 | 0.0115 | 0.0457 | 0.0329 | 0.1867 | 0.0208 | 0.0248 | 0.0179 | 0.0289 | 0.0502 |
| **3R** | 0.0789 | 0.0353 | 0.1108 | 0.0652 | 0.1821 | 0.0727 | 0.0341 | 0.0200 | 0.0689 | 0.0749 |
| **genome-wide** | 0.0820 | 0.0419 | 0.1119 | 0.0914 | 0.1564 | 0.0651 | 0.0370 | 0.0429 | 0.0736 | 0.0786 |
| | ***partial soft*** *(proportion)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.0816 | 0.0274 | 0.0533 | 0.1026 | 0.0468 | 0.0385 | 0.1127 | 0.0987 | 0.0674 | 0.0702 |
| **2R** | 0.0643 | 0.0388 | 0.0673 | 0.0530 | 0.0455 | 0.0280 | 0.0987 | 0.1114 | 0.0587 | 0.0634 |
| **3L** | 0.0576 | 0.0257 | 0.0645 | 0.0833 | 0.0484 | 0.0203 | 0.0742 | 0.1004 | 0.0610 | 0.0593 |
| **3R** | 0.0682 | 0.0285 | 0.0632 | 0.1540 | 0.0492 | 0.0251 | 0.0983 | 0.1182 | 0.0657 | 0.0756 |
| **genome-wide** | 0.0678 | 0.0312 | 0.0628 | 0.0979 | 0.0473 | 0.0279 | 0.0971 | 0.1088 | 0.0653 | 0.0676 |
| | ***partial soft – linked*** *(proportion)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.1723 | 0.1157 | 0.2884 | 0.4666 | 0.1327 | 0.1322 | 0.3085 | 0.2866 | 0.2294 | 0.2379 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **2R** | 0.1501 | 0.2584 | 0.2687 | 0.4003 | 0.1806 | 0.1097 | 0.2546 | 0.3147 | 0.2565 | 0.2421 |
| **3L** | 0.1534 | 0.1259 | 0.2874 | 0.4603 | 0.1588 | 0.0874 | 0.1885 | 0.2433 | 0.1737 | 0.2132 |
| **3R** | 0.1571 | 0.1435 | 0.2951 | 0.3931 | 0.1887 | 0.1225 | 0.2225 | 0.2550 | 0.2056 | 0.2222 |
| **genome-wide** | 0.1572 | 0.1723 | 0.2836 | 0.4230 | 0.1696 | 0.1141 | 0.2446 | 0.2787 | 0.2085 | 0.2304 |
| | ***neutral*** *(count)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 1,224.2 | 2,022.7 | 940.1 | 120.6 | 941.3 | 936.4 | 137.4 | 676.2 | 938.3 | 874.9 |
| **2R** | 1,919.5 | 1,866.2 | 1,508.9 | 53.4 | 1,115.0 | 1,828.0 | 187.0 | 1,231.0 | 1,369.9 | 1,213.6 |
| **3L** | 1,233.9 | 1,827.4 | 1,157.4 | 46.4 | 679.6 | 1,505.3 | 1,003.8 | 1,140.5 | 1,149.0 | 1,074.3 |
| **3R** | 2,021.8 | 2,633.2 | 1,401.3 | 71.9 | 1,146.7 | 1,848.1 | 1,189.6 | 1,609.5 | 1,505.4 | 1,490.3 |
| **genome-wide** | 6,399.3 | 8,349.5 | 5,007.8 | 292.3 | 3,882.6 | 6,117.8 | 2,517.8 | 4,657.1 | 4,832.5 | 4,653.0 |
| | ***completed hard*** *(count)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 0.0 | 0.0 | 0.0 | 1.9 | 1.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.5 |
| **2R** | 4.0 | 2.0 | 1.0 | -0.1 | 1.0 | 1.0 | 4.0 | 1.0 | 1.0 | 1.7 |
| **3L** | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.0 | 0.0 | 0.0 | 0.4 |
| **3R** | 0.0 | 3.0 | 2.0 | -0.1 | 2.0 | 1.0 | 1.0 | 3.0 | 1.5 | 1.5 |
| **genome-wide** | 4.0 | 6.0 | 3.0 | 1.7 | 4.0 | 2.0 | 7.0 | 5.0 | 4.0 | 4.1 |
| | ***completed hard – linked*** *(count)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 1.0 | -3.0 | 5.0 | 11.0 | 0.0 | 5.0 | 0.0 | 2.0 | 1.5 | 2.6 |
| **2R** | 19.0 | 4.3 | 6.0 | 17.0 | 7.0 | 4.0 | 8.0 | 11.0 | 7.5 | 9.5 |
| **3L** | 2.0 | 1.3 | 1.0 | 8.0 | 1.0 | 3.0 | 1.0 | 2.0 | 1.7 | 2.4 |
| **3R** | 10.0 | 4.7 | 8.0 | 3.0 | 1.0 | 0.0 | 0.0 | 5.0 | 3.9 | 4.0 |
| **genome-wide** | 32.0 | 7.3 | 20.0 | 39.0 | 9.0 | 12.0 | 9.0 | 20.0 | 16.0 | 18.5 |
| | ***completed soft*** *(count)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 124.0 | 72.8 | 51.2 | 83.2 | 87.8 | 132.8 | 178.3 | 138.4 | 105.9 | 108.6 |
| **2R** | 284.4 | 125.9 | 114.5 | 128.2 | 268.5 | 207.0 | 353.0 | 271.7 | 237.7 | 219.1 |
| **3L** | 93.8 | 48.2 | 57.5 | 120.3 | 62.2 | 74.3 | 133.9 | 116.8 | 84.1 | 88.4 |
| **3R** | 123.6 | 82.6 | 63.8 | 146.9 | 96.1 | 144.4 | 242.9 | 175.0 | 134.0 | 134.4 |
| **genome-wide** | 625.8 | 329.6 | 287.0 | 478.6 | 514.5 | 558.6 | 908.0 | 701.9 | 536.6 | 550.5 |
| | ***completed soft – linked*** *(count)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 251.7 | 231.8 | 377.5 | 753.3 | 372.2 | 912.2 | 1,019.5 | 580.9 | 479.2 | 562.4 |
| **2R** | 493.5 | 726.9 | 601.4 | 1,341.8 | 952.7 | 1,273.8 | 1,998.8 | 764.6 | 858.7 | 1,019.2 |
| **3L** | 150.6 | 211.2 | 316.6 | 1,064.0 | 207.4 | 410.8 | 526.9 | 380.2 | 348.4 | 408.4 |
| **3R** | 256.6 | 343.6 | 517.4 | 1,441.4 | 263.1 | 853.2 | 971.7 | 629.3 | 573.4 | 659.5 |
| **genome-wide** | 1,152.4 | 1,513.6 | 1,812.9 | 4,600.6 | 1,795.3 | 3,450.0 | 4,516.9 | 2,355.1 | 2,084.0 | 2,649.6 |
| | ***partial hard*** *(count)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 68.8 | 29.0 | 133.1 | 102.8 | 348.2 | 89.0 | 108.0 | 121.3 | 105.4 | 125.0 |
| **2R** | 150.1 | 109.0 | 258.5 | 415.9 | 387.7 | 161.0 | 121.0 | 134.8 | 155.5 | 217.2 |
| **3L** | 33.8 | 18.0 | 34.8 | 33.9 | 345.0 | 17.0 | 28.0 | 29.9 | 31.8 | 67.5 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **3R** | 83.0 | 53.0 | 155.6 | 78.9 | 608.6 | 120.0 | 74.0 | 44.4 | 80.9 | 152.2 |
| **genome-wide** | 335.6 | 209.0 | 582.0 | 631.4 | 1,689.4 | 387.0 | 331.0 | 330.3 | 361.3 | 562.0 |
| | ***partial hard – linked*** *(count)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 186.1 | 63.0 | 399.1 | 268.9 | 337.0 | 152.3 | 140.0 | 234.3 | 210.2 | 222.6 |
| **2R** | 415.3 | 347.0 | 626.5 | 763.9 | 569.0 | 387.7 | 163.0 | 245.8 | 401.5 | 439.8 |
| **3L** | 127.1 | 29.0 | 118.8 | 99.0 | 399.0 | 48.0 | 59.0 | 46.9 | 79.0 | 115.8 |
| **3R** | 282.9 | 139.0 | 448.6 | 292.9 | 665.0 | 276.6 | 131.0 | 81.4 | 279.8 | 289.7 |
| **genome-wide** | 1,011.4 | 578.0 | 1,593.0 | 1,424.7 | 1,970.0 | 864.5 | 493.0 | 608.3 | 938.0 | 1,067.9 |
| | ***partial soft*** *(count)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 202.9 | 77.4 | 154.2 | 319.4 | 119.0 | 103.3 | 308.2 | 281.7 | 178.5 | 195.8 |
| **2R** | 269.1 | 175.7 | 315.8 | 264.0 | 194.0 | 125.2 | 432.9 | 516.5 | 266.5 | 286.7 |
| **3L** | 119.7 | 64.8 | 167.9 | 250.2 | 103.4 | 46.8 | 176.7 | 262.6 | 143.8 | 149.0 |
| **3R** | 244.7 | 112.3 | 256.0 | 692.3 | 179.5 | 95.5 | 377.7 | 480.2 | 250.3 | 304.8 |
| **genome-wide** | 836.4 | 430.2 | 893.8 | 1,525.9 | 596.0 | 370.8 | 1,295.5 | 1,541.0 | 865.1 | 936.2 |
| | ***partial soft – linked*** *(count)* *corrected for false discovery* | | | | | | | | | |
| **2L** | 428.4 | 326.3 | 834.8 | 1,452.9 | 337.6 | 355.1 | 843.7 | 818.1 | 623.2 | 674.6 |
| **2R** | 628.0 | 1,169.9 | 1,261.4 | 1,991.9 | 770.1 | 491.3 | 1,116.3 | 1,458.7 | 1,143.1 | 1,110.9 |
| **3L** | 319.2 | 317.0 | 747.9 | 1,383.3 | 339.4 | 201.7 | 448.7 | 636.3 | 394.1 | 549.2 |
| **3R** | 563.5 | 564.6 | 1,194.3 | 1,766.8 | 689.1 | 466.3 | 855.1 | 1,036.2 | 772.1 | 892.0 |
| **genome-wide** | 1,939.0 | 2,377.9 | 4,038.5 | 6,594.8 | 2,136.2 | 1,514.3 | 3,263.8 | 3,949.3 | 2,820.8 | 3,226.7 |

