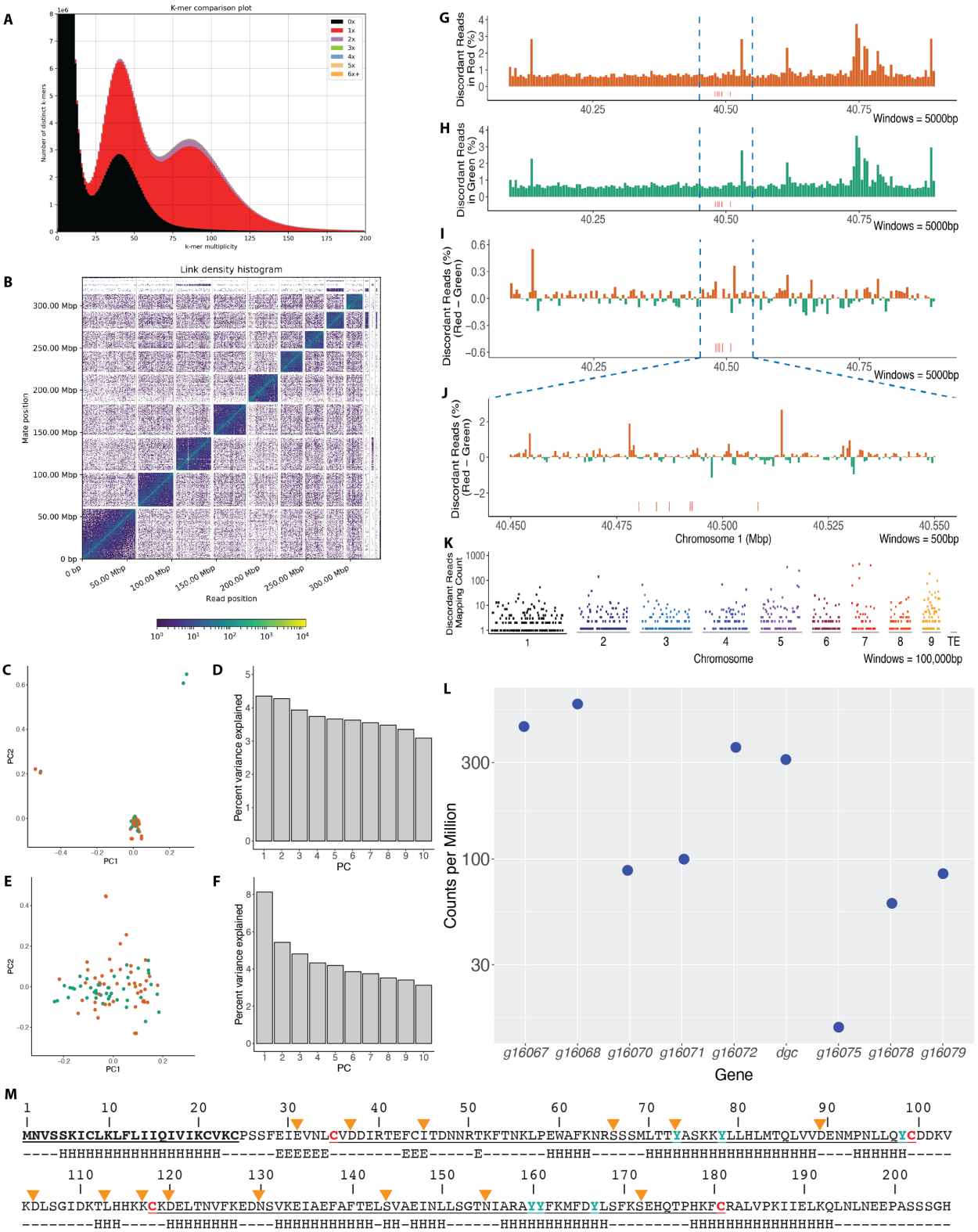



**Figure S1.** ***H. cornu*** **genome quality statistics and supporting genetic evidence that the 11**

**SNPs located near *dgc* are the only genetic variants strongly associated with gall color. Related to Figures 2 , 3 and 5.**

