



A genome-wide survey reveals a diverse array of enhancers coordinates the *Drosophila* innate immune response

Lianne B Cohen, Tamara Hadzic, Caitlin Sauer, et al.

Genome Res. published online April 15, 2026

Access the most recent version at doi:[10.1101/gr.281432.125](https://doi.org/10.1101/gr.281432.125)

| | |
|---------------------------------|--|
| P<P | Published online April 15, 2026 in advance of the print journal. |
| Accepted Manuscript | Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version. |
| Open Access | Freely available online through the <i>Genome Research</i> Open Access option. |
| Creative Commons License | This manuscript is Open Access. This article, published in <i>Genome Research</i> , is available under a Creative Commons License (Attribution-NonCommercial 4.0 International license), as described at http://creativecommons.org/licenses/by-nc/4.0/ . |
| Email Alerting Service | Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here . |



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Published by Cold Spring Harbor Laboratory Press

1 **A genome-wide survey reveals a diverse array of enhancers coordinates the *Drosophila***
2 **innate immune response**

3

4 Lianne B. Cohen^{1,2}, Tamara Hadzic^{2,3}, Caitlin Sauer^{1,2}, Julia R. Gibbs^{1,2}, Zeba Wunderlich^{1,2,3},

5 4*

6 1. Department of Biology, Boston University, Boston, MA 02215, USA

7 2. Biological Design Center, Boston University, Boston, MA 02215, USA

8 3. Program in Bioinformatics, Boston University, Boston, MA 02215, USA

9 4. Department of Biomedical Engineering, Boston University, Boston MA 02215, USA

10 *Corresponding author: zeba@bu.edu

11

12 **Abstract**

13 To defend against microbes, animals regulate a complex immune response. The *Drosophila*
14 innate immune system deploys a large transcriptional induction of signaling proteins, anti-
15 microbial effectors, and other critical immune factors. This transcriptional response is encoded
16 in enhancers, *cis*-regulatory sequences that modulate gene expression by binding transcription
17 factors (TFs). While enhancers and transcription factor binding sites (TFBS) have been
18 identified for several immune responsive genes in *Drosophila*, most enhancers that regulate
19 immune-induced genes are unknown. By identifying enhancers, we can understand how their
20 composition controls expression and contributes to infection outcome. We employ STARR-seq
21 (Self Transcribing Active Regulatory-Region sequencing) in a hemocyte-like cell line to identify
22 immune-specific enhancers across the *D. melanogaster* genome and perform ATAC-seq in
23 hemocytes extracted from adult flies to assess the chromatin state of these enhancers before
24 and after immune stimulus. We identify hundreds of enhancers responsive to IMD stimulation,
25 one of the two primary immune signaling pathways in *Drosophila*. As expected, immune
26 enhancers are enriched for motifs of Relish, an NF- κ B factor, and Kay/Jra, a bZip heterodimer

27 pair, involved in the Imd and JNK pathways respectively, compared to enhancers active in
28 unstimulated cells. However, when grouping enhancers by their target gene's expression timing
29 or functional role or by the enhancers' chromatin accessibility pre- or post-stimulus, different
30 groups of TFBS motifs are enriched, suggesting distinct regulatory logic for different parts of the
31 immune response. Identification and characterization of the diverse array of enhancers that
32 regulate the innate immune response expands our understanding of how animals fight
33 infections.

34 **Key words**

35 *Drosophila* immunity/Immune-responsive enhancer/STARR-seq

36

37 **Running Title**

38 Genome-wide survey of fly immune enhancers

39

40 **Introduction**

41 When encountering pathogenic microbes, animals must regulate an effective immune response
42 to survive infections. The fruit fly *Drosophila melanogaster* relies on its innate immune system to
43 defend against invading pathogens, without the aid of an adaptive immune system (Lemaitre
44 and Hoffmann 2007; Buchon et al. 2014). From the fat body, a liver-like organ, and hemocytes,
45 *Drosophila* blood cells, flies produce a large transcriptional response, inducing over 1,000 genes
46 upon challenge with bacteria and fungi (De Gregorio et al. 2001; Troha et al. 2018; Ramirez-
47 Corona et al. 2021). The resulting proteins carry out a variety of functions, including killing
48 pathogens, relaying signals within and between cells, redistributing metabolic resources, and
49 balancing the costs and the benefits of a persistent immune response.

50

51 The immune transcriptional response in flies is regulated by several intertwined pathways that
52 modulate the activity of key transcription factors (TFs). The Imd pathway responds to DAP-type
53 peptidoglycan, predominantly from gram-negative bacteria, and activates the NF- κ B TF Relish.
54 The Toll pathway activates an NF- κ B TF, Dif, and is activated by fungi and Lys-type
55 peptidoglycan from gram-positive bacteria (Valanne et al. 2011). The pathways interact in
56 several ways – some genes can be induced by either pathway, heterodimers between Dif and
57 Relish have been reported, and some immune stimuli appear to stimulate both pathways (Tanji
58 et al. 2007, 2010; Troha et al. 2018; Valanne et al. 2010). JNK signaling is triggered by Imd
59 stimulation via Tak1, a TGF β -associated kinase, activating the Kay/Jra heterodimer, also called
60 AP-1 (Tafesh-Edwards and Eleftherianos 2020). Mutation of key genes in these pathways
61 results in loss of immune-responsive transcription and decreased infection survival (De Gregorio
62 et al. 2002; Bond and Foley 2009; Brun et al. 2006). RNA-seq experiments revealed that these
63 pathways can orchestrate temporal and pathogen specificity within the immune response (Troha
64 et al. 2018; Schlamp et al. 2021).

65

66 Transcriptional programs are encoded in enhancers – *cis*-regulatory sequences composed of
67 TF binding sites that, when bound, regulate expression of target genes. Tens of immune
68 regulatory sequences have been identified by searching for NF- κ B and other motifs upstream of
69 genes that encode antimicrobial peptides and other highly induced immune genes (Senger et al.
70 2004; Busse et al. 2007). However, the vast majority of enhancers that control the ~1,000
71 immune-responsive genes are unknown. Identifying immune enhancers is critical to
72 understanding how innate immunity coordinates a complex response upon infection. Flies
73 provide an innate-only immune system to study these pathways, which are conserved not only
74 across insects, but also in mammals (Dushay and Eldon 1998; Yu et al. 2022).

75

76 To identify immune regulated enhancers genome wide, we used STARR-seq (Self Transcribing
77 Active Regulatory-Region sequencing) (Arnold et al. 2013; Muerdter et al. 2015). This high-
78 throughput activity-based assay identifies sequences that act as enhancers to drive their own
79 transcription. We performed this assay in S2* cells, a *D. melanogaster* hemocyte-like cell line
80 that is more immune responsive than its S2 parental line (Samakovlis et al. 1992; Cherbas et al.
81 2011). We focused on the immune response to a gram-negative bacterium, as it induces an
82 immune response in a cell autonomous manner that is reproducible in cell culture. To induce
83 this immune response in S2* cells, we pretreated cells with the steroid hormone 20E, before
84 addition of heat-killed *Serratia marcescens* (HKSM), a gram-negative bacterium. In whole
85 organisms, 20E has been shown to regulate development as well as immunity, and treatment of
86 S2* cells with 20E induces expression of PGRP-LC, a critical receptor in the Imd pathway (Rus
87 et al. 2013). By analyzing the enhancers revealed by STARR-seq, we aim to relate TF binding
88 site content to enhancer and target gene function. And by comparing these data to chromatin
89 accessibility in hemocytes before and after immune stimulation, we aim to characterize the *in*
90 *vivo* chromatin dynamics of immune-responsive enhancers.

91

92 **Results**

93

94 **A STARR-seq assay reveals thousands of enhancers active in immune-stimulated cells**

95 To identify the enhancers that regulate the *Drosophila* innate immune response, we performed a
96 STARR-seq experiment in S2* cells (Samakovlis et al. 1992; Cherbas et al. 2011). Briefly, a
97 genome-wide library of plasmids containing 500-750 bp fragments, under control of the
98 *Drosophila* synthetic core promoter (DSCP), was introduced into cells. Cells were treated with
99 the hormone 20-hydroxyecdysone (20E) and/or heat-killed *S. marcescens* (HKSM) to create
100 three treatment conditions, Control (water and PBS), 20E (20E and PBS) and IMD (20E and
101 HKSM) (Fig. 1A). Samples were processed following the STARR-seq protocol and peaks of

102 enhancer activity were called with the STARRPeaker program (Lee et al. 2020). Peaks from
103 different replicates of the same treatment show greater than 0.95 Pearson's correlation between
104 them. Samples treated in the 20E and IMD treatment groups also show a high level of
105 correlation, distinguishing them from Control treated samples (Fig. 1B).

106

107 To devise a set of enhancers that agreed across the replicates, we created the consensus
108 enhancers, the union of enhancers that overlap in at least 2 of the 3 replicates (Supplemental
109 Fig. S1A). The consensus enhancer set contains 2,388 Control enhancers, 3,080 20E
110 enhancers, and 2,934 IMD enhancers. Enhancers have an average length of 725, 746, and 743
111 bps for Control, 20E and IMD data sets, respectively, with the longest enhancers reaching over
112 2000 bps in length (Supplemental Fig. S1B). The consensus enhancers were used for the
113 remainder of the analysis.

114

115 We compared our consensus enhancers to previously-described enhancers to validate the
116 enhancers identified by this STARR-seq experiment. We first binned enhancers by their
117 genomic region and found the majority (63.1%-64.3%) fall within intronic regions of the genome,
118 with a smaller group (25.3-25.7%) in intergenic regions. Similar genomic distributions of
119 *Drosophila* enhancers were previously found using STARR-seq (Arnold et al. 2013) or enhancer
120 chromatin marks H3K4me1 and H3K27ac (The modENCODE Consortium et al. 2010) (Fig. 1C).
121 Additionally, since S2* are derived from S2 cells, we compared the enhancers in the
122 unperturbed sample (Control) to enhancers identified in unperturbed S2 cells. 72%
123 (1,727/2,388) of Control enhancers from the S2* cells have an overlap of at least 100 bp with
124 STARR-seq enhancers found in S2 cells (Arnold et al. 2013). Overall, the consensus enhancers
125 resemble previously identified *Drosophila* enhancers, in both sequence and genomic regions.

126

127 To determine if we identified enhancers that regulated immune genes, we assigned enhancers
128 to genes. Since the vast majority of *Drosophila* enhancers are close to their target genes, for
129 each gene we assigned enhancers that occurred in the sequence spanning from 15 kb
130 upstream to 5 kb downstream of the gene body, as in (Arnold et al. 2013; Kvon et al. 2014).
131 Previous work has finely mapped the time-course of transcriptional regulation of immune genes
132 in the first 24 hrs of Imd stimulation in whole adult flies (Schlamp et al. 2021) and clustered the
133 genes by their expression timing. We found IMD enhancers for just under half of these immune
134 genes (252/551), fairly evenly distributed across the time clusters (Fig. 1D). When looking at a
135 core list of Imd-associated effector peptides, including antimicrobial peptides, we found IMD
136 enhancers near almost all of these genes (28/32) (Fig. 1E) (Westlake et al. 2024). Our IMD
137 enhancers overlap with Relish and other NF- κ B sites that have been previously identified to
138 regulate immune genes, including *Mtk*, *PGRP-SD*, *CecC*, *AttC*, *AttD*, and *DptB* (Uvell and
139 Engstrom 2003, Senger et al. 2004, Busse et al. 2007). Therefore, we conclude that enhancers
140 found in the IMD set likely control immune genes.