**Table S4.** Enrichments of selective sweeps within DNA regions for each empirical mosquito population dataset.

| | | AOM | BFM | BFS | CMS | GAS | GNS | GWA | UGS |
|---|---|---|---|---|---|---|---|---|---|
| | | ***all four sweep states*** | | | | | | | |
| **gene** | **count** | 1177 | 820 | 1173 | 1530 | 1636 | 1028 | 1698 | 1839 |
| | **enrichment** | 1.11 | 1.24 | 1.19 | 1.13 | 1.08 | 1.14 | 1.15 | 1.18 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **mRNA** | **count** | 1176 | 819 | 1171 | 1529 | 1634 | 1028 | 1697 | 1836 |
| | **enrichment** | 1.11 | 1.24 | 1.19 | 1.13 | 1.08 | 1.14 | 1.15 | 1.18 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **exon** | **count** | 919 | 700 | 951 | 1230 | 1280 | 824 | 1365 | 1506 |
| | **enrichment** | 1.18 | 1.44 | 1.31 | 1.23 | 1.14 | 1.24 | 1.25 | 1.31 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **CDS** | **count** | 887 | 685 | 930 | 1198 | 1231 | 798 | 1325 | 1468 |
| | **enrichment** | 1.19 | 1.48 | 1.36 | 1.26 | 1.16 | 1.26 | 1.28 | 1.35 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **5' UTR** | **count** | 338 | 265 | 358 | 445 | 446 | 302 | 507 | 537 |
| | **enrichment** | 1.24 | 1.56 | 1.41 | 1.28 | 1.16 | 1.30 | 1.34 | 1.34 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **3' UTR** | **count** | 312 | 247 | 343 | 414 | 416 | 278 | 440 | 502 |
| | **enrichment** | 1.34 | 1.70 | 1.59 | 1.38 | 1.26 | 1.40 | 1.37 | 1.47 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| | | ***completed hard*** | | | | | | | |
| **gene** | **count** | 1 | 5 | 2 | 2 | 2 | 1 | 4 | 4 |
| | **enrichment** | 0.45 | 1.68 | 1.30 | 1.71 | 0.93 | 0.95 | 1.15 | 1.51 |
| | ***p*-value** | 0.9278 | 0.1034 | 0.5170 | 0.5043 | 0.7459 | 0.7780 | 0.5054 | 0.2250 |
| **mRNA** | **count** | 1 | 5 | 2 | 2 | 2 | 1 | 4 | 4 |
| | **enrichment** | 0.45 | 1.69 | 1.30 | 1.71 | 0.93 | 0.95 | 1.15 | 1.52 |
| | ***p*-value** | 0.9276 | 0.1020 | 0.5143 | 0.5043 | 0.7436 | 0.7770 | 0.5038 | 0.2237 |
| **exon** | **count** | 1 | 3 | 2 | 2 | 2 | 1 | 4 | 4 |
| | **enrichment** | 0.59 | 1.35 | 1.78 | 2.45 | 1.28 | 1.28 | 1.50 | 2.09 |
| | ***p*-value** | 0.8484 | 0.3910 | 0.3126 | 0.2786 | 0.5070 | 0.6298 | 0.2734 | 0.0762 |
| **CDS** | **count** | 1 | 3 | 2 | 2 | 2 | 1 | 4 | 4 |
| | **enrichment** | 0.62 | 1.42 | 1.87 | 2.61 | 1.35 | 1.34 | 1.58 | 2.21 |
| | ***p*-value** | 0.8288 | 0.3574 | 0.2881 | 0.2549 | 0.4743 | 0.6087 | 0.2387 | 0.0647 |
| **5' UTR** | **count** | 0 | 2 | 1 | 0 | 1 | 0 | 3 | 1 |
| | **enrichment** | 0.00 | 2.56 | 2.55 | 0.00 | 1.80 | 0.00 | 3.17 | 1.48 |
| | ***p*-value** | 1.0000 | 0.1804 | 0.3424 | 1.0000 | 0.4523 | 1.0000 | 0.0652 | 0.5172 |
| **3' UTR** | **count** | 0 | 2 | 1 | 0 | 1 | 1 | 0 | 2 |
| | **enrichment** | 0.00 | 2.94 | 2.92 | 0.00 | 2.13 | 4.15 | 0.00 | 3.50 |
| | ***p*-value** | 1.0000 | 0.1430 | 0.3015 | 1.0000 | 0.3906 | 0.2259 | 1.0000 | 0.1017 |
| | | ***completed soft*** | | | | | | | |
| **gene** | **count** | 485 | 345 | 283 | 376 | 425 | 521 | 715 | 617 |
| | **enrichment** | 1.25 | 1.70 | 1.83 | 1.57 | 1.41 | 1.36 | 1.41 | 1.47 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **mRNA** | **count** | 485 | 345 | 283 | 376 | 425 | 521 | 715 | 617 |
| | **enrichment** | 1.25 | 1.70 | 1.83 | 1.58 | 1.41 | 1.36 | 1.41 | 1.48 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **exon** | **count** | 424 | 335 | 276 | 356 | 394 | 459 | 665 | 582 |
| | **enrichment** | 1.47 | 2.24 | 2.39 | 1.99 | 1.76 | 1.62 | 1.77 | 1.88 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **CDS** | **count** | 416 | 334 | 274 | 351 | 389 | 453 | 653 | 575 |
| | **enrichment** | 1.52 | 2.35 | 2.51 | 2.08 | 1.83 | 1.68 | 1.83 | 1.96 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **5' UTR** | **count** | 157 | 126 | 96 | 124 | 141 | 172 | 239 | 213 |
| | **enrichment** | 1.55 | 2.40 | 2.37 | 2.01 | 1.78 | 1.74 | 1.83 | 1.96 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| **3' UTR** | **count** | 153 | 131 | 98 | 131 | 137 | 163 | 245 | 208 |
| | **enrichment** | 1.77 | 2.93 | 2.86 | 2.46 | 2.03 | 1.93 | 2.20 | 2.26 |
| | ***p*-value** | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| | | ***partial hard*** | | | | | | | |
| **gene** | **count** | 180 | 99 | 358 | 375 | 845 | 210 | 189 | 206 |
| | **enrichment** | 0.99 | 0.89 | 1.13 | 1.09 | 0.98 | 1.01 | 1.05 | 1.12 |
| | ***p*-value** | 0.5989 | 0.9251 | 0.0018 | 0.0364 | 0.7931 | 0.4632 | 0.2261 | 0.0305 |
| **mRNA** | **count** | 180 | 99 | 357 | 374 | 844 | 210 | 189 | 206 |
| | **enrichment** | 0.99 | 0.89 | 1.13 | 1.08 | 0.98 | 1.01 | 1.05 | 1.12 |
| | ***p*-value** | 0.5896 | 0.9219 | 0.0020 | 0.0392 | 0.7849 | 0.4540 | 0.2215 | 0.0294 |
| **exon** | **count** | 116 | 74 | 250 | 290 | 612 | 141 | 127 | 153 |
| | **enrichment** | 0.86 | 0.90 | 1.08 | 1.13 | 0.96 | 0.92 | 0.96 | 1.13 |
| | ***p*-value** | 0.9707 | 0.8584 | 0.1037 | 0.0141 | 0.8605 | 0.8407 | 0.6916 | 0.0487 |
| **CDS** | **count** | 110 | 69 | 244 | 282 | 584 | 130 | 122 | 150 |
| | **enrichment** | 0.86 | 0.88 | 1.11 | 1.16 | 0.97 | 0.90 | 0.98 | 1.18 |
| | ***p*-value** | 0.9685 | 0.8856 | 0.0415 | 0.0060 | 0.8130 | 0.9043 | 0.6225 | 0.0169 |
| **5' UTR** | **count** | 39 | 30 | 108 | 103 | 210 | 50 | 48 | 53 |
| | **enrichment** | 0.82 | 1.01 | 1.30 | 1.09 | 0.97 | 0.92 | 1.02 | 1.10 |
| | ***p*-value** | 0.9118 | 0.4890 | 0.0046 | 0.2032 | 0.6747 | 0.7255 | 0.4641 | 0.2741 |
| **3' UTR** | **count** | 42 | 26 | 101 | 105 | 189 | 42 | 34 | 50 |
| | **enrichment** | 1.04 | 1.03 | 1.44 | 1.33 | 1.01 | 0.91 | 0.86 | 1.25 |
| | ***p*-value** | 0.4317 | 0.4620 | 0.0003 | 0.0053 | 0.4343 | 0.7325 | 0.8156 | 0.0725 |
| | | ***partial soft*** | | | | | | | |
| **gene** | **count** | 511 | 371 | 530 | 777 | 364 | 296 | 790 | 1012 |
| | **enrichment** | 1.05 | 1.08 | 1.04 | 1.01 | 1.06 | 0.96 | 1.01 | 1.07 |
| | ***p*-value** | 0.0684 | 0.0298 | 0.1025 | 0.3927 | 0.0761 | 0.8089 | 0.3293 | 0.0019 |
| **mRNA** | **count** | 510 | 370 | 529 | 777 | 363 | 296 | 789 | 1009 |
| | **enrichment** | 1.05 | 1.08 | 1.04 | 1.01 | 1.06 | 0.97 | 1.01 | 1.07 |
| | ***p*-value** | 0.0692 | 0.0323 | 0.1045 | 0.3704 | 0.0809 | 0.7996 | 0.3228 | 0.0023 |
| **exon** | **count** | 378 | 288 | 423 | 582 | 272 | 223 | 569 | 767 |
| | **enrichment** | 1.06 | 1.13 | 1.12 | 1.02 | 1.07 | 0.98 | 0.99 | 1.10 |
| | ***p*-value** | 0.0903 | 0.0066 | 0.0021 | 0.2958 | 0.0884 | 0.6337 | 0.6480 | 0.0014 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **CDS** | **count** | 360 | 279 | 410 | 563 | 256 | 214 | 546 | 739 |
| | **enrichment** | 1.06 | 1.16 | 1.15 | 1.05 | 1.06 | 0.99 | 1.00 | 1.12 |
| | ***p*-value** | 0.0885 | 0.0023 | 0.0008 | 0.1591 | 0.1288 | 0.5577 | 0.5153 | 0.0004 |
| **5' UTR** | **count** | 142 | 107 | 153 | 218 | 94 | 80 | 217 | 270 |
| | **enrichment** | 1.14 | 1.21 | 1.17 | 1.11 | 1.07 | 1.02 | 1.09 | 1.11 |
| | ***p*-value** | 0.0535 | 0.0244 | 0.0253 | 0.0802 | 0.2633 | 0.4512 | 0.0966 | 0.0348 |
| **3' UTR** | **count** | 117 | 88 | 143 | 178 | 89 | 72 | 161 | 242 |
| | **enrichment** | 1.10 | 1.17 | 1.28 | 1.05 | 1.18 | 1.07 | 0.94 | 1.17 |
| | ***p*-value** | 0.1448 | 0.0834 | 0.0015 | 0.2652 | 0.0557 | 0.2863 | 0.7917 | 0.0051 |

