Transcriptional expression changes during compensatory plasticity in the prothoracic ganglion of the adult cricket *Gryllus bimaculatus*

Felicia Wang, Harrison Fisher, Maeve Morse, Lisa L. Ledwidge, Jack O'Brien, Sarah E. Kingston^, Justin Beckman, Jasmine J. Johnson, Lyn S. Miranda Portillo, Tabarak N. Al Musawi, Alexandra W. Rubenstein, David A. Michaelson, Hadley Wilson Horch*,

Department of Biology, Bowdoin College, 6500 College Station, Brunswick, Maine 04672 USA.

^School of Marine Sciences and Darling Marine Center, University of Maine, 193 Clarks Cove Rd., Walpole, ME 04573 USA;

*Correspondence to:
Dr. Hadley Wilson Horch, Department of Biology, Bowdoin College 6500 College Station, Brunswick ME 04011 USA. Phone: 207-798-4128; FAX: 207-725-3405; Email: hhorch@bowdoin.edu

Keywords:

RNASeq, Transcriptome, Dendritic Plasticity, GO term analysis, Differential expression, Injury

Abstract

Most adult organisms are limited in their capacity to recover from neurological damage. The auditory system of the Mediterranean field cricket, *Gryllus bimaculatus*, presents a compelling model for investigating neuroplasticity due to its unusual capabilities for structural reorganization into adulthood. Specifically, the dendrites of the central auditory neurons of the prothoracic ganglion sprout in response to the loss of auditory afferents. Deafferented auditory dendrites grow across the midline, a boundary they normally respect, and form functional synapses with the contralateral auditory afferents, restoring tuning-curve specificity. The molecular pathways underlying these changes are entirely unknown. Here, we used a multiple k-mer approach to re-assemble a previously reported prothoracic ganglion transcriptome that included ganglia collected one, three, and seven days after unilateral deafferentation in adult, male animals. We used EdgeR and DESeq2 to perform differential expression analysis and we examined Gene Ontologies to further understand the potential molecular basis of this compensatory anatomical plasticity. Enriched GO terms included those related to protein translation and degradation, enzymatic activity, and Toll signaling. Extracellular space GO terms were also enriched and included the upregulation of several protein yellow family members one day after deafferentation. Investigation of these regulated GO terms help to provide a broader understanding of the types of pathways that might be involved in this compensatory growth and can be used to design hypotheses around identified molecular mechanisms that may be involved in this unique example of adult structural plasticity.

46 **Background**

47 Most adult organisms, especially mammals, are limited in their capacity to adapt and
48 recover from neurological damage (1,2). The Mediterranean field cricket, *Gryllus bimaculatus*,
49 provides a model of neuroplasticity due to its demonstrated ability to compensate for neuronal
50 damage with novel dendritic growth and synapse formation, even into adulthood. Specifically,
51 the central auditory system, much of which resides in the prothoracic ganglion, reorganizes in
52 response to deafferentation caused by unilateral transection of auditory afferents in the adult (3–
53 5).

54 In *G. bimaculatus*, auditory information is transduced by the auditory organs, located on
55 the prothoracic limbs. Auditory afferents receive the sensory stimuli and convey this information
56 into the prothoracic ganglion where they form synapses with several different auditory neurons
57 (6,7). These neurons exist as mirror image pairs and their dendritic arbors remain localized
58 ipsilateral to the auditory input, typically not projecting contralaterally across the midline (8).
59 However, previous research has shown that after amputation of the prothoracic leg in the adult,
60 which removes the auditory organ and severs the afferents, the deafferented dendrites of the
61 ipsilateral auditory neurons sprout across the midline and form functional synapses with the
62 intact auditory afferents on the contralateral side. This reorganization is evident whether
63 deafferentation occurs in larvae (9,10) or adults (3–5). Various aspects of the physiological
64 consequences of this compensatory behavior have been studied (3,9,10), however little is known
65 about the molecular pathways and mechanisms underlying this growth.

66 Although the genome has only just become publicly available (11), various *de novo*
67 transcriptomes have been created for use in this species (12–14). Recently, a *de novo*
68 transcriptome of the prothoracic ganglion was assembled in an attempt to understand the

69 molecular basis of the compensatory response (15). This transcriptome was built with RNA from
70 individual prothoracic ganglia of both control and deafferented adult male crickets. Initially, this
71 transcriptome was assembled and mined for the presence of developmental guidance molecules.
72 These guidance molecules are known to play a well-conserved role in regulating the specific
73 growth of axonal and dendritic projections during the development of many species (16,17).
74 While these molecules have mainly been studied for their role in development, it has also been
75 suggested that alterations in their expression may influence the ability of axons and dendrites to
76 recover from injury in adulthood (15,18–20). Mining this cricket transcriptome revealed that
77 many well-conserved developmental guidance molecules, including slit, ephrins, netrins, and
78 semaphorins, were present within the adult prothoracic transcriptome (15). However, it is still
79 unknown whether the expression of these transcripts, or any other transcripts, are significantly
80 altered during this compensatory growth process.

81 The goal of this study was to better understand the underlying molecular control of the
82 compensatory growth behavior observed in the cricket. We assembled a new, more
83 representative and less redundant transcriptome of the cricket prothoracic ganglion using
84 multiple k-mer values during the assembly process. We also utilized the Evidential Gene
85 *tr2aacdsmRNA* classifier to reduce redundancies (21). This new transcriptome was used to
86 analyze changes in expression levels one, three, and seven days post-deafferentation. The
87 identified genes were then analyzed using GO annotation analysis to determine the classes of
88 genes that are differentially regulated over the course of the injury response. By performing this
89 analysis, it was possible to discover changes in gene expression that occur during the
90 compensatory growth response, allowing for insights into possible pathways or key molecules
91 critical to this process.

Results and Discussion

Transcriptome Assembly and Analysis

This transcriptomic study focused on the cricket, *Gryllus bimaculatus*, whose nervous system has been shown to have an unusual level of adult structural plasticity (3–5). We deafferented sensory neurons, including the auditory neurons, in the prothoracic ganglion of the adult cricket, by unilateral amputation of the prothoracic leg at the femur. Control amputations consisted of removal of the distal tip of the tarsus. We harvested prothoracic ganglia one, three, and seven days post-amputation. These time points were designed to capture transcriptional changes in response to the injury (one day), during initial sprouting (one and three days), growth across the midline (three and seven days), and *de novo* synapse formation (3,22).

Although a *G. bimaculatus* prothoracic ganglion transcriptome from this tissue was previously assembled, analysed, and mined (15), the present study re-assembled a new transcriptome based on those original cleaned and trimmed RNA-Seq reads. Five individual *de novo* transcriptomes were built in Trinity using five different k-mer lengths (21, 25, 27, 30, and 32). Transcriptome construction with longer k-mer lengths produces more reliable contigs, though biased toward highly expressed transcripts. In a complementary fashion, using a shorter k-mer length produces a more exhaustive set of contigs though also one more prone to noise (23,24). This trade-off between bias and noise induced by the choice of k-mer length suggests how a higher quality *de novo* assembly can be derived by integrating multiple k-mer lengths into an analysis (23). By combining results across k-mer lengths, we ensured that a complementary selection of contigs was included in the analysis (23,25,26). Correspondingly, we combined the five assemblies into a single reference transcriptome and subsequently filtered redundancies and fragments.

The individual assemblies had an N50 ranging from 1,219 - 2,341, with the longer k-mer assemblies yielding a longer N50 (Table 1). The median, average, and maximum contig length also increased as the k-mer length was increased. The total number of Trinity "genes" ranged from 283,278 to 351,829, with higher k-mer assemblies yielding fewer predicted genes. The GC content for each assembly remained roughly constant, between 40-41%. The overall alignment was greater than 98%, with multimapping percentages between 90.04-93.33% (Table 1). This high multimapping percentage can likely be attributed to the Trinity assembly process, which is conservative in its separation and identification of unique transcripts, producing high intra-assembly redundancy (27).

124 **Table 1: Individual k-mer assembly details.**

	k-mer = 21	k-mer = 25	k-mer = 27	k-mer = 30	k-mer = 32
Total # bases assembled	293,992,611	404,116,670	408,831,054	406,965,539	403,174,726
Total # assembled contigs	405,638	438,593	431,712	415,901	407,158
Total # Trinity "genes"	351,829	302,633	297,584	288,135	283,278
Average contig length (bp)	724.77	921.39	946	978.52	990.22
Median contig length (bp)	376	397	397	399	400
Maximum contig length (bp)	37,575	44,287	44,328	44,352	44,331
N50 (bp)	1219	2000	2141	2272	2341
GC count for assembly (%)	40.94	40.39	40.32	40.19	40.14
Overall alignment	98.61	98.58	98.66	98.71	98.74
Reads mapped 1 time (%)	4.4	3.69	2.71	2.08	2.55
Reads mapped >1 time (%)	90.04	90.95	92.21	93.33	92.9

125 Table 1. Summary metrics for five different de novo transcriptomes built with five different k-mer
126 lengths.
127

128 The five transcriptomes were combined to generate a transcriptome with a total of 2,099,002
129 contigs (Figure 1), presumably many of which were redundant. We used the EvidentialGene
130 *tr2aacdsmRNA* classifier to filter the redundancies within our transcriptome, which were present
131 due to both intra- and inter-assembly redundancies (21). The EvidentialGene program employs
132 an algorithm that operates on the open reading frames of the contigs to generate a non-redundant
133 transcriptome containing the optimal set of transcripts based on biological relevance and coding
134 potential (21,28). This program is often used in 'over-assembly' procedures where multiple
135 assemblies are combined (29–31). With our multi-k-mer assembly, EvidentialGene produced a

136 main ‘okay’ set, containing 55,895 contigs, and an alternative ‘okalt’ set, containing 143,364
137 contigs, which were combined to produce a final transcriptome with a total of 199,357 contigs,
138 reducing the overall number of contigs by 90.5%. BLAST searches across all the contigs yielded
139 matches for 127,324 transcripts, a 63.87% BLAST hit rate for the entire transcriptome. The
140 number of Trinity predicted genes after running EvidentialGene dropped slightly to 132,972. The
141 multimapping percentage was reduced from approximately 90% to around 21%.

142 To check the accuracy of the sequences predicted in the transcriptome, we used Sanger
143 sequencing to independently confirm the sequences of six randomly selected transcriptome
144 transcripts. Of these six, four of them were predicted to contain complete open reading frames
145 (ORFs), and two were missing the 3’ end. We analyzed 14,299 nucleotides of 15,478 predicted
146 base pairs (92%). The number of substitutions (16), insertions (84), and deletions (0) were noted;
147 overall, these differences accounted for approximately ~0.1% of the sequenced nucleotides (data
148 not shown). A few additional randomly selected sequences were highly repetitive and were not
149 amenable to Sanger sequencing; we did not proceed with an analysis of any of these candidates.

150 **Differential expression during compensatory plasticity**

151 To determine genes that were differentially regulated during compensatory plasticity, the
152 reads for each of the 16 Illumina libraries, which excluded the two outliers and three backfill
153 libraries (see Methods), were mapped back to our multiple k-mer transcriptome creating a counts
154 matrix. Pairwise comparisons of normalized counts data from deafferented *vs*. control crickets
155 were performed at each time point using both algorithms, EdgeR and DESeq2 (See Supplemental
156 Materials). The distribution of differentially expressed genes was initially visualized using
157 volcano plots (Figure 2). These plots revealed slightly different distributions of upregulated
158 versus downregulated genes between the two programs. Overall, however, we saw strong

159 correlations between these two programs for all time points (Figure 3), with the exception of a
160 few of the high fold-change candidates. Within this range, EdgeR was consistently more
161 conservative than DESeq2, which was especially true for a small number of upregulated
162 candidates (Figure 3a-c).

163 The majority of the transcripts were upregulated in the 2 to 4-fold range at one day (59%
164 of the transcripts), three days (55% of the transcripts), and seven days (45% of the transcripts).
165 The next largest group of transcripts was upregulated at 0 to 2-fold at one day (33% of
166 transcripts), three days (41% of transcripts), and seven days (39% of transcripts). For those
167 candidates that were downregulated, a majority of them at one day (63%) and three days (86%)
168 were downregulated less than 2-fold. At seven days, the bulk of candidates (70%) were
169 downregulated 2 to 4-fold. Analysis of 10 of the transcripts at each time point with the largest
170 fold changes revealed that most were unidentified and did not match anything in the NCBI
171 database when BLASTed. A few of these transcripts did have BLAST hits, such as a mucin-
172 5AC-like (down at one day), larval cuticle protein-3-like (down at seven days), and hypothetical
173 accessory gland protein (up at three and seven days).

174 Using EdgeR, 261 genes were found to be downregulated at one day post-deafferentation,
175 1,675 genes were downregulated at three days post-deafferentation, and 580 genes were
176 downregulated at seven days post-deafferentation (Figure 4a). Additionally, 2,234 genes were
177 determined to be upregulated at one day post-deafferentation, 1,860 genes upregulated at three
178 days post-deafferentation, and 290 genes upregulated at seven days post-deafferentation (Figure
179 4b).

180 A similar pairwise comparison of deafferented versus control crickets was performed
181 using the DESeq2 software and revealed that 985 genes were downregulated at one day post-

182 deafferentation, 3,049 genes were downregulated at three days, and 448 genes were
183 downregulated at seven days (Figure 4c). Additionally, 3,589 genes were upregulated at one day
184 post-deafferentation, 1,424 genes were upregulated at three days, and 535 genes were
185 upregulated at seven days (Figure 4d).

186 From these sets, simple comparisons were created to determine the number of genes
187 upregulated and downregulated across multiple timepoints. In EdgeR, there were four genes
188 downregulated at one and three days, two genes at one and seven days, two genes at three and
189 seven days, and 0 genes differentially downregulated across all three time points (Figure 4a). For
190 the upregulated genes, there were 174 genes differentially regulated at one and three days, 18
191 genes at one and seven days, 18 genes at three and seven days, and 32 genes at all three time
192 points (Figure 4b). Comparing the DESeq2 sets of genes across multiple timepoints showed that
193 there were nine genes downregulated at one and three days, one gene downregulated at one and
194 seven days, four genes downregulated at three and seven days, and 0 genes downregulated at all
195 three timepoints (Figure 4c). Additionally, 168 genes were found to be upregulated at one and
196 three days, 73 genes at one and seven days, 23 genes at three and seven days, and 40 genes at all
197 three time points (Figure 4d).