141
142 As an illustration of the data, we found two novel enhancers in the first intron of the transcription
143 factor *CrebA* – one shared across conditions and one that is IMD-specific (Fig. 1F). *CrebA* is
144 upregulated by both the Toll and Imd pathways and is crucial for survival after infection with a
145 diverse set of bacteria because it prevents infection-associated ER stress (Troha et al. 2018).
146 To confirm the IMD-specific enhancer's activity and response to immune stimulus, we
147 expressed GFP under the control of this enhancer and a minimal promoter in S2* cells. We saw
148 an increase in mean GFP expression by flow cytometry in IMD treated cells as compared to 20E
149 and Control cells (Fig. 1G). The features of the enhancers identified by STARR-seq, including
150 the close association of IMD enhancers with known immune genes and validation of an
151 individual enhancer, suggest that our approach has identified a large set of *bona fide*
152 enhancers.

153

154 **Analysis of enhancers reveals a variety of TFs used to activate each category of**
155 **enhancers**

156 To find the TFs controlling immune enhancers, we searched for the motifs that are enriched in
157 the IMD enhancer set compared to the genomic background using i-cisTarget (Imrichová et al.
158 2015). We first identified motifs enriched across the entire set of IMD enhancers, which yielded
159 473 significant motifs. Many of these motifs are similar, so we grouped them into families by
160 sequence similarity using phylogenetic trees and selected the following TF families for further
161 analysis (Supplemental Fig. S2A). First, we selected the TFs with well-established roles in Imd
162 immunity, Relish and the heterodimer pair Kay/Jra (Tafesh-Edwards and Eleftherianos 2020),
163 and ecdysone response, EcR/Usp and Eip74EF, (Rus et al. 2013). Next, we identified enriched
164 motifs for TFs that have roles in hemocyte differentiation: GATA factors, forkhead TFs, and
165 Gcm. GATA motifs can be bound by several family members, so we examined RNA-seq data
166 generated alongside the STARR-seq experiment. *Srp* and *pnr* were both highly expressed in
167 S2* cells and hemocytes in line with their roles in hemocyte development (Spahn et al. 2014;
168 Minakhina et al. 2011) (Supplemental Fig. S3A,B). Likewise, among the forkhead TFs, *jumu* is
169 the most highly expressed and is required for hemocyte differentiation and phagocytosis in
170 larvae (Hao et al. 2018). Gcm is a zinc finger TF that is expressed in embryonic hemocytes and
171 regulates their development (Bazzi et al. 2018).

172

173 To broaden our search for TFs, we used two approaches. First, we focused on motifs that were
174 enriched in IMD enhancers i-cisTarget and whose associated TF is expressed in S2* cells
175 (Supplemental Fig. S3A), but without a reported immune role. These include Trl, the GAGA
176 factor in *Drosophila*, as well as the bZIP TFs, Xbp1 and Xrp1. Second, we searched for motifs
177 enriched in enhancers associated with genes in the four immune time clusters. We saw a
178 number of basic helix-loop-helix (bHLH) motifs enriched in enhancers associated with the late

179 time cluster (Supplemental Fig. S2B). We included SREBP because it is activated by gram-
180 negative bacteria in the intestines and regulated by Imd in adipocytes, and Crp since it was
181 highly enriched via i-cisTarget. The motif for Hnf4, a zinc-finger nuclear hormone receptor
182 expressed in fat bodies and in S2* cells, was also included (Barry and Thummel 2016). While
183 not comprehensive, with this set of 13 TF motifs, we found at least three binding sites of any TF
184 motif in each of the IMD enhancers (Fig. 2A, C), indicating that this TF set has a role in the
185 majority of the 2,934 IMD enhancers.

186

187 Since our set of IMD enhancers also includes those that are constitutively active or induced
188 upon ecdysone addition, to identify TFs with specific immune roles, we compared the density of
189 TFBS between each treatment condition to calculate their enrichment (Fig. 2B). When
190 comparing 20E to the Control, we found the 20E enhancers are most significantly enriched for
191 EcR/Usp sites ($p < 0.05$; Fisher's exact test) as expected, given the role of this complex in the
192 ecdysone response. We also saw enrichment for Trl, Hnf4, GATA, and Gcm sites. In the IMD
193 condition, we expected to find enhancers that are activated from both individual responses to
194 20E and bacteria, as well as synergistic interactions between the two stimuli. Accordingly, in the
195 IMD vs. Control comparison, we found a significant enrichment in motifs for Relish, as well as
196 the sites enriched in the 20E vs. Control comparison. To isolate the response solely to immunity
197 , we compared IMD enhancers to 20E enhancers and found significant enrichments in motifs for
198 Relish and Kay/Jra. In the comparisons to Control enhancers, motif depletion indicates that
199 these sites are more common in enhancers active in the Control condition. Below, we test the
200 hypothesis that some of these TFs play a role in specific subsets of immune enhancers with
201 distinct functions, rather than all IMD enhancers (Fig. 3).

202

203 We hypothesized that if these 13 TFs are functionally regulating the IMD enhancers, we would
204 observe a correlation between the number of TFBS and the enhancer's activity. We observed

205 that as the total number of TFBS per enhancer increases, enhancer activity score generally
206 increases, with statistically significant differences in activity scores for enhancers with <10 vs.
207 11-15 TFBS, 16-20 vs. 21-25 TFBS, and 26-30 vs. 31+ TFBS ($p < 0.05$, Mann-Whitney U test,
208 Fig. 2C; 20E and Control in Supplemental Fig. S4A, B). The top two enriched motifs, Relish and
209 Kay/Jra show a general increase in activity score with more binding sites, although Kay/Jra
210 plateaus around 4 TFBS (Fig. 2D). For the other 11 TFs in IMD enhancers, the relationship
211 between activity score and number of binding sites varies; some TFs show a positive
212 relationship, while others have no correlation (Supplemental Fig. S4C). Overall, we found a
213 modest correlation between the total number of binding sites and activity as shown in Fig. 2C,
214 suggesting a functional role for these TF motifs in enhancer activity.

215

216 **Enhancers TF binding site content varies based on the function of the genes they**

217 **regulate**

218 We hypothesized that different sets of enhancers may be activated by distinct sets of TFs,
219 depending on each set's function, pathway, or timing of expression. To allow for sufficient
220 enhancer sample sizes for these gene subsets, we used all enhancers active in the IMD
221 condition, including as enhancers that are also active in other conditions, and enhancer-centric
222 gene assignments to link these enhancers to potential target genes (see Methods). We then
223 used enrichment analysis to identify TFBS that are more prevalent in enhancers associated with
224 different gene sets.

225

226 First, we grouped enhancers by immune pathways based on their assigned genes (Schlamp et
227 al. 2021). As expected, we saw Relish sites significantly enriched in Imd pathway related
228 enhancers, as well as significant Kay/Jra enrichment in enhancers associated with the JNK
229 pathway and wound response. ROS associated enhancers are enriched for Eip74EF, Gcm and
230 Xbp1 motifs. Wound response enhancers are enriched for hemocyte related TF binding sites:

231 Gcm and Forkhead TFs. We also saw Relish also enriched in Toll pathway related enhancers
232 suggesting some interactions between the Toll and Imd pathways (Fig. 3A). The enrichment of
233 TFBS for Relish and Kay/Jra in their corresponding pathways demonstrates the utility of
234 analyzing the IMD enhancer data set in smaller sets.

235

236 We next hypothesized that the enhancers that regulate genes induced at different times may
237 rely on different TFBS. Since transcripts produced from the STARR-seq plasmid are stabilized
238 via a poly(A) tail, we collected transcripts from enhancers that were active at any point within the
239 first 24 hrs of immune induction. Using the time-course data from Schlamp *et al.* 2021, we
240 labeled enhancers with different time clusters based on the expression of their assigned genes.
241 We found that Relish sites are enriched in enhancers that regulate genes expressed early to
242 mid immune induction (0-16 hrs) (Fig. 3B). Late expressing genes, 12-24 hrs, are enriched for
243 bHLH sites, SREBP and Crp, as well as GATA sites. SREBP is an activator of lipogenesis,
244 typically triggered by low lipid levels, including during infection (Charroux and Royet 2022).
245 Since Imd activation depletes lipid stores, its late activation may be in response to a metabolic
246 shift. Therefore, this analysis reveals that there is distinct regulatory logic at different stages of
247 the immune response, with early-acting enhancers regulated by the canonical Imd-responsive
248 TF, Relish and later acting enhancers relying on bHLH activators, suggesting a mechanism for
249 how genes are temporally regulated.

250

251 To determine if the role of a gene, regardless of pathway, affects how it is regulated, we
252 investigated the composition of TFBS using two sets of gene annotations. Using broad immune
253 role annotations, e.g. antiviral response (*Dcr-2*, *AGO2*), recognition (*PGRP* genes, *GNBP*
254 genes), signaling (*Relish*, *Dif*, *key*), or effector (AMPs, *Lysozyme* genes, *Bomanin* genes),
255 (Schlamp *et al.* 2021), we found effectors are controlled significantly by several TFs: Crp, Gcm,
256 Hnf4, and Relish. The signaling and recognition stages of the immune response are also

257 significantly controlled by Relish (Fig. 3C). We also used more detailed functional annotations
258 from a comprehensive set of genes involved in the response to infection (Westlake et al. 2024).
259 We saw several TFs enriched within functional groups. We found that Relish is the most highly
260 enriched transcription factor in enhancers assigned to effectors, followed by Hnf4 (Fig. 3D).
261 Relish and GATA motifs are enriched among enhancers associated with hematopoiesis. Xbp1,
262 a known component to the unfolded protein response (UPR) (Huang et al. 2017), is enriched in
263 both reactive oxygen species (ROS) and phagocytosis related enhancers. Overall, we find that
264 our set of immune-induced enhancers encompasses many facets of the immune response,
265 including timing, pathway and gene function.

266

267 **Enhancers can be further divided into activity classes**

268 Up to this point, we have analyzed all the enhancers active in the IMD condition together, but
269 some enhancers are active in the Control and/or 20E conditions, while others are only active
270 upon IMD stimulation. To discern between these types of enhancers, we defined seven *activity*
271 *classes*: Constitutive, Control Only, 20E Only, IMD Only, Control + 20E, Control + IMD, and IMD
272 + 20E. We found that the Constitutive enhancers comprised the largest activity class with 1,344
273 enhancers; the second largest group was IMD + 20E enhancers with 1,106 enhancers. Many
274 enhancers are likely active in both the 20E and IMD conditions because both treatments include
275 20E, but we also found 372 enhancers that are active only in the IMD condition (Fig. 4A). We
276 found relatively few enhancers active in the Control + IMD and Control + 20E groups, which is
277 biologically sensible – since the 20E and IMD conditions share the 20E stimulus, we expect
278 most enhancers active in the 20E condition to also be active in the IMD condition. To investigate
279 the contributions of each TF in determining an enhancer’s activity class, we trained two logistic
280 regression models on either all or a subset of the enhancers and found a modest ability to
281 discern between activity classes based on TF motif content (Supplemental Fig. S5). We saw

282 Relish, Trl, and Kay/Jra binding sites contribute the most to IMD Only enhancers, while
283 EcR/Usp binding sites contribute most to enhancers active in both IMD and 20E conditions.
284
285 Using these activity class assignments, we tested a proposed grammar between Relish and
286 Serpent (GATA) sites in immune-responsive enhancers (Supplemental Fig. S5E). Previous work
287 found that Relish and Serpent sites are often within 50 bps of each other and positioned in the
288 same relative orientation in the promoter proximal regions of ~50 immune genes. We found that
289 Relish and GATA sites are closer to each other in IMD only and IMD + 20E enhancers than
290 Constitutive enhancers ($p < 0.001$, two-sided Mann-Whitney U test), and the density of Relish-
291 GATA pairs is higher in IMD only enhancers compared to Constitutive enhancers (Supplemental
292 Fig. S5F, G). However, there was not a significantly higher proportion of more Relish-GATA
293 pairs in the same orientation in IMD Only and IMD + 20E enhancers compared to Constitutive
294 enhancers ($p = 0.156$, two proportion z-test). Therefore, our data support the idea that closely
295 spaced Relish-GATA pairs are associated with immune-responsiveness, but there was less
296 support for the importance of the relative orientation of these sites. .