**(A)** The K-mer spectra comparing kmers from the 10X genome sequencing and the Dovetail assembled genome was calculated using KAT[S1]. There are two peaks for multiplicity 1X, as expected when sequencing a heterozygous individual, and the vast majority of sequence reads map uniquely to the assembled genome 1X, revealing few duplications.
**(B)** The HiC contact map reveals that the great majority of read pairs map to approximately the same location within each of the nine major scaffolds. The X and Y axes indicate the mapping position for the first and second read of each read pair and the color indicates the number of read pairs mapping to each bin. Only scaffolds greater than 1 Mbp are shown.
**(C-F)** The first two components of a principal components analysis (PCA) of genome-wide variation **(C)** and of variation in an 800 kbp window centered on *dgc* **(E)** reveal no large scale association of genetic variation with gall color. The top ten principal components of the genome-wide **(D)** and targeted **(F)** PCA each explain less than 5% and 9%, respectively, of the genetic variance.
**(G-K)** Distribution of mapped, unpaired Illumina sequencing reads from the ~800 kb re-sequenced region centered on *dgc* provides no evidence for common large-scale genomic aberrations associated with *dgc* in aphids from either green or red galls. Discordant reads from fundatrices making red **(G)** and green **(H)** galls from the entire ~800kbp re-sequenced region are rare and are not specifically associated with gall color **(I)**. Close examination of the genomic region harboring SNPs strongly associated with gall color (red vertical lines) shows that discordant reads are rare and not linked to the associated SNPs **(J)**. Discordant reads do not map preferentially anywhere in the genome **(K)**.
**(L)** Expression levels of genes in the *bicycle* gene paralog group that includes *dgc* from a single individual that inhabited a red gall and carried a $dgc^{Green}/dgc^{Green}$ genotype at all 11 SNPs. *Dgc* is expressed at similar levels as *g16072*, similar to the pattern observed for individuals that make green galls (Figure 3C). Thus, it is likely that this individual, and possibly the remaining rare fundatrices that induce red galls with $dgc^{Green}/dgc^{Green}$ genotypes, do not do so through reduction in *dgc* expression, but instead carry genetic variation at one or more other genes that induce red galls.
**(M)** Predicted amino-acid sequence of *dgc* with several secondary structure predictions, including *SignalP* 5.0 prediction[S2] of an N-terminal secretion signal from positions 1-24 (bold, underlined) and *jpred4* secondary structure prediction[S3] below primary sequence (H = helical, E = extended). Cysteines and tyrosines are colored red and blue, respectively, and the two CYC motifs are underlined. Orange triangles above protein sequence indicate intron-exon boundaries.