198 Finally, simple comparisons were performed between EdgeR and DESeq2 with genes
199 determined to be differentially expressed at each of the three times points. For downregulated
200 genes, 180 were identified at one-day post deafferentation, 1,604 at three days, and 367 at seven
201 days. For upregulated genes, 2,099 were identified at one-day post deafferentation, 1,043 at three
202 days, and 272 at seven days (Figure 5). The genes found to be differentially regulated by both
203 programs were used for further analysis.

204 DESeq2 and EdgeR are the leading programs used for the analysis of RNA-Seq data,
205 with thousands of reports both using these methods for analyzing differential expression and
206 comparing their computational methods (32,33). While both operate under the hypothesis that
207 the majority of genes are not differentially expressed, they employ different computational
208 methods, especially with respect to the normalization process, to determine differentially
209 expressed genes (33). EdgeR and DESeq2 both use a normalization by distribution method, but
210 EdgeR relies on the Trimmed Mean of the M-values method, whereas DESeq2 uses a Relative
211 Log Expression method (34–37). Since different methods rely on differing assumptions in order
212 to identify differentially expressed genes, the results will vary slightly. One experiment
213 comparing EdgeR and DESeq2 found relatively similar lists of differentially expressed genes
214 produced by the two programs, with EdgeR producing more conservative, smaller gene lists (32).
215 In this study we decided to use two different programs to conduct the differential expression
216 analysis in order to create a smaller, more conservative set of genes for future functional
217 analyses. Out of the six comparisons between EdgeR and DESeq2 (upregulated and
218 downregulated at one, three, and seven days post injury), four out of the six resulted in EdgeR
219 producing a smaller set of genes than DESeq2 (Figure 5), in line with the study by Raplee and
220 colleagues (32). Although the two programs generated varying numbers of differentially
221 regulated genes, similar patterns in relative numbers were observed. Both programs showed a
222 decrease in the number of genes upregulated across time. For the downregulated genes, a peak in
223 the number of differentially regulated genes was found at three days post injury. This similarity
224 was expected given the relative similarity and previous studies of both analysis programs.

225 **BLAST and Gene Ontology Annotations**

226 Once we had identified a conservative set of transcripts predicted to be differentially
227 regulated, we used BLAST2GO (38) to try to identify them. Not all the transcripts inputted into
228 the BLAST2GO program resulted in BLAST hits and/or GO annotations. At one day
229 downregulated, 71% of genes had both BLAST and GO results and an additional 10% had only
230 BLAST hits. At three days downregulated, 36% of genes had both BLAST and GO results and
231 an additional 6% had only BLAST hits. At seven days downregulated, 31% of genes had both
232 BLAST and GO results and an additional 6% had only BLAST hits. For the one day upregulated,
233 53% of genes had both BLAST and GO results and an additional 10% had only BLAST hits. At
234 three days upregulated, 59% of genes had both BLAST and GO results and an additional 15%
235 had only BLAST hits. At seven days upregulated, 50% of genes had both BLAST and GO results
236 and an additional 15% had only BLAST hits (Figure 6).

237 For the six lists of differentially expressed transcripts, there was a range between 37-81%
238 of transcripts having BLAST hits against the nr database and 23-62% against the manually
239 curated and annotated Swiss-Prot database. After mapping with GO terms, this was reduced to
240 about 31-59%. This left close to half of the differentially regulated transcripts with no functional
241 information. These transcripts could represent uncharacterized proteins, which may or may not
242 be playing an important role in the compensatory growth response. Since we performed a
243 BLASTx looking at proteins, it is also possible that these transcripts are non-coding RNAs.
244 Although polyA selection was used as part of the RNA-Seq process, this may not be completely
245 efficient in removing all non-coding RNAs, specifically long non-coding RNAs (39,40). Finally,
246 it is also possible that there were issues within these transcripts themselves, either due to an error

247 during the assembly process or the sequences being too short to be matched with confidence.
248 Regardless, we completed no further analysis of these transcripts.

249 One set of intriguing proteins that were found to be upregulated at one day were members
250 of the Protein Yellow family. Our transcriptome predicted the upregulation of 15 different
251 Protein Yellow transcripts that appear to consist of at least eight different splice variants of
252 Protein Yellow-f (data not shown). We used qPCR to confirm the upregulation of these
253 transcripts in independent samples and found weak support for this upregulation (p=0.1; data not
254 shown). The primers used were designed to target all eight splice variants identified in our
255 transcriptome, but because all the splice variants of Protein Yellow-f in the cricket are not well-
256 characterized, we may have inadvertently captured additional splice variants in this analysis that
257 are not differentially expressed. Future experiments confirming the enrichment of each of these
258 splice variants along with functional validation will be necessary.

259 Protein Yellow belongs to the Major Royal Jelly protein family and are secreted proteins
260 found in the extracellular region (41). Protein Yellow was first characterized in *Drosophila*
261 *melanogaster* for its role in pigmentation (42). Other research in honeybees revealed the
262 importance of Royal Jelly proteins in development and social behavior in addition to a possible
263 role in the CNS (41,43). However, the function of these Yellow/Royal Jelly proteins is not
264 completely understood (42). While the role of these proteins in crickets is unclear, they were
265 statistically differentially regulated and would be an interesting molecular family to investigate
266 for their role in deafferentation-induced plasticity.

267 **GO Term Distributions**

268 Based on a preliminary grouping of GO terms by the three root classes, it appeared that
269 several classes of GO terms were found to be associated with our differentially expressed genes

270 (Figure 7). While the top five represented GO terms encompassed most of the GO terms in the
271 cellular component category, there was a much broader range of GO terms represented in the
272 molecular function and biological process categories

273 Web Gene Ontology Annotation Plot (44,45) was used to plot a broader distribution of
274 GO terms and visually compare annotations among timepoints (Figure 8). Cellular component,
275 molecular function, and biological process were displayed on traditional WEGO histograms
276 (Figure 8a, b). The percentage of genes indicates the percentage of the genes within a given list
277 that were annotated with the given GO term or one of the child nodes of the term. GO terms with
278 higher percentage representation included terms describing membrane-related components as
279 well as terms related to catalytic activity, binding, and metabolic and cellular processes.

280 **Gene Ontology Categories**

281 We examined whether any of the candidate GO terms we identified here were associated
282 with injury-related plasticity paradigms identified in other species. For example, perhaps
283 successful regeneration after injury depends on the recapitulation of developmental proteins that
284 promote neurogenesis (46) or guide axons and dendrites (18). If these molecular strategies were
285 important for the plasticity observed in the cricket, we would predict that we might see changes
286 in the expression of transcription factors involved in neurogenesis and/or the regulation of
287 several classes of guidance cues normally involved in midline control in insect embryos. When
288 searching our differentially regulated candidates, a few genes downregulated at three days were
289 associated with GO terms that were related to neurogenesis (**GO:0007465**: R7 cell fate
290 commitment and **GO:0045466**: R7 cell differentiation). We found only one transcript that was
291 annotated with an axon guidance-related GO term, which was identified as a "twitchin-like

292 protein," (Table 2). Twitchin/Unc-22 is a large protein kinase thought to be important in muscle
293 development and function (47).

294 An initial study of our original single k-mer transcriptome explored this developmental
295 recapitulation hypothesis by mining the transcriptome for the presence of various guidance
296 molecules (15). Though transcripts corresponding to the signaling families, slit, netrin, ephrin,
297 and semaphorin were identified within this adult transcriptome, the BLAST and Gene Ontology
298 analyses performed here did not identify an abundant number of guidance molecules as
299 differentially regulated. Despite this result, however, it is important to note that the transcriptome
300 and differential expression analysis were performed on the whole prothoracic ganglion, which
301 could mask important changes that occur in single cells after deafferentation, such as in
302 ascending neuron 1 and 2 (AN-1 and AN-2). Single cell RNA-Seq analysis of the ANs could
303 help to determine whether there are changes in expression occurring on a smaller anatomical
304 level.

305 Based on results from different types of injury model systems in other organisms,
306 additional functional categories that we hypothesized could change during the compensatory
307 growth response were those related to apoptosis (48,49) and Wnt signaling (50). In our
308 differentially expressed gene sets we did not find enrichment in terms related to apoptosis.
309 Searching our results for Wnt-related GO terms, revealed a few genes annotated with Wnt
310 pathway members at 3 days post deafferentation (Table 2). These genes had a top BLAST hit of
311 atrial natriuretic peptide-converting enzyme isoform X1, Frizzled-4, and secreted frizzled-related
312 protein 5-like.

313 We looked for the presence of a number of additional groups of proteins that influence
314 neuronal morphogenesis, plasticity, or remodeling (Table 2). For example, the matrix

315 metalloproteases (MMPs) are required for axonal guidance (51) as well as dendritic remodeling
316 during metamorphosis in *Drosophila melanogaster* (52). Importantly, the expression of some
317 MMPs appear to contribute to poor recovery after spinal cord injury in mammals (53). We did
318 not find enrichment in any of these terms at any time point (Table 2), indicating that the injury-
319 induced anatomical plasticity in the cricket may rely on different pathways than have been
320 identified in other species. Furthermore, it is notable that factors, such as MMPs, that can restrict
321 growth or contribute to pruning in other organisms are not upregulated upon injury in the cricket.
322 Several GO terms associated with the candidates found in a previous subtraction
323 hybridization study (54) were also found to be differentially expressed in the present experiment,
324 often showing significant changes in expression at the three- and seven-day post-deafferentation
325 time point (Table 2, bold). These include oxidoreductase, alpha-amylase, endoglucanase, and
326 alcohol dehydrogenase. As noted by Horch and colleagues (54), many of these enzymes have
327 been observed in the hemolymph of insects and play a role in metabolism and immune response.
328 Although it is possible that these findings are due to contaminants during the extraction of the
329 prothoracic ganglion, the results would imply that the extractions differed between control and
330 experimental animals in multiple experiments. Given that the differential expression of several
331 enzyme transcripts was found both in this study and in our former suppression subtractive
332 hybridizaiton (SSH; Ref. 54) study, it is less likely that these enzymes are artifact or
333 contamination effects. Particularly, several differentially regulated transcripts were associated
334 with oxidoreductase activity across all time points. The BLAST hits of these transcripts showed
335 some of the enzymes to be retinal dehydrogenases. Retinal dehydrogenase along with alcohol
336 dehydrogenase, another regulated GO term, are involved with the production of retinoic acid.
337 Retinoic acid has been shown to be involved with development, regeneration, synaptic plasticity,

338 and neurite outgrowth (55–57) implying that regulation of retinoic acid production may influence
339 these processes. Another class of oxidoreductases that appeared abundant within the BLAST hits
340 was the cytochrome P450 family. Cytochrome P450 is a superfamily of monooxygenase
341 enzymes and several families of cytochrome P450 exist in insects. These enzymes are known to
342 have a variety of functional roles in insects including growth and development (58). Cytochrome
343 P450 has also been shown to regulate ecdysone signaling in insects, including crickets (59,60).
344 Ecdysone signaling is crucial for developmental processes and morphogenesis, but has also been
345 shown to be important in the dendritic remodeling of *Drosophila melanogaster* sensory neurons
346 (52,59). While these protein families represent some of the transcripts annotated with
347 "oxidoreductase activity", given the wide range of such transcripts, it is difficult to discern the
348 role of all of the regulated proteins.

350 **Table 2: Evaluation of the number and fold-change of GO terms of interest.**

GO ID	GO Term	Down 1	Down 3	Down 7	Up 1	Up 3	Up 7
GO:0007411	**axon guidance**	N/A	N/A	N/A	N/A	**1 (3.10)**	N/A
GO:0048813	dendrite morphogenesis	N/A	N/A	N/A	N/A	N/A	N/A
GO:0022008	**neurogenesis**	N/A	**6 (1.00 ± 0.01)**	N/A	N/A	**1 (3.10)**	N/A
GO:0008219	cell death	N/A	N/A	N/A	N/A	N/A	N/A
GO:0099544	perisynaptic space	N/A	N/A	N/A	N/A	N/A	N/A
GO:1990773	MMP secretion	N/A	N/A	N/A	N/A	N/A	N/A
GO:0198738	**cell-cell signaling by wnt**	N/A	**4 (-1.99 ± 0.34)**	N/A	N/A	**1 (1.28)**	**1 (1.20)**
GO:0004185	serine-type carboxypeptidase activity	N/A	N/A	N/A	N/A	N/A	N/A
GO:0003735	ribosomal constituent	N/A	N/A	N/A	N/A	N/A	N/A
GO:0003743	translation initiation factor activity	N/A	N/A	N/A	N/A	N/A	N/A
GO:0008135	**translation factor activity, RNA binding**	N/A	N/A	N/A	N/A	N/A	**12 (2.85 ± 0.24)**
GO:0008106	**alcohol dehydrogenase (NADP+) activity**	N/A	N/A	N/A	**3 (1.26 +/- 0.01)**	N/A	N/A
GO:0004556	**alpha-amylase activity**	N/A	N/A	**8 (-2.77 ± 0.27)**	N/A	**13 (2.65 ± 0.61)**	N/A
GO:0015927	trehalase activity	N/A	N/A	N/A	N/A	N/A	N/A
GO:0016491	**oxidoreductase activity**	**4 (-2.38 ± 2.24)**	**112 (-1.40 ± 0.43)**	**14 (-2.82 ± 0.45)**	**175 (2.45 ± 0.92)**	**124 (2.38 ± 0.90)**	**7 (3.48 ± 1.69)**
GO:0008810	**cellulase activity (endoglucanase)**	N/A	N/A	**3 (-2.43 +/- 0.30)**	N/A	**1 (1.97)**	N/A
GO:0003796	lysozyme activity	N/A	N/A	N/A	N/A	N/A	N/A
GO:0004129	cytochrome-c oxidase activity	N/A	N/A	N/A	N/A	N/A	N/A
GO:0016567	**protein ubiquitination**	**1 (-1.60)**	**1 (-1.24)**	N/A	**3 (2.53 ± 0.29)**	**11 (2.15 ± 0.76)**	**22 (1.91 ± 0.10)**
GO:0019783	ubiquitin-like protein specific protease activity	N/A	N/A	N/A	N/A	N/A	N/A

351 Table 2: Number of transcripts associated with the selected GO terms, including any child term of the
352 GO term, at each time point. For each represented GO term, the average +/- standard deviation of the
353 log2foldchange across the transcripts is given in parentheses. GO Terms in bold indicate significant
354 changes in expression. GO IDs in italics were selected because similar transcripts were present in a prior
355 suppression subtractive hybridization study (54).
356

358 **Gene Ontology Enrichment Analysis**

359 Metascape (61) was used to determine enriched GO terms across the differentially
360 expressed gene lists. Differentially expressed genes were first reBLASTed against the curated
361 Swiss-Prot database to retrieve appropriate gene identifiers. Similar ratios of BLAST hit
362 percentages across timepoints were observed against Swiss-Prot as with the nr database,
363 however, the percentage of genes with BLAST hits was lower when BLASTed against Swiss-
364 Prot versus the nr database (Table 3).

