Proximity labeling reveals a new *in vivo* network of interactors for the histone demethylase KDM5

Matanel Yheskel[1], Simone Sidoli[2], and Julie Secombe[1,3,4]

[1] Department of Genetics, Albert Einstein College of Medicine, Bronx, New York, United States of America.

[2] Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York, United States of America.

[3] Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, New York, United States of America.

[4] Lead contact and corresponding author: julie.secombe@einsteinmed.edu

Running title: KDM5-interacting proteins identified using proximity labeling

Key words: KDM5, histone demethylase, TurboID, proximity labeling, mass spectrometry, chromatin modifiers, insulators

Summary

KDM5 family proteins are multi-domain regulators of transcription that when dysregulated contribute to cancer and intellectual disability. KDM5 proteins can regulate transcription through their histone demethylase activity in addition to demethylase-independent gene regulatory functions that remain less characterized. To expand our understanding of the mechanisms that contribute to KDM5-mediated transcription regulation, we used TurboID proximity labeling to identify KDM5-interacting proteins. Using *Drosophila melanogaster*, we enriched for biotinylated proteins from KDM5-TurboID-expressing adult heads, and generated a novel control for DNA-adjacent background in the form of dCas9:TurboID. We identified both known and novel candidate proteins, including members of the SWI/SNF and NURF chromatin remodeling complexes, the non-specific lethal (NSL) complex, Mediator, and several insulator proteins. Combined, our data shed new light on potential demethylase-independent activities of KDM5. In the context of KDM5 dysregulation, these interactions may play key roles in the alteration of evolutionarily conserved transcriptional programs implicated in human disorders.

40

41 **Introduction**

42 Lysine demethylase 5 (KDM5) family proteins are multidomain transcriptional regulators
43 able to recognize and enzymatically modify chromatin.[1,2] The best characterized function of KDM5
44 proteins is their histone demethylase activity, which cleaves a chromatin mark that is found at
45 most active promoters, trimethylated lysine 4 on histone H3 (H3K4me3).[1,3–6] KDM5 proteins are
46 evolutionarily conserved, with four paralogous genes in mammals encoding KDM5A-KDM5D,
47 while animals with smaller genomes such as nematodes and flies possess a single *kdm5* gene.
48 The importance of KDM5 function is emphasized by the observation that changes to the
49 expression of this family of proteins is associated with two clinical outcomes: cancer and
50 intellectual disability.[7–9] KDM5A and KDM5B are amplified or overexpressed in a range of
51 cancers, including breast, ovarian, skin, and lung.[8,10–13] KDM5A/B appear to play several roles in
52 tumorigenesis, including promoting cell cycle progression and regulating the metabolism of
53 cancer stem cells.[14–16] In contrast to the gain of function seen in cancer cells, loss of function
54 variants in the autosomal paralogs KDM5A, KDM5B, and the X-linked KDM5C have been
55 observed in individuals with intellectual disability (ID).[17–21] KDM5 proteins have an evolutionarily
56 conserved role in regulating critical gene expression programs in neurons as evidenced by
57 morphological and functional neuronal phenotypes in KDM5B and KDM5C knockout mice. [22–24]
58 Similarly, flies and nematodes with *kdm5* mutations display altered neuroanatomical development
59 and/or neurotransmission. [25–28]

60

61 KDM5 catalytic function is mediated by the joint activity of the Jumonji N (JmjN) and JmjC
62 domains and is classically thought to result in transcriptional repression. In addition, KDM5
63 proteins possess other potential gene regulatory domains, including plant homeodomain domain
64 (PHD) motifs that can recognize H3K4me2/3 or H3K4me0, and a potential DNA binding A/T
65 interaction domain (ARID).[29–34] These binding domains likely function in-concert with the histone

demethylase activity of KDM5 by, for example, recruiting it to target promoters through the activity of individual or combinations of accessory domains. Conversely, non-enzymatic functions of these domains and/or other motifs of KDM5 that have no currently known function, such as the C_5HC_2 domain, could regulate transcription through distinct mechanisms. There is ample evidence that KDM5 proteins can regulate transcription independently of their demethylase activity. For instance, KDM5 is essential for viability in flies in a manner that is independent of its histone demethylase activity.[33,35] In addition, both demethylase-dependent and independent functions of KDM5 are critical for *Drosophila* neuronal development and function. [36,37] Interestingly, some missense alleles of KDM5C observed in individuals with intellectual disability diminish its enzymatic activity, while others do not.[23,38–40] Similarly, demethylase dependent and independent activities of KDM5 proteins are likely to be important for their contributions to the etiology of and spread of cancers.[41,42] For example, KDM5B demethylase-independent functions in breast cancer promote metastatic potential to the lung.[41,42] Thus, even though KDM5 proteins derive their name from their enzymatic function, other conserved motifs contribute to their gene regulatory activities, although these activities remain much less characterized.

Understanding the repertoire of gene regulatory mechanisms utilized by KDM5 family proteins requires a comprehensive understanding of the proteins they can interact with. Traditional immunoprecipitation coupled with mass spectrometry (IP-MS) approaches have been used to identify proteins that form complexes with KDM5 family proteins in both mammals and *Drosophila*.[40,43–45] These experiments have revealed several conserved interactions, most notably with histone deacetylase 1 (HDAC1) and other proteins known to associate with this chromatin modifier.[46,47] To expand our understanding of the proteins that function with KDM5 to mediate its gene regulatory activities, we used TurboID-mediated proximity labeling.[48] This has been shown to be a powerful technique to identify weak or transient interactions that may otherwise be disturbed during the process of traditional IP experiments.[48–55] This technique takes advantage of

92 the promiscuous biotin ligase activity of TurboID, which results in the biotinylation of lysine
93 residues within 10 nm of its active site. When expressed as a chimeric fusion to a protein of
94 interest, interacting proteins will be biotin-labeled.[56,48] Covalently modified proteins are then
95 recovered with streptavidin beads and prepared for liquid chromatography-tandem mass
96 spectrometry (LC-MS/MS). By expressing KDM5 that was N- or C-terminally tagged with TurboID
97 *in vivo*, we recovered about half of previously identified interactions in *Drosophila*, and almost all
98 interactions known to be conserved in mammalian cells, clearly demonstrating the robustness of
99 this technique. Furthermore, we have discovered a novel interactome for KDM5 that suggests
100 roles in the function of the SWI/SNF, non-specific lethal (NSL), NURF, and Mediator complexes,
101 in addition to chromatin insulation.
102
103 **Results**
104 **Chimeric TurboID-KDM5 proteins are functional and broadly biotinylate**
105 To identify KDM5 interactors *in vivo*, we created constructs in which KDM5 was N- or C-terminally
106 tagged with TurboID to maximally identify proteins that could function with KDM5. Because our
107 long-term goal is to further develop our *Drosophila* model of KDM5-induced intellectual disability
108 (ID), we chose to carry out our TurboID studies using adult heads to enrich for neuronal tissue,
109 using the general workflow shown in Fig 1A. Generating a TurboID system that closely mimics
110 endogenous *kdm5* expression has been shown to be important for delivering more specific
111 biotinylation compared to overexpression.[57] Based on our prior generation of a UASp-*kdm5*
112 transgene that is expressed at approximately endogenous levels when crossed to a range of Gal4
113 drivers, we generated transgenic flies harboring HA-tagged UASp-TurboID:*kdm5* and UASp-
114 *kdm5*:TurboID (Fig. 1B).[58] Therefore, we generated both N-terminal (NT-KDM5) and C-terminal
115 (CT-KDM5) TurboID fusions of KDM5 to understand the full breadth of its interactome and to
116 highlight terminus-specific interactions. To test the functionality and expression of the chimeric
117 KDM5-TurboID proteins, we expressed them ubiquitously in a $kdm5^{140}$ null mutant background.[33]