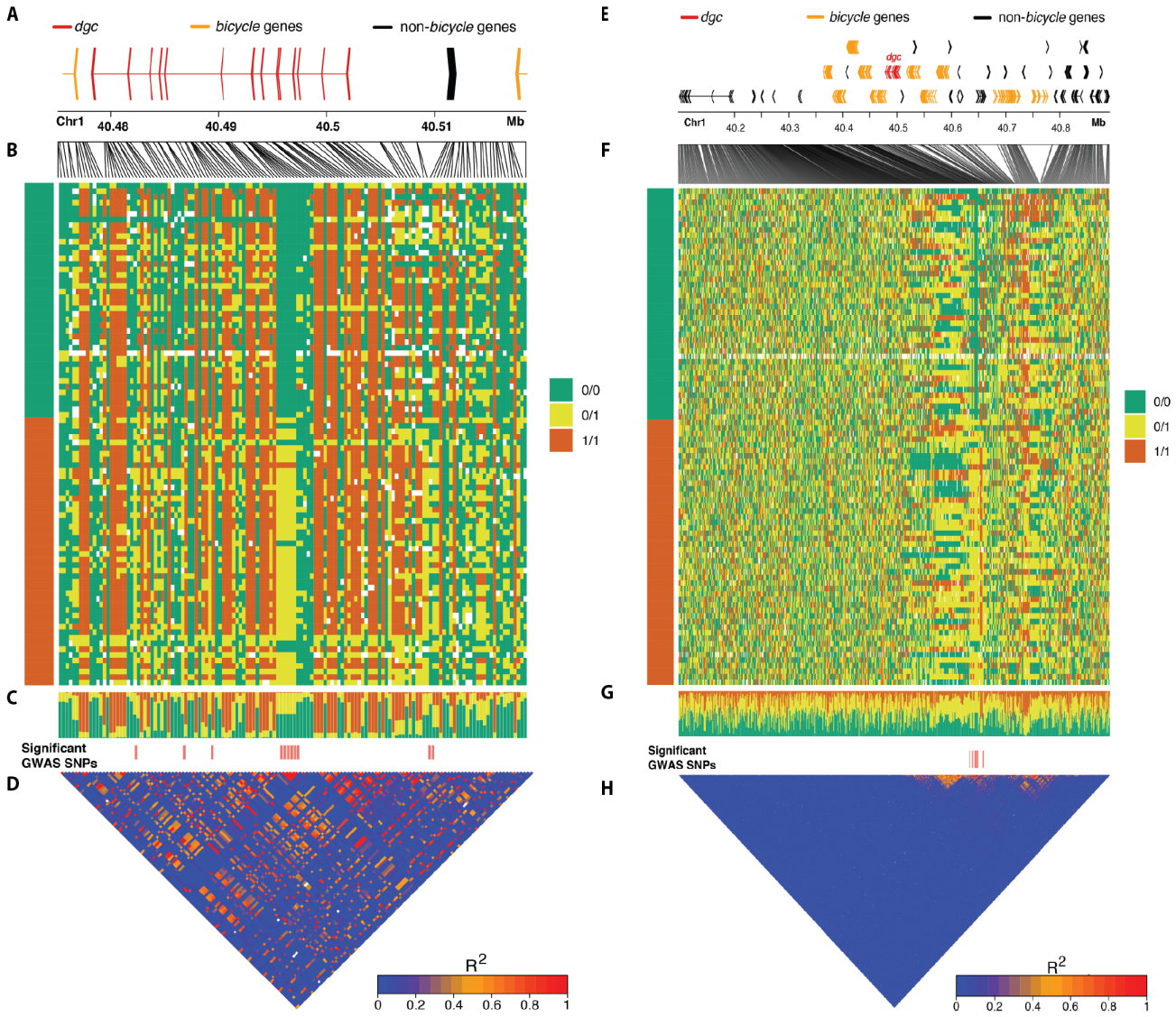


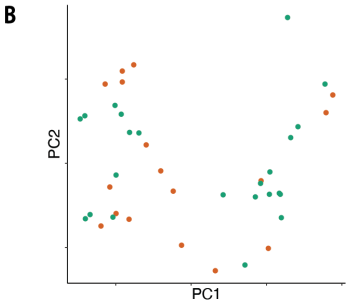

**Figure S2. Genotypes of SNPs in the *dgc* gene region. Related to Figure 2.**
**(A)** Gene models for the region ~40.475-40.52 of Chromosome 1. *Dgc* is labelled in red and other *bicycle* genes are labelled in yellow.
**(B)** Genotypes for each of the 90 individuals (42 from green galls, 48 from red galls; 4 low coverage individuals were excluded) from the original GWAS inferred from approximately 60X sequencing coverage for this region. Each row represents the genotypes at each of 138 SNPs. Genotypes are color coded as homozygous reference allele (green), heterozygous (yellow), or homozygous non-reference allele (orange).
**(C)** Frequencies of each genotype at each variable position in plot **(B)**.
**(D)** Linkage disequilibrium in *dgc* gene region. Strong linkage disequilibrium is observed between the 11 SNPs linked to gall color (labelled as "Significant GWAS SNPs" above plot), but not between these SNPs and other variable sites.

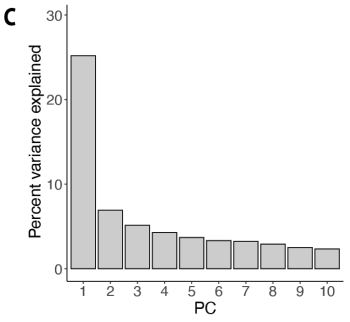
**(E)** Gene models for the region ~40.1-40.9 of Chromosome 1. *Dgc* is labelled in red, other *bicycle* genes are labelled in yellow.
**(F)** Genotypes for each of the 90 individuals (42 from green galls, 48 from red galls; 4 low coverage individuals were excluded) from the original GWAS inferred from approximately 60X sequencing coverage for this region. Each row represents the genotypes at each of 698 SNPs. Genotypes are color coded as homozygous reference allele (green), heterozygous (yellow), or homozygous non-reference allele (orange). Data were thinned to exclude any SNPs within 500bp of each other.
**(G)** Frequencies of each genotype at each variable position in plot **(F)**.
**(H)** Strong linkage disequilibrium is observed between the 11 SNPs linked to gall color (labelled as “Significant GWAS SNPs” in plot above), but not between these SNPs and other variable sites.

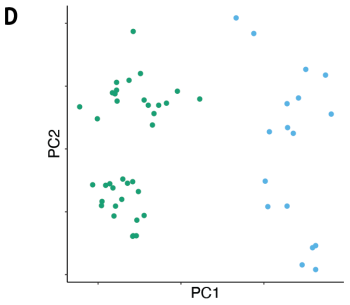
A

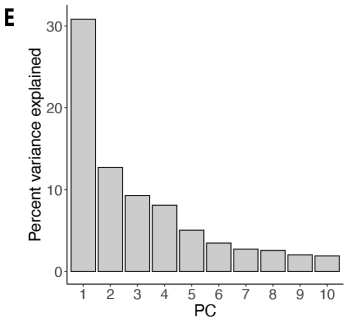
| | | chr1:40,496,912 | | Binomial |
|---|---|---|---|---|
| Sample | Gall color | A | T | Test P |
| DSG_7 | Green | 212 | 206 | 0.81 |
| DSG_8 | Green | 206 | 135 | 0.0001 |
| DSG_9 | Green | 497 | 386 | 0.0002 |
| DSG_10 | Green | 644 | 403 | 9.357e-14 |
| DSG_11 | Green | 291 | 339 | 0.0611 |
| DSG_13 | Green | 572 | 448 | 0.0001 |
| DSR_11 | Red | 4 | 11 | 0.1185 |

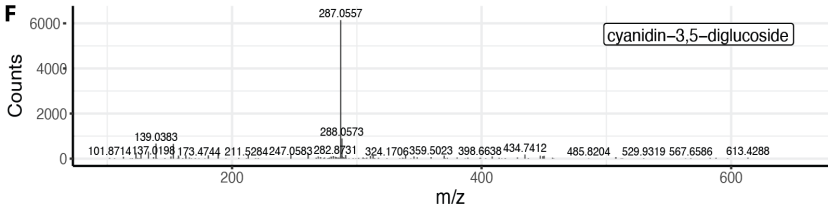
| | | chr1:40,501,949 | | Binomial |
|---|---|---|---|---|
| Sample | Gall color | A | T | Test P |
| DSG_7 | Green | 306 | 296 | 0.7138 |
| DSG_13 | Green | 697 | 633 | 0.0840 |
| DSR_11 | Red | 48 | 13 | 7.665e-06 |