Significant values (*i.e.* $p < 0.05$) underlined.

**Table S5.** Enrichments of selective sweeps within groupings of insecticide resistance genes for each empirical mosquito population dataset.

| | | AOM | BFM | BFS | CMS | GAS | GNS | GWA | UGS |
|---|---|---|---|---|---|---|---|---|---|
| | | ***all four sweep states*** | | | | | | | |
| metabolism | **count** | 67 | 58 | 60 | 88 | 69 | 57 | 75 | 95 |
| | enrichment | 1.57 | 2.16 | 1.50 | 1.65 | 1.25 | 1.61 | 1.35 | 1.57 |
| | ***p*-value** | 0.0011 | <0.0001 | 0.0048 | <0.0001 | 0.0464 | 0.0018 | 0.0076 | 0.0001 |
| target site | **count** | 9 | 9 | 5 | 7 | 7 | 7 | 8 | 9 |
| | enrichment | 1.95 | 2.90 | 1.16 | 1.38 | 1.12 | 1.69 | 1.28 | 1.44 |
| | ***p*-value** | 0.0132 | 0.0007 | 0.4434 | 0.2087 | 0.4439 | 0.0823 | 0.2509 | 0.1096 |
| behavior | **count** | 18 | 19 | 15 | 20 | 20 | 18 | 18 | 20 |
| | enrichment | 1.73 | 2.63 | 1.42 | 1.51 | 1.37 | 2.04 | 1.24 | 1.26 |
| | ***p*-value** | 0.0507 | 0.0039 | 0.1619 | 0.1041 | 0.1518 | 0.0187 | 0.2480 | 0.2231 |
| cuticular | **count** | 2 | 1 | 2 | 6 | 5 | 2 | 5 | 7 |
| | enrichment | 0.82 | 0.60 | 0.87 | 1.77 | 1.49 | 1.02 | 1.62 | 1.94 |
| | ***p*-value** | 0.7046 | 0.8049 | 0.6813 | 0.1215 | 0.2408 | 0.5782 | 0.1921 | 0.0603 |
| | | ***completed hard*** | | | | | | | |
| metabolism | **count** | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 1 |
| | enrichment | 0.00 | 0.00 | 26.63 | 0.00 | 0.00 | 0.00 | 0.00 | 8.64 |
| | ***p*-value** | 1.0000 | 1.0000 | 0.0121 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.0932 |
| target site | **count** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | enrichment | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | ***p*-value** | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| behavior | **count** | 0 | 0 | 2 | 0 | 0 | 1 | 2 | 3 |
| | enrichment | 0.00 | 0.00 | 80.00 | 0.00 | 0.00 | 68.97 | 45.56 | 94.64 |
| | ***p*-value** | 1.0000 | 1.0000 | 0.0047 | 1.0000 | 1.0000 | 0.0109 | 0.0076 | 0.0020 |
| cuticular | **count** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | enrichment | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | ***p*-value** | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| | | ***completed soft*** | | | | | | | |
| metabolism | **count** | 32 | 23 | 18 | 23 | 23 | 32 | 45 | 33 |
| | enrichment | 1.93 | 2.68 | 2.62 | 2.12 | 1.78 | 1.97 | 2.13 | 1.87 |
| | ***p*-value** | 0.0046 | 0.0008 | 0.0038 | 0.0036 | 0.0179 | 0.0021 | <0.0001 | 0.0053 |
| target site | **count** | 5 | 5 | 0 | 3 | 4 | 5 | 5 | 3 |
| | enrichment | 2.59 | 4.51 | 0.00 | 2.26 | 2.41 | 2.45 | 1.85 | 1.42 |
| | ***p*-value** | 0.0298 | 0.0025 | 1.0000 | 0.1322 | 0.0701 | 0.0390 | 0.1085 | 0.3543 |
| behavior | **count** | 11 | 4 | 2 | 10 | 12 | 9 | 7 | 9 |
| | enrichment | 2.79 | 1.79 | 1.13 | 3.91 | 3.89 | 2.29 | 1.29 | 1.98 |
| | ***p*-value** | 0.0180 | 0.2175 | 0.4666 | 0.0040 | 0.0025 | 0.0512 | 0.3263 | 0.0908 |
| cuticular | **count** | 1 | 0 | 0 | 0 | 1 | 2 | 5 | 4 |
| | enrichment | 1.06 | 0.00 | 0.00 | 0.00 | 1.54 | 2.24 | 4.42 | 3.84 |
| | ***p*-value** | 0.5901 | 1.0000 | 1.0000 | 1.0000 | 0.4597 | 0.2233 | 0.0075 | 0.0307 |
| | | ***partial hard*** | | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| metabolism | **count** | 11 | 8 | 18 | 36 | 32 | 10 | 8 | 24 |
| | enrichment | 1.40 | 1.77 | 1.39 | 2.58 | 1.02 | 1.27 | 1.16 | 3.27 |
| | ***p*-value** | 0.1944 | 0.1293 | 0.1453 | 0.0002 | 0.4736 | 0.2850 | 0.3799 | 0.0001 |
| target site | **count** | 0 | 0 | 1 | 4 | 3 | 1 | 0 | 1 |
| | enrichment | 0.00 | 0.00 | 0.66 | 2.70 | 0.80 | 1.11 | 0.00 | 1.07 |
| | ***p*-value** | 1.0000 | 1.0000 | 0.8033 | 0.0479 | 0.7734 | 0.6117 | 1.0000 | 0.6268 |
| behavior | **count** | 1 | 2 | 2 | 1 | 8 | 1 | 1 | 0 |
| | enrichment | 0.52 | 1.50 | 0.59 | 0.27 | 0.93 | 0.50 | 0.62 | 0.00 |
| | ***p*-value** | 0.7638 | 0.3395 | 0.7553 | 0.9470 | 0.5637 | 0.7603 | 0.6956 | 1.0000 |
| cuticular | **count** | 0 | 1 | 0 | 1 | 5 | 0 | 0 | 1 |
| | enrichment | 0.00 | 4.10 | 0.00 | 1.43 | 2.37 | 0.00 | 0.00 | 1.95 |
| | ***p*-value** | 1.0000 | 0.2027 | 1.0000 | 0.4852 | 0.0684 | 1.0000 | 1.0000 | 0.3747 |
| | | ***partial soft*** | | | | | | | |
| metabolism | **count** | 25 | 27 | 25 | 31 | 19 | 16 | 27 | 45 |
| | enrichment | 1.22 | 1.86 | 1.14 | 1.04 | 1.30 | 1.22 | 0.86 | 1.15 |
| | ***p*-value** | 0.2060 | 0.0095 | 0.2976 | 0.4337 | 0.1829 | 0.2658 | 0.7766 | 0.2125 |
| target site | **count** | 4 | 4 | 4 | 2 | 0 | 2 | 3 | 5 |
| | enrichment | 1.59 | 2.18 | 1.55 | 0.68 | 0.00 | 1.15 | 0.79 | 1.15 |
| | ***p*-value** | 0.2227 | 0.0966 | 0.2479 | 0.8377 | 1.0000 | 0.5334 | 0.7865 | 0.4586 |
| behavior | **count** | 6 | 13 | 9 | 9 | 2 | 7 | 9 | 10 |
| | enrichment | 1.19 | 3.36 | 1.56 | 1.27 | 0.55 | 2.23 | 1.11 | 0.95 |
| | ***p*-value** | 0.3837 | 0.0043 | 0.1779 | 0.3076 | 0.7908 | 0.0840 | 0.4277 | 0.5751 |
| cuticular | **count** | 1 | 0 | 2 | 5 | 0 | 0 | 0 | 3 |
| | enrichment | 0.82 | 0.00 | 1.58 | 2.36 | 0.00 | 0.00 | 0.00 | 1.32 |
| | ***p*-value** | 0.6906 | 1.0000 | 0.3598 | 0.0728 | 1.0000 | 1.0000 | 1.0000 | 0.3991 |