365 **Table 3: Comparison of the percentage of genes with BLAST hits in the nr database vs.**
366 **Swiss-Prot.**

	nr	Swiss-Prot
Down 1	81%	59%
Down 3	42%	36%
Down 7	37%	23%
Up 1	62%	59%
Up 3	73%	62%
Up 7	65%	57%

367

368 Enrichment analysis by Metascape showed the most enriched terms at three days post-
369 deafferentation across both upregulated and downregulated genes. Examining the multi-list
370 analysis, there were 22 enriched GO terms found including those related to morphogenesis,
371 extracellular space, and neuron fate commitment (Figure 9). No enriched GO terms were found
372 in the upregulated gene set at seven days post-deafferentation.

373 One category of interest that was revealed in this enrichment analysis was the GO Term
374 "Regulation of Toll-signaling pathway." Our results indicate over 40 different Toll-signaling-
375 related transcripts were differentially expressed, including Toll receptors and the serine
376 proteases, Snake and Spirit. The differential regulation of a relatively large number of transcripts
377 related to Toll-signaling encourages us to generate future hypotheses focused on the role of this
378 pathway in the injury-induced plasticity of the auditory system. At that point, validation of
379 sequence, function, and expression levels will be necessary, especially because several of the
380 identified candidates are likely splice variants of individual genes.

381 Toll receptors are most commonly associated with their function in immunity and
382 development, however, research in *Drosophila melanogaster* suggests that they may also play a
383 role in regulating cell number, connectivity, and synaptogenesis (62). Activation of Toll
384 receptors can regulate cell number through either neuroprotective or pro-apoptotic functions,
385 depending on the receptor type. These functions of Toll receptors were shown to exist in both
386 development and adulthood (63). Toll receptors, specifically Toll-6 and Toll-7, have also been
387 shown to have neurotrophic receptor-like functions through their ability to bind multiple ligands,
388 including neurotrophin-like proteins in invertebrates (64). Neurotrophins are known to regulate
389 cell proliferation and neuronal survival and development, thereby suggesting an important role
390 for Toll receptors in neuronal systems (63,64). Furthermore, in *Drosophila melanogaster* the
391 receptor Toll-8 was shown to positively regulate synaptic growth through a retrograde
392 neurotrophin-like signaling mechanism (65). These studies show that Toll signaling may play an
393 important role in regulating structural plasticity in invertebrates and, given their enrichment in
394 our differentially regulated gene set, may be crucial to the dendritic reorganization observed in
395 the cricket.

396

397 **Conclusions**

398 Unilateral tympanal organ removal in the cricket, *Gryllus bimaculatus*, leads to a robust
399 reorganization of dendrites in the auditory system of the prothoracic ganglion. This novel growth
400 and *de novo* synapse formation restores the ability of the deafferented neurons to respond to
401 sound. Our transcriptomic analyses identified thousands of transcripts up- and down-regulated
402 after deafferentation. We highlight transcriptional changes related to protein translation and
403 degradation, enzymatic activity, and Toll signaling that appear to be enriched after
404 deafferentation. The data presented here allows the development of targeted hypotheses that
405 could elucidate the mechanisms responsible for the deafferentation-induced synaptic plasticity in
406 the auditory system of crickets. The mechanisms at play here can be compared and contrasted
407 with those identified in the terminal ganglion of the cricket after unilateral loss of a cercus (66).

408

409 **Methods**

410 ***Animals, injury, and library preparation***

411 Prothoracic ganglia from approximately 60 adult, male Mediterranean field crickets,
412 *Gryllus bimaculatus* were harvested and 21 individual ganglia were ultimately used as the
413 sources of RNA for this transcriptome (15). Male crickets that were adults for 3-5 days received
414 either a control amputation of the distal segment of the left tarsus ("foot chop" control crickets),
415 or the left prothoracic leg was severed mid-femur removing the auditory organ and deafferenting
416 the ipsilateral central auditory neurons ("deafferented" experimental crickets). Males were
417 chosen due to the potential sexual dimorphism in rates of dendritic growth after deafferentation
418 (22). Several crickets were prepared for backfill as previously described (15). Prothoracic

419 ganglia were removed from crickets 1, 3, or 7 days after amputation at the femur or tarsus, or 18
420 hours post-backfill, and total RNA was purified as previously described (Figure 10).
421 The QIAGEN RNeasy Lipid Tissue Minikit was used to purify total RNA from each
422 sample individually. RNA concentrations were assessed after TURBO DNA-free treatment
423 (Ambion by Life Technologies) with a nanospectrophotometer (Nanodrop, Thermo Fisher
424 Scientific) or a fluorometer (Qubit, Thermo Fisher Scientific). An Agilent 2100 Bioanalyzer
425 (Applied Biosystems, Carlsbad, CA) was used to further assess sample quality. Based on
426 evaluation of RNA quality and concentration of individual ganglion samples, the best 3 samples
427 for each condition were selected for sequencing. Standard Illumina paired-end library protocols
428 were used to prepare samples. The Illumina Hiseq 2500 platform, running v4 chemistry to
429 generate ~25M paired end reads of 100bp in length for each sample, was used to sequence the
430 RNA (15).
431 ***Transcriptome assembly***
432 Trinity software (Trinity-v2.6.5) was run using previously processed and filtered reads of
433 prothoracic ganglion libraries (15). A multi-k-mer assembly was created by building five *de novo*
434 transcriptomes using a single k-mer length (21, 25, 27, 30, 32) and subsequently combining
435 them. The following parameters were used: minimum contig length of 200, library normalization
436 with maximum read coverage 50, RF strand specific read orientation, maximum memory,
437 250GB, and 32 CPUs. Individual assemblies were analyzed using the *TrinityStats.pl* script and
438 alignment statistics were obtained using Bowtie2 (v 2.3.4.1). The PRINSEQ interactive program
439 (67) was used to generate additional summary statistics on each assembly.
440 A k-mer number identity was added to each contig's Trinity ID, all five assemblies were
441 concatenated, and the Evidential Gene program was used to create a single non-redundant

442 assembly. Evidential Gene relies on the Transdecoder.LongOrfs method, identifies the longest
443 ORFs, removes fragments, and uses a BLAST on self to identify highly similar sequences (98%).
444 The main (okay) and alternative (okalt) sets output from Evidential Gene were combined into a
445 final FASTA file used as the transcriptome for all subsequent analyses. Bam files, sorted bam
446 files, bam index files, and idxstats.txt files were created using samtools (68). This assembly is
447 publicly available at NCBI (Bioproject: PRJNa376023, SUB8325660). The metajinomics python
448 mapping tools (69) were used to generate a counts matrix.

449 ***Coverage Analysis***

450 Samtools was used to extract the sequencing depth at every base position for each contig
451 in every cricket sample, and a python script was used to extract the mean and standard deviation
452 of depth for each contig. The package plotly in R (70) was used to plot the depth of each cricket
453 sample. Graphs were visually compared to determine outliers. Two outliers, 1C1 and 7C2, were
454 removed.

455 EdgeR and DESeq were used to run the differential expression analysis, (35,37). The
456 raw read counts generated for each of the libraries, excluding the outliers and the backfill
457 conditions, were used as input to both programs. Similar filtering and normalization functions
458 were used in both programs to exclude any contigs that did not have at least one count per
459 million in at least two libraries. Comparisons between control and deafferented samples were
460 performed at each time point to create lists of upregulated and downregulated genes with a p
461 value cutoff of 0.05. Pairwise comparison results were then compared across time points and
462 were then compared and visualized between the two programs. Another set of lists containing the
463 genes overlapping the two programs was created for continued analysis. The EnhancedVolcano
464 package available in R was used to visualize differential gene expression in volcano plots (71).

465 ***PCR Confirmation***

466 Six sequences were randomly selected for amplification and Sanger sequencing in order
467 to validate the assembly. Sets of primers were designed to obtain the sequence of most of each
468 sequence as predicted by the Trinity assembly. Primers, available on request, were designed for
469 TRINITY21_DN57089_c8_g2_i1.p1 (hypothetical coiled-coil domain protein),
470 TRINITY21_DN58301_c9_g1_i8.p1 (eukaryotic translation initiation factor 4 gamma 3-like),
471 TRINITY25_DN131062_c0_g1_i1.p1 (protein unc-13 homolog 4B),
472 TRINITY27_DN140563_c0_g1_i5.p1 (syntaxin-binding protein 5),
473 TRINITY32_DN141398_c1_g1_i7.p1 (cytochrome P450 301a1),
474 TRINITY21_DN54942_c12_g1_i5.p1 (kinesin light chain). cDNA derived from RNA purified
475 form independent control and deafferented prothoracic ganglion samples was used for PCR. PCR
476 amplicons were gel purified and sequenced, and sequences were aligned and analyzed in
477 Geneious Prime Software (Version 2019.2.3).

478 ***qPCR Validation of Protein Yellow***

479 Quantitative PCR was used to validate predicted expression changes in Protein Yellow
480 candidates. RNA was extracted as described above and reverse transcribed with oligo-dT primers
481 and Superscript III. qPCR primers against Protein Yellow-f (Left:
482 GCGTCTGGCAGAACAGCTCC and Right: CGTGGATGAAGGAGGCGGTG) were designed
483 using a modified version of Primer 3 (version 2.3.7) within Geneious and were chosen so that all
484 potential splice variants would be quantified. Reactions were run on an a QuantSudio3
485 (ThermoFisher) with Power SYBR Green Master Mix (Thermofisher) following manufacturer's
486 protocols and with 1M betaine in the master mix due to high GC content (72%) of the amplicon.
487 Annealing temperatures of primers were validated by completing a qPCR using a temperature
488 gradient. PCR efficiencies were checked by running 6, 2-fold serial dilutions of cDNA template,
489 with resulting slope value of -3.213 indicating acceptable efficiency. 2 samples per condition

490 were run in triplicate. Expression values were normalized relative to 2 reference genes (Arm and
491 EF1a), which were identified using RefFinder (72) as the most stable among 6 different
492 reference gene candidates. PCR Miner (73) was used to assess differential expression.
493
494 ***BLAST Searches***

495 A Perl script was used to extract differentially expressed sequences. The NCBI BLASTx
496 local tool (74) was used to identify proteins similar to the translated nucleotide query sequences.
497 An E-value cutoff of 1e-3 was used and max target sequences was set to 1, and max hits per
498 sequence was set to 1, resulting in the output of only the top hit. Query sequences were
499 BLASTed against the entire non-redundant database downloaded from the NCBI website on
500 August 2, 2018.

501 ***Gene Ontology Analysis***

502 The program, BLAST2GO provided GO annotations for differentially regulated genes
503 (38) using the following parameters: BLASTx-fast against the nr database, number of blast
504 BLAST hits = 20, E-value of 1.0 e -3, word size of 6, hsp length cutoff of 33, with default
505 mapping and annotation settings. GO terms found to be associated with various genes were
506 manually grouped according to GO subtype (cellular component, biological process, or
507 molecular function) and plotted to view the distribution across time points. The web-based
508 CateGOrizer program was used to batch analyze each set of GO terms and determine the number
509 of GO terms under higher order GO classes of interest (75). WEGO 2.0 (Web Gene Ontology
510 Annotation Plot) with a GO level of 2 was also used to plot histograms of the GO annotations for
511 the differentially regulated genes (45). To further analyze the differentially expressed genes, an
512 enrichment analysis was performed with Metascape (61). A multiple gene list analysis looking at

513 the enrichment of the three classes of Gene Ontology terms was performed using *Drosophila*
514 *melanogaster* as the analysis species.
515
516 **Availability of data and material**
517 Initial description of assembly of this transcriptome in Fisher et al., 2018. Re-assembly was
518 completed as described above and is publicly available at NCBI (Bioproject: PRJNa376023,
519 SUB8325660)
520
521 **Declarations:**
522 **Ethics approval and consent to participate**
523 Not applicable.
524
525 **Availability of data and materials**
526 Transcriptomic data are available on NCBI (Bioproject: PRJNa376023
527 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNa376023) and details on the multi-k-mer
528 assembly (GFMG02000000) can be found here:
529 https://www.ncbi.nlm.nih.gov/nuccore/GFMG00000000
530
531 **Competing interests**
532 The authors declare that they have no competing interests
533
534 **Funding**

535 Research reported in this project was supported by an Institutional Development Award (IDeA)
536 from the National Institute of General Medical Sciences of the National Institutes of Health
537 under grant number P20GM103423.
538

539 **Authors' contributions**

540 HF, assisted by LL, collected tissue and HF completed the original transcriptome assembly; FW
541 and MM reassembled using multiple K-mers, compacted the new transcriptome using Evigene
542 and did the differential expression analysis; JB, JJ, LMP, and TAM completed the sanger
543 sequencing analysis; DM identified the most stable reference genes; AR and JB completed
544 Protein Yellow phylogenies, LL completed the qPCR, JO and SK consulted on the statistical
545 differential expression analysis; FW wrote the paper; HWH obtained funding for this project and
546 contributed to the writing. All authors read and approved the final manuscript
547

548 **Acknowledgements**

549 We thank Marko Melendy for animal care support and Meera Prasad for consulting on revisions
550 to this manuscript.

552 **References**

553 1. Prigge CL, Kay JN. Dendrite morphogenesis from birth to adulthood. Current Opinion in
554 Neurobiology. 2018 Dec 1;53:139–45.

555 2. Sampaio-Baptista C, Sanders Z-B, Johansen-Berg H. Structural Plasticity in a with motor
556 learning and stroke rehabilitation. Annual Review of Neuroscience. 2018;41(1):25–40.

557 3. Brodfuehrer PD, Hoy RR. Effect of auditory deafferentation on the synaptic connectivity of
558 a pair of identified interneurons in adult field crickets. J Neurobiol. 1988;19(1):17–38.