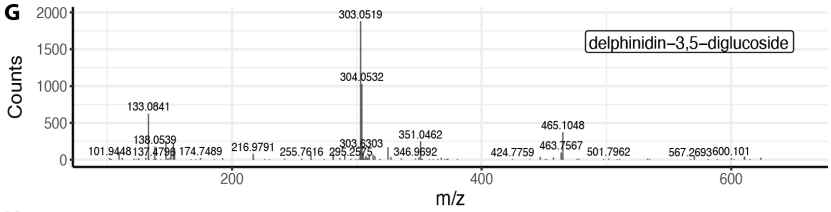


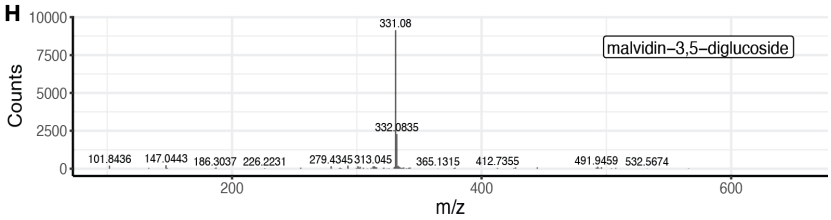



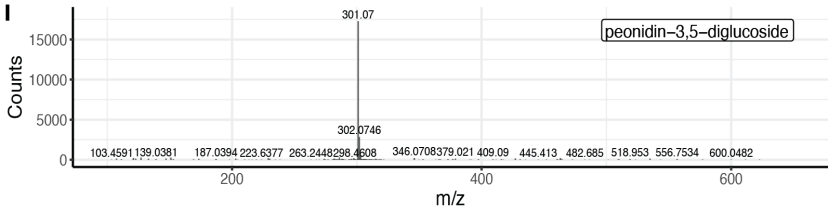



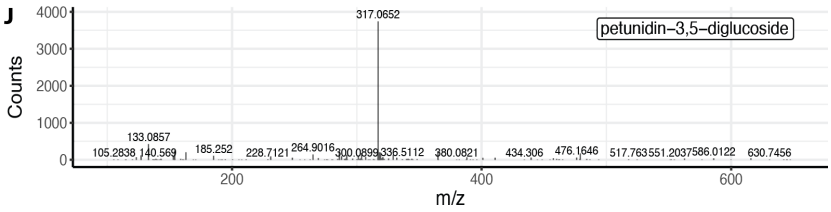



**Figure S3. Expression of *dgc* transcripts, genome-wide gene expression analysis of *H. virginiana* galls versus leaf, and MS/MS spectra of anthocyanins detected in red galls. Related to Figures 3 and 4.**

**(A)** Number of fundatrix salivary gland RNA-seq reads containing alternative *dgc* exonic SNPs at two genomic locations for aphids collected from green (blue font) and red (red font) galls. Four out of eight green and one out of two red samples deviate from the expectation of equal expression from both alleles in an individual. There is no evidence for allelic imbalance between fundatrices making green and red galls. In addition, both alternative alleles in red samples are expressed at substantially lower levels than the same alleles in green samples. Thus, a single copy of the SNPs that repress *dgc* expression (Figure 3C) appears to repress *dgc* from both alleles in heterozygous individuals.

**(B-E)** Principal components analysis **(B,D)** and percent variance explained of first 10 principal components **(C,E)** for the RNA-seq data of red (red dots) and green (green dots) galls **(B,C)** and galls (green dots) and leaves (blue dots) **(D,E)**.

**(F-J)** MS/MS spectra of anthocyanins detected in the red gall extract, which demonstrate the presence of the aglycone product ions previously reported for cyanidin-3,5-diglucoside (G), delphinidin-3,5-diglucoside (H), malvidin-3,5-diglucoside (I), peonidin-3,5-diglucoside (J), and petunidin-3,5-diglucoside (K).