Significant values (*i.e.* $p < 0.05$) underlined.

**Table S6.** Statistically significant enrichments of selective sweeps within GO term annotations for each empirical mosquito population

| | GO term | function | A O M count | enrichment | *q* -value | |
|---|---|---|---|---|---|---|
| *all four sweep states :molecular function* | | | | | | *all four sweep states :cellular components* |
| | GO:0005524 | ATP binding | 78 | 1.46 | 0.03875 | |
| | GO:0003676 | nucleic acid binding | 74 | 1.53 | 0.02583 | |
| | GO:0005488 | binding | 31 | 1.78 | 0.02583 | |
| | GO:0003723 | RNA binding | 29 | 1.80 | 0.02583 | |
| *completed soft :cellular components* | | | | | | |
| | GO:0005886 | plasma membrane | 27 | 2.15 | 0.03780 | *:molecular function* |
| | GO:0030286 | dynein complex | 6 | 4.21 | 0.03780 | |
| | GO:0005847 | mRNA cleavage and p | 4 | 8.14 | 0.03780 | |
| | GO:0005737 | cytoplasm | 39 | 1.90 | 0.03780 | |
| | GO:0005694 | chromosome | 6 | 5.54 | 0.02835 | |
| | GO:0016021 | integral component o | 121 | 1.56 | <0.01260 | |
| | GO:0005634 | nucleus | 60 | 1.80 | <0.01260 | |
| | GO:0016020 | membrane | 42 | 1.91 | <0.01260 | |
| *:molecular function* | | | | | | |
| | GO:0005524 | ATP binding | 40 | 1.90 | 0.01598 | *completed soft* |
| | GO:0005488 | binding | 18 | 2.50 | 0.01598 | *:biological processes* |
| | GO:0005515 | protein binding | 99 | 1.68 | <0.01598 | |
| | GO:0003676 | nucleic acid binding | 50 | 2.60 | <0.01598 | |
| | GO:0003723 | RNA binding | 19 | 3.04 | <0.01598 | |
| | | | | | | *:cellular components* |

*:molecular function*

dataset.

| GO term | function | B F M count | enrichment | *q* -value |
|---|---|---|---|---|
| GO:0030286 | dynein complex | 7 | 3.60 | 0.03720 |
| GO:0005737 | cytoplasm | 57 | 1.75 | 0.00744 |
| GO:0016020 | membrane | 63 | 1.77 | 0.00744 |
| GO:0005634 | nucleus | 83 | 1.54 | <0.00744 |
| GO:0016021 | integral component of | 199 | 1.57 | <0.00744 |
| GO:0005887 | integral component of | 25 | 2.37 | <0.00744 |
| GO:0003723 | RNA binding | 22 | 2.08 | 0.03875 |
| GO:0005524 | ATP binding | 55 | 1.67 | 0.02214 |
| GO:0005506 | iron ion binding | 18 | 3.24 | 0.02214 |
| GO:0005515 | protein binding | 181 | 1.86 | <0.01550 |
| GO:0008270 | zinc ion binding | 57 | 1.88 | <0.01550 |
| GO:0003676 | nucleic acid binding | 78 | 2.43 | <0.01550 |
| GO:0005488 | binding | 29 | 2.49 | <0.01550 |
| GO:0016787 | hydrolase activity | 15 | 3.14 | <0.01550 |
| GO:0007165 | signal transduction | 20 | 4.39 | <0.03513 |
| GO:0006508 | proteolysis | 29 | 4.84 | <0.03513 |
| GO:0000398 | mRNA splicing, via sp | 8 | 7.27 | <0.03513 |
| GO:0005852 | eukaryotic translation | 3 | 11.92 | 0.03860 |
| GO:0005847 | mRNA cleavage and p | 3 | 13.63 | 0.02573 |
| GO:0005763 | mitochondrial small r | 3 | 13.57 | 0.02413 |
| GO:0005840 | ribosome | 7 | 4.66 | 0.01654 |
| GO:0005654 | nucleoplasm | 5 | 11.08 | 0.01654 |

| | GO term |
|---|---|
| *all four sweep states :biological processes* | |
| | GO:0006508 |
| | GO:0000398 |
| *:cellular components* | |
| | GO:0005840 |
| | GO:0005737 |
| | GO:0005622 |
| | GO:0016021 |
| | GO:0016020 |
| | GO:0005634 |
| | GO:0005654 |
| *:molecular function* | |
| | GO:0005524 |
| | GO:0005509 |
| | GO:0005515 |
| | GO:0003676 |
| | GO:0005488 |
| | GO:0003723 |
| | GO:0004252 |
| *completed soft :biological processes* | |
| | GO:0045944 |
| | GO:0006508 |
| | GO:0007165 |
| | GO:0007018 |
| | GO:0000398 |
| *:cellular components* | |