297

298 **Enhancers' activity classes are associated with their target genes' immune role**

299 The assignment of enhancers to activity classes allowed us to make more refined gene-
300 enhancer assignments based on RNA-seq data taken alongside the STARR-seq data. We
301 matched enhancers in each activity class to genes by first identifying genes expressed in the
302 conditions corresponding to the activity class. For example, for the Constitutive class, we
303 considered genes expressed in the Control, 20E, and IMD conditions, while for the IMD Only
304 class, we considered genes expressed in the IMD condition, regardless of their expression in
305 other conditions (see Methods and Supplemental Fig. S6A). We did not restrict, for example,
306 IMD Only enhancers to genes only expressed in the IMD condition because it is possible that
307 the gene is controlled by multiple enhancers with different activity, e.g. one IMD-inducible

308 enhancer and a second constitutive enhancer (Fig. 1F). We also do not restrict IMD Only
309 enhancers to genes *differentially* expressed in the IMD condition because it is possible that a
310 gene is controlled by a IMD-inducible enhancer and another enhancer active only in the control
311 condition. Using these expressed genes, we then made gene-enhancer assignments using the
312 same 15 kb upstream/5 kb downstream windows as before. This approach filtered out 6,832
313 gene-enhancer pairs with discordant expression-activity patterns and retained 8,518 gene-
314 enhancer assignments (Supplemental Fig. S6B, C). In addition to multiple enhancers assigned
315 to a gene, we also found multiple genes assigned to a single enhancer (Supplemental Fig.
316 S6D). There is an increasing number of examples in fly and mammalian genomes where a
317 single enhancer can control multiple genes making these assignments reasonable (Levo et al.
318 2022, Mohrs et al. 2001). As an orthogonal check, we found genes differentially expressed
319 between the conditions were assigned to enhancers in the appropriate activity class
320 (Supplemental Fig. S6E). Therefore, we have confidence that filtering for genes based on
321 expression yields a high-quality set of gene-enhancer assignments.

322
323 We hypothesized that genes with different roles in the immune system might be controlled by
324 enhancers with different activity classes. We compared the enhancers controlling immune
325 recognition, signaling, and effector genes (Fig. 4B) and found that effector genes had the largest
326 proportion (38%) of IMD Only enhancers, significantly more than the proportion of IMD Only
327 enhancers among all enhancers ($p < 0.05$, one proportion z-test). Signaling genes had the
328 largest proportion of Constitutive enhancers (44%), slightly larger than all enhancers, ($p < 0.05$,
329 one proportion two-sample z-test), likely because these components have more consistent
330 expression across conditions. We also compared the enhancer activity classes for genes sorted
331 by their time of induction and found that genes induced in the Early and Mid time clusters have
332 significantly more IMD Only enhancers than genes induced later in the infection (Fig. 4C).
333 Although the later time clusters include genes that are immune induced, they may be controlled

334 by a larger proportion of Constitutive enhancers for several reasons. For example, some genes
335 may be activated by secondary signals, e.g. changes in metabolism, that have roles in other
336 pathways. Other genes may be expressed in tissues not well-modeled by the hemocyte-like S2*
337 cells. This analysis indicates that, depending on their immune role or expression timing, specific
338 gene groups are associated with different enhancer activity classes.

339

340 This analysis also allows us to investigate the distribution of enhancer-gene distances. We
341 hypothesized that Constitutive enhancers were more likely to overlap with the transcription start
342 site (TSS) of their target genes, since distal enhancers are often found to be condition- or
343 tissue-specific (Zabidi et al. 2015). Consistent with our hypothesis, Constitutive enhancers
344 overlapped with the TSS of their target genes (the 0 bp bin) somewhat more often than the
345 inducible IMD + 20E enhancers ($p < 0.05$, 2-sided z-test) (Fig. 4D). But overall, we saw little
346 difference in the full distribution of gene-enhancer distances based on activity class, with more
347 than 60% of enhancers within 10,000 bps of the TSS of their target gene and 90% of enhancers
348 within 15,000 bps.

349

350 **Individual enhancer reporters validate activity classes**

351 To further validate the STARR-seq measurements, we tested the activity of eleven enhancer
352 reporters (Fig. 5A). Enhancers are named for the gene with the closest TSS that is expressed in
353 the same conditions as the enhancer (with the exception of *Dso* enhancer, see Supplemental
354 Fig. S7A). We chose nine immune-responsive enhancers from two groups: enhancers that
355 aligned with previously identified immune enhancers and novel enhancers that were in the IMD
356 Only or IMD + 20E enhancer classes. We included one enhancer (*CG9837*) in the Constitutive
357 class. We introduced the reporters into S2* cells individually and measured each reporter's
358 activity Control, 20E, and IMD conditions using flow cytometry (Fig. 5B). As expected, the
359 reporters for the *PGRP-SD* and *Mtk* enhancers, which closely match previously identified

360 enhancers from Busse et al. and Senger et al., respectively, significantly upregulate GFP upon
361 IMD stimulation, as compared to the Control condition. Among the remaining seven reporters in
362 the IMD Only or IMD + 20E classes, six enhancers upregulate transcription upon IMD
363 stimulation (*CrebA*, *lectin-26Cb*, *RpS26*, *Lmpt*, *CG5758*, *Cog8*), and one enhancer does not
364 (*Dso*). The *CG9837* enhancer reporter drives high activity in both the Control and IMD
365 conditions (Fig. 5D), validating that it is a strong constitutive enhancer, and the negative control
366 (DSCP) drove low expression in both conditions, indicating that it is a minimal, weak promoter.
367

368 We were surprised to find an IMD-induced enhancer between *Dso1* and *Dso2* genes, since
369 these genes are induced by the Toll pathway (Levy et al. 2004; Cohen et al. 2020,
370 Supplemental Fig. S7A). The *Dso* enhancer activates expression in response to Toll stimulus,
371 either on its own or when both the IMD and Toll pathways are induced in the Dual condition
372 (Fig. 5C), indicating this is a Toll-responsive enhancer. Five of the eight IMD-responsive
373 enhancers, *Mtk*, *CrebA*, *CG5758*, *lectin-46Cb* and *RpS26*, also respond to Toll and Dual
374 induction beyond the negative control (Supplemental Fig. S7B, C), indicating that some
375 enhancers can respond to both the IMD and Toll pathways. Synthesizing the data across these
376 experiments, we see several groups of enhancers: those that are only IMD-responsive (*PGRP-*
377 *SD*, *Lmpt*, *Cog8*), those that are only Toll-responsive (*Dso*), those that are Toll- and IMD-
378 responsive (*Mtk*, *CrebA*, *CG5758*, *lectin-46Cb* and *RpS26*), and those that are constitutive
379 (*CG9837*), and find that STARR-seq identified enhancers generally recapitulate their activity in
380 individual enhancer-reporter constructs.

381

382 **Most enhancers do not change chromatin accessibility upon immune stimulation**

383 The STARR-seq assay identifies sequences that can drive expression in the cell type and
384 conditions tested, independent of chromatin context. In a whole organism, both sequence
385 information and chromatin structure determine an enhancer's ability to regulate gene

386 expression. To compare STARR-seq enhancers to the chromatin state in the animal itself, we
387 conducted ATAC-seq in hemocytes isolated from adult flies in either control or immune-
388 stimulated conditions. To mimic the cell culture induction and avoid contaminating bacterial DNA
389 in the ATAC-seq library, we performed the immune induction using heat-killed *S. marcescens*
390 (HKSM) and measured chromatin accessibility 24 hours after treatment (Fig. 6A).

391
392 In the original STARR-seq paper, roughly half of the enhancers were found to be accessible in
393 the genome of the same S2 cell line used for the assay (Arnold et al. 2013). Here, we compared
394 STARR-seq enhancers to the equivalent cell type in intact animals, and therefore expected that
395 the overlap between enhancers and open chromatin might be slightly lower. However, we found
396 that half (49.8%) of all STARR-seq identified enhancers are in regions that are accessible in
397 either or both conditions in hemocytes (Supplemental Fig. S8A). We then divided enhancers
398 into accessibility groups: those that are open in both control and stimulated conditions (always
399 open), those closed in both conditions (always closed), opened by HKSM stimulus, or closed by
400 HKSM stimulus. More than >90% of enhancers maintain their chromatin structure for the
401 duration of the experiment, either remaining open or closed (Fig. 6B). In enhancers grouped by
402 activity class, we found that Constitutive enhancers are more likely to be always open than in
403 IMD Only and IMD + 20E enhancers (Fig. 6C; $p < 0.002$, z-test, Bonferroni multiple test
404 correction). The fraction of enhancers opened by HKSM is marginally higher in the IMD + 20E
405 enhancers than Constitutive enhancers ($p = 0.099$, z-test). Similarly, when enhancers are
406 grouped by treatment, there is a slight increase in HKSM opened enhancers in 20E enhancers
407 vs. Control ($p = 0.055$, z-test) (Supplemental Fig. S8A). We next asked whether the chromatin
408 state of enhancers differed by the immune role of the associate target genes. Enhancers
409 assigned to effectors or signaling components have a larger proportion of enhancers that are
410 always in open chromatin compared to all enhancers, ($p < 0.01$, z-test, Bonferroni multiple test
411 correction) (Fig. 6D). Enhancers assigned to early and late expressing immune genes also have

412 a higher likelihood of always lying in open chromatin ($p < 0.05$, $p < 0.001$, z-test, Bonferroni
413 multiple test correction). In general we find that, regardless of activity class, enhancers assigned
414 to immune genes tend to be in open chromatin regions more than all enhancers. Overall, the
415 chromatin state of most enhancers identified via STARR-seq is unchanged in the 24 hours after
416 treatment with HKSM.

417
418 To understand what determines the accessibility of different groups of enhancers, we measured
419 the density of the 13 TF motifs previously selected (Fig. 2) in each accessibility group. We found
420 that compared to all enhancers, enhancers that are opened upon HKSM treatment are enriched
421 for EcR/Usp, Hnf4, and Trl TFBS (Fig. 6E). Trl, a pioneer TF known to open closed chromatin
422 (Chetverina et al. 2021), is strongly enriched in HKSM opened and always open enhancers and
423 strongly depleted in HKSM closed enhancers. EcR/Usp was found to be significantly enriched in
424 the HKSM opened enhancers, but not in any other accessibility group. Previous work has
425 identified Hnf4 as a key regulator of metabolism and lipid storage in development (Barry and
426 Thummel 2016; Storelli et al. 2019). Here we saw its binding sites enriched in enhancers that
427 either opened or closed upon immune stimulation, possibly indicating that Hnf4 is performing its
428 metabolic switch role in reaction to infection. Relish is not enriched in any accessibility group,
429 indicating that its role is primarily activating transcription without influencing chromatin state. We
430 looked for additional motifs enriched in the accessibility groups through i-cisTarget and found
431 motifs for CrebA, an immune TF (Troha et al. 2018), and a putative zinc finger TF differentially
432 enriched between the accessibility groups (Supplemental Fig. S8B). In sum, we see that TFBS
433 content is correlated with accessibility of enhancers and that Trl and EcR/Usp may play a role in
434 opening chromatin for enhancers in response to immune stimulation.

435

436

437

438 **Discussion**

439 In this study, we have identified immune responsive enhancers across the entire *Drosophila*
440 genome. Immune enhancers are enriched in motifs for Relish and Kay/Jra, the terminal
441 transcription factors of the Imd and JNK pathways, respectively. We also find that distinct
442 classes of immune enhancers differ by their TFBS composition and activity class. While most
443 enhancers' accessibility does not change upon infection in hemocytes, enhancers that are
444 opened upon immune stimulation may be controlled by ecdysone signaling with an enrichment
445 of EcR/Usp binding sites.