559 4. Horch HW, Sheldon E, Cutting CC, Williams CR, Riker DM, Peckler HR, et al. Bilateral
560 consequences of chronic unilateral deafferentation in the auditory system of the cricket
561 *Gryllus bimaculatus*. Dev Neurosci. 2011;33(1):21–37.

562 5. Schmitz B. Neuroplasticity and phonotaxis in monaural adult female crickets (*Gryllus*
563 *bimaculatus* de Geer). Journal of Comparative Physiology A: Neuroethology.
564 1989;164(3):343–58.

565 6. Popov AV, Markovich AM, Andjan AS. Auditory interneurons in the prothoracic ganglion
566 of the cricket, *Gryllus bimaculatus* deGeer. J Comp Physiol. 1978 Jun 1;126(2):183–92.

567 7. Poulet JFA, Hedwig B. Tympanic membrane oscillations and auditory receptor activity in
568 the stridulating cricket *Gryllus bimaculatus*. Journal of Experimental Biology.
569 2001;204(7):1281–93.

570 8. Wohlers DW, Huber F. Topographical organization of the auditory pathway within the
571 prothoracic ganglion of the cricket *Gryllus campestris* L. Cell Tissue Res.
572 1985;239(3):555–65.

573 9. Hoy RR, Nolen TG, Casaday GC. Dendritic sprouting and compensatory synaptogenesis in
574 an identified interneuron following auditory deprivation in a cricket. Proc Natl Acad Sci
575 USA. 1985;82(22):7772–6.

576 10. Schildberger K, Wohlers DW, Schmitz B. Morphological and physiological changes in
577 central auditory neurons following unilateral foreleg amputation in larval crickets. Journal
578 of Comparative Neurology. 1986;158:291–300.

579 11. Ylla G, Nakamura T, Itoh T, Kajitani R, Toyoda A, Tomonari S, et al. Cricket genomes: the
580 genomes of future food [Internet]. Genomics; 2020 Jul [cited 2020 Aug 6]. Available from:
581 http://biorxiv.org/lookup/doi/10.1101/2020.07.07.191841

582 12. Bando T, Ishimaru Y, Kida T, Hamada Y, Matsuoka Y, Nakamura T, et al. Analysis of
583 RNA-Seq data reveals involvement of JAK/STAT signaling during leg regeneration in the
584 cricket *Gryllus bimaculatus.* Development. 2013;140(5):959–64.

585 13. Zeng V, Extavour CG. ASGARD: an open-access database of annotated transcriptomes for
586 emerging model arthropod species. Database (Oxford). 2012;2012:bas048.

587 14. Zeng V, Ewen-Campen B, Horch HW, Roth S, Mito T, Extavour CG. Developmental gene
588 discovery in a hemimetabolous insect: *De novo* assembly and annotation of a transcriptome
589 for the cricket *Gryllus bimaculatus*. Dearden PK, editor. PLoS ONE. 2013;8(5):e61479.

590 15. Fisher HP, Pascual MG, Jimenez SI, Michaelson DA, Joncas CT, Quenzer ED, et al. De
591 novo assembly of a transcriptome for the cricket *Gryllus bimaculatus* prothoracic ganglion:
592 An invertebrate model for investigating adult central nervous system compensatory
593 plasticity. PLoS ONE. 2018;13(7):e0199070.

594 16. Dickson BJ. Molecular mechanisms of axon guidance. Science. 2002 Dec
595 6;298(5600):1959–64.

596 17. Tessier-Lavigne M, Goodman CS. The Molecular Biology of Axon Guidance. Science.
597 1996 Nov 15;274(5290):1123–33.

598 18. Harel NY, Strittmatter SM. Can regenerating axons recapitulate developmental guidance
599 during recovery from spinal cord injury? Nat Rev Neurosci. 2006;7(8):603–16.

600 19. Kikuchi K, Kishino A, Konishi O, Kumagai K, Hosotani N, Saji I, et al. *In vitro* and *in vivo*
601 characterization of a novel semaphorin 3A inhibitor, SM-216289 or xanthofulvin. J Biol
602 Chem. 2003;278(44):42985–91.

603 20. Yu H-H, Araj HH, Ralls SA, Kolodkin AL. The transmembrane semaphorin Sema I is
604 required in *Drosophila* for embryonic motor and CNS axon guidance. Neuron.
605 1998;20:207–20.

606 21. Gilbert D. EvidentialGene: mRNA Transcript Assembly Software [Internet].
607 EvidentialGene: tr2aacds, mRNA Transcript Assembly Software. 2013 [cited 2020 Jan 8].
608 Available from: http://arthropods.eugenes.org/EvidentialGene/trassembly.html

609 22. Pfister A, Johnson A, Ellers O, Horch HW. Quantification of dendritic and axonal growth
610 after injury to the auditory system of the adult cricket *Gryllus bimaculatus*. Front Physiol.
611 2013;3:367.

612 23. Surget-Groba Y, Montoya-Burgos JI. Optimization of de novo transcriptome assembly
613 from next-generation sequencing data. Genome Res. 2010 Oct 1;20(10):1432–40.

614 24. Zhao Q-Y, Wang Y, Kong Y-M, Luo D, Li X, Hao P. Optimizing de novo transcriptome
615 assembly from short-read RNA-Seq data: a comparative study. BMC Bioinformatics.
616 2011;12(Suppl 14):S2.

617 25. Haznedaroglu BZ, Reeves D, Rismani-Yazdi H, Peccia J. Optimization of de novo
618 transcriptome assembly from high-throughput short read sequencing data improves
619 functional annotation for non-model organisms. BMC Bioinformatics. 2012;13(1):170.

620 26. Mamrot J, Legaie R, Ellery SJ, Wilson T, Seemann T, Powell DR, et al. De novo
621 transcriptome assembly for the spiny mouse (Acomys cahirinus). Sci Rep. 2017
622 Dec;7(1):8996.

623 27. Cerveau N, Jackson DJ. Combining independent de novo assemblies optimizes the coding
624 transcriptome for nonconventional model eukaryotic organisms. BMC Bioinformatics. 2016
625 Dec;17(1):525.

626 28. Gilbert DG. Longest protein, longest transcript or most expression, for accurate gene
627 reconstruction of transcriptomes? [Internet]. Bioinformatics; 2019 Nov [cited 2020 Jan 9].
628 Available from: http://biorxiv.org/lookup/doi/10.1101/829184

629 29. Gan HM, Austin C, Linton S. Transcriptome-guided identification of carbohydrate active
630 enzymes (CAZy) from the Christmas Island Red Crab, *Gecarcoidea natalis* and a vote for
631 the inclusion of transcriptome-derived crustacean CAZys in comparative studies. Mar
632 Biotechnol. 2018;20(5):654–65.

633 30. Richardson MF, Sequeira F, Selechnik D, Carneiro M, Vallinoto M, Reid JG, et al.
634 Improving amphibian genomic resources: a multitissue reference transcriptome of an iconic
635 invader. GigaScience [Internet]. 2018 Jan 1 [cited 2020 Jan 9];7(1). Available from:
636 https://academic.oup.com/gigascience/article/doi/10.1093/gigascience/gix114/4662864

637 31. Rivera-García L, Rivera-Vicéns RE, Veglia AJ, Schizas NV. De novo transcriptome
638 assembly of the digitate morphotype of *Briareum asbestinum* (Octocorallia: Alcyonacea)
639 from the southwest shelf of Puerto Rico. Marine Genomics. 2019;47:100676.

640 32. Raplee ID, Evsikov AV, Marín de Evsikova C. Aligning the aligners: Comparison of RNA
641 sequencing data alignment and gene expression quantification tools for clinical breast
642 cancer research. Journal of Personalized Medicine. 2019;9(2):18.

643 33. Varet H, Brillet-Guéguen L, Coppée J-Y, Dillies M-A. SARTools: A DESeq2- and EdgeR-
644 based R pipeline for comprehensive differential analysis of RNA-seq data. PLOS ONE.
645 2016 Jun 9;11(6):e0157022.

646 34. Evans C, Hardin J, Stoebel DM. Selecting between-sample RNA-Seq normalization
647 methods from the perspective of their assumptions. Briefings in Bioinformatics. 2018 Sep
648 28;19(5):776–92.

649 35. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
650 RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

651 36. Maza E, Frasse P, Senin P, Bouzayen M, Zouine M. Comparison of normalization methods
652 for differential gene expression analysis in RNA-Seq experiments. Communicative &
653 Integrative Biology. 2013;6(6):e25849.

654 37. Robinson MD, McCarthy DJ, Smyth GK. EdgeR: a Bioconductor package for differential
655 expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.

656 38. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-
657 throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids
658 Research. 2008;36(10):3420–35.

659 39. Doughty T, Kerkhoven E. Extracting novel hypotheses and findings from RNA-seq data.
660 FEMS Yeast Res [Internet]. 2020 Mar 1 [cited 2020 Jun 11];20(2). Available from:
661 http://academic.oup.com/femsyr/article/20/2/foaa007/5721245

662 40. Zhao S, Zhang Y, Gamini R, Zhang B, von Schack D. Evaluation of two main RNA-seq
663 approaches for gene quantification in clinical RNA sequencing: polyA+ selection versus
664 rRNA depletion. Scientific Reports. 2018 Mar 19;8(1):4781.

665 41. Kucharski R, Maleszka R, Hayward DC, Ball EE. A royal jelly protein Is expressed in a
666 subset of Kenyon Cells in the mushroom bodies of the honey bee brain.
667 Naturwissenschaften. 1998;85(7):343–6.

668 42. Ferguson LC, Green J, Surridge A, Jiggins CD. Evolution of the insect yellow gene family.
669 Molecular Biology and Evolution. 2011;28(1):257–72.

670 43. Albert Š, Klaudiny J. The MRJP/YELLOW protein family of <i>Apis mellifera<i/i>
671 Identification of new members in the EST library. Journal of Insect Physiology.
672 2004;50(1):51–9.

673 44. Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, et al. WEGO: a web tool for plotting
674 GO annotations. Nucleic Acids Research. 2006;34(Web Server):W293–7.

675 45. Ye J, Zhang Y, Cui H, Liu J, Wu Y, Cheng Y, et al. WEGO 2.0: a web tool for analyzing
676 and plotting GO annotations, 2018 update. Nucleic Acids Research. 2018;46(W1):W71–5.

677 46. Quadrato G, Di Giovanni S. Waking up the sleepers: Shared transcriptional pathways in
678 axonal regeneration and neurogenesis. Cell Mol Life Sci. 2013;70(6):993–1007.

679 47. Matsunaga Y, Qadota H, Furukawa M, Choe H (Helen), Benian GM. Twitchin kinase
680 interacts with MAPKAP kinase 2 in *Caenorhabditis elegans* striated muscle. Mol Biol Cell.
681 2015;26(11):2096–111.

682 48. Baldi A, Calia E, Ciampini A, Riccio M, Vetuschi A, Persico A, et al. Deafferentation-
683 induced apoptosis of neurons in thalamic somatosensory nuclei of the newborn rat: critical
684 period and rescue from cell death by peripherally applied neurotrophins. Eur J Neurosci.
685 2000;12:2281–90.

686 49. Garcia-Valenzuela E, Gorczyca W, Darzynkiewicz Z, Sharma SC. Apoptosis in adult
687 retinal ganglion cells after axotomy. J Neurobiol. 1994 Apr;25(4):431–8.

688 50. Herman PE, Papatheodorou A, Bryant SA, Waterbury CKM, Herdy JR, Arcese AA, et al.
689 Highly conserved molecular pathways, including Wnt signaling, promote functional
690 recovery from spinal cord injury in lampreys. Sci Rep. 2018;8(1):18.

691 51. Miller CM, Page-McCaw A, Broihier HT. Matrix metalloproteinases promote motor axon
692 fasciculation in the Drosophila embryo. Development. 2008 Jan 1;135(1):95–109.

52. Kuo CT, Jan LY, Jan YN. Dendrite-specific remodeling of Drosophila sensory neurons requires matrix metalloproteases, ubiquitin-proteasome, and ecdysone signaling. Proc Natl Acad Sci USA. 2005;102(42):15230–5.

53. Zhang H, Chang M, Hansen CN, Basso DM, Noble-Haeusslein LJ. Role of matrix metalloproteinases and therapeutic benefits of their inhibition in spinal cord injury. Neurotherapeutics. 2011;8(2):206–20.

54. Horch HW, McCarthy SS, Johansen SL, Harris JM. Differential gene expression during compensatory sprouting of dendrites in the auditory system of the cricket *Gryllus bimaculatus*. Insect Molecular Biology. 2009;18(4):483–96.

55. Aoto J, Nam CI, Poon MM, Ting P, Chen L. Synaptic signaling by all-trans retinoic acid in homeostatic synaptic plasticity. Neuron. 2008;60(2):308–20.

56. Dmetrichuk JM, Carlone RL, Spencer GE. Retinoic acid induces neurite outgrowth and growth cone turning in invertebrate neurons. Dev Biol. 2006;294:39–49.

57. Maden M. Retinoic acid in the development, regeneration and maintenance of the nervous system. Nature Reviews Neuroscience. 2007;8(10):755–65.

58. Scott JG, Wen Z. Cytochromes P450 of insects: the tip of the iceberg. Pest Management Science. 2001;57(10):958–67.

59. Chavez VM, Marques G, Delbecque JP, Kobayashi K, Hollingsworth M, Burr J, et al. The Drosophila disembodied gene controls late embryonic morphogenesis and codes for a cytochrome P450 enzyme that regulates embryonic ecdysone levels. Development. 2000;127(19):4115–26.

60. Liebrich W (Universitat U, Durnberger BB, Hoffmann KH. Ecdysone 20-monooxygenase in a cricket (*Gryllus bimaculatus* ensifera, Gryllidae): activity throughout adult life cycle. Comparative biochemistry and physiology : A [Internet]. 1991 [cited 2021 Feb 5]; Available from: https://agris.fao.org/agris-search/search.do?recordID=US9150408

61. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature Communications. 2019;10(1):1523.

62. Anthoney N, Foldi I, Hidalgo A. Toll and Toll-like receptor signalling in development. Development [Internet]. 2018 [cited 2020 Apr 17];145(9). Available from: https://dev.biologists.org/content/145/9/dev156018

63. Li G, Forero MG, Wentzell JS, Durmus I, Wolf R, Anthoney NC, et al. A Toll-receptor map underlies structural brain plasticity. Bellen HJ, Banerjee U, editors. eLife. 2020 Feb 18;9:e52743.