| | | | | | | |
|---|---|---|---|---|---|---|
| GO:0005622 | intracellular | 20 | 2.96 | 0.00772 | | GO:0005840 |
| GO:0016021 | integral component of | 88 | 2.10 | <0.00772 | | GO:0030286 |
| GO:0016020 | membrane | 30 | 2.52 | <0.00772 | | GO:0016020 |
| GO:0005634 | nucleus | 49 | 2.74 | <0.00772 | | GO:0016021 |
| GO:0005737 | cytoplasm | 29 | 2.75 | <0.00772 | | GO:0005634 |
| | | | | | | GO:0005737 |
| GO:0016491 | oxidoreductase activit | 12 | 3.70 | 0.04368 | | GO:0005622 |
| GO:0004812 | aminoacyl-tRNA ligas | 4 | 8.16 | 0.03609 | | GO:0005654 |
| GO:0003735 | structural constituent | 8 | 4.60 | 0.02750 | *:molecular function* | |
| GO:0004177 | aminopeptidase activi | 3 | 16.04 | 0.02357 | | GO:0008234 |
| GO:0003682 | chromatin binding | 6 | 5.64 | 0.01904 | | GO:0030170 |
| GO:0000166 | nucleotide binding | 8 | 5.66 | 0.01904 | | GO:0008270 |
| GO:0004252 | serine-type endopepti | 17 | 4.73 | 0.00750 | | GO:0005515 |
| GO:0005506 | iron ion binding | 11 | 6.25 | 0.00750 | | GO:0005524 |
| GO:0005085 | guanyl-nucleotide exc | 4 | 7.35 | 0.00750 | | GO:0003676 |
| GO:0046872 | metal ion binding | 19 | 2.93 | <0.00750 | | GO:0003723 |
| GO:0005515 | protein binding | 108 | 3.28 | <0.00750 | | GO:0005488 |
| GO:0008270 | zinc ion binding | 33 | 3.30 | <0.00750 | | GO:0016887 |
| GO:0005524 | ATP binding | 38 | 3.46 | <0.00750 | | GO:0003777 |
| GO:0003723 | RNA binding | 15 | 4.26 | <0.00750 | | GO:0035091 |
| GO:0016887 | ATPase activity | 10 | 4.37 | <0.00750 | | |
| GO:0005488 | binding | 18 | 4.53 | <0.00750 | | |
| GO:0003676 | nucleic acid binding | 51 | 4.79 | <0.00750 | | |

| B function | F count | S enrichment | *q* -value |
|---|---|---|---|
| proteolysis | 57 | 2.15 | <0.04905 |
| mRNA splicing, via sp | 15 | 2.97 | <0.04905 |
| ribosome | 17 | 2.41 | 0.02126 |
| cytoplasm | 73 | 1.54 | 0.01240 |
| intracellular | 51 | 1.71 | 0.01240 |
| integral component of | 263 | 1.44 | <0.00930 |
| membrane | 77 | 1.49 | <0.00930 |
| nucleus | 119 | 1.53 | <0.00930 |
| nucleoplasm | 9 | 4.93 | <0.00930 |
| ATP binding | 70 | 1.44 | 0.04429 |
| calcium ion binding | 39 | 1.64 | 0.02583 |
| protein binding | 207 | 1.47 | <0.01550 |
| nucleic acid binding | 87 | 1.85 | <0.01550 |
| binding | 32 | 1.93 | <0.01550 |
| RNA binding | 31 | 1.99 | <0.01550 |
| serine-type endopept | 39 | 2.49 | <0.01550 |
| positive regulation of | 7 | 7.89 | 0.04288 |
| proteolysis | 20 | 4.39 | <0.02680 |
| signal transduction | 19 | 5.27 | <0.02680 |
| microtubule-based m | 7 | 8.48 | <0.02680 |
| mRNA splicing, via sp | 9 | 11.10 | <0.02680 |

| | GO term | C function | M count |
|---|---|---|---|
| *all four sweep states* | | | |
| *:biological processes* | | | |
| | GO:0006508 | proteolysis | 70 |
| | GO:0000398 | mRNA splicing, via sp | 16 |
| *:cellular components* | | | |
| | GO:0005634 | nucleus | 128 |
| | GO:0016020 | membrane | 91 |
| | GO:0016021 | integral component of | 307 |
| | GO:0005681 | spliceosomal complex | 9 |
| *:molecular function* | | | |
| | GO:0005524 | ATP binding | 86 |
| | GO:0016887 | ATPase activity | 22 |
| | GO:0005515 | protein binding | 250 |
| | GO:0008270 | zinc ion binding | 91 |
| | GO:0003676 | nucleic acid binding | 105 |
| | GO:0004252 | serine-type endopept | 49 |
| *completed soft* | | | |
| *:biological processes* | | | |
| | GO:0006511 | ubiquitin-dependent p | 7 |
| | GO:0016579 | protein deubiquitinat | 6 |
| | GO:0000398 | mRNA splicing, via sp | 10 |
| *:cellular components* | | | |
| | GO:0005634 | nucleus | 38 |
| | GO:0005681 | spliceosomal complex | 5 |
| | GO:0016021 | integral component of | 107 |
| | GO:0016020 | membrane | 41 |
| | GO:0005886 | plasma membrane | 25 |
| *:molecular function* | | | |

| | | | |
|---|---|---|---|
| ribosome | 7 | 5.72 | 0.01945 |
| dynein complex | 5 | 9.22 | 0.01111 |
| membrane | 23 | 2.54 | 0.00648 |
| integral component of | 77 | 2.37 | <0.00648 |
| nucleus | 42 | 3.03 | <0.00648 |
| cytoplasm | 30 | 3.67 | <0.00648 |
| intracellular | 20 | 3.74 | <0.00648 |
| nucleoplasm | 7 | 20.40 | <0.00648 |
| | | | |
| cysteine-type peptida | 5 | 18.48 | 0.04658 |
| pyridoxal phosphate b | 4 | 11.15 | 0.03416 |
| zinc ion binding | 27 | 3.51 | <0.00949 |
| protein binding | 91 | 3.54 | <0.00949 |
| ATP binding | 31 | 3.59 | <0.00949 |
| nucleic acid binding | 31 | 3.69 | <0.00949 |
| RNA binding | 12 | 4.40 | <0.00949 |
| binding | 15 | 4.93 | <0.00949 |
| ATPase activity | 11 | 5.83 | <0.00949 |
| microtubule motor ac | 7 | 9.20 | <0.00949 |
| phosphatidylinositol b | 4 | 13.99 | <0.00949 |

| | | |
|---|---|---|
| GO:0030620 | U2 snRNA binding | 1 |
| GO:0004843 | ubiquitin-specific pro | 4 |
| GO:0005515 | protein binding | 101 |
| GO:0003676 | nucleic acid binding | 38 |
| GO:0005524 | ATP binding | 39 |
| GO:0008270 | zinc ion binding | 36 |
| GO:0005488 | binding | 15 |
| GO:0003723 | RNA binding | 15 |
| GO:0016887 | ATPase activity | 13 |

| S enrichment | *q* -value | | GO term | G function | A count | S enrichment | *q* -value |
|---|---|---|---|---|---|---|---|
| | | *all four sweep states* | | | | | |
| | | *:cellular components* | | | | | |
| 1.94 | <0.04905 | | GO:0005634 | nucleus | 148 | 1.32 | <0.03720 |
| 2.66 | <0.04905 | *:molecular function* | | | | | |
| | | | GO:0005524 | ATP binding | 96 | 1.36 | 0.03875 |
| 1.31 | 0.03720 | | GO:0005515 | protein binding | 248 | 1.26 | <0.02583 |
| 1.39 | 0.02480 | | GO:0003676 | nucleic acid binding | 94 | 1.46 | <0.02583 |
| 1.31 | <0.0186 | | GO:0008270 | zinc ion binding | 90 | 1.49 | <0.02583 |
| 3.51 | <0.0186 | *completed soft* | | | | | |
| | | *:cellular components* | | | | | |
| 1.42 | 0.03875 | | GO:0016021 | integral component of | 96 | 1.54 | 0.01895 |
| 2.06 | 0.01550 | | GO:0016020 | membrane | 35 | 2.00 | 0.01895 |
| 1.40 | <0.0155 | | GO:0030286 | dynein complex | 6 | 4.88 | 0.01895 |
| 1.64 | <0.0155 | | GO:0005737 | cytoplasm | 34 | 2.10 | 0.01895 |
| 1.67 | <0.0155 | | GO:0005829 | cytosol | 12 | 3.58 | 0.01895 |
| 2.30 | <0.0155 | | GO:0005634 | nucleus | 70 | 2.56 | <0.01895 |
| | | *:molecular function* | | | | | |
| | | | GO:0005549 | odorant binding | 14 | 5.02 | 0.04700 |
| 6.04 | <0.03727 | | GO:0017137 | Rab GTPase binding | 4 | 6.22 | 0.04494 |
| 7.20 | <0.03727 | | GO:0003723 | RNA binding | 13 | 2.63 | 0.04261 |
| 8.34 | <0.03727 | | GO:0005509 | calcium ion binding | 19 | 2.16 | 0.03995 |
| | | | GO:0016787 | hydrolase activity | 8 | 4.06 | 0.03688 |
| 1.82 | 0.01640 | | GO:0004525 | ribonuclease III activi | 3 | 11.07 | 0.02663 |
| 10.00 | 0.01025 | | GO:0004197 | cysteine-type endope | 7 | 12.69 | 0.00726 |
| 2.11 | <0.01025 | | GO:0005515 | protein binding | 101 | 2.12 | <0.00726 |
| 2.88 | <0.01025 | | GO:0005524 | ATP binding | 39 | 2.30 | <0.00726 |
| 3.22 | <0.01025 | | GO:0046872 | metal ion binding | 23 | 2.41 | <0.00726 |
| | | | GO:0003677 | DNA binding | 32 | 2.78 | <0.00726 |