118 Western blot analysis using adult heads showed that NT-KDM5 and CT-KDM5 were expressed
119 at levels similar to those observed from an endogenously HA-tagged KDM5 (Fig. 1C). Importantly,
120 N- and C-terminally tagged KDM5 proteins were able to restore viability to *kdm5*^{140} null mutant
121 flies, which normally die prior to adulthood (Fig. 1D).[35] Tagging KDM5 with TurboID therefore does
122 not interfere with its function. Flies in which TurboID-KDM5 was the only source of KDM5
123 (*kdm5*^{140};Ubi-Gal4>UASp;TurboID:*kdm5*) were used for all subsequent experiments to maximize
124 the number of interactors identified.
125
126 **Determining the proper controls to identify the KDM5 proximitome**
127 To confidently identify proteins that function with KDM5, appropriate controls are critical.
128 Due to the novelty of the technique, there is no standard set of controls for proximity labeling
129 experiments. Many studies simply enrich over endogenous biotinylation and bead background.[59–
130 62] Other studies expressed forms of TurboID alone that were localized to the specific cellular
131 compartment that the protein of interest resided, such as the cellular membrane.[48,49,51,52] As a
132 non-TurboID-expressing wild-type control with a similar genetic background, we used a fly strain
133 in which endogenous *kdm5* is removed and HA-tagged KDM5 is expressed using its endogenous
134 promoter from a transgene inserted at the same locus as the TurboID constructs
135 (*kdm5*^{140}*;gkdm5*^{WT}). We will refer to these as control animals. We also generated a transgene
136 able to express nuclear localized, HA-tagged, TurboID alone using the same UAS promoter used
137 for CT-*kdm5* and NT-*kdm5,* in an effort to assay general nuclear background (Fig. 1B). To
138 compare the levels of TurboID alone to TurboID-KDM5 we expressed these transgenes using
139 Ubi-Gal4 in a wild-type and *kdm5*^{140} background, respectively. Unsurprisingly, given the small
140 size of TurboID (36kDa) compared to NT-KDM5 and CT-KDM5 (240kDa), anti-HA western blot
141 from adult heads showed significantly higher levels of expression for TurboID alone, possibly
142 creating high levels of background biotinylation in this strain (Fig. 1C). Because biotin is essential
143 for animal viability and thus included in the standard fly food used for crosses and stock

144 maintenance, we assessed the ability of all TurboID transgenes to biotinylate proteins when
145 expressed using Ubi-Gal4 by probing with infrared-conjugated streptavidin. Compared to control
146 flies, similar levels of biotin-conjugated proteins were observed in heads expressing TurboID, NT-
147 KDM5 and CT-KDM5, demonstrating their ability to biotinylate *in vivo* (Fig 1C).
148
149 To identify proteins preferentially biotinylated by NT-KDM5 and CT-KDM5 compared to
150 control and TurboID alone, we carried out streptavidin-bead pulldowns in quadruplicate followed
151 by liquid chromatography and mass spectrometry (LC-MS/MS). This experiment (Experiment 1)
152 identified a total of 1332 proteins, 476 of which are found in the nucleus where we have previously
153 shown KDM5 to be localized.[27] Principal component analysis (PCA) of normalized nuclear protein
154 abundances showed that TurboID alone, NT-KDM5, and CT-KDM5 clustered together, but were
155 distinct from controls (Fig S1A). We therefore compared the proteins identified in NT-KDM5 and
156 CT-KDM5 to control heads which revealed enrichment of 172 and 184 proteins, respectively,
157 using a *p*-value cutoff of 0.05 (Fig S1B, C; Table S1). 136 proteins were commonly enriched by
158 N- and C-terminally tagged KDM5, suggesting that we can robustly detect proteins in proximity to
159 KDM5 (Fig. S1D). To assess the quality of our data, we determined how many known KDM5
160 interactors were identified in our analyses. Fifteen proteins have been established to form a
161 complex with *Drosophila* KDM5 through IP-MS studies or targeted co-IP experiments (Table S2).
162 Suggesting the robustness with which the TurboID approach identifies *bona fide* KDM5-
163 associated proteins, 7 known interactors were identified by NT-KDM5 and 6 by CT-KDM5 (47%
164 and 40%, respectively). We additionally assessed biotinylated protein enrichment of NT-KDM5
165 and CT-KDM5 compared to TurboID alone (Fig S1E, F). As expected, based on the higher level
166 of expression of this protein compared to TurboID-tagged KDM5, a high level of background was
167 observed in these flies. This resulted in fewer proteins being enriched in NT-KDM5 and CT-KDM5
168 (32 and 61, respectively), reduced overlap between the datasets, and a reduction in the number
169 of known interactors identified (Fig S1G, H). Interestingly, TurboID alone appears to show bias in

170 its biotinylation of nuclear proteins, as comparing TurboID to control revealed significant
171 enrichment of 199 proteins, a majority of which are involved in chromatin-mediated transcriptional
172 regulation (Fig. S1I).
173
174 Because of concerns related to the use of TurboID alone as a control, we generated a
175 transgene encoding an enzymatically inactive form of the Cas9 enzyme (dCas9) fused to an HA-
176 tagged nuclear localized TurboID (UASp-dCas9:TurboID). This dCas9:TurboID fusion protein is
177 more similar in size to the KDM5 fusion proteins, being 194kDa and 235kDa, respectively,
178 compared to 36kDa for TurboID alone. In addition, Cas9 and dCas9 can scan the DNA, potentially
179 making this fusion an appropriate control for chromatin binding proteins such as KDM5 by
180 restricting biotinylation to DNA-adjacent proteins.[63] Ubi-Gal4-mediated expression of this
181 transgene revealed that dCas9:TurboID was expressed at similar levels to the KDM5-TurboID
182 fusion proteins and was able to biotinylate (Fig. 1C). Repeating the proximity labeling experiment,
183 we carried out triplicate streptavidin-bead pulldowns from heads of control, TurboID,
184 dCas9:TurboID, NT-KDM5 and CT-KDM5 flies (Experiment 2). MS analyses identified 1,146
185 proteins across all samples, 203 of which were nuclear. PCA from this second experiment showed
186 that NT-KDM5 and CT-KDM5 clustered together, indicating that these datasets are more alike to
187 each other than to any of the controls. Like our first experiment, TurboID alone clustered with NT-
188 KDM5 and CT-KDM5, and clustered distinctly from control and dCas9:TurboID samples (Fig 2A).
189 Using these data, we compared NT-KDM5 and CT-KDM5 to control, dCas9:TurboID, and to
190 TurboID alone (Fig. 2B; Table S3). Proteins enriched in the KDM5 samples compared to control
191 or TurboID gave similar results to those obtained in the first experiment (Fig 2B). Using control
192 flies as reference, 82 proteins were identified using NT-KDM5 and 68 for CT-KDM5, with 61 of
193 these proteins being identified in both datasets (Fig 2C-E). Compared to TurboID alone, only 29
194 and 24 proteins were enriched for NT-KDM5 and CT-KDM5, with 13 overlapping between the two
195 datasets (Fig. 2F-H). Importantly, we find comparing NT-KDM5 and CT-KDM5 with