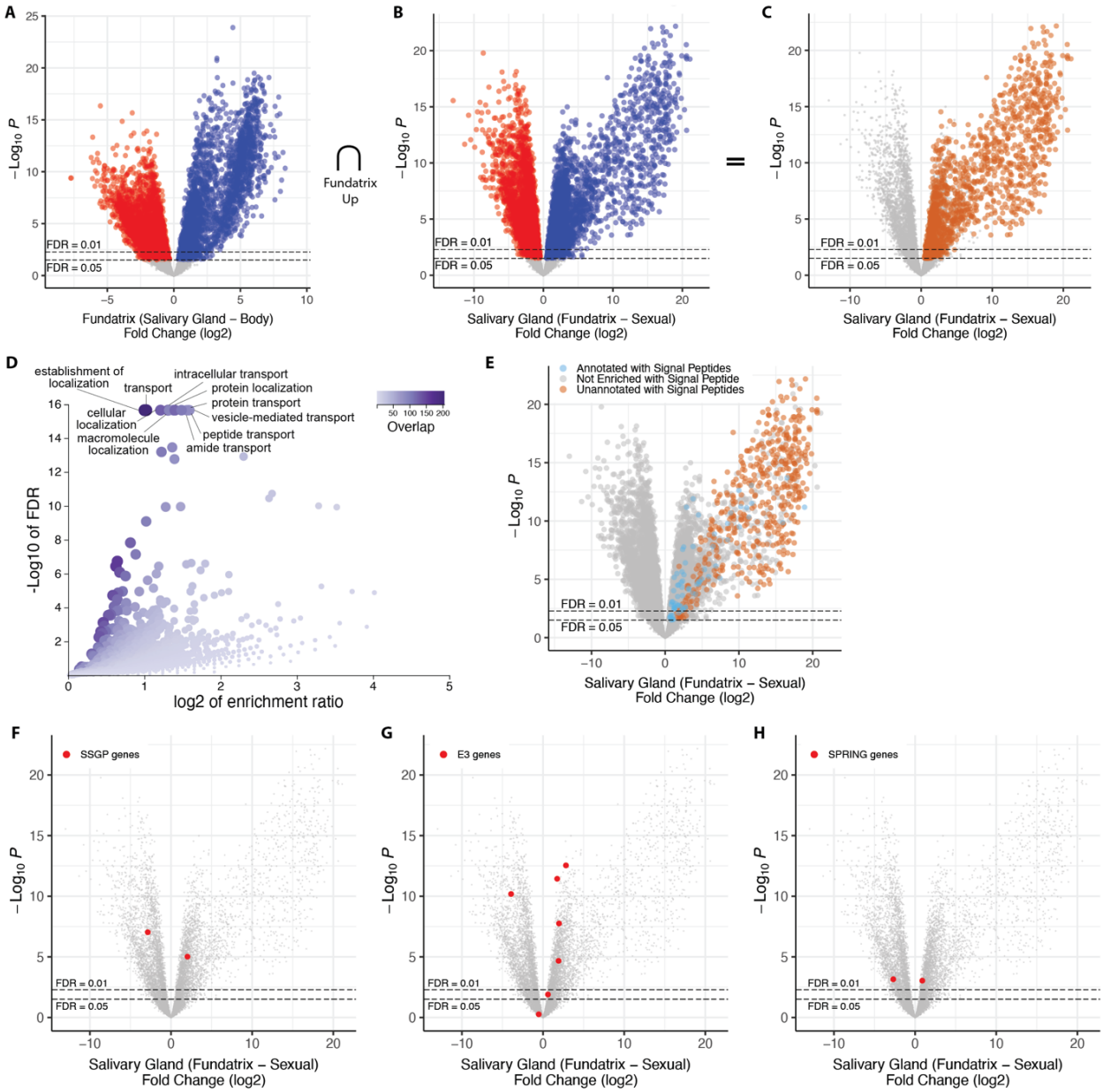



**Figure S4. Further analysis of *H. cornu* fundatrix salivary gland enriched genes and *bicycle* genes. Related to Figure 5 and STAR Methods.**

**(A-C)** Differential expression comparing gene expression in fundatrix salivary glands versus fundatrix whole bodies **(A)** and salivary glands of fundatrices versus sexuals **(B)** shown as volcano plots. Both comparisons reveal that a large number of genes are strongly over-expressed in fundatrix salivary glands. Intersection of the genes over-expressed in these two comparisons yielded the focal collection of genes over-expressed specifically in fundatrix salivary glands **(C)** (also see Figure 5A). Genes differentially over- or under-expressed at FDR < 0.05 are shown in **(A)** and **(B)** in blue or red, respectively.

**(D)** Volcano plot of Gene Ontology over-representation analysis of "Annotated" genes from Figure 5A showing KEGG pathway terms for biological processes that were over-represented

with a raw P value < 0.05. No GO terms were significantly under-represented with a raw P value < 0.05.
**(E)** Volcano plot of differential expression results shown in panel **(B)** with "Unannotated" and "Annotated" genes with signal peptides over-expressed in fundatrix salivary glands labeled in red and blue, respectively.
**(F-H)** Volcano plots of differential expression of previously proposed gall effector genes in fundatrix versus sexual salivary glands for SSGP genes **(F)**, ubiquitin E3-ligase genes **(G)**, and secreted RING-domain genes **(H)** indicated as red dots.