| | |
|---|---|
| 3333.33 | 0.02950 |
| 9.50 | 0.02213 |
| 2.49 | <0.01264 |
| 2.85 | <0.01264 |
| 2.97 | <0.01264 |
| 3.07 | <0.01264 |
| 3.28 | <0.01264 |
| 3.41 | <0.01264 |
| 4.87 | <0.01264 |

| | | | | |
|---|---|---|---|---|
| GO:0008270 | zinc ion binding | 42 | 2.81 | <0.00726 |
| GO:0003700 | sequence-specific DN | 19 | 2.82 | <0.00726 |
| GO:0005488 | binding | 17 | 2.89 | <0.00726 |
| GO:0003676 | nucleic acid binding | 51 | 3.23 | <0.00726 |
| GO:0004674 | protein serine/threon | 9 | 3.84 | <0.00726 |
| GO:0008234 | cysteine-type peptida | 7 | 14.78 | <0.00726 |

| | GO term | function | G count | N enrichment | S *q* -value | |
|---|---|---|---|---|---|---|
| *all four sweep states* | | | | | | *all four sweep states* |
| *:cellular components* | | | | | | *:cellular components* |
| | GO:0005634 | nucleus | 101 | 1.40 | 0.03720 | |
| | GO:0016021 | integral component of | 237 | 1.38 | <0.03720 | |
| *:molecular function* | | | | | | |
| | GO:0003676 | nucleic acid binding | 65 | 1.54 | 0.02583 | |
| | GO:0005515 | protein binding | 182 | 1.40 | <0.02583 | |
| | GO:0005524 | ATP binding | 70 | 1.55 | <0.02583 | *:molecular function* |
| *completed soft* | | | | | | |
| *:biological processes* | | | | | | |
| | GO:0006886 | intracellular protein t | 11 | 3.34 | 0.04136 | |
| | GO:0000398 | mRNA splicing, via spl | 8 | 4.10 | 0.04136 | |
| | GO:0006379 | mRNA cleavage | 4 | 12.77 | 0.04136 | |
| | GO:0006396 | RNA processing | 9 | 4.98 | 0.04136 | |
| | GO:0045944 | positive regulation of | 11 | 4.92 | <0.04136 | |
| *:cellular components* | | | | | | |
| | GO:0005739 | mitochondrion | 13 | 2.85 | 0.03820 | *completed soft* |
| | GO:0016020 | membrane | 40 | 1.79 | 0.03820 | *:biological processes* |
| | GO:0005737 | cytoplasm | 38 | 1.91 | 0.03820 | |
| | GO:0016021 | integral component of | 131 | 1.66 | <0.01910 | |
| | GO:0005634 | nucleus | 72 | 2.16 | <0.01910 | |
| *:molecular function* | | | | | | |
| | GO:0005515 | protein binding | 116 | 1.90 | <0.01630 | |
| | GO:0008270 | zinc ion binding | 40 | 2.14 | <0.01630 | *:cellular components* |
| | GO:0003676 | nucleic acid binding | 45 | 2.34 | <0.01630 | |
| | GO:0003677 | DNA binding | 36 | 2.54 | <0.01630 | |
| | GO:0005524 | ATP binding | 53 | 2.56 | <0.01630 | |

*:molecular function*

| GO term | function | G W A count | enrichment | *q* -value |
|---|---|---|---|---|
| GO:0045202 | synapse | 10 | 2.14 | 0.03720 |
| GO:0005737 | cytoplasm | 96 | 1.38 | 0.03720 |
| GO:0005654 | nucleoplasm | 9 | 3.16 | 0.03720 |
| GO:0005634 | nucleus | 146 | 1.29 | 0.01860 |
| GO:0016021 | integral component of | 327 | 1.21 | <0.01860 |
| GO:0004497 | monooxygenase activi | 22 | 2.15 | 0.04844 |
| GO:0016887 | ATPase activity | 25 | 1.75 | 0.02214 |
| GO:0004252 | serine-type endopepti | 44 | 1.83 | 0.02214 |
| GO:0008270 | zinc ion binding | 92 | 1.47 | 0.01550 |
| GO:0003677 | DNA binding | 68 | 1.44 | <0.01550 |
| GO:0005515 | protein binding | 302 | 1.49 | <0.01550 |
| GO:0005524 | ATP binding | 112 | 1.59 | <0.01550 |
| GO:0003676 | nucleic acid binding | 110 | 1.67 | <0.01550 |
| GO:0006357 | regulation of transcrip | 16 | 2.27 | 0.04108 |
| GO:0006508 | proteolysis | 36 | 2.37 | 0.04108 |
| GO:0007018 | microtubule-based m | 9 | 3.39 | 0.04108 |
| GO:0016579 | protein deubiquitinat | 7 | 3.83 | 0.04108 |
| GO:0000278 | mitotic cell cycle | 5 | 6.54 | 0.04108 |
| GO:0005886 | plasma membrane | 33 | 1.91 | 0.03192 |
| GO:0005829 | cytosol | 15 | 2.45 | 0.02785 |
| GO:0012505 | endomembrane syste | 5 | 5.10 | 0.02681 |
| GO:0005694 | chromosome | 6 | 4.57 | 0.02553 |

| | GO term |
|---|---|
| *all four sweep states* | |
| *:cellular components* | |
| | GO:0005737 |
| | GO:0016020 |
| | GO:0016021 |
| | GO:0005634 |
| *:molecular function* | |
| | GO:0004004 |
| | GO:0003690 |
| | GO:0005549 |
| | GO:0003723 |
| | GO:0005524 |
| | GO:0046872 |
| | GO:0005515 |
| | GO:0003676 |
| | GO:0008270 |
| | GO:0005488 |
| | GO:0004252 |
| *completed hard* | |
| *:cellular components* | |
| | GO:0016020 |
| *completed soft* | |
| *:biological processes* | |
| | GO:0007155 |
| | GO:0006396 |
| | GO:0006508 |
| | GO:0007165 |
| | GO:0006355 |

| | | | | | | |
|---|---|---|---|---|---|---|
| GO:0005622 | intracellular | 30 | 1.88 | 0.01436 | | GO:0007018 |
| GO:0005813 | centrosome | 6 | 4.65 | 0.01436 | | GO:0000398 |
| GO:0016021 | integral component of | 169 | 1.63 | <0.00638 | *:cellular components* | |
| GO:0016020 | membrane | 54 | 1.86 | <0.00638 | | GO:0005737 |
| GO:0005737 | cytoplasm | 51 | 1.91 | <0.00638 | | GO:0005622 |
| GO:0005634 | nucleus | 86 | 1.95 | <0.00638 | | GO:0005739 |
| GO:0005739 | mitochondrion | 18 | 3.05 | <0.00638 | | GO:0016021 |
| GO:0005654 | nucleoplasm | 9 | 8.25 | <0.00638 | | GO:0016020 |
| | | | | | | GO:0005634 |
| GO:0005509 | calcium ion binding | 27 | 1.84 | 0.04640 | | GO:0005886 |
| GO:0004252 | serine-type endopepti | 24 | 2.61 | 0.04640 | | GO:0030286 |
| GO:0003777 | microtubule motor ac | 8 | 3.27 | 0.04060 | *:molecular function* | |
| GO:0036459 | ubiquitinyl hydrolase | 6 | 4.36 | 0.02953 | | GO:0004386 |
| GO:0046872 | metal ion binding | 29 | 1.86 | 0.01624 | | GO:0030515 |
| GO:0005488 | binding | 22 | 2.25 | 0.00902 | | GO:0008270 |
| GO:0005515 | protein binding | 166 | 2.05 | <0.00902 | | GO:0046872 |
| GO:0003677 | DNA binding | 39 | 2.10 | <0.00902 | | GO:0003690 |
| GO:0008270 | zinc ion binding | 57 | 2.30 | <0.00902 | | GO:0003677 |
| GO:0005524 | ATP binding | 64 | 2.34 | <0.00902 | | GO:0003824 |
| GO:0003824 | catalytic activity | 31 | 2.35 | <0.00902 | | GO:0005524 |
| GO:0016887 | ATPase activity | 15 | 2.61 | <0.00902 | | GO:0005515 |
| GO:0003676 | nucleic acid binding | 70 | 2.72 | <0.00902 | | GO:0003676 |
| GO:0000287 | magnesium ion bindir | 7 | 5.36 | <0.00902 | | GO:0016887 |
| | | | | | | GO:0003723 |
| | | | | | | GO:0005488 |
| | | | | | | GO:0003777 |
| | | | | | *partial hard* | |
| | | | | | *:cellular components* | |
| | | | | | | GO:0000798 |