196 dCas9:TurboID yielded data very similar to that seen using control animals, despite providing a
197 higher biotinylation background. 66 and 59 proteins were enriched in NT-KDM5 and CT-KDM5,
198 respectively, with an overlap of 48 proteins (Fig. 2I-K).
199
200 To complete our characterization of dCas9:TurboID as a tool to identify enriched DNA-
201 adjacent proteins, we compared these data to both TurboID alone and to control. Similar to data
202 from the first experiment comparing TurboID and control, TurboID alone shows a preference for
203 biotinylating a large number of chromatin-related proteins, even compared to dCas9:TurboID (Fig.
204 S2A). Comparing dCas9:TurboID to control revealed enrichment in a relatively small number of
205 proteins that were enriched for transcriptional-regulatory proteins consistent with the ability of
206 dCas9:TurboID to biotinylate targets while scanning DNA (Fig S2B, C). Confirming the challenges
207 related to using TurboID alone, the consistency with which proteins were enriched in NT-KDM5
208 or CT-KDM5 compared to TurboID alone was very low, with very little agreement across
209 experiments using a *p*-value cutoff <0.05 or <0.1 (Fig S2D, E). Moreover, the number of previously
210 identified interactors remained low in experiment two when comparing to TurboID alone, with only
211 5 and 3 being identified in NT-KDM5 and CT-KDM5, respectively (Fig. S2F). We therefore suggest
212 that endogenous biotinylation or dCas9:TurboID are superior to TurboID as controls for proximity
213 labeling experiments where the protein of interest is nuclear-specific and DNA-adjacent.
214
215 **Proximity labeling identifies new potential KDM5 interacting complexes**
216 To build a high confidence list of proteins which interact with KDM5, we combined data from
217 experiment 1 comparing NT-KDM5 and CT-KDM5 to control (2 datasets), and experiment 2 in
218 which NT-KDM5 and CT-KDM5 were compared to control (2 datasets) as well as dCas9:TurboID
219 (2 datasets). To do this, we began by filtering for enriched nuclear proteins across all six datasets
220 using a *p*-value<0.1. We first filtered for proteins identified in at least two of six datasets. Then to
221 include the possibility of terminus-exclusive interactors, we required that at each terminus proteins

222 had to be identified in 0 (exclusive to other terminus) or in 2 of 3 datasets (Fig. 3A). Demonstrating
223 the power of this approach, this includes 7 of the 15 (47%) known *Drosophila* KDM5 interactors.
224 This ratio increases to 7 of 8 (88%) for those interactors that have been shown in both *Drosophila*
225 and mammals. In addition, we identified several proteins not previously been found to be
226 *Drosophila* KDM5 interactors, but have been purified using KDM5A, KDM5B, KDM5C and/or
227 KDM5D in mammalian cells. These include the nucleosome remodeler Mi-2, the chromatin
228 assembly factor 1 (Caf1;Caf1-55), the actyl-lysine binding protein Zymnd8 and the
229 heterochromatin-associated protein HP1c (Fig 3B). [6,46,47] With these stringent filtering criteria, we
230 identified a total of 87 proteins (Fig. 3C). These high confidence interactors are highlighted in red
231 in volcano plots comparing NT-KDM5 and CT-KDM5 to dCas9:TurboID (Fig. 3D, E). All
232 comparisons are shown in Supplemental Figure 3.
233
234 Our proximity-labeling studies using Turbo-KDM5 revealed a broader interactome than
235 previously described in the literature. To better understand the relationships between the proteins
236 identified in our study, we generated a protein interaction map using Cytoscape and STRING (Fig.
237 4A).[64,65] Gene Ontology (GO) analysis shows that many of these proteins have roles in the
238 regulation of gene expression, chromatin modification, and chromatin remodeling (Fig. 4B). In
239 addition to confirming the strong link between KDM5 and Sin3/HDAC1-containing complexes,
240 these analyses also highlighted interactions with new protein complexes. Among these, we find
241 proteins such as BEAF-32, Chromator (Chro), Putzig (Pzg), and Cp190 that function in regulating
242 genomic architecture, suggesting a unstudied role for KDM5 in this process.[66–71] To confirm our
243 MS data, we performed western blots on input and streptavidin pulldowns and probed with
244 antibodies specific for BEAF-32 and CP190. This confirmed enrichment of these two proteins in
245 both NT-KDM5 and CT-KDM5 compared to dCas9:TurboID-expressing and control flies (Fig. 4C).
246 The extensive biotinylation of chromatin-related proteins, such as BEAF-32 and CP190, by
247 TurboID alone is also evident in this pulldown, as discussed previously. In addition, we identified

248 proteins critical for forming the transcriptional pre-initiation complex (TPIC), which is consistent
249 with the promoter-proximal binding of KDM5 proteins across species.[27,46,72–74,74] We analyzed a
250 recent crystal structure of the human TPIC and found that distinct surfaces interacted with KDM5,
251 directly revealing the specificity of biotinylation using TurboID-KDM5. Specifically, three adjacent
252 proteins in the mediator complex (MED1, MED14 and MED17) and three adjacent subunits of
253 TFIID (Taf4, Taf6, and Taf9) were identified in our analyses (Fig. 4D).[75] This suggests that KDM5
254 may play a direct role in enhancer-promoter communications that regulate the transcriptional
255 activity of target genes.
256
257 **KDM5 and newly identified interactors occupy overlapping genomic binding sites in**
258 ***Drosophila* and human cells**
259 To further explore the relationship between KDM5 and newly identified interactors in
260 *Drosophila*, we compared their genomic binding with those of interactors using publicly available
261 ChIP-seq datasets. We used published KDM5 ChIP-seq data from whole adult flies to interrogate
262 ChIP-Atlas as a means to identify datasets from any *Drosophila* cell type that significantly overlap
263 (via permutation 100X).[76] We then overlapped these with our high confidence interactor list to
264 reveal a total of 27 overlapping datasets. 30 proteins in our interactor list had available data on
265 ChIP Atlas. (Fig. 5A, B). These overlapping datasets included known interactors such as Sin3A,
266 in addition to BEAF-32 that we confirmed to be in proximity to KDM5, both of which show that
267 KDM5 flanks their DNA-binding sites (Fig. 5C, D). Representatives of other newly identified
268 proteins that are members of several different complexes were also examined, including the DNA
269 replication-related element factor Dref, the chromatin remodeler Iswi, the WD40-repeat protein
270 will die slowly (wds) and the acetyl-histone binding Brd2/Brd4 ortholog fs(1)h (Fig. 5E-H). The
271 similar binding profiles of all these proteins can also be seen in representative genome browser
272 tracks (Fig. 5I)
273