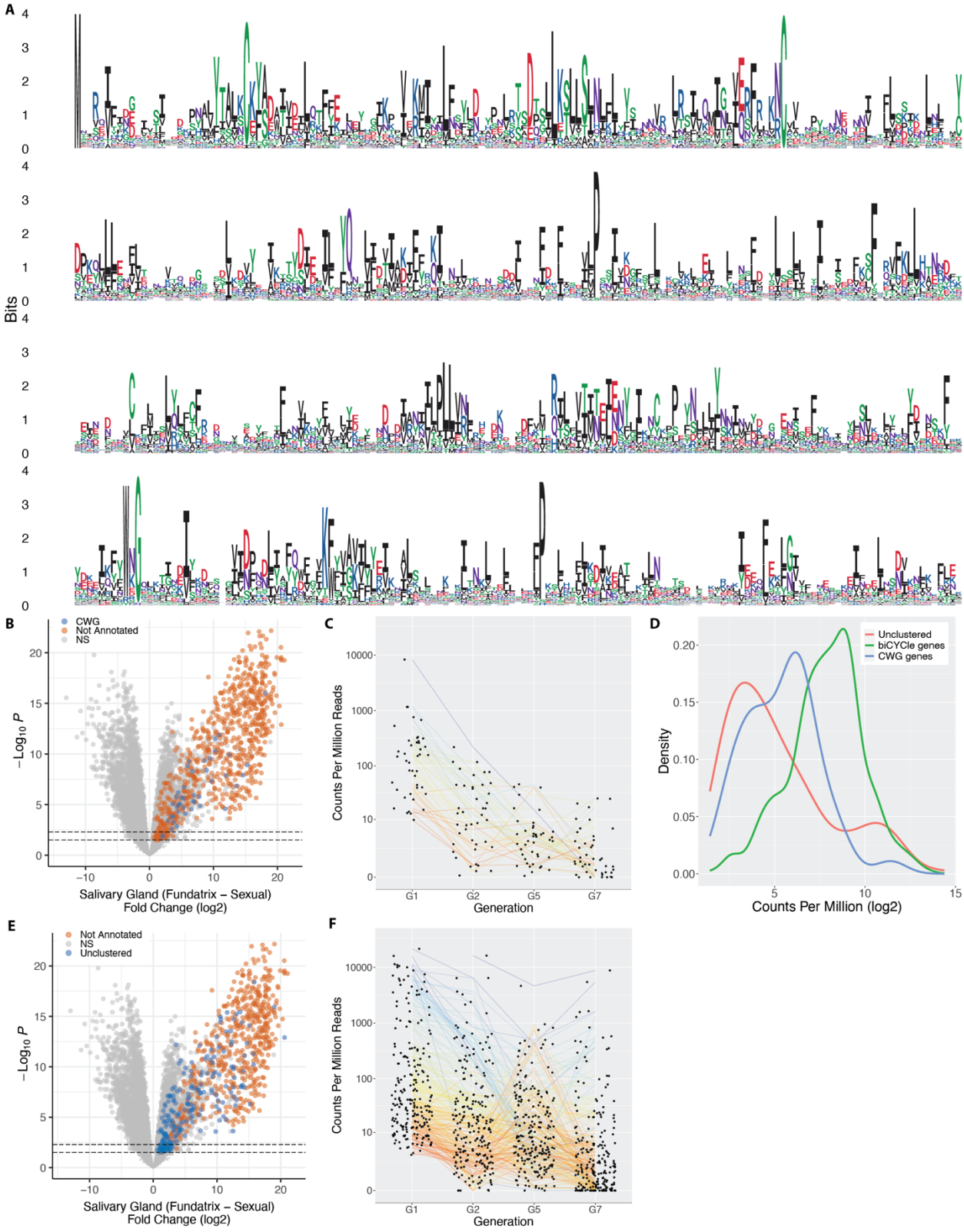

A
Bits
B
CWG
Not Annotated
NS
−Log10 P
Salivary Gland (Fundatrix – Sexual)
Fold Change (log2)
C
Counts Per Million Reads
Generation
G1
G2
G5
G7
D
Unclustered
biCYCle genes
CWG genes
Density
Counts Per Million (log2)
E
Not Annotated
NS
Unclustered
−Log10 P
Salivary Gland (Fundatrix – Sexual)
Fold Change (log2)
F
Counts Per Million Reads
Generation
G1
G2
G5
G7

**Figure S5. Further analysis of unannotated fundatrix salivary gland enriched genes. Related to Figure 5.**

**(A)** Logo plot of the aligned predicted protein sequences from CWG genes (also see Figure 5B). Sequence alignment used for logo was filtered to highlight conserved positions.

**(B, E)** Volcano plots of differential expression of fundatrix versus sexual salivary glands with *CWG* genes **(B)** and Unclustered genes **(E)** labelled in purple for clusters identified through hierarchical clustering (Figure 5B).

**(C, F)** Expression levels of *CWG* **(C)** and Unclustered **(F)** genes in the salivary glands of individuals across four generations of the *H. cornu* life cycle. Lines connect the same gene across generations and are color coded by relative expression levels in G1, with blue to red representing most to least strongly expressed, respectively.

**(D)** Distribution of expression levels of *bicycle* genes (green), *CWG* genes (purple), and remaining unannotated genes (red) in fundatrix salivary glands. The *bicycle* genes are, on average, the most strongly differentially expressed category of unannotated genes.