446

447 As was expected, Relish is the most enriched TF in immune responsive enhancers, since the
448 Imd pathway culminates in its activation. Here, we determine that Relish is a general activator of
449 all Imd enhancers, but is preferentially found in the immune specific enhancers of early
450 response genes and effectors. The Imd pathway activates the JNK pathway and concordantly,
451 we find Kay/Jra binding sites enriched in immune enhancers and particularly in late acting
452 enhancers (Tafesh-Edwards and Eleftherianos 2020).

453

454 The major pathways controlling the humoral immune response, Toll and Imd, are each activated
455 by different microbes, but there has been evidence of interactions between the pathways. For
456 example, wounding and damage associated signals activate both immune pathways (Troha et
457 al. 2018; Schlamp et al. 2021). Heterodimers between Relish and Dif, an NF- κ B protein
458 downstream of Toll, can form and activate transcription (Tanji et al. 2007, 2010; Morris et al.
459 2016). Also, certain genes like *Mtk*, *Drs*, and *CrebA* can be activated by both Toll and Imd
460 (Busse et al. 2007; De Gregorio et al. 2002; Troha et al. 2018; Valanne et al. 2010). Consistent
461 with this observation, we see that both Imd and Toll activate the enhancers for *Mtk* and *CrebA*.
462 But, there is no evidence of synergistic enhancer activity when both stimuli are present
463 (Supplemental Fig. S7B). Regulation by both pathways could be encoded in enhancers in

464 multiple ways. The enhancers could contain distinct binding sites for Relish and Dif, common
465 sites that can be bound by Relish or Dif, or sites that can be bound by Relish-Dif heterodimers.
466 A previously-studied *Mtk* enhancer has Relish-specific sites and a common site that binds both
467 Dif and Relish (Busse et al. 2007). We also identified one Dif-specific site in our expanded *Mtk*
468 STARR-seq enhancer. In addition to three Relish-specific sites, the *CrebA* enhancer contains a
469 “GGGAATTCT” site, which closely resembles a common NF- κ B site, and a “GGGAACCACT”
470 site, a 1 bp mismatch with the Relish-Dif heterodimer site (Tanji et al. 2007, 2010). Relish’s
471 enrichment among the enhancers of Toll pathway components (Fig. 3A) suggests that other
472 immune enhancers integrate signals from multiple pathways. Overall, we find additional
473 evidence that some immune enhancers can be activated by both Toll and Imd pathways via
474 distinct Relish and Dif sites, common binding sites, and heterodimer sites.

475
476 Beyond the roles for Relish and Kay/Jra, we find the patterns of motif enrichment for Xbp1, Trl,
477 Crp, SREBP, and Hnf4. Xbp1 is enriched in enhancers associated with ROS and phagocytosis.
478 Upon infection, a biphasic ROS response stimulates hemocytes to differentiate (Myers et al,
479 2018). Perhaps, an increase in secreted peptides during an immune response activates the
480 UPR and Xbp1 could in turn regulate the ROS and phagocytic pathways. Trl motifs were not
481 enriched in enhancers associated with any gene classes based on immune function or immune
482 expression timing (Fig. 3), but were enriched in both always open and HKSM opened enhancers
483 in hemocytes (Fig. 6E). This suggests that Trl’s role in the immune response is remodeling the
484 chromatin landscape, consistent with its role in development (Gaskill et al. 2021). Crp, a known
485 anti-apoptosis factor (Sopko et al. 2015; Atkins et al. 2016), could function to maintain the
486 hemocyte population upon infection, with motifs enriched in mid- and late-expressed immune
487 genes and among enhancers associated with effector and signaling genes. Both Hnf4 and
488 SREBP have roles in regulating lipid biosynthesis and metabolism, and *SREBP* is
489 downregulated in *Hnf4* mutants, suggesting a link between these genes (Vonolfen et al. 2024).

490 There is evidence that these TFs are differentially regulated during infection or inflammation
491 from studies that focus on different cell types (Charroux and Royet 2022). Given the large
492 metabolic impact of the immune response (Bland 2023), further studies into the activity of these
493 TFs in different tissues and stages of infection may reveal how expression is regulated after the
494 immediate response driven by canonical immune-responsive TFs.

495
496 Our data emphasizes the significant role of the hormone ecdysone in modulating the
497 transcriptional immune response. Ecdysone is the master regulator of *Drosophila*
498 developmental transitions, but it is also required for the Imd response in both cells and whole
499 animals (Meister and Richards 1996; Rus et al. 2013). Ecdysone induces the expression of the
500 peptidoglycan receptor PGRP-LC, increasing the responsiveness of cells to gram-negative
501 bacteria. We find that just under half (45.3%) of IMD + 20E enhancers contain both EcR/Usp
502 and Relish motifs. The co-occurrence of Relish and EcR/Usp sites suggest that ecdysone also
503 has a role in activating Relish-dependent immune enhancers. Additionally, chromatin
504 accessibility changes in hemocytes from adult flies upon immune stimulus reveal that EcR/Usp
505 is highly enriched in enhancers opened by HKSM. EcR/Usp recruits TRR and the MLR
506 COMPASS complex, which monomethylates H3K4, priming enhancers for an immune response
507 (Zraly et al. 2021). Therefore, we find that ecdysone and EcR/Usp are crucial regulators of
508 transcriptional immune response likely by both increasing enhancer activation and chromatin
509 accessibility.

510
511 Our *in vivo* ATAC-seq experiments in hemocytes revealed that roughly 90% of enhancers do
512 not qualitatively change their accessibility status during the course of the experiment, while the
513 remaining ~10% change. Though there may be quantitative changes in accessibility, these
514 changes do not necessarily correlate with expression (Merrill et al. 2022). This relatively stable
515 chromatin profile is in line with the model that the majority of enhancers available to a cell type

516 acquire their enhancer histone marks (e.g. H3K4me1) and accessibility during lineage
517 specification, while a subset, so-called “latent enhancers,” acquire these marks and accessibility
518 after a stimulus. In human macrophages, ~15% of enhancers are latent and only gain enhancer
519 histone marks upon stimulus (Ostuni et al. 2013). We suggest our HKSM opened enhancers are
520 also latent enhancers (Fig. 6B).

521
522 Although we identified hundreds of immune-responsive enhancers and distinct patterns of TF
523 motif enrichment within them, there are some limitations to this study. Though powerful,
524 STARR-seq may yield both false positives, e.g. due to the lack of chromatin context for the
525 reporter plasmids and because electroporation may trigger cellular responses, and false
526 negatives, e.g. due to limited reporter size or promoter choice. In our study, 9 of 10 individual
527 enhancers reporter measurements aligned with their high throughput STARR-seq results. We
528 found one false positive: the *Dso* enhancer appeared as a IMD + 20E enhancer in the STARR-
529 seq data, but the individual reporter was not responsive to Toll but not IMD induction. We also
530 find evidence of false negatives: previously described IMD-responsive enhancers upstream of
531 *Dpt*, *CecA*, *PGRP-SB1*, and *AttA* are not part of the STARR-seq data (Busse et al 2007, Uvell
532 and Engstrom 2003, Senger et al 2004). The lack of signal in these regions could be due to
533 differences in induction protocols or STARR-seq limitations. Lastly TFBS analysis was based on
534 13 selected motifs. The accuracy of TF assignments depends on the quality of the motif, and we
535 sometimes cannot distinguish between the binding of closely related TF paralogs. Furthermore,
536 enhancers are likely regulated by other transcription factors, including those with poor or
537 missing motifs.

538
539 The diverse array of enhancers characterized in this study reflects the demands of an innate
540 immune system. In animals without an adaptive immune system, the innate transcriptional
541 response must encompass everything needed to defend against invading pathogens, while

542 balancing a limited metabolic budget and maintaining housekeeping functions. Identification of
543 immune enhancers across both the expanse of the genome and functional roles of the immune
544 system adds to our understanding of how the immune response is controlled and the multiple
545 types of enhancers involved.

546

547 **Methods**

548 Genome-wide reporter library construction

549 To generate a genome-wide reporter library, genomic DNA was isolated from both male and
550 female *iso-1* flies, fragmented via sonication, and size selected (500-750 bps). To prepare and
551 clone these fragments into the the pSTARR-seq fly plasmid (AddGene, #71499), we used the
552 protocol outlined in Neumayr, *et al* (Supplemental Methods).

553

554 STARR-seq in S2* cells

555 To test the genomic library's enhancer activity, we performed STARR-seq in S2* cells. S2* cells
556 were a gift from Steven Wasserman and grown according to standard methods in complete
557 Schneider's media at 28°C. For a three-condition experiment, 2.8×10^9 S2* cells were
558 electroporated with the STARR-seq library (see Supplemental Methods). Cells were split into
559 three treatment groups: Control, 20E and IMD, treated with either water (Control) or a final
560 concentration of 40 nM 20E (Sigma-Aldrich) (20E and IMD) for 24 hrs, and then PBS (Control
561 and 20E) or heat-killed *S. marcescens* at a final concentration of 0.4 OD (IMD) for 24 hrs. RNA
562 extraction and library preparation were completed as in Neumayr et al. 2019. Paired end
563 Illumina sequencing was performed for 3 biological replicates and input library.

564

565 Alongside the STARR-seq replicates, RNA-seq samples were created in triplicate.

566 Untransfected S2* cells were treated identically to STARR-seq samples. Total RNA was sent to
567 Genewiz for poly(A) mRNA selection and library preparation.

568

569 Computational analysis570 **Identifying STARR-seq peaks**

571 STARR-seq reads were aligned to the dm6 genome using Bowtie 2 and SAMtools (Langmead
572 and Salzberg 2012, Danecek et al 2021). deepTools was used to convert alignment files to
573 bigWig coverage tracks (Ramírez et al. 2016). STARRPeaker was used to call peaks (p -value
574 threshold < 0.05) (Lee et al. 2020). Each peak's activity score ($100 \times (\text{output}/\text{normalized-input})$),
575 where output is the number of reads in the peak, and normalized-input is the coverage of the
576 peak in the input library. Enhancers can be larger than the maximum library fragment size (750
577 bp) since STARRpeaker identifies peaks that may come from multiple overlapping fragments.
578 We built consensus peaks of the replicates in each treatment group. If a peak was in at least
579 two of three replicates (with at least 1 bp overlap), the consensus enhancer was defined as the
580 union of the peaks that occur at that location and assigned the average activity score from the
581 peaks that contribute to them.

582

583 Since many enhancers appear across multiple treatment conditions, we defined seven activity
584 classes: Constitutive (Control, 20E, IMD), Control Only, 20E Only, IMD Only, IMD and 20E, 20E
585 and Control, and IMD and Control. Enhancers were merged via a union operation across
586 conditions if they overlapped by at least 60%.

587

588 **RNA-seq analysis**

589 RNA-seq reads were aligned with Bowtie 2 and counts were generated with Subread's
590 featureCounts (Langmead and Salzberg 2012, Liao et al. 2014). To identify any outliers,
591 principal component analysis (PCA) was completed. Sample 1, a control sample, was
592 determined to be an outlier and removed, leaving 2 control samples, 3 20E samples, and 3 IMD
593 samples.

594

595 Transcription factor assignments

596 To identify TFBS within all the IMD enhancers and within time-specific sets of IMD enhancers,
597 we used i-cisTarget (version 6.0) to find motifs with a normalized enrichment score above 3.0
598 (Imrichová et al. 2015). Motifs were grouped into motif trees using the R Bioconductor package
599 universalmotif (Tremblay 2024). TFBS were identified on both DNA strands with the FIMO tool
600 of the MEME suite using p -value $< 10^{-3}$ (Grant et al. 2011).

601

602 Gene assignments

603 Gene centric enhancer assignments (used in Fig. 1) were made by scanning 15 kb upstream
604 and 5 kb downstream of a gene and assigning any enhancer in this window to the gene, as in
605 Arnold, et al. Since enhancers function independent of orientation, enhancer-centric
606 assignments (Fig. 3) were made by scanning 15 kb upstream to 15 kb downstream of an
607 enhancer and assigning any genes in this window to the enhancer. More restrictive enhancer-
608 gene assignments were made by sorting genes into activity classes based on expression in the
609 RNA-seq data (expression threshold: TPM > 1). For each activity class, genes were matched to
610 the allowed set of enhancers by scanning 15 kb upstream and 5 kb downstream of the gene
611 (Supplemental Methods).