274 To assess the extent to which our high confidence KDM5 interactors might be
275 evolutionarily conserved, we investigated genomic co-occupancy in human cells. We first
276 converted our 87 *Drosophila* high confidence interactors to their human ortholog(s), which
277 resulted in a total of 138 proteins due to humans possessing multiple paralogous proteins for
278 some *Drosophila* proteins (DIOPT v8.0[77] score > 8/15). Due to the strong link between KDM5B
279 and breast cancer, and the wealth of ChIP-seq datasets available in cell lines derived from this
280 cancer type, we used KDM5B data from MCF-7 cells for these analyses.[78] Using peaks called
281 from this ChIP-seq data, we interrogated all available breast cancer cell line datasets again using
282 ChIP-Atlas (100X permutation). This revealed that 26 candidate interactors had binding profiles
283 that significantly overlapped with KDM5B binding (Fig. 6A, B). Interestingly, these included the
284 KDM5 paralogs KDM5A and KDM5C, suggesting that there may be overlapping or redundant
285 function for these proteins. Similar to our studies using *Drosophila*, some proteins identified were
286 known interactors, such as SIN3A that is known show similar genomic binding to KDM5B (Fig.
287 6C).[79] Other proteins overlapped with our genomic binding studies in *Drosophila*, including Brd4
288 (fs(1)h) and WDR5 (wds), while other proteins were identified because datasets were available in
289 human cells and not *Drosophila* (Fig. 6 C,D). These included the Mediator subunit MED1 and the
290 small ubiquitin-like modifier SUMO2 (*Drosophila* Sumo) (Fig. 6E,F). A combined genome browser
291 emphasizes the colocalization of these proteins with KDM5B (Fig. 6G). Thus, the high confidence
292 KDM5 interactors identified in *Drosophila* may be important for the function of KDM5B and other
293 KDM5 paralogs in mammalian cells.
294
295 **Identified KDM5 interactors are implicated in neurodevelopmental disorders**
296 Based on the association between genetic variants in KDM5A, KDM5B, and KDM5C and
297 intellectual disability (ID) and autism spectrum disorder (ASD), KDM5-interactors could potentially
298 be implicated in neurological disorders.[80–85] To examine this in more detail, we used the Simons
299 Foundation Autism Research Initiative (SFARI) genes as an up-to-date source of genes with

300 significant causal links to ASD.[86] The 1231 genes in this database are scored based on the level
301 of confidence of association, with a score of 1 being the strongest link, in addition to whether ASD
302 occurs as part of a syndrome (S). Using our list of 138 human ortholog-converted KDM5
303 interactors, we found that 26 of these overlap with SFARI ASD-associated proteins (p=8e-08; Fig.
304 7A-B). Some of these proteins have clear links to each other, such as TAF4 and TAF6 that are
305 components of TFIID, while others are associated with numerous other aspects of transcriptional
306 regulation. To look more broadly into the link between KDM5 interactors and neurodevelopmental
307 disorders, we used the Developmental Brain Disorder Gene Database (DBD) which is a curated
308 list of genes implicated in disorders such as ID, ASD, attention deficit hyperactivity disorder
309 (ADHD) and schizophrenia.[87] Twenty-four human-converted orthologs represented in DBD have
310 been shown to contribute to ID, ASD, ADHD, and Schizophrenia (Fig. 7C). Unsurprisingly given
311 the frequency that ID and ASD co-occur, 14 proteins were identified in both datasets, including
312 KDM5B, KDM5C and BRD4. In addition, 9 ID-associated proteins were identified, including
313 MED17, the NURD chromatin remodeling complex component GATAD2A and the C-terminal
314 binding protein (CtBP) transcriptional repressor. Combined, these analyses potentially greatly
315 expand our understanding of the KDM5-network of proteins and provide new avenues for
316 investigating the links between KDM5 family proteins and the etiology of neurodevelopmental
317 disorders. A summary of our KDM5 interaction data highlighting proteins with known roles in
318 transcriptional regulation is shown in Figure 7D.
319

320 **Discussion**

321 Here we describe the interactome of *Drosophila* KDM5 in the adult head using TurboID-
322 mediated proximity labeling. To identify the broadest selection of potential interactors, we N- and
323 C-terminally TurboID-tagged KDM5 as some interactors were expected to be in proximity to both
324 termini while others might be terminus specific. Importantly, TurboID-KDM5 chimeric proteins
325 were functional, as they were able to rescue the lethality caused by a $kdm5^{140}$ null allele. Using

326 NT-KDM5 and CT-KDM5, we performed two experiments to optimize the experimental controls,
327 as there are none currently established for this relatively new technique. Our study revealed that
328 expression of TurboID alone led to high background levels of biotinylation, particularly of
329 chromatin-related proteins. In contrast, using control (endogenous biotinylation and bead
330 background) and dCas9:TurboID provided similar and more reasonable background to identify
331 proteins enriched by expression of TurboID-KDM5. Because we carried out two separate MS
332 experiments, we were able to stringently filter our data to retain only those proteins that were
333 nuclear localized and showed high reproducibility and rigor. This led to the identification of 87
334 KDM5 high confidence interactors, 11 of which were previously described in either *Drosophila* or
335 mammals. Notably, while we refer to proteins identified in our TurboID analyses as interactors,
336 we acknowledge that proximity labeling does not necessarily detect direct interactions. However,
337 TurboID biotinylates lysine residues within 10 nm of its active site[56], which is equivalent to about
338 27 bp of B-DNA (8 bp per 3.4 nm). However unlikely, this short biotinylation radius can result in
339 the identification of proteins that are nearby but in a distinct complex that is not physically touching
340 KDM5. While these proteins are not in the same complex(es) as KDM5, they could still function
341 with KDM5 to regulate gene expression. For simplicity, we will refer TurboID-enriched proteins as
342 interactors.
343
344 Several lines of evidence allow us to have confidence in our described KDM5 interactome.
345 In addition to the reproducibility with which proteins were detected by NT- and CT-KDM5 across
346 the two experiments, many of the proteins identified had previously known physical connectivity
347 with each other. Thus, rather than identifying individual components of complexes, we identified
348 proteins well known to complex with each other, such as the SWI/SNF and NURF chromatin
349 remodeling complexes. Interestingly, a functional link between KDM5 and these complexes is
350 supported by studies in mouse embryonic stem cells which showed that a loss of KDM5B altered
351 nucleosome position surrounding the transcription start site (TSS), although the mechanism was

352 not revealed. We also identified the insulator proteins BEAF-32, Chromator, Putzig, and CP190
353 which complex together.[68,69] Our previous investigation in a fly strain harboring an allele
354 associated with human intellectual disability identified enrichment of BEAF-32 binding sites at
355 dysregulated genes. [28] Functionally, it is also notable that mutations in *kdm5*, *BEAF-32*, and *putzig*
356 all modify position effect variegation (PEV) suggesting that these proteins may function together
357 to regulate chromatin compaction and/or organization.[88–90] For some TurboID-identified proteins
358 such as the Mediator complex components (MED1, MED14, and MED17), as well as the TFIID
359 proteins (Taf4, Taf6, and Taf9), published structural data provides clear support for their link to
360 KDM5.[75,91] The Mediator proteins and Taf proteins neighbor each other, respectively, in the hTPIC
361 crystal structure. Moreover, we found that KDM5 interactions could be mapped to distinct surfaces
362 at the hTPIC providing clear evidence for KDM5 localization with respect to key transcriptional
363 initiation machinery. This specificity additionally suggests that the enrichment of these
364 transcriptional preinitiation proteins is not simply due to the binding of KDM5 proteins near the
365 promoter region of its target genes. If that were the case, then the entire preinitiation complex
366 would have been identified in our datasets, including TBP and RNA Pol II. Binding specificity in
367 this context can provide insight into how KDM5 carries out its functions at target gene promoters
368 to influence gene expression. Although limited by the number of available ChIP-seq datasets
369 available, corroborating evidence for our interactors also comes from the extensive overlap in
370 genomic binding observed in *Drosophila* and/or mammalian cells.
371
372 KDM5s relationship with other gene regulatory complexes provides insight into the links
373 between KDM5 dysregulation and human disorders. Many of the interacting complexes identified
374 in our study have been implicated in tumorigenesis, including NSL and SWI/SNF.[92–95]
375 Furthermore, like KDM5, MED1 has been implicated as a transcriptional coactivator that mediates
376 breast cancer metastasis and treatment resistance.[96,97] Identification of KDM5 interactors may
377 provide insight to novel mechanisms of KDM5-mediated transcriptional regulation which underlie

tumor development and progression. Changes to protein interactions could also contribute to the intellectual disability seen in individuals with genetic variants in KDM5A, KDM5B, or KDM5C. Indeed, for mutations that do not alter histone demethylase activity, this may be a contributor to cognitive dysfunction. Our analyses of KDM5 interactors revealed an enrichment in proteins found to be altered in neurodevelopmental disorders whose clinical presentations overlap with those seen for KDM5 genes. KDM5 and interactor mutations could therefore influence neurodevelopment through common pathways. Altogether, our study suggests that KDM5 likely functions through numerous transient interactions with interconnected complexes to regulate gene expression in a context-dependent manner.