727 64. McIlroy G, Foldi I, Aurikko J, Wentzell JS, Lim MA, Fenton JC, et al. Toll-6 and Toll-7
728 function as neurotrophin receptors in the *Drosophila melanogaster* CNS. Nature
729 Neuroscience. 2013;16(9):1248–56.

730 65. Ballard SL, Miller DL, Ganetzky B. Retrograde neurotrophin signaling through Tollo
731 regulates synaptic growth in Drosophila. J Cell Biol. 2014 Mar 31;204(7):1157–72.

732 66. Prasad MP, Detchou DKE, Wang F, Ledwidge LL, Kingston SE, Wilson Horch H.
733 Transcriptional expression changes during compensatory plasticity in the terminal ganglion
734 of the adult cricket Gryllus bimaculatus. BMC Genomics. 2021 Oct 14;22(1):742.

735 67. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets.
736 Bioinformatics. 2011 Mar 15;27(6):863–4.

737 68. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence
738 alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.

739 69. Choi J. Metajinomics mapping tool. 2017; Available from:
740 https://github.com/metajinomics/mapping_tools

741 70. Sievert C. Interactive Web-Based Data Visualization with R, plotly, and shiny [Internet].
742 Chapman and Hall/CRC; 2020. Available from: https://plotly-r.com

743 71. Blighe K, Rana S, Lewis M. EnhancedVolcano: Publication-ready volcano plots with
744 enhanced colouring and labeling. [Internet]. 2020 [cited 2020 Jan 7]. Available from:
745 https://github.com/kevinblighe/EnhancedVolcano

746 72. Xie F, Xiao P, Chen D, Xu L, Zhang B. miRDeepFinder: a miRNA analysis tool for deep
747 sequencing of plant small RNAs. Plant Mol Biol. 2012 Sep 1;80(1):75–84.

748 73. Zhao S, Fernald RD. Comprehensive Algorithm for Quantitative Real-Time Polymerase
749 Chain Reaction. Journal of Computational Biology. 2005 Oct 1;12(8):1047–64.

750 74. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool.
751 Journal of Molecular Biology. 1990;215(3):403–10.

752 75. Hu Z, Bao J, Reecy J. CateGOrizer: A web-based program to batch analyze gene ontology
753 classification categories. Online Journal of Bioinformatics. 2008;9:108–12.

754

755

757 **Figure legends**

758 Figure 1. Summary of workflow for multi k-mer assembly.

759 Figure 2. Volcano plots of differential gene expression in *G. bimaculatus* prothoracic ganglia at
760 one day (a,b), 3 days (c,d) and 7 days (e,f) after deafferentation. The horizontal dotted line marks
761 a p-value of 0.05, and the vertical dotted line marks no predicted fold change. Each point
762 represents a contig determined to be differentially regulated by EdgeR (a,c,e) or DESeq2 (b,d,f).
763 Blue points represent the contigs determined to be significantly regulated.
764
765 Figure 3: Correlation of fold-changes predicted by EdgeR and DESeq2 for upregulated
766 transcripts (a,b,c) and downregulated transcripts (d,e,f) at one(a,d), three (b,e), and seven (c,f)
767 days.
768
769 Figure 4. Differentially regulated genes across three time points. Similar patterns in relative
770 numbers of differentially regulated genes are observed between the two programs. a) EdgeR
771 identified downregulated genes b) EdgeR identified upregulated genes c) DESeq2 identified
772 downregulated genes d) DESeq2 identified upregulated genes.
773
774 Figure 5. Differentially regulated genes compared across the two analytical programs, DESeq2
775 and EdgeR. The number of genes found to be differentially regulated by both programs varies by
776 condition.
777
778 Figure 6. Percentage of sequences with no BLAST hits, BLAST hits, and BLAST hits with
779 additional GO term mapping. Distribution of sequences varies across times points and regulation.
780
781 Figure 7. GO term analysis organized into the three root classes: cellular component (CC),
782 molecular function (MF), and biological process (BP). The top 5 represented GO terms across all
783 time points in each class are represented. Many highly represented GO terms were found in the
784 CC class whereas a broader range of GO terms were found in the MF and BP classes.
785
786 Figure 8. WEGO histograms with the distribution of Gene Ontology terms grouped by cellular
787 component, molecular function, or biological process. a) GO terms associated with upregulated
788 genes across all time points. b) GO terms associated with downregulated genes across all time
789 points. Percentages are noted on the left and the number of genes within the given list that were
790 annotated with the GO term/child term are noted on the right. On the right axis, the top numbers
791 (turquoise) correspond to the one-day data, the middle numbers, (magenta) corresponds to the
792 three-day data, and the bottom numbers correspond to the seven-day data (teal).
793
794 Figure 9. Heatmap of enrichment terms as determined by Metascape. Colored by p-value as
795 indicated at the top.
796
797 Figure 10. Summary of experimental design. 21 PTG ganglia were removed from "deafferented"
798 or "foot chop control" animals 1, 3, or 7 days-post injury. Three additional animals were
799 backfilled 18 hrs prior to PTG removal. RNASeq data from these animals were included in the
800 assembly but not in the differential expression.

801

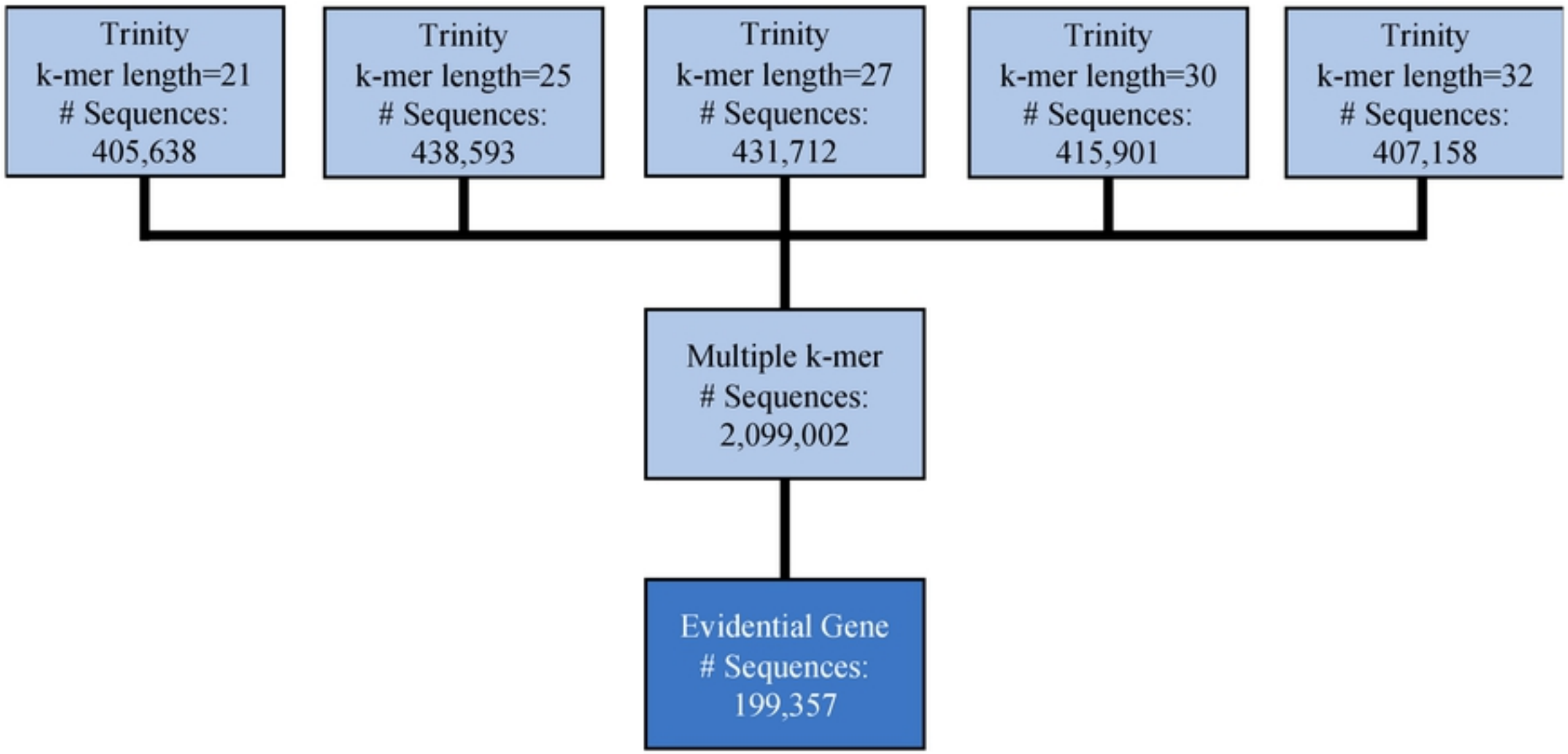

Trinity
k-mer length=21
Sequences:
405,638
Trinity
k-mer length=25
Sequences:
438,593
Trinity
k-mer length=27
Sequences:
431,712
Trinity
k-mer length=30
Sequences:
415,901
Trinity
k-mer length=32
Sequences:
407,158
Multiple k-mer
Sequences:
2,099,002
Evidential Gene
Sequences:
199,357

Figure 1

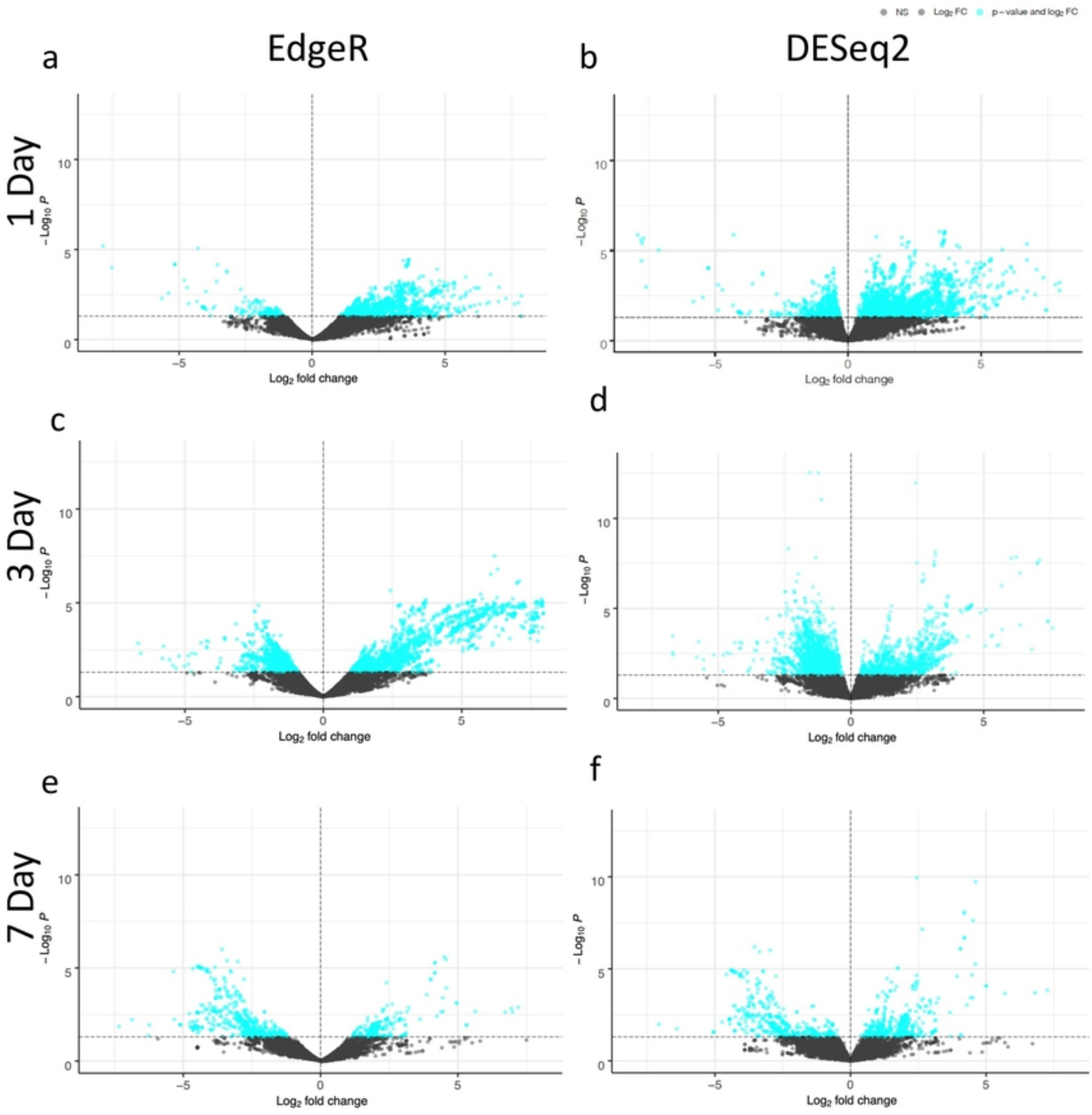

NS
Log2 FC
p-value and log2 FC
a
EdgeR
b
DESeq2
1 Day
c
d
3 Day
e
f
7 Day
−Log10 P
Log2 fold change