612

613 Construction and flow cytometry of enhancer reporters

614 Enhancers were amplified from *iso-1* flies and cloned upstream of the Drosophila Synthetic
615 Core Promoter (DSCP) (Pfeiffer et al. 2008; Arnold et al. 2013) driving EGFP with the origin of
616 replication and AmpR gene from pAc5.1/V5-His-A (Invitrogen). See Supplemental Table 1 for
617 regions cloned and Supplemental Methods for additional details.

618

619 To allow for Toll stimulation, plasmids were electroporated into S2* cells stably transfected with
620 pMT-Torso-pelle, which contains Pelle, a Toll pathway component fused to the transmembrane
621 domain of Torso and under control of the CuSO₄-activated metallothionein promoter (Galindo et
622 al. 1995, Silverman et al. 2000). Cells were split in five treatment conditions: Control, 20E, IMD-
623 induced, Toll-induced, or Toll- and IMD- (dual) induced (Supplemental Methods).

624

625 ATAC-seq with adult fly hemocytes

626 We injected 60-70 2-7 day old male OregonR flies with ~50 nL of heat-killed *Serratia*
627 *marcescens* at OD=0.5 in PBS and incubated them at 25°C for 24 hrs, alongside uninjected
628 age-matched controls. To quickly extract hemocytes to ensure ATAC-seq data quality,
629 centrifugation was used (Cabrera et al 2023, Jones and Eleftherianos 2023). Hemolymph
630 including hemocytes was collected by placing 20-25 anesthetized flies in 0.5 mL tubes with 3
631 20G needle holes at the bottom and then covered with glass beads. Flies were spun at 6,000×g
632 at 4°C for 1 min into ATAC-seq lysis buffer (Grandi et al. 2022). Tagmentation and library prep
633 protocols were adapted from Grandi et al. 2022 (see Supplemental Methods). Paired end
634 Illumina sequencing was completed on three biological replicates per condition.

635

636 ATAC-seq analysis

637 ATAC-seq reads were trimmed using cutadapt (Martin 2011), aligned with Bowtie 2 (Langmead
638 and Salzberg 2012). Peaks were called using MACS2 (Zhang et al. 2008), using narrow peak
639 calling and a *p*-value of 0.01. Consensus ATAC-seq peaks were built from the widest margins of
640 peaks present in at least 2 replicates. STARR-seq enhancers were labeled open if >50 bps of
641 enhancer were in consensus ATAC-seq peaks.

642

643

644

645 **Data Access**

646 All raw and processed sequencing data generated in this study have been submitted to the
647 NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) under accession
648 number GSE308695.

649

650 **Competing Interest Statement**

651 The authors declare no competing interests.

652

653 **Acknowledgements**

654 We thank Benedetta D'Elia for help in developing the ATAC-seq protocol and Ila Rosen for help
655 in preliminary plasmid construction. We thank Alexander Stark, Cosmas Arnold, Bernardo de
656 Almeida, and Rui Catarino for helpful advice on the STARR-seq protocol. We thank Steven
657 Wasserman for the S2* cells. This work was funded by NSF Award 2223888 (to ZW). TH was
658 supported by NIH Award T32GM150533 and CS was supported by Boston University's
659 Undergraduate Research Opportunities Program and the Boston University STEM Pathways
660 Program (Department of Defense DoD STEM FY20 Award HQ00342110008). The content is
661 solely the responsibility of the authors and does not necessarily represent the official views of
662 any of the funders. Figures created with BioRender.com

663

664 Author Contributions: L.B.C. and Z.W. conceived the project and wrote the paper. L.B.C.
665 performed the experiments and analyzed the data except for the following: T.H. developed
666 activity classes and expression filtered gene assignment. C.S. created and tested the enhancer
667 reporters. J.R.G. analyzed ATAC-seq data. L.B.C. and Z.W. supervised all experiments and
668 analysis. Z.W. obtained funding for the project. All authors have read, edited, and approved the
669 final manuscript.

670 **References**

- 671 Arnold, C. D., Gerlach, D., Stelzer, C., Boryń, Ł. M., Rath, M., & Stark, A. (2013). Genome-wide
672 quantitative enhancer activity maps identified by STARR-seq. *Science*.
673 <https://doi.org/10.1126/science.1232542>
- 674 Atkins, M., Potier, D., Romanelli, L., Jacobs, J., Mach, J., Hamaratoglu, F., Aerts, S., & Halder,
675 G. (2016). An Ectopic Network of Transcription Factors Regulated by Hippo Signaling
676 Drives Growth and Invasion of a Malignant Tumor Model. *Current Biology*.
677 <https://doi.org/10.1016/j.cub.2016.06.035>
- 678 Barry, W. E., & Thummel, C. S. (2016). The Drosophila HNF4 nuclear receptor promotes
679 glucose-stimulated insulin secretion and mitochondrial function in adults. *ELife*.
680 <https://doi.org/10.7554/eLife.11183>
- 681 Bazzi, W., Cattenoz, P. B., Delaporte, C., Dasari, V., Sakr, R., Yuasa, Y., & Giangrande, A.
682 (2018). Embryonic hematopoiesis modulates the inflammatory response and larval
683 Hematopoiesis in Drosophila. *ELife*. <https://doi.org/10.7554/eLife.34890>
- 684 Bland, M. L. (2023). Regulating metabolism to shape immune function: Lessons from
685 Drosophila. In *Seminars in Cell and Developmental Biology*.
686 <https://doi.org/10.1016/j.semcdb.2022.04.002>
- 687 Bond, D., & Foley, E. (2009). A quantitative RNAi Screen for JNK Modifiers Identifies Pvr as a
688 Novel Regulator of Drosophila Immune Signaling. *PLoS Pathogens*.
689 <https://doi.org/10.1371/journal.ppat.1000655>

- 690 Brun, S., Vidal, S., Spellman, P., Takahashi, K., Tricoire, H., & Lemaitre, B. (2006). The
691 MAPKKK Mekk1 regulates the expression of Turandot stress genes in response to septic
692 injury in *Drosophila*. *Genes to Cells*. <https://doi.org/10.1111/j.1365-2443.2006.00953.x>
- 693 Buchon, N., Silverman, N., & Cherry, S. (2014). Immunity in *Drosophila melanogaster*-from
694 microbial recognition to whole-organism physiology. In *Nature Reviews Immunology*.
695 <https://doi.org/10.1038/nri3763>
- 696 Busse, M. S., Arnold, C. P., Towb, P., Katrivesis, J., & Wasserman, S. A. (2007). A κ B
697 sequence code for pathway-specific innate immune responses. *EMBO Journal*.
698 <https://doi.org/10.1038/sj.emboj.7601798>
- 699 Cabrera, K., Hoard, D. S., Gibson, O., Martinez, D. I., & Wunderlich, Z. (2023). *Drosophila*
700 immune priming to *Enterococcus faecalis* relies on immune tolerance rather than
701 resistance. *PLoS pathogens*, 19(8), e1011567.
702 <https://doi.org/10.1371/journal.ppat.1011567>
- 703 Charroux, B., & Royet, J. (2022). Gut-derived peptidoglycan remotely inhibits bacteria
704 dependent activation of SREBP by *Drosophila* adipocytes. *PLoS Genetics*.
705 <https://doi.org/10.1371/journal.pgen.1010098>
- 706 Cherbas, L., Willingham, A., Zhang, D., Yang, L., Zou, Y., Eads, B. D., Carlson, J. W., Landolin,
707 J. M., Kapranov, P., Dumais, J., Samsonova, A., Choi, J. H., Roberts, J., Davis, C. A.,
708 Tang, H., Van Baren, M. J., Ghosh, S., Dobin, A., Bell, K., ... Cherbas, P. (2011). The
709 transcriptional diversity of 25 *Drosophila* cell lines. *Genome Research*.
710 <https://doi.org/10.1101/gr.112961.110>

- 711 Chetverina, D., Erokhin, M., & Schedl, P. (2021). GAGA factor: a multifunctional pioneering
712 chromatin protein. In *Cellular and Molecular Life Sciences*. [https://doi.org/10.1007/s00018-](https://doi.org/10.1007/s00018-021-03776-z)
713 [021-03776-z](https://doi.org/10.1007/s00018-021-03776-z)
- 714 Cohen, L. B., Lindsay, S. A., Xu, Y., Lin, S. J. H., & Wasserman, S. A. (2020). The Daisho
715 Peptides Mediate Drosophila Defense Against a Subset of Filamentous Fungi. *Frontiers in*
716 *Immunology*, 11. <https://doi.org/10.3389/fimmu.2020.00009>
- 717 Danecek P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A.,
718 Keane, T., McCarthy, S. A., Davies, R. M., Li, H. (2021) Twelve years of SAMtools and
719 BCFtools. *GigaScience*, Volume 10, Issue 2, giab008,
720 <https://doi.org/10.1093/gigascience/giab008>
- 721 De Gregorio, E., Spellman, P. T., Rubin, G. M., & Lemaitre, B. (2001). Genome-wide analysis of
722 the Drosophila immune response by using oligonucleotide microarrays. *Proceedings of the*
723 *National Academy of Sciences of the United States of America*.
724 <https://doi.org/10.1073/pnas.221458698>
- 725 De Gregorio, E., Spellman, P. T., Tzou, P., Rubin, G. M., & Lemaitre, B. (2002). The Toll and
726 Imd pathways are the major regulators of the immune response in Drosophila. *EMBO*
727 *Journal*. <https://doi.org/10.1093/emboj/21.11.2568>
- 728 Dushay, M. S., & Eldon, E. D. (1998). Drosophila immune responses as models for human
729 immunity. *American Journal of Human Genetics*. <https://doi.org/10.1086/301694>
- 730 Galindo, R.L., Edwards, D.N., Gillespie, S.K.H., Wasserman, S.A. (1995). Interaction of the
731 pelle kinase with the membrane-associated protein tube is required for transduction of the
732 dorsoventral signal in *Drosophila* embryos. *Development*; 121 (7): 2209–2218. doi:
733 <https://doi.org/10.1242/dev.121.7.2209>

- 734 Gaskill, M. M., Gibson, T. J., Larson, E. D., & Harrison, M. M. (2021). GAF is essential for
735 zygotic genome activation and chromatin accessibility in the early *Drosophila* embryo.
736 *ELife*. <https://doi.org/10.7554/eLife.66668>
- 737 Grant, C. E, Timothy L. Bailey, T. L., & Noble, W. S. (2011) FIMO: Scanning for occurrences of
738 a given motif. *Bioinformatics* 27(7):1017-1018.
739 <https://doi.org/10.1093/bioinformatics/btr064>
- 740 Grandi, F. C., Modi, H., Kampman, L., & Corces, M. R. (2022). Chromatin accessibility profiling
741 by ATAC-seq. In *Nature Protocols*. <https://doi.org/10.1038/s41596-022-00692-9>
- 742 Hao, Y., Yu, S., Luo, F., & Jin, L. H. (2018). Jumu is required for circulating hemocyte
743 differentiation and phagocytosis in *Drosophila*. *Cell Communication and Signaling*.
744 <https://doi.org/10.1186/s12964-018-0305-3>
- 745 Huang, H. W., Zeng, X., Rhim, T., Ron, D., & Ryoo, H. D. (2017). The requirement of IRE1 and
746 XBP1 in resolving physiological stress during *Drosophila* development. *Journal of Cell*
747 *Science*. <https://doi.org/10.1242/jcs.203612>
- 748 Imrichová, H., Hulselmans, G., Atak, Z. K., Potier, D., & Aerts, S. (2015). I-cisTarget 2015
749 update: Generalized cis-regulatory enrichment analysis in human, mouse and fly. *Nucleic*
750 *Acids Research*. <https://doi.org/10.1093/nar/gkv395>
- 751 Jones, K., & Eleftherianos, I. (2023). A Simple Protocol for Isolating Hemolymph from Single
752 *Drosophila melanogaster* Adult Flies. *Methods and protocols*, 6(5), 100.
753 <https://doi.org/10.3390/mps6050100>
- 754 Kounatidis, I., & Ligoxygakis, P. (2012). *Drosophila* as a model system to unravel the layers of
755 innate immunity to infection. *Open Biology*. <https://doi.org/10.1098/rsob.120075>