ACKNOWLEDGMENTS

We thank members of the Secombe Lab, Melissa Castiglione, Blair Scheider, Michael Rogers, and Bethany Terry, as well as Michelle Schumacher for their feedback throughout this project and edits on the manuscript. We also thank the laboratory of Maria Cristina Gambetta for providing the anti-CP190 antibody and appreciate the availability of stocks from the Bloomington *Drosophila* Stock Center (NIH P40OD018537). We are additionally grateful to the Cancer Center Support Grant P30CA013330. This research was supported by the NIH Ruth L. Kirschstein National Research Service Award F31GM146347 and the Einstein MSTP Training Grant T32GM007288 to M.Y., NIH R01GM112783 and support from the Irma T. Hirschl Trust to J.S. The Sidoli lab gratefully acknowledges for financial support AFAR (Sagol Network GerOmics award), Deerfield (Xseed award), Relay Therapeutics, Merck, the NIH Office of the Director (1-S10-OD030286-01), the Einstein-Mount Sinai Diabetes Research Center, and the Einstein Cancer Center (P30-CA013330-47).

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as underrepresented ethnic minority in science. One or more of the authors self-identifies as a member of the LGBTQ+ community. One or more of the authors received support from a program designed to increase minority representation in science.

AUTHOR CONTRIBUTIONS

Conceptualization, M.Y., J.S.; Methodology, M.Y., S.S.; Investigation, M.Y. and J.S.; Writing – original draft, J.S. and M.Y., Writing – Reviewing and Editing, J.S., M.Y., and S.S..; Funding acquisition, J.S., M.Y and S.S., Supervision, J.S. and S.S.

416 **FIGURE LEGENDS**

417 **Figure 1: TurboID-tagged KDM5 proteins are functional and biotinylate endogenous**
418 **proteins**

419 (A) Schematic of the workflow used for purifying and fingerprinting biotinylated proteins from
420 adult heads using LC-MS/MS mass spectrometry.

421 (B) Schematic of the four UASp constructs generated able to express HA-tagged TurboID
422 alone, dCas9:TurboID, NT-KDM5 and CT-KDM5.

423 (C) Western blot using adult heads showing levels of expression of KDM5 using anti-HA,
424 biotinylation using a streptavidin-680 conjugate and the loading control alpha-tubulin.
425 Genotypes: *kdm5*^{140} ; *gkdm5:HA* (a wild-type strain; Control), Ubi-Gal4/+ ; UASp-
426 *TurboID*/+ (TurboID), Ubi-Gal4/+ ; UASp-*dCas9:TurboID* (dCas9:TurboID), *kdm5*^{140}, Ubi-
427 Gal4/*kdm5*^{140} ; UASp-NT- *kdm5*/+ (NT-KDM5) and *kdm5*^{140}, Ubi-Gal4/*kdm5*^{140} ; UASp-CT-
428 *kdm5* /+ (CT-KDM5).

429 (D) Rescue of *kdm5*^{140}-induced lethality by ubiquitous expression of UAS-NT-*KDM5* or UAS-
430 CT-*KDM5* using Ubi-Gal4. Genotype of male and female adult flies shown is *kdm5*^{140}, Ubi-
431 Gal4/*kdm5*^{140} ; UASp-NT- *kdm5* /+ (NT-KDM5) and *kdm5*^{140}, Ubi-Gal4/*KDM5*^{140} ; UASp-
432 CT- *kdm5* / + (CT-KDM5).

433

434 **Figure 2: Identifying candidate KDM5 interactors through TurboID**

435 (A) Principal component analysis of normalized nuclear protein abundances from control (Ctl;
436 grey), dCas9:TurboID (T-dCas9; green), TurboID (black), NT-KDM5 (red) and CT-KDM5
437 (blue).

438 (B) Summary of known *Drosophila* KDM5 interactors, whether the interaction is conserved in
439 mammals (mouse or human) and the identification of these proteins in experiment 2
440 compared to control (Ctl) and dCas9:TurboID. Three interactors identified in mammalian
441 cells but not previously in *Drosophila* are also included (Caf1-55, Mi-2 and Zmynd8).

442 (C) Volcano plot showing data comparing biotinylated proteins enriched by NT-KDM5 to
443 control.
444 (D) Volcano plot showing data comparing CT-KDM5 to control.
445 (E) Venn diagram showing overlap of the proteins identified using NT-KDM5 and CT-KDM5
446 compared to control (Ctl).
447 (F) Volcano plot showing data comparing NT-KDM5 to TurboID.
448 (G) Volcano plot showing data comparing CT-KDM5 to TurboID.
449 (H) Venn diagram showing overlap of the proteins identified comparing NT-KDM5 and CT-
450 KDM5 to TurboID.
451 (I) Volcano plot showing data comparing NT-KDM5 to dCas9:TurboID.
452 (J) Volcano plot showing data comparing CT-KDM5 to dCas9:TurboID.
453 (K) Venn diagram showing overlap of the proteins identified using NT-KDM5 and CT-KDM5
454 compared to dCas9:TurboID.
455 For all volcano plots shown, known interactors are indicated with text. Red dots indicate
456 significantly enriched proteins ($p<0.05$) and the dotted line on Y-axis indicates $p=0.05$.
457
458 **Figure 3: Identification of a high confidence KDM5 interactors**
459 (A) Filtering workflow for identification of high confidence KDM5 interactors by combining data
460 from experiments 1 and 2.
461 (B) Summary of known *Drosophila* KDM5 interactors, whether the interaction is conserved in
462 mammals (mouse or human) and the identification of these proteins in experiment 1 and
463 experiment 2 compared to control and dCas9:TurboID. Three interactors identified in
464 mammalian cells but not previously in *Drosophila* are also included (Caf1-55, Mi-2 and
465 Zmynd8).