Figure 2

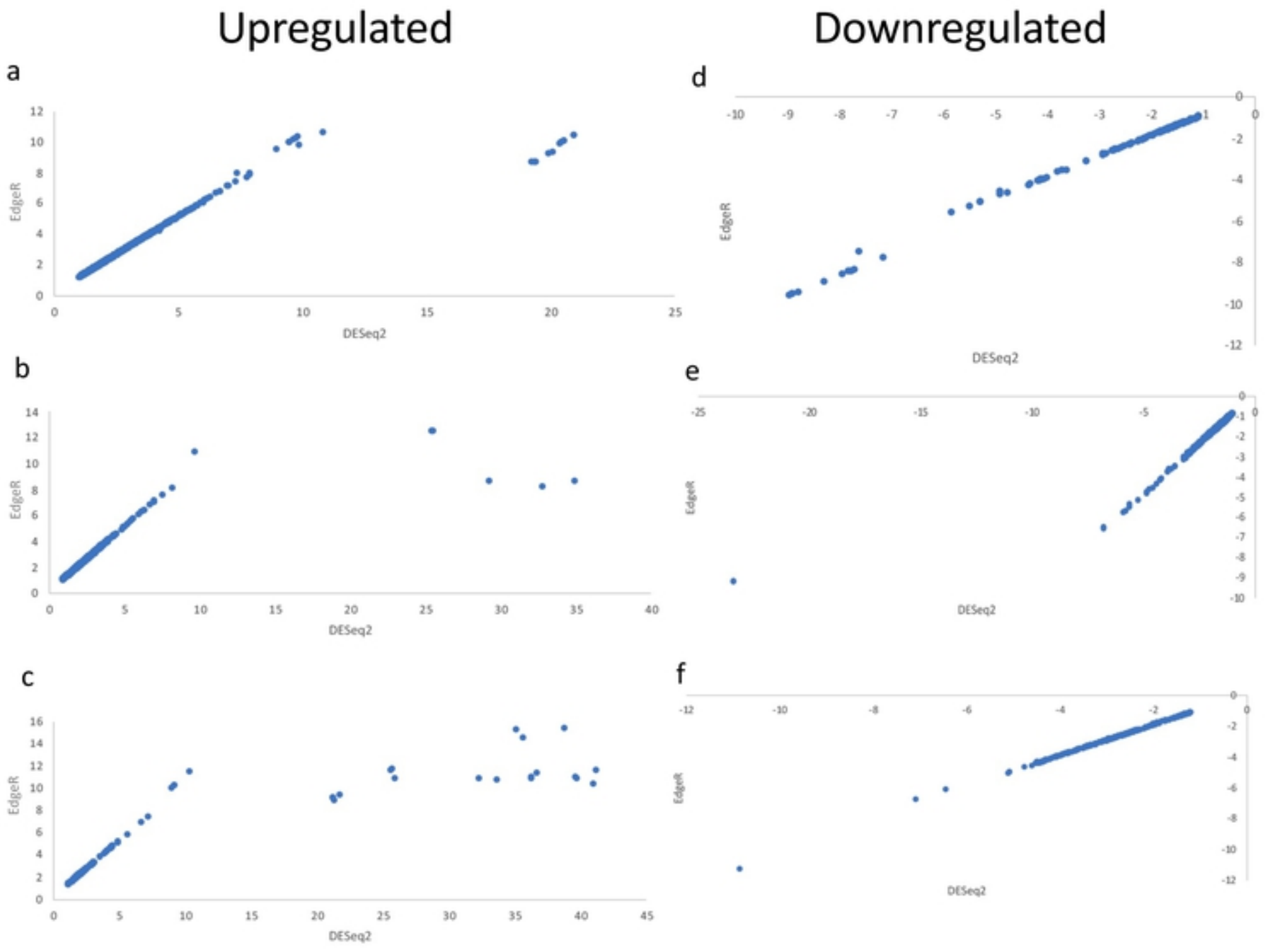
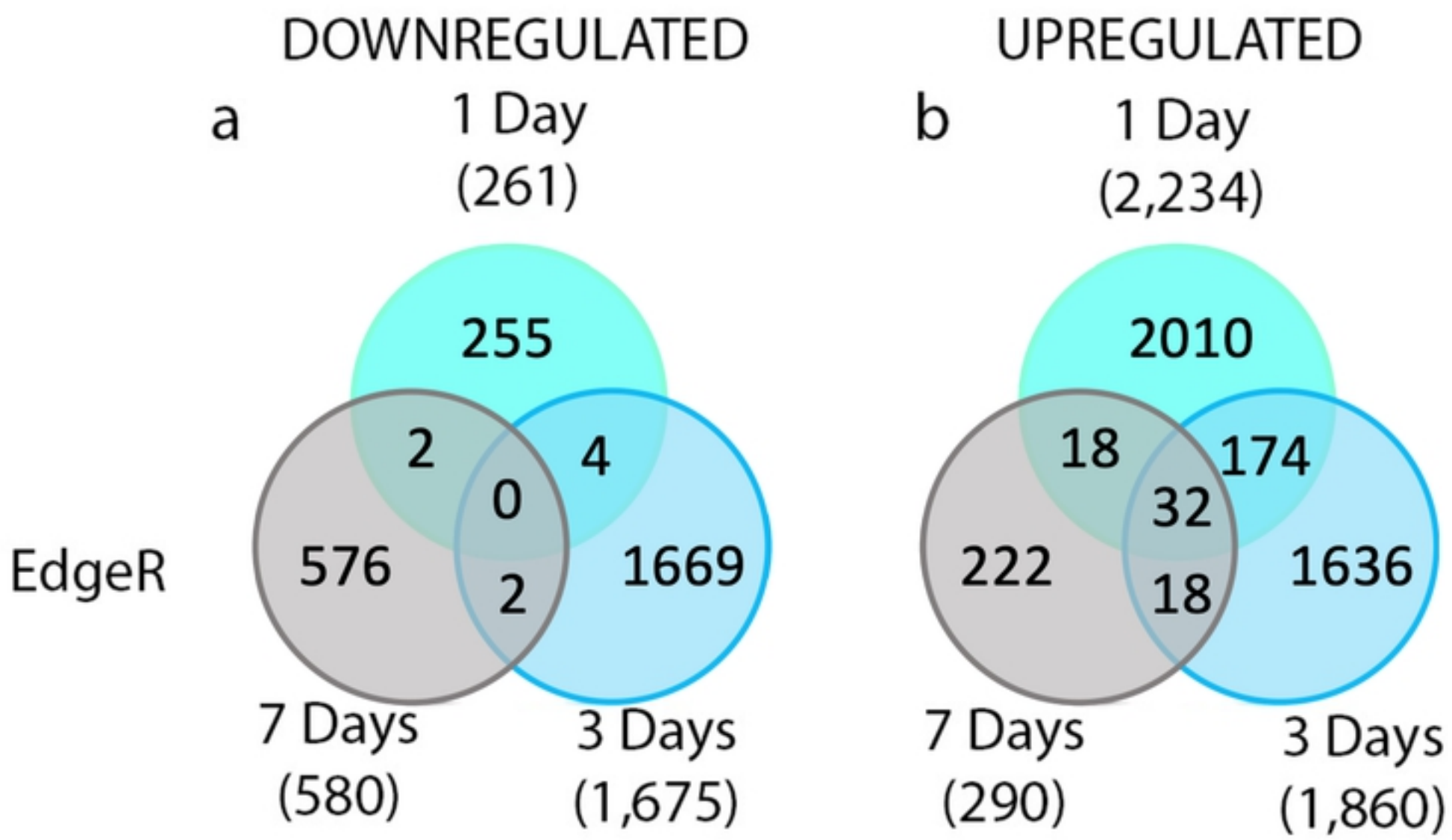


Figure 3


DOWNREGULATED
a
1 Day
(261)
255
2
4
0
576
2
1669
7 Days
(580)
3 Days
(1,675)
UPREGULATED
b
1 Day
(2,234)
2010
18
174
32
222
18
1636
7 Days
(290)
3 Days
(1,860)
EdgeR

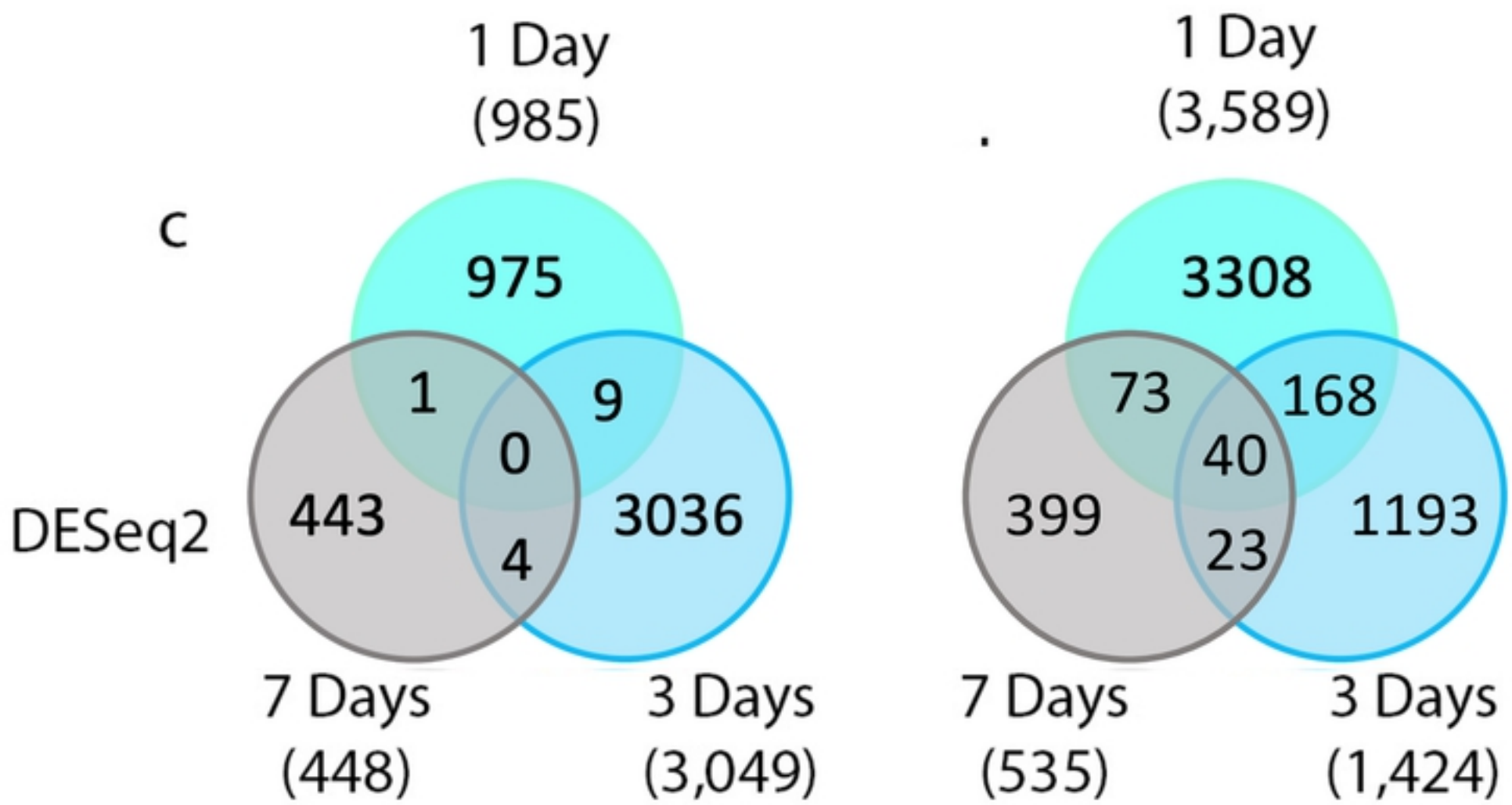

1 Day
(985)
c
975
1
9
0
443
4
3036
7 Days
(448)
3 Days
(3,049)
1 Day
(3,589)
3308
73
168
40
399
23
1193
7 Days
(535)
3 Days
(1,424)
DESeq2

Figure 4

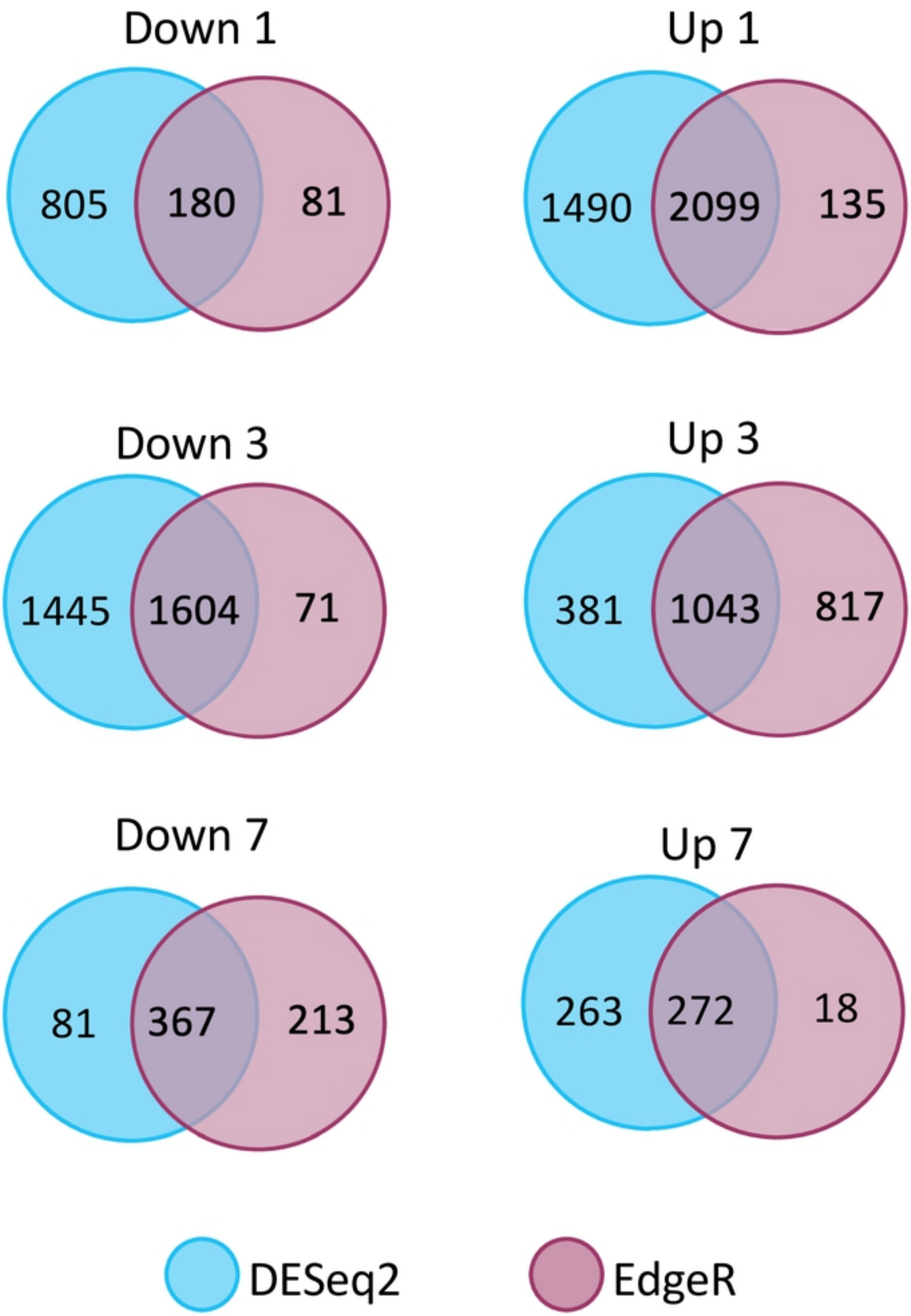

Down 1
805
180
81
Up 1
1490
2099
135
Down 3
1445
1604
71
Up 3
381
1043
817
Down 7
81
367
213
Up 7
263
272
18
DESeq2
EdgeR