- 756 Kvon, E. Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J. O., Pagani, M., Schernhuber, K.,
757 Dickson, B. J., & Stark, A. (2014). Genome-scale functional characterization of *Drosophila*
758 developmental enhancers in vivo. *Nature*. <https://doi.org/10.1038/nature13395>
- 759 Langmead, B. & Salzberg, S. (2012) Fast gapped-read alignment with Bowtie 2. *Nature*
760 *Methods*. 9:357-359. <https://doi.org/10.1038/nmeth.1923>
- 761 Lee, D., Shi, M., Moran, J., Wall, M., Zhang, J., Liu, J., Fitzgerald, D., Kyono, Y., Ma, L., White,
762 K. P., & Gerstein, M. (2020). STARRPeaker: uniform processing and accurate
763 identification of STARR-seq active regions. *Genome Biology*.
764 <https://doi.org/10.1186/s13059-020-02194-x>
- 765 Lemaitre, B., & Hoffmann, J. (2007). The host defense of *Drosophila melanogaster*. In *Annual*
766 *Review of Immunology*. <https://doi.org/10.1146/annurev.immunol.25.022106.141615>
- 767 Levo, M., Raimundo, J., Bing, X.Y., Sisco, Z., Batut, P. J., Ryabichko, S., Gregor, T., Levine,
768 M.S. (2022). Transcriptional coupling of distant regulatory genes in living embryos. *Nature*
769 605, 754–760. <https://doi.org/10.1038/s41586-022-04680-7>
- 770 Levy, F., Rabel, D., Charlet, M., Bulet, P., Hoffmann, J. A., & Ehret-Sabatier, L. (2004).
771 Peptidomic and proteomic analyses of the systemic immune response of *Drosophila*.
772 *Biochimie*, 86(9–10), 607–616. <https://doi.org/10.1016/j.biochi.2004.07.007>
- 773 Liao, Y., Smyth, G. K. & Shi, W. (2014). featureCounts: an efficient general purpose program for
774 assigning sequence reads to genomic features. *Bioinformatics*, 30(7):923-30.
775 <https://doi.org/10.1093/bioinformatics/btt656>
- 776 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing
777 reads. *EMBnet journal*, 17(1), pp. 10-12. doi:<https://doi.org/10.14806/ej.17.1.200>

- 778 Meister, M., & Richards, G. (1996). Ecdysone and insect immunity: The maturation of the
779 inducibility of the dipteracin gene in *Drosophila* larvae. *Insect Biochemistry and Molecular*
780 *Biology*. [https://doi.org/10.1016/0965-1748\(95\)00076-3](https://doi.org/10.1016/0965-1748(95)00076-3)
- 781 Merrill, C. B., Montgomery, A. B., Pabon, M. A., Shabalín, A. A., Rodan, A. R., & Rothenfluh, A.
782 (2022). Harnessing changes in open chromatin determined by ATAC-seq to generate
783 insulin-responsive reporter constructs. *BMC genomics*, 23(1), 399.
784 <https://doi.org/10.1186/s12864-022-08637-y>
- 785 Minakhina, S., Tan, W., & Steward, R. (2011). JAK/STAT and the GATA factor Pannier control
786 hemocyte maturation and differentiation in *Drosophila*. *Developmental Biology*.
787 <https://doi.org/10.1016/j.ydbio.2011.01.035>
- 788 Mohrs, M., Blankespoor, C., Wang, Z.E., Loots, G., Afzal, V., Hadeiba, H., Shinkai, K. Rubin,
789 E., Locksley, R. M. (2001). Deletion of a coordinate regulator of type 2 cytokine expression
790 in mice. *Nat Immunol* 2, <https://doi.org/10.1038/ni0901-842>
- 791 Morris, O., Liu, X., Domingues, C., Runchel, C., Chai, A., Basith, S., Tenev, T., Chen, H., Choi,
792 S., Pennetta, G., Buchon, N., & Meier, P. (2016). Signal Integration by the I κ B Protein
793 Pickle Shapes *Drosophila* Innate Host Defense. *Cell Host and Microbe*.
794 <https://doi.org/10.1016/j.chom.2016.08.003>
- 795 Muerdter, F., Boryń, Ł. M., & Arnold, C. D. (2015). STARR-seq - Principles and applications. In
796 *Genomics*. <https://doi.org/10.1016/j.ygeno.2015.06.001>
- 797 Myers, A. L., Harris, C. M., Choe, K. M., & Brennan, C. A. (2018). Inflammatory production of
798 reactive oxygen species by *Drosophila* hemocytes activates cellular immune defenses.
799 *Biochemical and Biophysical Research Communications*.
800 <https://doi.org/10.1016/j.bbrc.2018.09.126>

- 801 Neumayr, C., Pagani, M., Stark, A., & Arnold, C. D. (2019). STARR-seq and UMI-STARR-seq:
802 Assessing Enhancer Activities for Genome-Wide-, High-, and Low-Complexity Candidate
803 Libraries. *Current Protocols in Molecular Biology*. <https://doi.org/10.1002/cpmb.105>
- 804 Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A.,
805 Prosperini, E., Ghisletti, S., & Natoli, G. (2013). Latent enhancers activated by stimulation
806 in differentiated cells. *Cell*. <https://doi.org/10.1016/j.cell.2012.12.018>
- 807 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M.,
808 Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D.,
809 Brucher, M., Perrot, M. & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python.
810 *JMLR* 12, pp. 2825-2830.
- 811 Pfeiffer, B. D., Jenett, A., Hammonds, A. S., Ngo, T. T. B., Misra, S., Murphy, C., Scully, A.,
812 Carlson, J. W., Wan, K. H., Laverty, T. R., Mungall, C., Svirskas, R., Kadonaga, J. T., Doe,
813 C. Q., Eisen, M. B., Celniker, S. E., & Rubin, G. M. (2008). Tools for neuroanatomy and
814 neurogenetics in *Drosophila*. *Proceedings of the National Academy of Sciences of the*
815 *United States of America*. <https://doi.org/10.1073/pnas.0803697105>
- 816 Quinlan, A. R., Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic
817 features. *Bioinformatics*. 15;26(6):841-2. doi: 10.1093/bioinformatics/btq033.
- 818 Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar,
819 F., Manke, T. (2016) deepTools2: a next generation web server for deep-sequencing data
820 analysis. *Nucleic Acids Res.* 44(W1):W160-5. doi: 10.1093/nar/gkw257.
- 821 Ramirez-Corona, B. A., Fruth, S., Ofoegbu, O., & Wunderlich, Z. (2021). The mode of
822 expression divergence in *Drosophila* fat body is infection-specific. *Genome Research*.
823 <https://doi.org/10.1101/gr.269597.120>

- 824 Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., &
825 Mesirov, J. P. (2011). Integrative genomics viewer. *Nat Biotechnol* 29, 24–26.
826 <https://doi.org/10.1038/nbt.1754>
- 827 Rus, F., Flatt, T., Tong, M., Aggarwal, K., Okuda, K., Kleino, A., Yates, E., Tatar, M., &
828 Silverman, N. (2013). Ecdysone triggered PGRP-LC expression controls *Drosophila* innate
829 immunity. *EMBO Journal*. <https://doi.org/10.1038/emboj.2013.100>
- 830 Russell, T. A., Ayaz, A., Davidson, A. D., Fernandez-Sesma, A., & Maringer, K. (2021). Imd
831 pathway-specific immune assays reveal nf-kb stimulation by viral rna pamps in aedes
832 aegypti aag2 cells. *PLoS Neglected Tropical Diseases*.
833 <https://doi.org/10.1371/journal.pntd.0008524>
- 834 Samakovlis, C., Asling, B., Boman, H. G., Gateff, E., & Hultmark, D. (1992). In vitro induction of
835 cecropin genes - an immune response in a *Drosophila* blood cell line. *Biochemical and*
836 *Biophysical Research Communications*, 188(3), 1169–1175. [https://doi.org/10.1016/0006-](https://doi.org/10.1016/0006-291X(92)91354-S)
837 [291X\(92\)91354-S](https://doi.org/10.1016/0006-291X(92)91354-S)
- 838 Schlamp, F., Delbare, S. Y. N., Early, A. M., Wells, M. T., Basu, S., & Clark, A. G. (2021).
839 Dense time-course gene expression profiling of the *Drosophila melanogaster* innate
840 immune response. *BMC Genomics*. <https://doi.org/10.1186/s12864-021-07593-3>
- 841 Senger, K., Armstrong, G. W., Rowell, W. J., Kwan, J. M., Markstein, M., & Levine, M. (2004).
842 Immunity Regulatory DNAs Share Common Organizational Features in *Drosophila*.
843 *Molecular Cell*. [https://doi.org/10.1016/S1097-2765\(03\)00500-8](https://doi.org/10.1016/S1097-2765(03)00500-8)
- 844 Silverman, N., Zhou, R., Stöven, S., Pandey, N., Hultmark, D., Maniatis, T. (2000). A *Drosophila*
845 IkappaB kinase complex required for Relish cleavage and antibacterial immunity. *Genes*
846 *Dev.*;14(19):2461-71. doi: 10.1101/gad.817800.

- 847 Sopko, R., Lin, Y. Bin, Makhijani, K., Alexander, B., Perrimon, N., & Brückner, K. (2015). A
848 Systems-Level Interrogation Identifies Regulators of Drosophila Blood Cell Number and
849 Survival. *PLoS Genetics*. <https://doi.org/10.1371/journal.pgen.1005056>
- 850 Spahn, P., Huelsmann, S., Rehorn, K. P., Mischke, S., Mayer, M., Casali, A., & Reuter, R.
851 (2014). Multiple regulatory safeguards confine the expression of the GATA factor serpent
852 to the hemocyte primordium within the drosophila mesoderm. *Developmental Biology*.
853 <https://doi.org/10.1016/j.ydbio.2013.12.012>
- 854 Storelli, G., Nam, H. J., Simcox, J., Villanueva, C. J., & Thummel, C. S. (2019). Drosophila
855 HNF4 Directs a Switch in Lipid Metabolism that Supports the Transition to Adulthood.
856 *Developmental Cell*. <https://doi.org/10.1016/j.devcel.2018.11.030>
- 857 Tafesh-Edwards, G., & Eleftherianos, I. (2020). JNK signaling in Drosophila immunity and
858 homeostasis. In *Immunology Letters*. <https://doi.org/10.1016/j.imlet.2020.06.017>
- 859 Tanji, T., Hu, X., Weber, A. N. R., & Ip, Y. T. (2007). Toll and IMD Pathways Synergistically
860 Activate an Innate Immune Response in Drosophila melanogaster . *Molecular and Cellular*
861 *Biology*. <https://doi.org/10.1128/mcb.01814-06>
- 862 Tanji, T., Yun, E. Y., & Ip, Y. T. (2010). Heterodimers of NF- κ B transcription factors DIF and
863 Relish regulate antimicrobial peptide genes in Drosophila. *Proceedings of the National*
864 *Academy of Sciences of the United States of America*.
865 <https://doi.org/10.1073/pnas.1009473107>
- 866 The modENCODE Consortium, Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, Eaton
867 ML, Landolin JM, Bristow CA, Ma L, et al. 2010. Identification of functional elements and
868 regulatory circuits by Drosophila modENCODE. *Science* 330: 1787-1797.