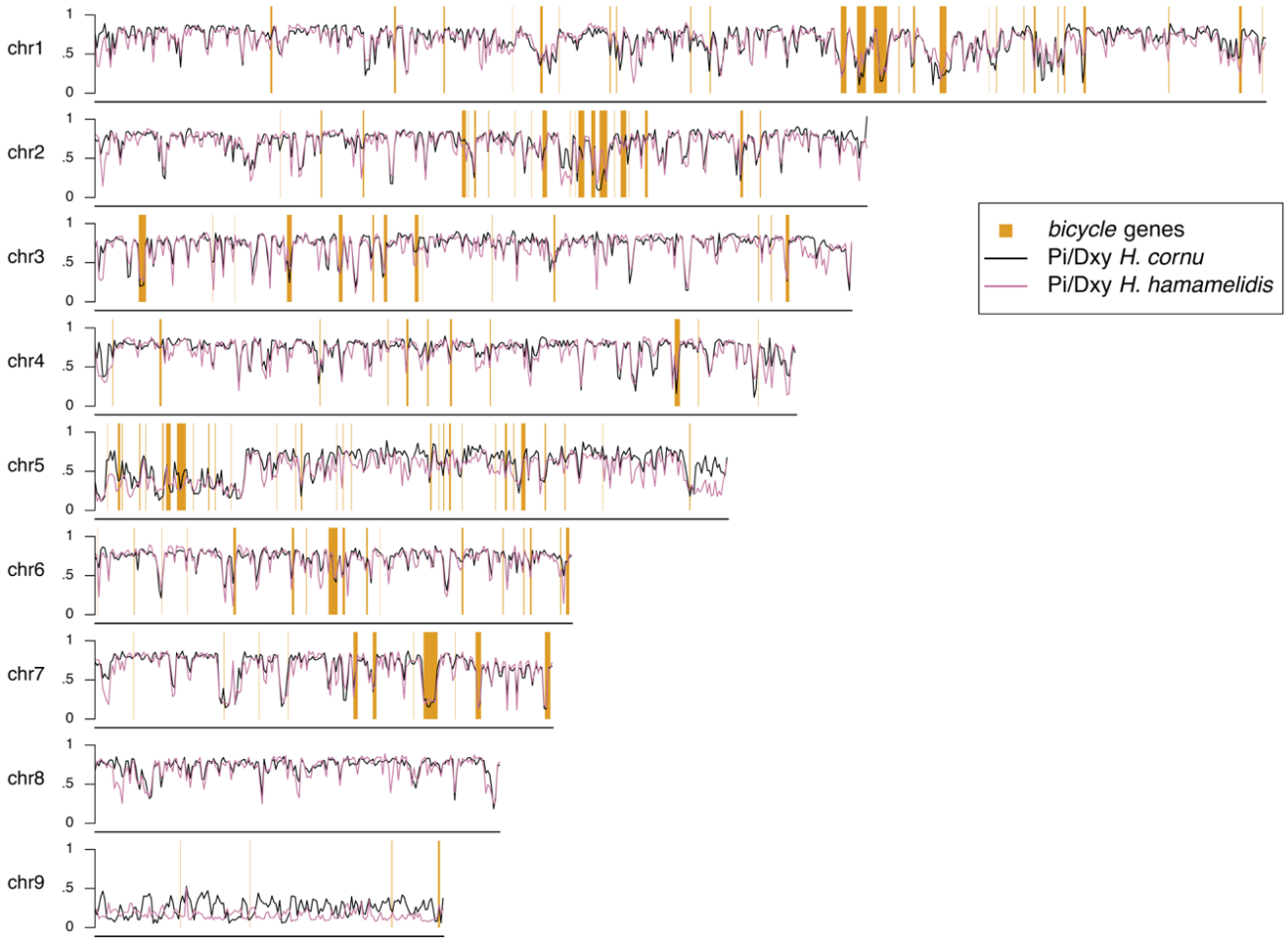



**Figure S6. Strongly reduced polymorphism to divergence levels are found in *bicycle* gene regions. Related to Figure 7.**
Genome-wide non-overlapping window moving averages of polymorphism (Pi) for *H. cornu* (black lines) and *H. hamamelidis* (purple lines) divided by divergence (Dxy) between the species. The locations of *bicycle* genes are indicated with orange rectangles.