Figure 5

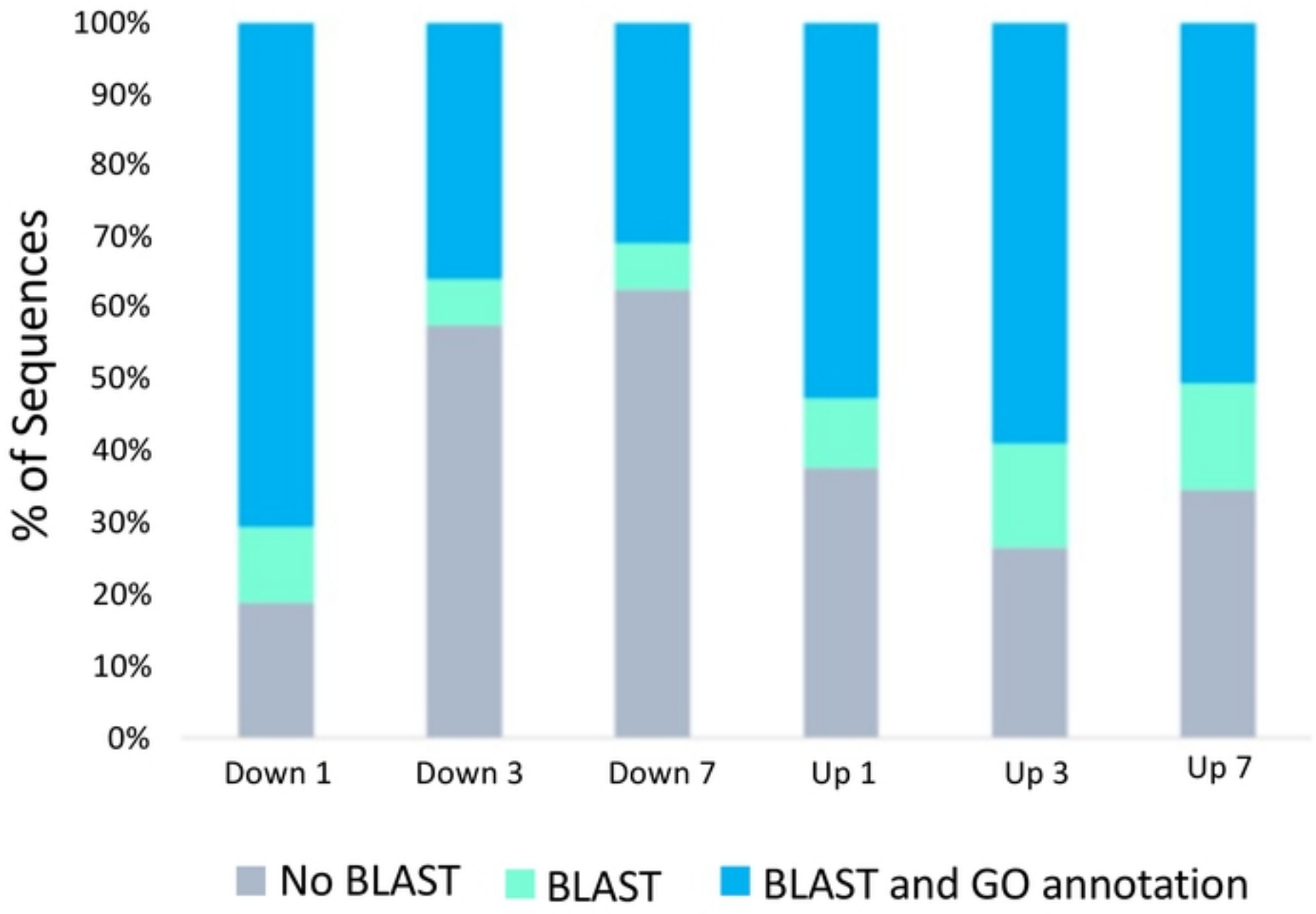
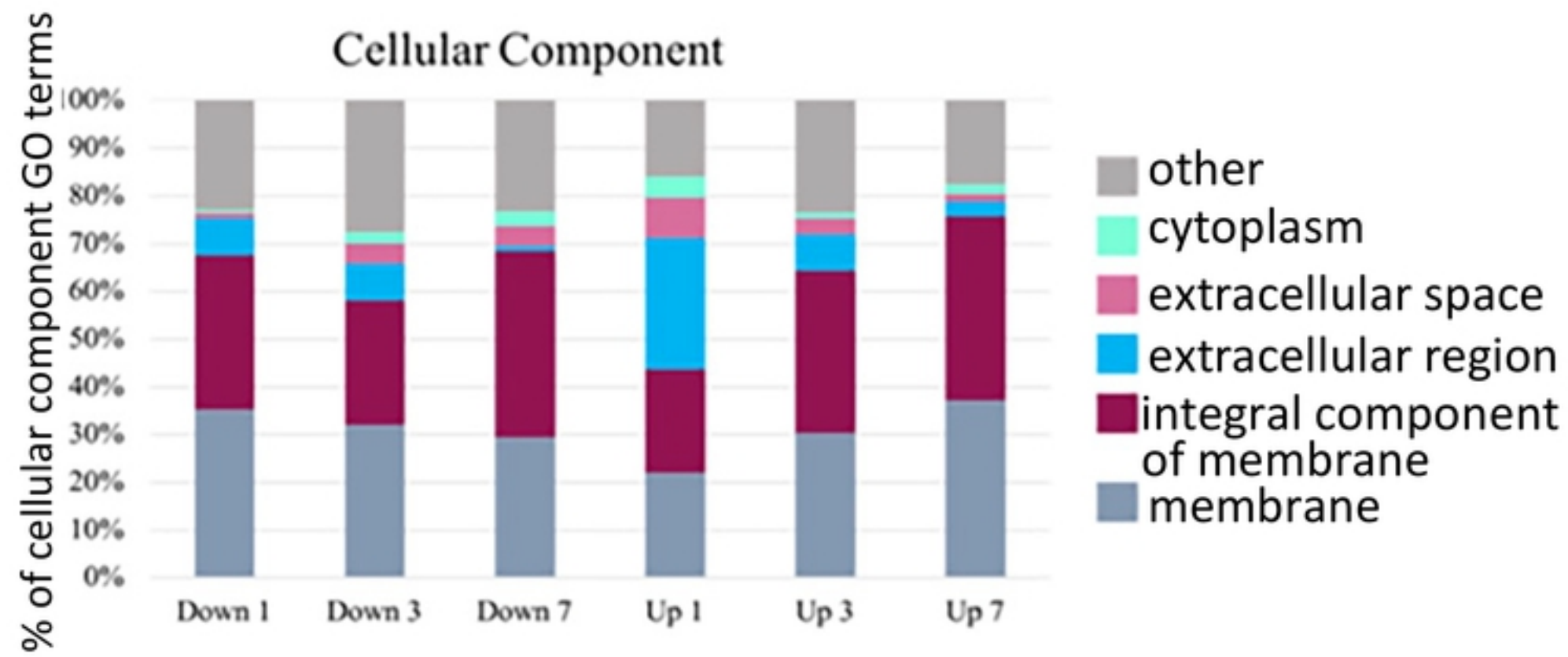
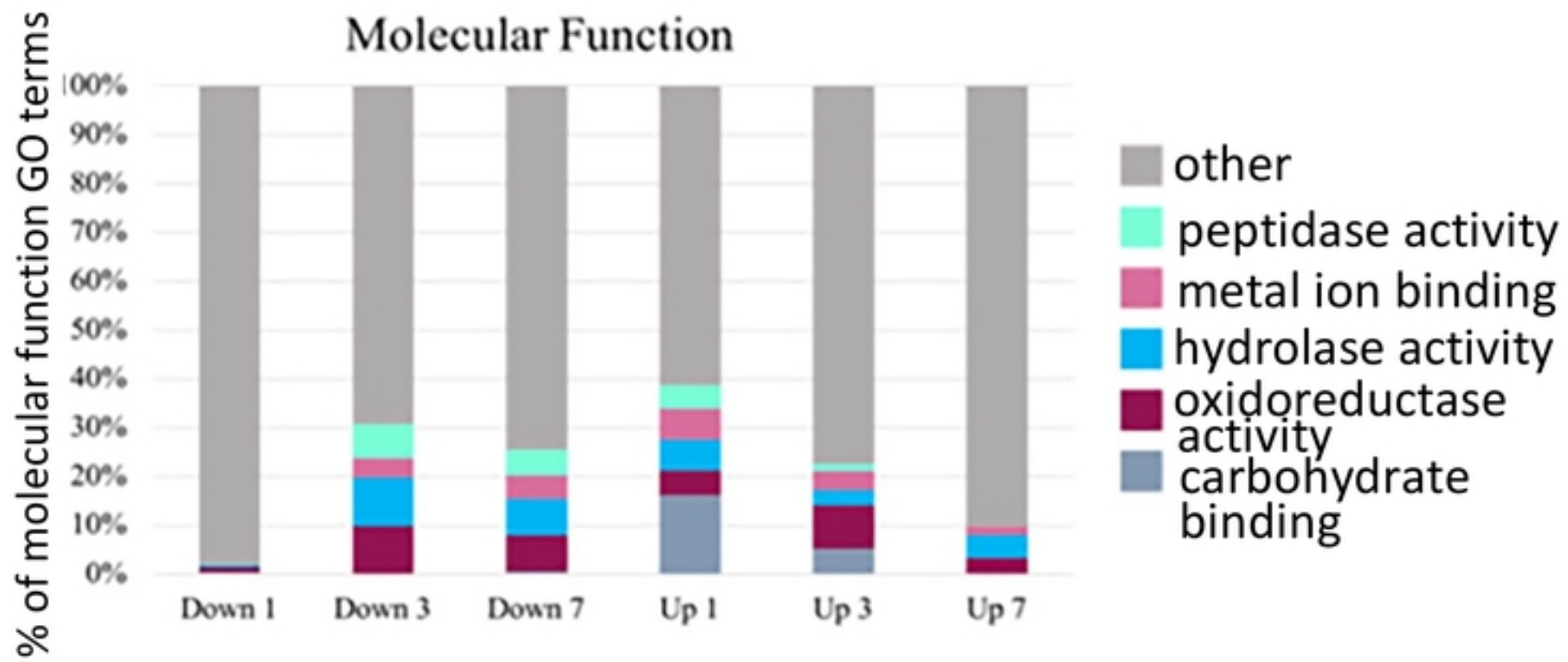

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
% of Sequences
Down 1
Down 3
Down 7
Up 1
Up 3
Up 7
No BLAST
BLAST
BLAST and GO annotation

Figure 6


Cellular Component
% of cellular component GO terms
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Down 1
Down 3
Down 7
Up 1
Up 3
Up 7
other
cytoplasm
extracellular space
extracellular region
integral component of membrane
membrane


Molecular Function
% of molecular function GO terms
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Down 1
Down 3
Down 7
Up 1
Up 3
Up 7
other
peptidase activity
metal ion binding
hydrolase activity
oxidoreductase activity
carbohydrate binding

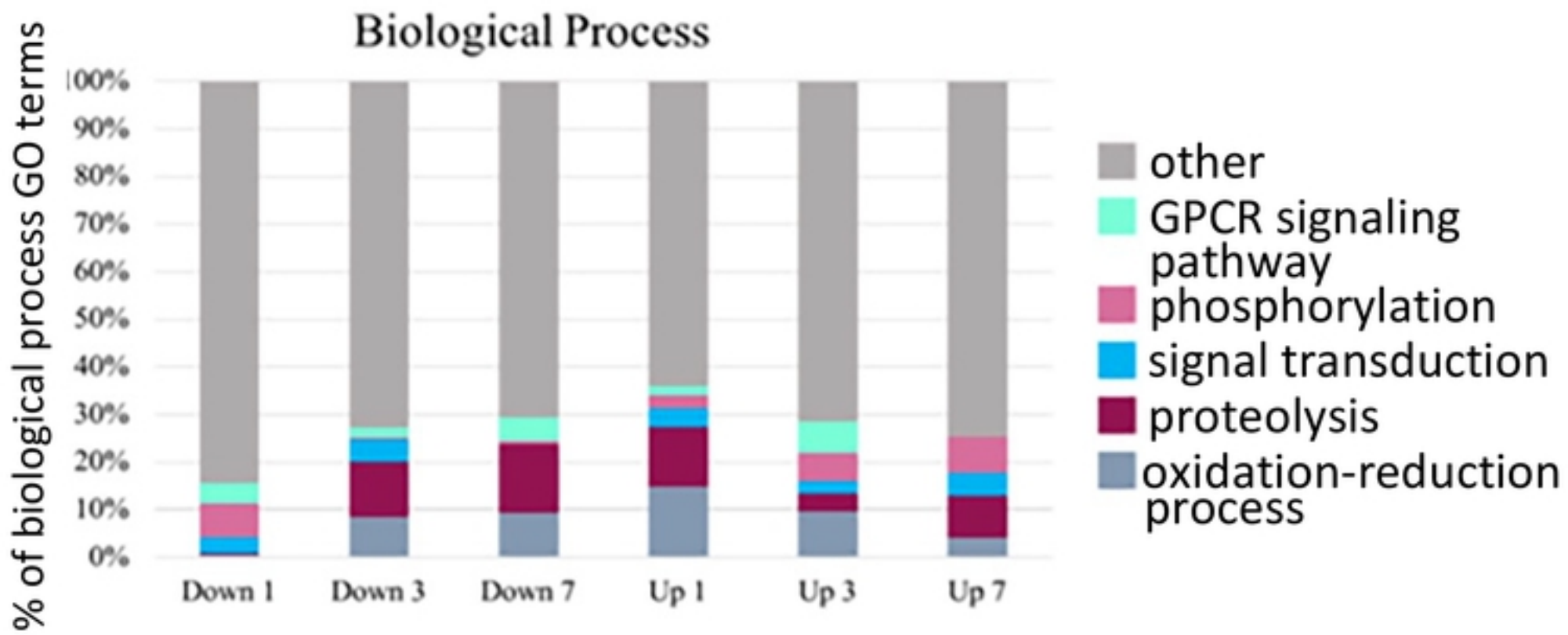

Biological Process
% of biological process GO terms
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Down 1
Down 3
Down 7
Up 1
Up 3
Up 7
other
GPCR signaling pathway
phosphorylation
signal transduction
proteolysis
oxidation-reduction process