- 869 Tremblay, B. J., (2024). universalmotif: An R package for biological motif analysis. *Journal of*
870 *Open Source Software*, 9(100), 7012, <https://doi.org/10.21105/joss.07012>
- 871 Troha, K., Im, J. H., Revah, J., Lazzaro, B. P., & Buchon, N. (2018). Comparative
872 transcriptomics reveals CrebA as a novel regulator of infection tolerance in *D.*
873 *melanogaster*. *PLoS Pathogens*. <https://doi.org/10.1371/journal.ppat.1006847>
- 874 Uvell H, Engström Y. (2003). Functional characterization of a novel promoter element required
875 for an innate immune response in *Drosophila*. *Mol Cell Biol*. Nov;23(22):8272-81. doi:
876 10.1128/MCB.23.22.8272-8281.2003. PMID: 14585984; PMCID: PMC262376.
- 877 Valanne, S., Myllymäki, H., Kallio, J., Schmid, M. R., Kleino, A., Murumägi, A., Airaksinen, L.,
878 Kotipelto, T., Kaustio, M., Ulvila, J., Esfahani, S. S., Engström, Y., Silvennoinen, O.,
879 Hultmark, D., Parikka, M., & Rämet, M. (2010). Genome-Wide RNA Interference in
880 *Drosophila* Cells Identifies G Protein-Coupled Receptor Kinase 2 as a Conserved
881 Regulator of NF- κ B Signaling . *The Journal of Immunology*.
882 <https://doi.org/10.4049/jimmunol.1000261>
- 883 Valanne, S., Wang, J.-H., & Rämet, M. (2011). The *Drosophila* Toll Signaling Pathway . *The*
884 *Journal of Immunology*. <https://doi.org/10.4049/jimmunol.1002302>
- 885 Vonolfen, M. C., Meyer Zu Altenschildesche, F. L., Nam, H.-J., Brodesser, S., Gyenis, A.,
886 Buellesbach, J., Lam, G., Thummel, C. S., & Storelli, G. (2024). *Drosophila* HNF4 acts in
887 distinct tissues to direct a switch between lipid storage and export in the gut. *Cell Reports*,
888 43(9), 114693. <https://doi.org/10.1016/j.celrep.2024.114693>
- 889 Westlake, H., Hanson, M., & Lemaitre, B. (2024). *The Drosophila immunity handbook*.
890 <https://doi.org/10.55430/6304TDIHVA01>

- 891 Yu, S., Luo, F., Xu, Y., Zhang, Y., & Jin, L. H. (2022). *Drosophila* Innate Immunity Involves
892 Multiple Signaling Pathways and Coordinated Communication Between Different Tissues.
893 In *Frontiers in Immunology*. <https://doi.org/10.3389/fimmu.2022.905370>
- 894 Zabidi, M. A., Arnold, C. D., Schernhuber, K., Pagani, M., Rath, M., Frank, O., & Stark, A.
895 (2015). Enhancer-core-promoter specificity separates developmental and housekeeping
896 gene regulation. *Nature*. <https://doi.org/10.1038/nature13994>
- 897 Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C.,
898 Myers, R. M., Brown, M., Li, W., & Liu, X. S. (2008) Model-based analysis of ChIP-Seq
899 (MACS). *Genome Biol.* 9(9):R137. doi: 10.1186/gb-2008-9-9-r137
- 900 Zraly, C. B., Zakkar, A., Perez, J. H., Ng, J., White, K. P., Slattery, M., & Dingwall, A. K. (2021).
901 The *Drosophila* MLR COMPASS complex is essential for programming cis-regulatory
902 information and maintaining epigenetic memory during development. *Nucleic Acids*
903 *Research*. <https://doi.org/10.1093/NAR/GKAA082>

904

905 Figure 1: Immune enhancers identified by STARR-seq (A) Schematic of STARR-seq experiment.
906 Fragmented genomic DNA is cloned into the pSTARR-seq plasmid, creating a library of putative
907 enhancer regions. This library is transfected into S2* cells, split into 3 different populations and treated
908 with either H2O/PBS, 20E/PBS, 20E/Heat-killed *Serratia marcescens* (HKSM). (B) Pearson's correlation
909 of each individual replicate across three treatments. (C) Distribution of enhancers in genomic regions (D)
910 Schlamp et al. (2021) identified four time clusters of gene expression in the first 24 hrs following immune
911 stimulation in adult flies; peak expression times for each time cluster are shown. Enhancers were
912 assigned to these immune genes and the fraction of immune genes identified by Schlamp et al. with an
913 immune enhancer is plotted. (E) Schematic of Imd pathway activated by DAP-type peptidoglycan binding
914 to peptidoglycan recognition proteins (PGRPs) leading to the cleavage of the NF- κ B Relish, allowing its

915 translocation into the nucleus where it activates transcription of antimicrobial peptides (AMPs). Pie chart
916 of the number of enhancers assigned to 32 Imd-associated AMPs and other effectors (Westlake et al.
917 2024) (F) IGV view of the *CrebA* gene locus, showing the STARR-seq output tracks and STARRPeaker
918 called peaks for all treatment conditions. STARR-seq identified two enhancers in the *CrebA* intron, one
919 shared across all conditions and one that is IMD specific (Robinson et al. 2011). (G) Overlay histogram of
920 GFP levels of S2* cells transfected with a GFP reporter of the IMD-specific *CrebA* enhancer analyzed by
921 flow cytometry.

922 Figure 2: Transcription factor binding content varies between immune enhancers (A) Number of binding
923 sites for each TF normalized by total length in basepairs of enhancers in each data set. Density is
924 reported as the number of TFBS motifs per basepair. (B) Log of odds ratio of TFBSs in three
925 comparisons. Enrichment is reported as the logarithm of the ratio of TFBS density in each of the two
926 enhancer groups being compared. Positive enrichment values are in blue and negative enrichment values
927 are in red, and significance is indicated by * ($p < 0.05$, Fisher's exact test). (C) Activity score (log scale)
928 from STARR-seq of IMD enhancers plotted by total binding sites. Individual enhancers are shown as
929 dots and medians and quartiles are displayed as box and whisker plots. The activity scores of enhancers
930 in adjacent bins were compared by Mann-Whitney *U* test with Bonferroni correction;* indicates $p < 0.05$,
931 non-significant comparisons are not shown. (D) Activity score (log scale) of IMD enhancers by number of
932 Relish or Kay/Jra binding sites.

933 Figure 3: TFBS profiles of IMD enhancers vary according to pathways, timing, and function. Log odds
934 ratio of TFBS in IMD enhancers grouped by pathway and process (Schlamp et al. 2021) (A) time cluster
935 (Schlamp et al. 2021) (B), broad immune role (Schlamp et al. 2021) (C), or functional group (Westlake et
936 al. 2024) (D) Positive enrichment values are in blue and negative enrichment values are in red, and
937 significance is indicated by * ($p < 0.05$, Fisher's exact test).

938 Figure 4: Activity classification of enhancers reveals specific properties of inducible and constitutive
939 enhancers. (A) Venn diagram of enhancers showing activity in treatment conditions, defining activity
940 classes. (B) Distribution of enhancer activity classes by broad immune role (Schlamp et al. 2021). There
941 is a significantly higher fraction of IMD Only enhancers in the effector group as compared to All
942 Enhancers, while the signaling group has a higher fraction of constitutive enhancers as compared to All

943 Enhancers (one-proportion z-tests, $p < 0.05$, indicated by *) The number of genes in each immune role is
944 listed in parentheses. (C) Distribution of enhancer activity classes by time cluster (Schlamp et al 2021).
945 The fraction of IMD Only enhancers is significantly larger in the Early and Mid clusters as compared to All
946 Enhancers (one-proportion z-tests, $p < 0.05$, indicated by *). The number of genes in each time cluster is
947 listed in parentheses. (D) Distributions of distances between enhancers and paired genes. Distance is
948 measured from the closest edge of the enhancer to the transcriptional start site (TSS). For intronic
949 enhancers, the distance will include the length of the introns and exons between the TSS and enhancer.
950 The fraction of Constitutive enhancers that overlap with their paired genes is significantly greater than the
951 fraction of inducible IMD + 20E enhancers that overlap with genes (2-sided proportion z-test, p -value $<$
952 0.05). The number of Enhancer-Gene pairs is listed in parentheses for each activity class.

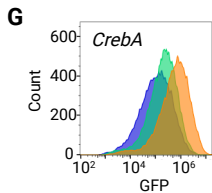
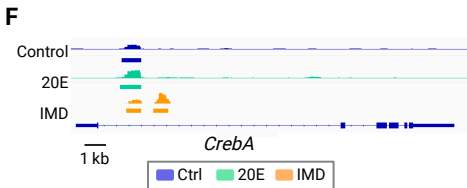
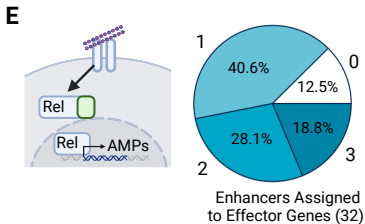
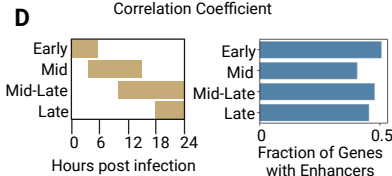
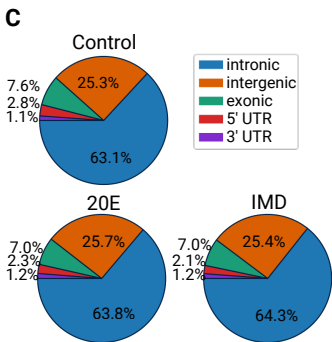
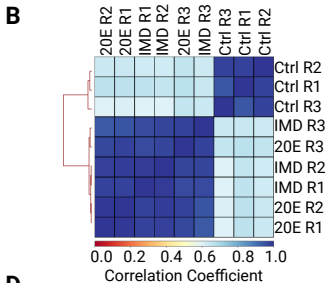
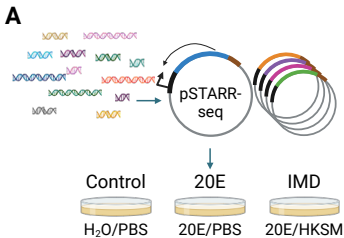
953 Figure 5: STARR-seq enhancers reporters demonstrate immune specific expression (A) Schematic of
954 enhancer reporter plasmids, with GFP under control of STARR-seq enhancers and minimal promoter,
955 DSCP. (B) Log_2 (fold change) of GFP expression measured by flow cytometry normalized to Control
956 treatment of enhancer reporters and promoter-only (DSCP) negative control. Bars represent the mean
957 and dots show three biological replicates. Error bars represent standard error of the mean. Gray dashed
958 line marks the expression level of DSCP control upon IMD stimulation. Significant gene induction is
959 labeled with * $p < 0.05$ (one sided Mann-Whitney U test). (C) Log_2 (fold change) of GFP expression
960 normalized to Control treatment of *daisho* enhancer reporter. Significant gene induction is labeled with *
961 $p < 0.05$ (one sided Mann-Whitney U test) (D) Log_2 of GFP expression levels for reporters comparing IMD
962 treatment to Control treatment. Both treatment conditions are normalized to untransfected S2* Torso-pelle
963 cells. Dashed line is $y=x$, denoting an equal response to IMD and control treatment.