| | U | G | S | |
|---|---|---|---|---|
| function | | count | enrichment | *q* -value |
| cytoplasm | | 98 | 1.37 | 0.03720 |
| membrane | | 107 | 1.40 | 0.01240 |
| integral component of | | 357 | 1.31 | <0.01240 |
| nucleus | | 159 | 1.39 | <0.01240 |
| ATP-dependent RNA h | | 11 | 2.46 | 0.04227 |
| double-stranded DNA | | 6 | 3.56 | 0.04227 |
| odorant binding | | 33 | 2.29 | 0.03444 |
| RNA binding | | 39 | 1.76 | 0.00969 |
| ATP binding | | 105 | 1.46 | <0.00969 |
| metal ion binding | | 58 | 1.51 | <0.00969 |
| protein binding | | 309 | 1.51 | <0.00969 |
| nucleic acid binding | | 104 | 1.51 | <0.00969 |
| zinc ion binding | | 104 | 1.64 | <0.00969 |
| binding | | 45 | 1.90 | <0.00969 |
| serine-type endopept | | 46 | 1.99 | <0.00969 |
| membrane | | 4 | 22.45 | <0.03900 |
| cell adhesion | | 13 | 3.41 | 0.04590 |
| RNA processing | | 8 | 3.95 | 0.04590 |
| proteolysis | | 34 | 2.92 | 0.04284 |
| signal transduction | | 21 | 2.40 | 0.02678 |
| regulation of transcrip | | 35 | 2.19 | <0.02678 |

| | | | |
|---|---|---|---|
| microtubule-based m | 9 | 4.42 | <0.02678 |
| mRNA splicing, via sp | 11 | 5.27 | <0.02678 |
| | | | |
| cytoplasm | 39 | 1.85 | 0.02438 |
| intracellular | 25 | 1.98 | 0.02438 |
| mitochondrion | 15 | 3.17 | 0.02438 |
| integral component o | 153 | 1.87 | <0.00780 |
| membrane | 53 | 2.31 | <0.00780 |
| nucleus | 81 | 2.32 | <0.00780 |
| plasma membrane | 31 | 2.36 | <0.00780 |
| dynein complex | 8 | 6.03 | <0.00780 |
| | | | |
| helicase activity | 5 | 5.35 | 0.03634 |
| snoRNA binding | 4 | 7.29 | 0.03262 |
| zinc ion binding | 40 | 2.07 | 0.00707 |
| metal ion binding | 26 | 2.12 | 0.00707 |
| double-stranded DNA | 5 | 9.64 | 0.00707 |
| DNA binding | 34 | 2.32 | <0.00707 |
| catalytic activity | 25 | 2.33 | <0.00707 |
| ATP binding | 52 | 2.40 | <0.00707 |
| protein binding | 159 | 2.49 | <0.00707 |
| nucleic acid binding | 57 | 2.74 | <0.00707 |
| ATPase activity | 15 | 3.32 | <0.00707 |
| RNA binding | 22 | 3.35 | <0.00707 |
| binding | 30 | 3.99 | <0.00707 |
| microtubule motor ac | 8 | 4.27 | <0.00707 |
| | | | |
| nuclear cohesin comp | 1 | INF | <0.03900 |

**Table S7.** Example command lines to generate *discoal* simulations.

| | |
|---|---|
| AOM | discoal 120 2000 55000 -Pt 100.667293 1006.672930 -Pre 1107.340223 3322.020670 -en 0.000214 0 1.0021 |
| BFM | discoal 138 2000 55000 -Pt 1330.793457 13307.934565 -Pre 14638.728022 43916.184065 -en 0.000161 0 0. |
| BFS | discoal 162 2000 55000 -Pt 1750.204699 17502.046985 -Pre 19252.251684 57756.755051 -en 0.000038 0 1. |
| CMS | discoal 182 2000 55000 -Pt 465.266905 4652.669049 -Pre 5117.935954 15353.807861 -en 0.000061 0 1.111 |
| GAS | discoal 112 2000 55000 -Pt 569.780857 5697.808570 -Pre 6267.589427 18802.768280 -en 0.000080 0 0.631 |
| GNS | discoal 62 2000 55000 -Pt 1999.973115 19999.731147 -Pre 21999.704262 65999.112785 -en 0.000820 0 0.9 |
| GWA | discoal 92 2000 55000 -Pt 974.582413 9745.824130 -Pre 10720.406543 32161.219629 -en 0.000242 0 1.001 |
| UGS | discoal 206 2000 55000 -Pt 1165.186611 11651.866111 -Pre 12817.052722 38451.158167 -en 0.000072 0 1. |

Location changes by setting the value of $x (between 0 and 1)
To simulate complete sweeps, remove the -Pc flag and values
To simulate hard sweeps, remove the -Pf flag and values
To simulate no sweeps, remove the -Pc, -Pf, -x, -ws, -Pa, and -Pu flags and according values

.85 -en 0.000363 0 1.003592 -en 0.000439 0 1.003717 -en 0.000517 0 1.020643 -en 0.000598 0 1.024216 -en 0.000680 0 1.
.985949 -en 0.000216 0 0.983492 -en 0.000271 0 1.000000 -en 0.000328 0 0.999114 -en 0.000385 0 1.000000 -en 0.00044
.001158 -en 0.000077 0 1.004693 -en 0.000117 0 1.017850 -en 0.000157 0 1.019335 -en 0.000198 0 1.022229 -en 0.00024
.333 -en 0.000096 0 2.040230 -en 0.000160 0 2.927333 -en 0.000252 0 3.603759 -en 0.000368 0 3.724538 -en 0.000489 0
.586 -en 0.000132 0 0.589447 -en 0.000181 0 0.159569 -en 0.000195 0 0.159078 -en 0.000209 0 0.154646 -en 0.000222 0
)92861 -en 0.001110 0 0.794064 -en 0.001350 0 0.540221 -en 0.001520 0 0.490645 -en 0.001679 0 0.473441 -en 0.001838
.700 -en 0.000367 0 1.007686 -en 0.001640 0 1.007342 -en 0.002131 0 1.007686 -en 0.002480 0 1.008099 -en 0.003237 0
.000968 -en 0.000096 0 1.008257 -en 0.000121 0 1.008772 -en 0.000146 0 1.022998 -en 0.000172 0 1.007552 -en 0.00019

024809 -en 0.000764 0 1.037077 -en 0.000851 0 1.040212 -en 0.000939 0 1.043685 -en 0.001029 0 1.054121 -en 0.00112
4 0 1.077979 -en 0.000508 0 1.071814 -en 0.000573 0 1.024200 -en 0.000636 0 1.049526 -en 0.000702 0 1.070901 -en 0.0
0 0 1.025899 -en 0.000282 0 1.030126 -en 0.000326 0 1.030824 -en 0.000369 0 1.031031 -en 0.000459 0 1.031171 -en 0.0
3.778291 -en 0.000613 0 3.793999 -en 0.000739 0 3.801097 -en 0.000867 0 3.813393 -en 0.000997 0 3.819256 -en 0.0011
0.153297 -en 0.000236 0 0.151153 -en 0.000250 0 0.149826 -en 0.000278 0 0.149120 -en 0.000511 0 0.149826 -en 0.0005
0 0.467565 -en 0.002002 0 0.463145 -en 0.002170 0 0.461279 -en 0.002344 0 0.456089 -en 0.002523 0 0.453183 -en 0.00
0.997880 -en 0.003438 0 0.560152 -en 0.003554 0 0.530789 -en 0.003667 0 0.530786 -en 0.003784 0 0.530789 -en 0.0039
7 0 1.003068 -en 0.000223 0 1.001495 -en 0.000249 0 1.000662 -en 0.000275 0 0.997039 -en 0.000302 0 0.997529 -en 0.0