466 (C) High confidence interactors based on their identification in experiment 1 (compared to
467 control) and experiment 2 (compared to control and dCas9:TurboID). Dark red box
468 indicates enrichment of $p<0.05$, pink indicates $p<0.1$.
469 (D) Volcano plot showing data comparing NT-KDM5 to dCas9:TurboID with high confidence
470 KDM5 interactors from panel C indicated with red dots. The dotted line on Y-axis indicates
471 $p=0.05$.
472 (E) Volcano plot showing data comparing CT-KDM5 to dCas9:TurboID with high confidence
473 KDM5 interactors from panel C indicated with red dots. The dotted line on Y-axis indicates
474 $p=0.05$.
475
476 **Figure 4: TurboID-tagged KDM5 biotinylates proteins involved in several aspects of gene**
477 **expression regulation.**
478 (A) STRING analyses of nuclear proteins that were significantly biotinylated by NT-KDM5 and
479 CT-KDM5. Grey lines indicate known physical interactions between proteins. Darker lines
480 indicate a higher confidence of interaction. Cytoscape was used to manually cluster
481 annotated proteins based on their STRING Cluster Enrichment and known functions.
482 Unconnected notes are not shown.
483 (B) Gene Ontology Biological Process (GO-BP) analyses of high confidence proteins.
484 (C) Anti-BEAF-32 and anti-CP190 blot from biotinylated proteins purified from adult heads of
485 the genotypes shown. Input shows the presence of BEAF-32 and CP190, respectively,
486 across all samples.
487 (D) Structure of the human pre-initiation complex (7ENA) showing proteins identified as high
488 confidence interactors in our TurboID data in red bubbles. DNA (yellow), TBP (pink
489 bubbles) and RNA polymerase II (cyan bubbles) are also shown.
490

491 **Figure 5: Proteins enriched in KDM5 TurboID experiments show overlapping genomic**
492 **binding profiles.**

493 (A) Venn diagram showing overlap between high confidence biotinylated proteins (red) and
494 datasets enriched when comparing published KDM5 ChIP-seq data from whole adult flies
495 to other *Drosophila* ChIP datasets (blue). A total of 27 datasets showed significant overlap
496 using ChIP-atlas. 30 interactors had ChIP-seq datasets in the ChIP Atlas database
497 (green).

498 (B) Volcano plot showing fold enrichment and p-values of the permutation analyses between
499 KDM5 and 27 ChIP-seq datasets of high confidence interactors.

500 (C) Genomic binding profiles of Sin3A ChIP from S2 cells and KDM5 showing similar
501 localization.

502 (D) Genomic binding profiles of BEAF-32 ChIP from S2 cells and KDM5 showing similar
503 localization.

504 (E) Genomic binding profiles of Dref ChIP from Kc167 cells and KDM5 showing similar
505 localization.

506 (F) Genomic binding profiles of Iswi ChIP from Kc167 cells and KDM5 showing similar
507 localization.

508 (G) Genomic binding profiles of wds ChIP from BG3 cells and KDM5 showing similar
509 localization.

510 (H) Genomic binding profiles of fs(1)h ChIP from Kc167 cells and KDM5 showing similar
511 localization.

512 (I) Representative genome browser image showing binding of KDM5, Sin3A, BEAF-32, Dref,
513 fs(1)h, Iswi and wds.

514

515 **Figure 6: Human orthologs of proteins enriched in *Drosophila* KDM5 TurboID experiments**
516 **show overlapping genomic binding profiles.**

517 (A) Venn diagram showing overlap between human orthologs of *Drosophila* KDM5 high
518 confidence interactors and their enrichment when comparing published KDM5B ChIP-seq
519 data to breast cancer ChIP-seq datasets using ChIP-Atlas. 26 datasets significantly
520 overlapped.
521 (B) Volcano plot showing fold enrichment and p-values of the permutation analyses between
522 KDM5B and 26 ChIP-seq datasets of human-ortholog converted high confidence
523 interactors.
524 (C) Genomic binding profiles of KDM5B, SIN3A and BRD4 showing similar genome-wide
525 binding. Binding is shown relative to KDM5B due to the much larger number of SIN3A and
526 BRD4 binding sites in the genome compared to KDM5B.
527 (D) Genomic binding profiles of WDR5 (ortholog of *Drosophila* wds) ChIP from MCF-10 cells
528 and KDM5B showing similar localization.
529 (E) Genomic binding profiles of MED1 ChIP from MCF-7 cells and KDM5B showing similar
530 localization.
531 (F) Genomic binding profiles of SUMO2 (ortholog of *Drosophila* Sumo) ChIP from MCF-7 cells
532 and KDM5B showing similar localization.
533 (G) Representative genome browser image showing binding of KDM5B, SIN3A, MED1,
534 BRD4, WDR5 and SUMO2.
535
536 **Figure 7: A subset of KDM5 interactors are implicated in neurodevelopmental disorders**
537 (A) Overlap between human orthologs of *Drosophila* KDM5 interactors and genes associated
538 with ASD using the SFARI database. S indicates syndromic ASD. Scores indicate
539 confidence of causal association, with 1 indicating strongest link.
540 (B) 26 candidate interacting proteins identified and their SFARI score.
541 (C) 24 candidate interacting proteins were identified as being implicated in ID, ASD and/or
542 schizophrenia using DBD.

543 (D) Model for KDM5 interactions with key transcriptional proteins that are likely to impact the
544 expression of downstream target genes.
545
546
547 **MATERIALS AND METHODS**
548 **Fly strains and care**
549 Fly crosses were maintained at 25°C with 50% humidity and a 12-hour light/dark cycle. Food (per
550 liter) contained 18g yeast, 22g molasses, 80g malt extract, 9g agar, 65 cornmeal, 2.3g methyl
551 para-benzoic acid, 6.35ml propionic acid. The number of male and female animals were equal
552 across the genotypes examined. The *kdm5*^{140} null allele has been previously described. [35]
553
554 **Cloning and Transgenesis**
555 The N- and C-terminally TurboID-tagged constructs were generated by cloning the coding region
556 of *kdm5* upstream or downstream of HA:TurboID from pCDNA3-TurboID
557 (RRID:Addgene_107171)[48] in the pUASpattB vector (RRID:DGRC_1358). UASp-HA:TurboID
558 with a NLS was generated by cloning HA:TurboID into the same UASpattB vector. UASp-
559 dCas:HA:Turbo:NLS was made by combining the dCas9 open reading frame from SID3s-dCas9-
560 KRAB (RRID:Addgene_106399) with HA:TurboID:NLS in the pUASpattB vector
561 (RRID:DGRC_1358). All transgenes were generated by injection into y^1 M{RFP[3xP3.PB]
562 GFP[E.3xP3]=vas-int.Dm}ZH-2A w^* ; M{3xP3-RFP.attP}ZH-86Fb at BestGene Inc.
563
564 **Western Blotting**
565 Western analyses were carried out as previously described. [35] Briefly, five 2- to 5-day old adult fly
566 heads were homogenized in 2x NuPAGE LDS sample buffer, sonicated for 10 mins, treated with
567 DTT, run on a 4-12% Bis-Tris 1 mm gel and transferred to a PVDF membrane. The following
568 primary antibodies were used: mouse anti-HA (1:1000, Cell Signaling Technology Cat# 2367,