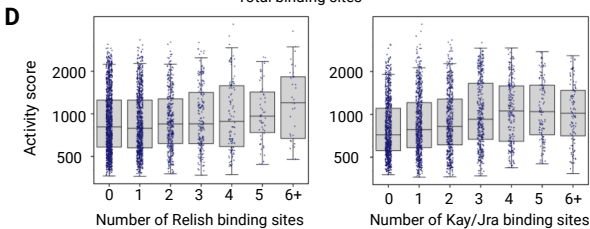
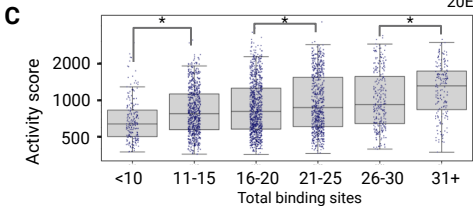
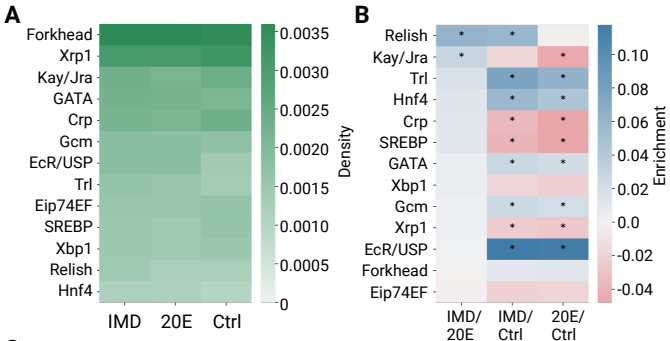
964

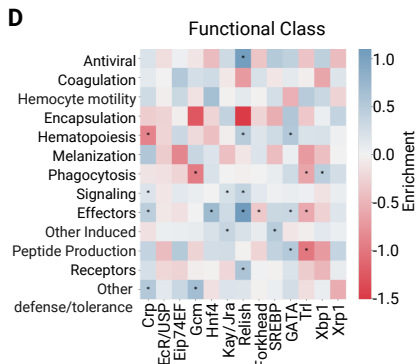
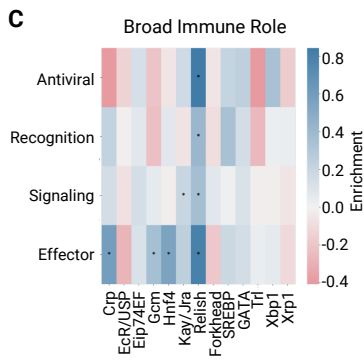
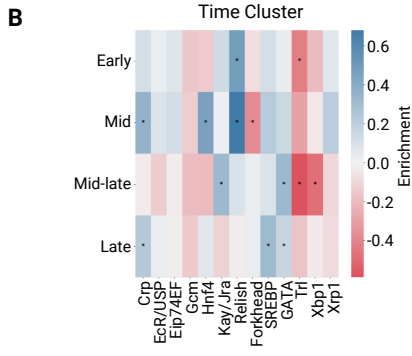
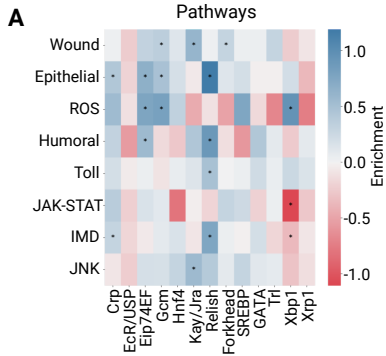
965 Figure 6: ATAC-seq reveals immune inducible enhancers are made accessible by EcR/USP, Hnf4 and Trl
966 (A) Diagram of ATAC-seq on immune stimulated hemocytes. Flies are injected with HKSM and along with
967 age matched controls are incubated at 25°C for 24 hours. Then flies are anesthetized, hemocytes are
968 extracted via centrifugation before lysis and transposition. Tn5 was used to insert adaptors into
969 accessible chromatin. These fragments are then isolated and sequenced to identify accessible peaks. (B)
970 Venn Diagram of accessibility groups of enhancers. Always closed enhancers are in blue and are

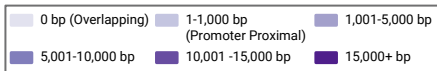
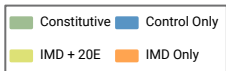
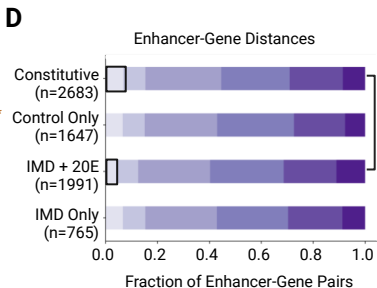
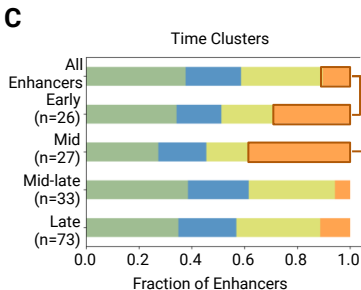
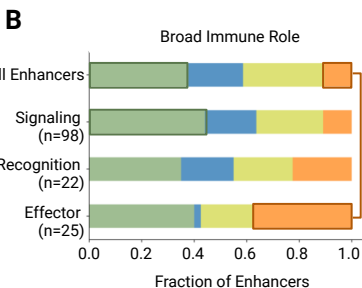
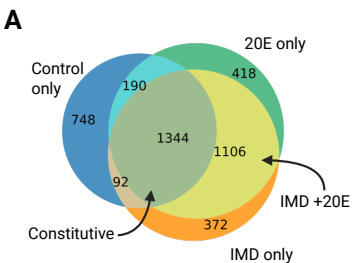
971 inaccessible in both the control and HKSM samples. The right circle is enhancers open in the control
972 sample and the left circle is enhancers open in the HKSM sample. The intersection has enhancers open
973 in both. The number of enhancers in each group is listed in parentheses. (C) Fraction of enhancers in
974 each accessibility group by activity class. Constitutively active enhancers are more likely to be always
975 open than immune responsive enhancers, *** indicates $p < 0.005$, z-test, Bonferroni multiple test
976 correction. Insert of the same data for HKSM opened enhancers, plotted separately for clarity. A greater
977 proportion of IMD + 20E enhancers are HKSM opened than Constitutive enhancers (* indicates $p < 0.1$,
978 Two sample z-test). (D) Fraction of enhancers in accessibility groups grouped by broad immune roles
979 (left) and time clusters (right). ** indicates $p < 0.05$, z-test, Bonferroni multiple test correction, *** indicates
980 $p < 0.005$, z-test, Bonferroni multiple test correction (E) Log odds ratio of TFBS in enhancers by
981 accessibility compared to all enhancers, for previously identified motifs. Positive enrichment values are in
982 blue and negative enrichment values are in red, and significance is indicated by * ($p < 0.05$, Fisher's test).

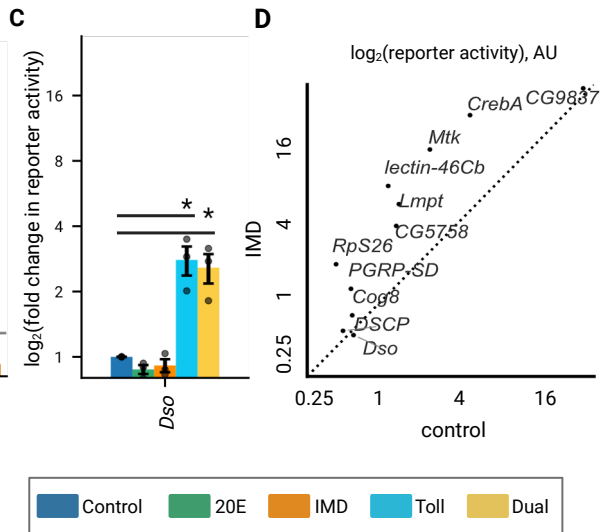
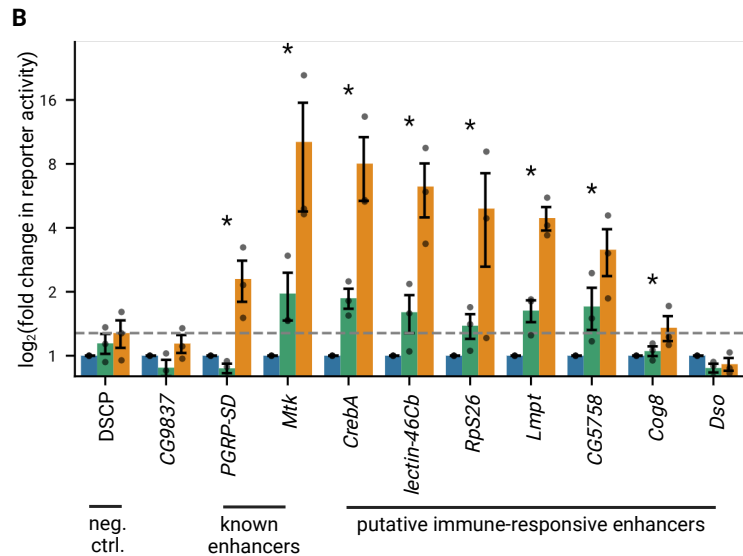
983

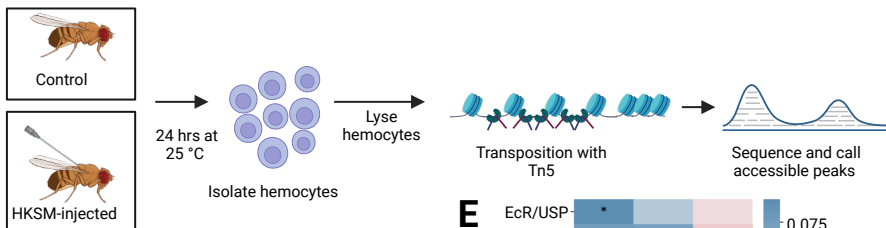
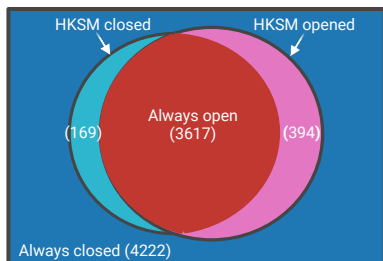
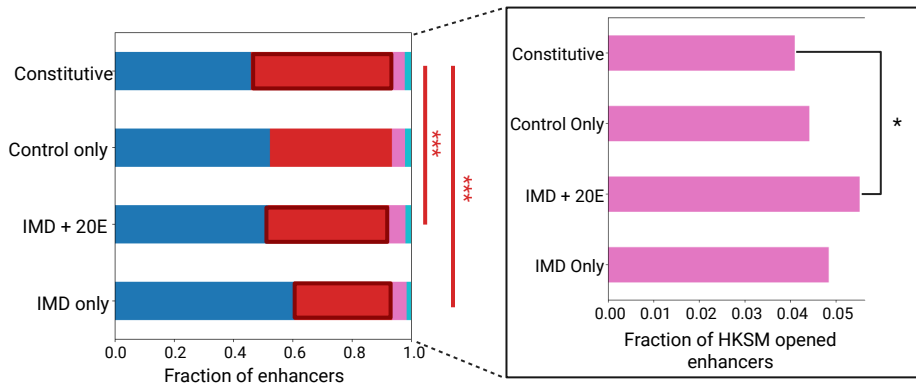
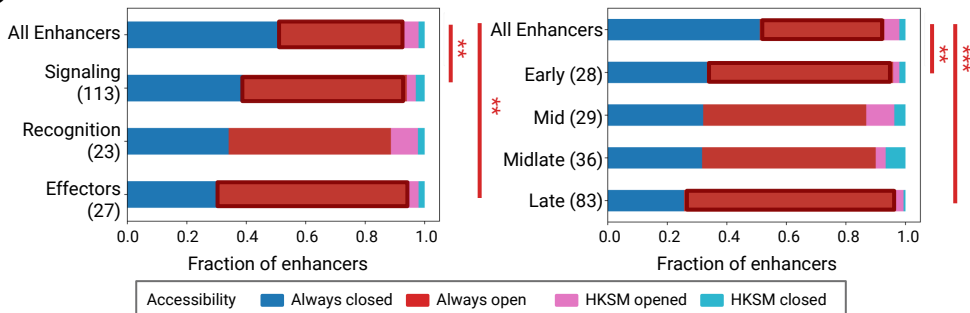










A**B****C****D****E**