Figure 7

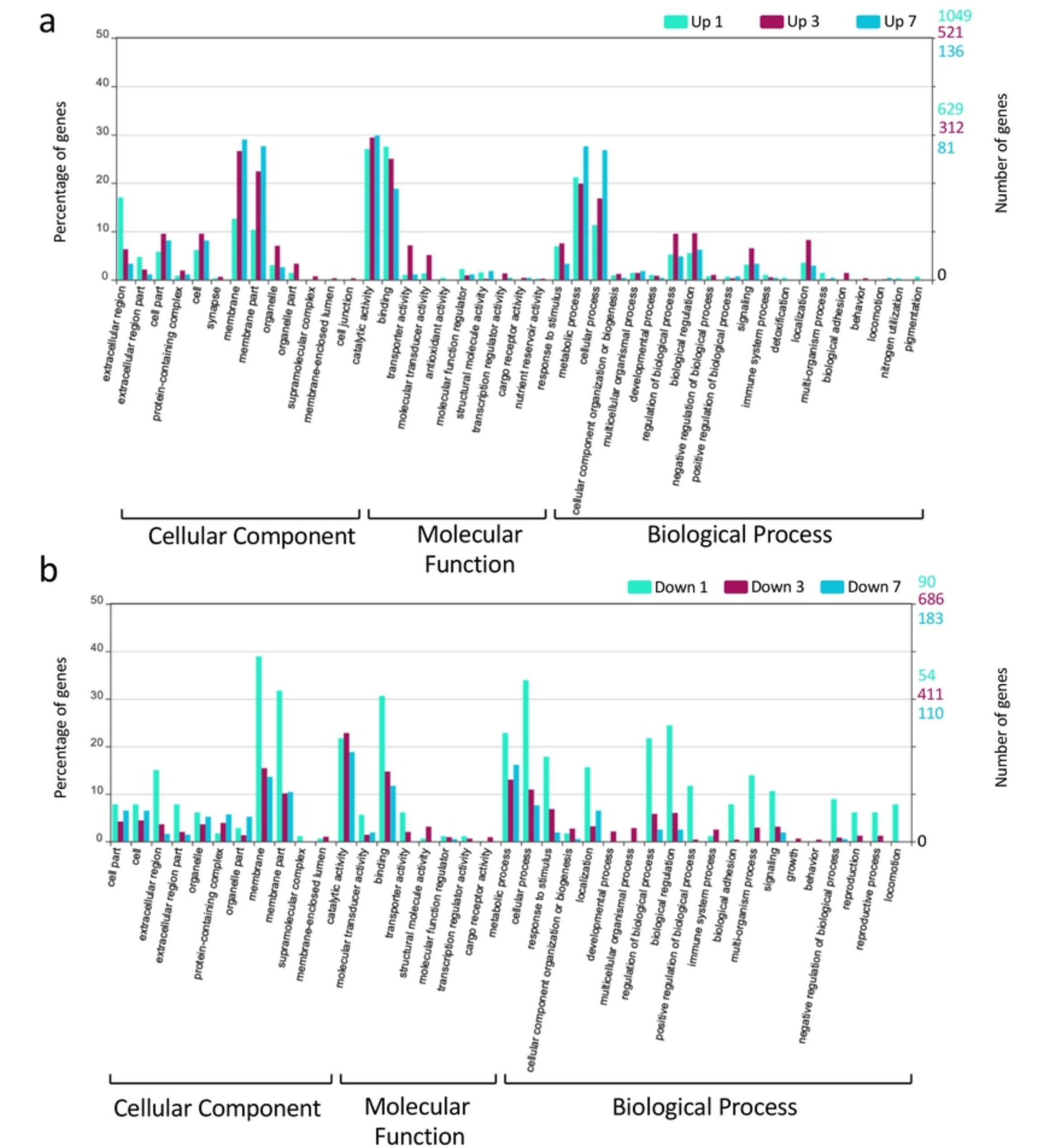

a
Up 1
Up 3
Up 7
1049
521
136
629
312
81
0
Percentage of genes
Number of genes
50
40
30
20
10
0
extracellular region
extracellular region part
cell part
protein-containing complex
cell
synapse
membrane
membrane part
organelle
organelle part
supramolecular complex
membrane-enclosed lumen
cell junction
catalytic activity
binding
transporter activity
molecular transducer activity
antioxidant activity
molecular function regulator
structural molecule activity
transcription regulator activity
cargo receptor activity
nutrient reservoir activity
response to stimulus
metabolic process
cellular process
cellular component organization or biogenesis
multicellular organismal process
developmental process
regulation of biological process
biological regulation
negative regulation of biological process
positive regulation of biological process
signaling
immune system process
detoxification
localization
multi-organism process
biological adhesion
behavior
locomotion
nitrogen utilization
pigmentation
Cellular Component
Molecular Function
Biological Process
b
Down 1
Down 3
Down 7
90
686
183
54
411
110
0
Percentage of genes
Number of genes
50
40
30
20
10
0
cell part
cell
extracellular region
extracellular region part
organelle
protein-containing complex
organelle part
membrane
membrane part
supramolecular complex
membrane-enclosed lumen
catalytic activity
molecular transducer activity
binding
transporter activity
structural molecule activity
molecular function regulator
transcription regulator activity
cargo receptor activity
metabolic process
cellular process
response to stimulus
cellular component organization or biogenesis
localization
developmental process
multicellular organismal process
regulation of biological process
biological regulation
positive regulation of biological process
immune system process
biological adhesion
multi-organism process
signaling
growth
behavior
negative regulation of biological process
reproduction
reproductive process
locomotion
Cellular Component
Molecular Function
Biological Process

Figure 8

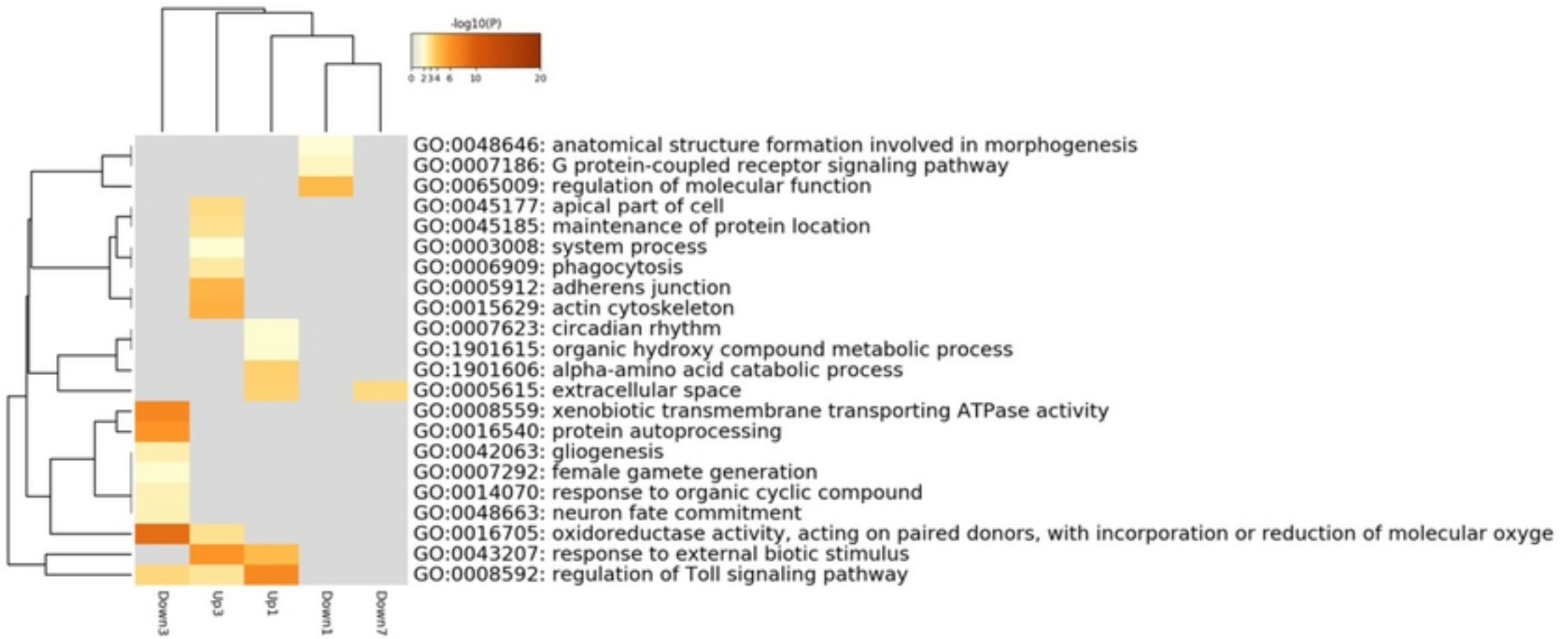

-log10(P)
0 2 3 4 6 10 20
GO:0048646: anatomical structure formation involved in morphogenesis
GO:0007186: G protein-coupled receptor signaling pathway
GO:0065009: regulation of molecular function
GO:0045177: apical part of cell
GO:0045185: maintenance of protein location
GO:0003008: system process
GO:0006909: phagocytosis
GO:0005912: adherens junction
GO:0015629: actin cytoskeleton
GO:0007623: circadian rhythm
GO:1901615: organic hydroxy compound metabolic process
GO:1901606: alpha-amino acid catabolic process
GO:0005615: extracellular space
GO:0008559: xenobiotic transmembrane transporting ATPase activity
GO:0016540: protein autoprocessing
GO:0042063: gliogenesis
GO:0007292: female gamete generation
GO:0014070: response to organic cyclic compound
GO:0048663: neuron fate commitment
GO:0016705: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxyge
GO:0043207: response to external biotic stimulus
GO:0008592: regulation of Toll signaling pathway
Down3
Up3
Up1
Down1
Down7

Figure 9

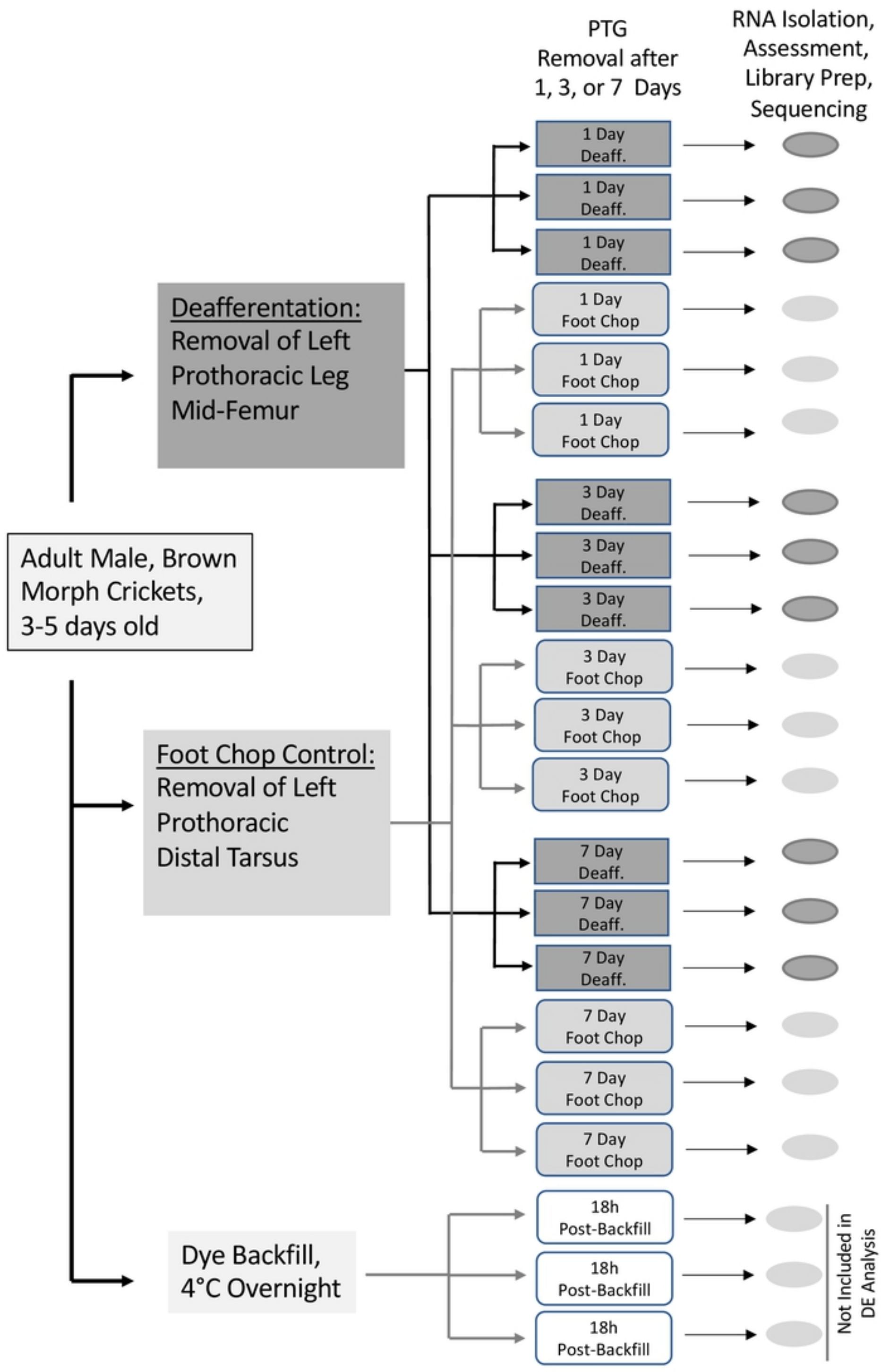

PTG
Removal after
1, 3, or 7 Days
RNA Isolation,
Assessment,
Library Prep,
Sequencing
Adult Male, Brown
Morph Crickets,
3-5 days old
Deafferentation:
Removal of Left
Prothoracic Leg
Mid-Femur
Foot Chop Control:
Removal of Left
Prothoracic
Distal Tarsus
Dye Backfill,
4°C Overnight
1 Day
Deaff.
1 Day
Foot Chop
3 Day
Deaff.
3 Day
Foot Chop
7 Day
Deaff.
7 Day
Foot Chop
18h
Post-Backfill
Not Included in
DE Analysis

Figure 10