569 RRID: AB_10691311), Streptavidin 680 (1:10,000, ThermoFisher, Streptavidin Alexa Fluor 680
570 conjugate), rabbit anti-alpha-Tubulin (1:5000, Cell Signaling Technology Cat# 2144,
571 RRID:AB_2210548), mouse anti-BEAF-32 (1:500, DSHB Cat# anti-BEAF, RRID:AB_1553420),
572 and rabbit anti-CP190 (1:5000, Kaushal et. al 2021[98], non-commercial). Secondary antibodies
573 used were IRDye® 680RD Donkey anti-Mouse IgG (1:5000; LI-COR Biosciences Cat# 925-
574 68072, RRID: AB_2814912) and IRDye® 800CW Donkey anti-Rabbit IgG (1:5000; LI-COR
575 Biosciences Cat# 926-32213, RRID: AB_621848). Blots were scanned and processed using a LI-
576 COR Odyssey Infrared scanner.
577
578 **Purifying and identifying proteins using TurboID**
579 **Biotinylated Protein Enrichment**
580 2-5 day-old flies were flash frozen in liquid nitrogen and decapitated and a total of ten heads were
581 used per sample. Heads were homogenized in 250 μL RIPA Buffer (Thermofisher 89901)
582 supplemented with Halt™ Protease Inhibitor Cocktail (Thermofisher, 78430) and centrifuged at 4
583 °C for 10 minutes at 15,000XG to remove debris. 100 μL of Pierce™ Streptavidin Magnetic Beads
584 (Thermofisher, 88817) were washed twice with RIPA and the cleared lysate was added. The
585 lysate-bead mixture was incubated with rotation at 4 °C overnight. The next day the lysate was
586 discarded, and beads were washed twice with RIPA, once with 1M KCl, once with 0.1 M Na_2HCO_3,
587 once with 1 M Urea in 10 mM Tris pH 8.0, and twice again with RIPA. For Western Blot analyses
588 all RIPA was removed and biotinylated proteins were eluted with 4X NuPAGE™ LDS Sample
589 Buffer (Invitrogen, NP0007) supplemented with 2 mM biotin and 20 mM DTT.
590 **On-bead protein digestion**
591 Proteins were digested directly on streptavidin beads. 5 mM DTT and 50 mM ammonium
592 bicarbonate (pH = 8) were added to the solution and left on the bench for about 1 hour for disulfide
593 bond reduction. Samples were then alkylated with 20 mM iodoacetamide in the dark for 30

594 minutes. Afterward, 500 ng of trypsin was added to the samples, which were digested at 37 °C
595 for 18 h. The peptide solution was dried in a vacuum centrifuge.

596 **Sample desalting**

597 Prior to mass spectrometry analysis, samples were desalted using a 96-well plate filter (Orochem)
598 packed with 1 mg of Oasis HLB C-18 resin (Waters). Briefly, the samples were resuspended in
599 100 µl of 0.1% TFA and loaded onto the HLB resin, which was previously equilibrated using 100
600 µl of the same buffer. After washing with 100 µl of 0.1% TFA, the samples were eluted with a
601 buffer containing 70 µl of 60% acetonitrile and 0.1% TFA and then dried in a vacuum centrifuge.

602 **LC-MS/MS Acquisition and Analysis**

603 Samples were resuspended in 10 µl of 0.1% TFA and loaded onto a Dionex RSLC Ultimate 300
604 (Thermo Scientific), coupled online with an Orbitrap Fusion Lumos (Thermo Scientific).
605 Chromatographic separation was performed with a two-column system, consisting of a C-18 trap
606 cartridge (300 µm ID, 5 mm length) and a picofrit analytical column (75 µm ID, 25 cm length)
607 packed in-house with reversed-phase Repro-Sil Pur C18-AQ 3 µm resin. Peptides were separated
608 using a 90 min gradient from 4-30% buffer B (buffer A: 0.1% formic acid, buffer B: 80% acetonitrile
609 + 0.1% formic acid) at a flow rate of 300 nL/min. The mass spectrometer was set to acquire
610 spectra in a data-dependent acquisition (DDA) mode. Briefly, the full MS scan was set to 300-
611 1200 m/z in the orbitrap with a resolution of 120,000 (at 200 m/z) and an AGC target of 5x10e5.
612 MS/MS was performed in the ion trap using the top speed mode (2 secs), an AGC target of 1x10e4
613 and an HCD collision energy of 35. Raw files were searched using Proteome Discoverer software
614 (v2.4, Thermo Scientific) using SEQUEST search engine and the UniProt database of Drosophila
615 melanogaster. The search for total proteome included variable modification of N-terminal
616 acetylation, and fixed modification of carbamidomethyl cysteine. Trypsin was specified as the
617 digestive enzyme with up to 2 missed cleavages allowed. Mass tolerance was set to 10 pm for
618 precursor ions and 0.2 Da for product ions. Peptide and protein false discovery rate was set to
619 1%. Following the search, data was processed as described by Aguilan et al.[99]. Briefly, proteins

620 were log2 transformed, normalized by the average value of each sample and missing values were
621 imputed using a normal distribution 2 standard deviations lower than the mean. Statistical
622 regulation was assessed using heteroscedastic T-test (if p-value < 0.05). Data were assumed to
623 be Gaussian distributed but this was not formally tested.
624
625 **Bioinformatic Analyses**
626 The Enrichment Analysis function on ChIP Atlas[76] was used to perform permutation tests.
627 Experiment types were set to "TFs and others", Cell type class was set to 'All types' and 'Breast'
628 for Drosophila and Human respectively. A 100X random permutation of each was used as the
629 control. For *Drosophila*, selected profiles were generated using: KDM5 (SRX1084165), Sin3A
630 (SRX1158165), BEAF-32 (SRX386677), Dref (SRX170984), Iswi (SRX5346167), wds
631 (SRX2399656), and fs(1)h (SRX203000). For Human, selected profiles generated using: KDM5B
632 (SRX3285561), SIN3A (SRX190318), MED1 (SRX673749), WDR5 (SRX3958389), BRD4
633 (SRX5089551), and SUMO2 (SRX3541112). Deeptools(3.5.1)[100] computeMatrix and
634 plotHeatmap functions were used to make profiles and heatmaps, and Pygenometracks(3.7)[101]
635 was used to visualize ChIP-seq tracks.
636

637

638 **References**

639 1. Pavlenko, E., Ruengeler, T., Engel, P. & Poepsel, S. Functions and interactions of mammalian
640 KDM5 demethylases. *Frontiers in genetics* **13**, 906662 (2022).
641 2. Holowatyj, A., Yang, Z.-Q. & Pile, L. A. Histone lysine demethylases in Drosophila
642 melanogaster. *Fly (Austin)* **9**, 36–44 (2015).
643 3. Lee, N. *et al.* The trithorax-group protein Lid is a histone H3 trimethyl-Lys4 demethylase.
644 *Nature Structural & Molecular Biology* **14**, 341–343 (2007).
645 4. Secombe, J., Li, L., Carlos, L. & Eisenman, R. N. The Trithorax group protein Lid is a
646 trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. *Genes &*
647 *Development* **21**, 537–551 (2007).
648 5. Klose, R. J. *et al.* The retinoblastoma binding protein RBP2 is an H3K4 demethylase. *Cell*
649 **128**, 889–900 (2007).
650 6. Klein, B. J. *et al.* The histone-H3K4-specific demethylase KDM5B binds to its substrate and
651 product through distinct PHD fingers. *Cell reports* **6**, 325–335 (2014).
652 7. Yang, G.-J. *et al.* The emerging role of KDM5A in human cancer. *J Hematol Oncol* **14**, 30
653 (2021).
654 8. Blair, L. P., Cao, J., Zou, M. R., Sayegh, J. & Yan, Q. Epigenetic regulation by lysine
655 demethylase 5 (KDM5) enzymes in cancer. *Cancers* **3**, 1383–1404 (2011).
656 9. Hatch, H. A. M. & Secombe, J. Molecular and cellular events linking variants in the histone
657 demethylase KDM5C to the intellectual disability disorder Claes-Jensen syndrome. *FEBS J*
658 (2021) doi:10.1111/febs.16204.