2 0 1.074898 -en 0.001219 0 1.079116 -en 0.001318 0 1.085179 -en 0.001419 0 1.089208 -en 0.001523 0 1.089378 -en 0.00
00770 0 1.020093 -en 0.000836 0 0.974589 -en 0.000900 0 0.893579 -en 0.000959 0 0.829158 -en 0.001015 0 0.741725 -en
00891 0 1.030126 -en 0.000943 0 1.027106 -en 0.000995 0 1.019246 -en 0.001047 0 1.019335 -en 0.001100 0 1.019086 -en
28 0 3.828034 -en 0.001261 0 3.836271 -en 0.001396 0 3.850718 -en 0.001534 0 3.864462 -en 0.001673 0 3.866261 -en 0.0
30 0 0.150585 -en 0.000550 0 0.158882 -en 0.000572 0 0.165277 -en 0.000595 0 0.170755 -en 0.000620 0 0.504559 -en 0.0
2708 0 0.451860 -en 0.002900 0 0.450422 -en 0.003099 0 0.448048 -en 0.003307 0 0.446177 -en 0.003747 0 0.445082 -en
04 0 0.530786 -en 0.004027 0 0.528428 -en 0.004154 0 0.526746 -en 0.004285 0 0.526125 -en 0.004703 0 0.524966 -en 0.0
00328 0 0.995950 -en 0.000382 0 0.994539 -en 0.000410 0 0.995950 -en 0.000437 0 0.997039 -en 0.000465 0 1.000000 -en

)1628 0 1.090078 -en 0.001846 0 1.095578 -en 0.001959 0 1.095843 -en 0.003823 0 1.095578 -en 0.004141 0 1.089208 -en
n 0.001066 0 0.697406 -en 0.001115 0 0.673163 -en 0.001163 0 0.648442 -en 0.001210 0 0.625999 -en 0.001256 0 0.61677
n 0.001154 0 1.019246 -en 0.001209 0 1.020601 -en 0.001265 0 1.018926 -en 0.001321 0 1.013910 -en 0.001378 0 1.00468
)01814 0 3.870850 -en 0.002102 0 3.876694 -en 0.002249 0 3.879188 -en 0.002398 0 3.891219 -en 0.002702 0 3.891609 -e
)00694 0 0.617331 -en 0.000787 0 0.670013 -en 0.000890 0 0.682974 -en 0.000998 0 0.694737 -en 0.001111 0 0.696173 -e
0.004229 0 0.433300 -en 0.004481 0 0.167095 -en 0.004583 0 0.167041 -en 0.004690 0 0.166713 -en 0.004802 0 0.165622
)04851 0 0.524897 -en 0.005004 0 0.524506 -en 0.005498 0 0.523930 -en 0.005857 0 0.523704 -en 0.006047 0 0.523703 -e
n 0.000494 0 0.999703 -en 0.000523 0 0.997529 -en 0.000552 0 1.000968 -en 0.000581 0 1.003490 -en 0.000611 0 1.00455

ı 0.004305 0 1.085229 -en 0.004473 0 1.064777 -en 0.004641 0 1.022083 -en 0.004807 0 0.971303 -en 0.004969 0 0.71761
72 -en 0.001303 0 0.612164 -en 0.001349 0 0.609098 -en 0.001397 0 0.606156 -en 0.001444 0 0.605458 -en 0.001493 0 0.6
33 -en 0.001435 0 1.000118 -en 0.001493 0 0.985278 -en 0.001551 0 0.967523 -en 0.001608 0 0.621680 -en 0.001646 0 0.5
en 0.002856 0 3.898662 -en 0.003014 0 3.903107 -en 0.003334 0 3.906573 -en 0.003498 0 3.918087 -en 0.004177 0 3.9235
en 0.001227 0 0.695425 -en 0.001346 0 0.670013 -en 0.001463 0 0.665223 -en 0.001583 0 0.628336 -en 0.001700 0 0.5947
2 -en 0.004920 0 0.165268 -en 0.005044 0 0.164789 -en 0.005175 0 0.164082 -en 0.005313 0 0.163800 -en 0.005459 0 0.16
en 0.006450 0 0.523432 -en 0.006886 0 0.522467 -en 0.007118 0 0.518847 -en 0.007358 0 0.518125 -en 0.007608 0 0.5151
54 -en 0.000671 0 1.006577 -en 0.000702 0 1.008257 -en 0.000734 0 1.008763 -en 0.000765 0 1.012032 -en 0.000797 0 1.0

.3 -en 0.005091 0 0.537087 -en 0.005185 0 0.499553 -en 0.005275 0 0.485986 -en 0.005553 0 0.481404 -en 0.005749 0 0.4
503039 -en 0.001543 0 0.602428 -en 0.001593 0 0.597894 -en 0.001644 0 0.597796 -en 0.001695 0 0.591886 -en 0.001747
599478 -en 0.001683 0 0.583576 -en 0.001719 0 0.576138 -en 0.001756 0 0.564006 -en 0.001792 0 0.559137 -en 0.001829
546 -en 0.004710 0 3.933192 -en 0.004894 0 3.941164 -en 0.006667 0 3.944081 -en 0.006879 0 3.941164 -en 0.007093 0 3.
745 -en 0.001813 0 0.551675 -en 0.001921 0 0.449593 -en 0.002011 0 0.182750 -en 0.002049 0 0.176691 -en 0.002087 0 0.
53448 -en 0.005614 0 0.163443 -en 0.005778 0 0.163309 -en 0.005954 0 0.163127 -en 0.006141 0 0.162466 -en 0.006342 0
.09 -en 0.007868 0 0.513952 -en 0.008140 0 0.513027 -en 0.008424 0 0.507032 -en 0.008718 0 0.503846 -en 0.009025 0 0.
011293 -en 0.000830 0 1.012032 -en 0.000863 0 1.019422 -en 0.000896 0 1.022639 -en 0.000930 0 1.022788 -en 0.000965

72265 -en 0.006647 0 0.471576 -en 0.006909 0 0.469288 -en 0.007188 0 0.461541 -en 0.007333 0 0.449950 -en 0.007479
0 0.591675 -en 0.001801 0 0.591238 -en 0.001910 0 0.589350 -en 0.002082 0 0.589006 -en 0.002141 0 0.588612 -en 0.002
0 0.555749 -en 0.001866 0 0.552085 -en 0.001904 0 0.551380 -en 0.001942 0 0.550215 -en 0.002059 0 0.550213 -en 0.002
933192 -en 0.007531 0 3.918087 -en 0.007755 0 3.903965 -en 0.007980 0 3.883352 -en 0.008208 0 3.866473 -en 0.008439
174670 -en 0.002125 0 0.174154 -en 0.002164 0 0.173619 -en 0.002205 0 0.172705 -en 0.002246 0 0.171831 -en 0.002289
0.162077 -en 0.006556 0 0.161210 -en 0.006786 0 0.159987 -en 0.007032 0 0.159833 -en 0.007298 0 0.159969 -en 0.0079
490114 -en 0.009340 0 0.468610 -en 0.009656 0 0.438019 -en 0.009967 0 0.417032 -en 0.010280 0 0.406419 -en 0.010603
0 1.023465 -en 0.000999 0 1.024238 -en 0.001035 0 1.026880 -en 0.001070 0 1.027391 -en 0.001107 0 1.029008 -en 0.001

0 0.488293 -en 0.007643 0 0.557396 -en 0.007838 0 0.562350 -en 0.008042 0 0.565078 -en 0.008255 0 0.569917 -en 0.00
2202 0 0.588174 -en 0.002264 0 0.586833 -en 0.002327 0 0.586283 -en 0.002391 0 0.579687 -en 0.002456 0 0.575413 -en
2226 0 0.550215 -en 0.002751 0 0.550213 -en 0.004234 0 0.550215 -en 0.004410 0 0.549391 -en 0.004502 0 0.549190 -en
9 0 3.827356 -en 0.008671 0 3.789089 -en 0.008904 0 3.782912 -en 0.009140 0 3.771942 -en 0.009380 0 3.733972 -en 0.0
9 0 0.171531 -en 0.002333 0 0.171340 -en 0.002378 0 0.171091 -en 0.002425 0 0.171054 -en 0.002473 0 0.170923 -en 0.0
904 0 0.159987 -en 0.009052 0 0.132045 -en 0.010355 0 0.129493 -en 0.010895 0 0.116267 -en 0.011449 0 0.107984 -en
3 0 0.389248 -en 0.010930 0 0.379522 -en 0.011268 0 0.370162 -en 0.011619 0 0.365981 -en 0.011988 0 0.365069 -en 0.0
1143 0 1.029808 -en 0.001180 0 1.030334 -en 0.001218 0 1.030881 -en 0.001256 0 1.032075 -en 0.001295 0 1.032478 -en