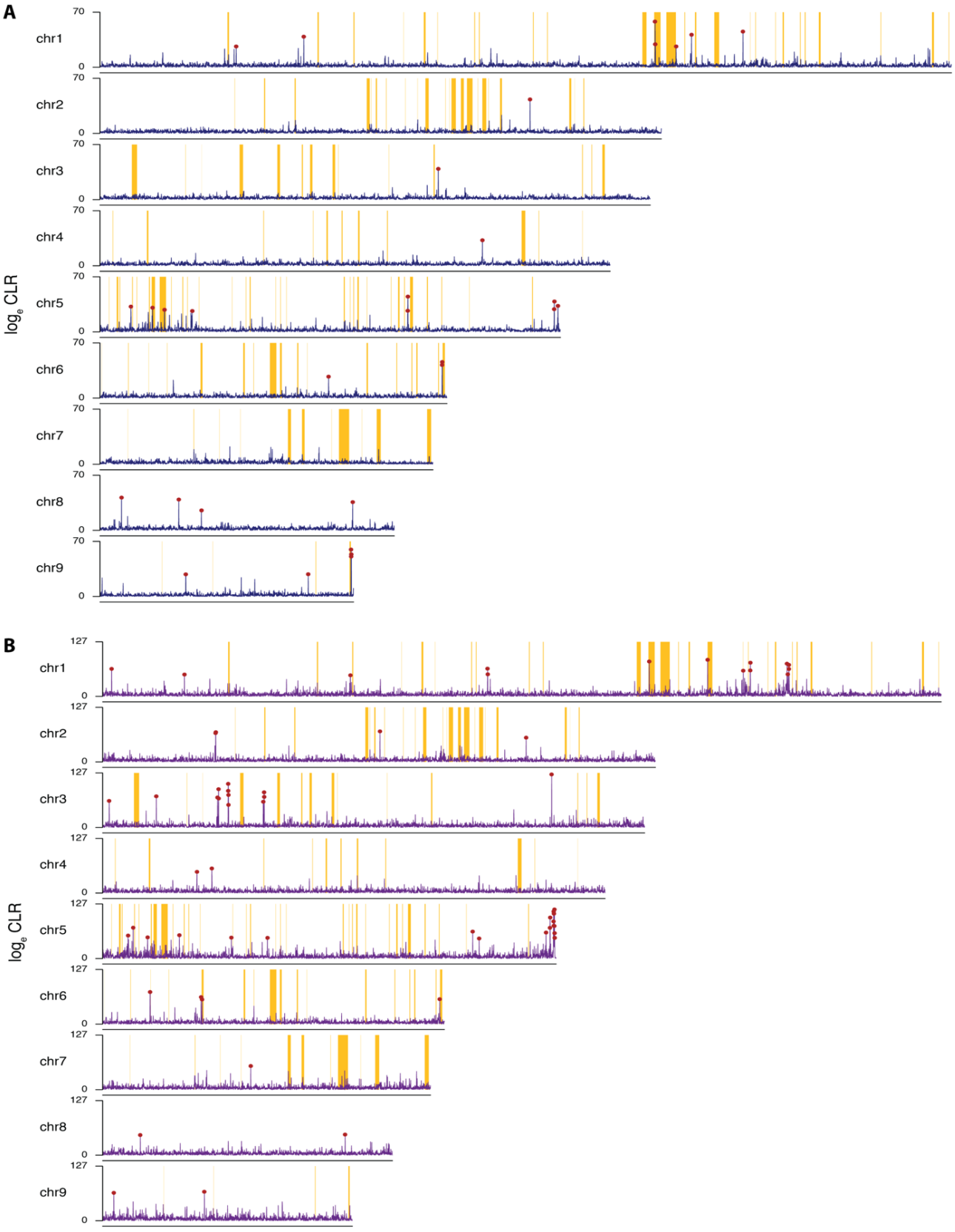

A
log$_e$ CLR
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
70
0
B
log$_e$ CLR
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
127
0

**Figure S7. Genome-wide distribution of signatures of recent selective sweeps detected by SweeD in *H. cornu* and *H. hamamelidis*. Related to Figure 7.**

Genome-wide distribution of the $\log_e$ of the composite likelihood ratio (CLR) of evidence for a recent selective sweep[S4] is shown in purple for *H. cornu* **(A)** and *H. hamamelidis* **(B)**. The location of significant signals ($P < 0.01$) calculated after simulations assuming neutrality (STAR Methods) is indicated with red circles. The locations of *bicycle* genes are indicated with orange rectangles.

| | $d_N/d_S$ | | |
|---|---|---|---|
| | <1 | >1 | Proportion >1 |
| non-*bicycle* | 5716 | 229 | 0.039 |
| *bicycle* | 10 | 58 | 0.853 |
| Proportion *bicycle* | 0.002 | 0.202 | |

**Table S1. $d_N/d_S$ test for significant positive selection on genes diverged between *H. cornu* and *H. hamamelidis*. Only genes estimated to have experienced significant positive or negative selection are included. Related to Figure 7.**

| | CMH alpha | CMH alpha CI | non-synonymous substitutions ± SD | synonymous substitutions ± SD | Mean # CDS |
|---|---|---|---|---|---|
| *bicycle* | 0.33 | 0.242-0.408 | 8.0 ± 6.3 | 1.7 ± 1.5 | 683 |
| *bicycle* dnds > 1 | 0.45 | 0.362-0.533 | 9.8 ± 6.7 | 1.3 ± 1.2 | 689 |
| *bicycle* dnds significantly > 1 | 0.62 | 0.450-0.733 | 15.9 ± 8.8 | 1.1 ± 1.3 | 692 |
| *CWG* | 0.38 | 0.246-0.498 | 15.4 ± 7.8 | 6.3 ± 3.4 | 2337 |
| unclustered | 0.27 | 0.110-0.410 | 3.4 ± 4.0 | 1.4 ± 2.0 | 710 |
| Annotated with signal peptide | 0.42 | 0.243-0.557 | 3.8 ± 4.8 | 2.5 ± 2.2 | 1414 |
| No annotation with signal peptide | 0.34 | 0.262-0.404 | 8.7 ± 6.9 | 2.2 ± 2.3 | 848 |
| Annotated without signal peptide | 0.41 | 0.349-0.461 | 3.1 ± 4.8 | 3.2 ± 4.1 | 1750 |
| No annotation without signal peptide | 0.33 | 0.189-0.439 | 4.5 ± 4.8 | 1.4 ± 1.8 | 710 |

**Table S2. McDonald-Kreitman *alpha* and gene statistics for different classes of genes overexpressed in *H. cornu* fundatrix salivary glands. Related to Figures 5 and 7.**

**Supplemental References**


S1. Mapleson, D., Garcia Accinelli, G., Kettleborough, G., Wright, J., and Clavijo, B.J. (2017). KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. Bioinformatics *33*, 574–576.

S2. Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. *37*, 420–423.

S3. Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: A protein secondary structure prediction server. Nucleic Acids Res. *43*, W389–W394.

S4. Pavlidis, P., Žˇivkovic, D., Stamatakis, A., and Alachiotis, N. (2013). SweeD: Likelihood-Based Detection of Selective Sweeps in Thousands of Genomes. Mol. Biol. Evol. *30*, 2224–2234.

S5. Cembrowski, M.S., Phillips, M.G., DiLisio, S.F., Shields, B.C., Winnubst, J., Chandrashekar, J., Bas, E., and Spruston, N. (2018). Dissociable Structural and Functional Hippocampal Outputs via Distinct Subiculum Cell Classes. Cell *173*, 1280–1292.e18.