659 10. Yamamoto, S. *et al.* JARID1B is a luminal lineage-driving oncogene in breast cancer.
660 *Cancer Cell* **25**, 762–777 (2014).
661 11. Hinohara, K. *et al.* KDM5 histone demethylase activity links cellular transcriptomic
662 heterogeneity to therapeutic resistance. *Cancer Cell* **34**, 939-953.e9 (2018).
663 12. Horton, J. R. *et al.* Characterization of a linked jumonji domain of the KDM5/JARID1
664 family of histone H3 lysine 4 demethylases. *The Journal of Biological Chemistry* **291**, 2631–
665 2646 (2016).
666 13. Arifuzzaman, S., Khatun, M. R. & Khatun, R. Emerging of lysine demethylases (KDMs):
667 From pathophysiological insights to novel therapeutic opportunities. *Biomedicine &*
668 *Pharmacotherapy (Biomedecine & Pharmacotherapie)* **129**, 110392 (2020).
669 14. He, R. *et al.* H3K4 demethylase KDM5B regulates cancer cell identity and epigenetic
670 plasticity. *Oncogene* **41**, 2958–2972 (2022).
671 15. Xhabija, B. & Kidder, B. L. KDM5B is a master regulator of the H3K4-methylome in
672 stem cells, development and cancer. *Semin Cancer Biol* **57**, 79–85 (2019).
673 16. Roesch, A. *et al.* Overcoming intrinsic multidrug resistance in melanoma by blocking the
674 mitochondrial respiratory chain of slow-cycling JARID1B(high) cells. *Cancer Cell* **23**, 811–
675 825 (2013).
676 17. Iwase, S. *et al.* The X-linked mental retardation gene SMCX/JARID1C defines a family
677 of histone H3 lysine 4 demethylases. *Cell* **128**, 1077–1088 (2007).
678 18. Hatch, H. A. M., O'Neil, M. H., Marion, R. W., Secombe, J. & Shulman, L. H.
679 Caregiver-reported characteristics of children diagnosed with pathogenic variants in KDM5C.
680 *American Journal of Medical Genetics. Part A* (2021) doi:10.1002/ajmg.a.62381.

681 19. Abidi, F. E. *et al.* Mutations in JARID1C are associated with X-linked mental retardation,
682 short stature and hyperreflexia. *Journal of Medical Genetics* **45**, 787–793 (2008).
683 20. Tzschach, A. *et al.* Novel JARID1C/SMCX mutations in patients with X-linked mental
684 retardation. *Human Mutation* **27**, 389 (2006).
685 21. Jensen, L. R. *et al.* Mutations in the JARID1C gene, which is involved in transcriptional
686 regulation and chromatin remodeling, cause X-linked mental retardation. *American Journal of*
687 *Human Genetics* **76**, 227–236 (2005).
688 22. Harrington, J., Wheway, G., Willaime-Morawek, S., Gibson, J. & Walters, Z. S.
689 Pathogenic KDM5B variants in the context of developmental disorders. *Biochim Biophys Acta*
690 *Gene Regul Mech* **1865**, 194848 (2022).
691 23. Iwase, S. *et al.* A Mouse Model of X-linked Intellectual Disability Associated with
692 Impaired Removal of Histone Methylation. *Cell Rep* **14**, 1000–1009 (2016).
693 24. Scandaglia, M. *et al.* Loss of Kdm5c Causes Spurious Transcription and Prevents the
694 Fine-Tuning of Activity-Regulated Enhancers in Neurons. *Cell Rep* **21**, 47–59 (2017).
695 25. Mariani, L., Lussi, Y. C., Vandamme, J., Riveiro, A. & Salcini, A. E. The H3K4me3/2
696 histone demethylase RBR-2 controls axon guidance by repressing the actin-remodeling gene
697 wsp-1. *Development* **143**, 851–863 (2016).
698 26. Belalcazar, H. M., Hendricks, E. L., Zamurrad, S., Liebl, F. L. W. & Secombe, J. The
699 histone demethylase KDM5 is required for synaptic structure and function at the Drosophila
700 neuromuscular junction. *Cell reports* **34**, 108753 (2021).
701 27. Hatch, H. A., Belalcazar, H. M., Marshall, O. J. & Secombe, J. A KDM5-Prospero
702 transcriptional axis functions during early neurodevelopment to regulate mushroom body
703 formation. *eLife* **10**, (2021).

704 28. Zamurrad, S., Hatch, H. A. M., Drelon, C., Belalcazar, H. M. & Secombe, J. A
705 Drosophila Model of Intellectual Disability Caused by Mutations in the Histone Demethylase
706 KDM5. *Cell Rep* **22**, 2359–2369 (2018).
707 29. Yao, W., Peng, Y. & Lin, D. The flexible loop L1 of the H3K4 demethylase JARID1B
708 ARID domain has a crucial role in DNA-binding activity. *Biochem Biophys Res Commun* **396**,
709 323–328 (2010).
710 30. Tu, S. *et al.* The ARID domain of the H3K4 demethylase RBP2 binds to a DNA
711 CCGCCC motif. *Nat Struct Mol Biol* **15**, 419–421 (2008).
712 31. Scibetta, A. G. *et al.* Functional analysis of the transcription repressor PLU-1/JARID1B.
713 *Mol Cell Biol* **27**, 7220–7235 (2007).
714 32. Villaseñor, R. *et al.* ChromID identifies the protein interactome at chromatin marks. *Nat*
715 *Biotechnol* **38**, 728–736 (2020).
716 33. Li, L., Greer, C., Eisenman, R. N. & Secombe, J. Essential functions of the histone
717 demethylase lid. *PLoS Genetics* **6**, e1001221 (2010).
718 34. Wang, G. G. *et al.* Haematopoietic malignancies caused by dysregulation of a chromatin-
719 binding PHD finger. *Nature* **459**, 847–851 (2009).
720 35. Drelon, C., Belalcazar, H. M. & Secombe, J. The histone demethylase KDM5 is essential
721 for larval growth in drosophila. *Genetics* **209**, 773–787 (2018).
722 36. Belalcazar, H. M., Hendricks, E. L., Zamurrad, S., Liebl, F. L. W. & Secombe, J. The
723 histone demethylase KDM5 is required for synaptic structure and function at the Drosophila
724 neuromuscular junction. *Cell Rep* **34**, 108753 (2021).

725 37. Hatch, H. A., Belalcazar, H. M., Marshall, O. J. & Secombe, J. A KDM5-Prospero
726 transcriptional axis functions during early neurodevelopment to regulate mushroom body
727 formation. *Elife* **10**, e63886 (2021).
728 38. Brookes, E. *et al.* Mutations in the intellectual disability gene KDM5C reduce protein
729 stability and demethylase activity. *Hum Mol Genet* **24**, 2861–2872 (2015).
730 39. Vallianatos, C. N. *et al.* Altered Gene-Regulatory Function of KDM5C by a Novel
731 Mutation Associated With Autism and Intellectual Disability. *Front Mol Neurosci* **11**, 104
732 (2018).
733 40. Tahiliani, M. *et al.* The histone H3K4 demethylase SMCX links REST target genes to X-
734 linked mental retardation. *Nature* **447**, 601–605 (2007).
735 41. Paroni, G. *et al.* HER2-positive breast-cancer cell lines are sensitive to KDM5 inhibition:
736 definition of a gene-expression model for the selection of sensitive cases. *Oncogene* **38**, 2675–
737 2689 (2019).
738 42. Cao, J. *et al.* Histone demethylase RBP2 is critical for breast cancer progression and
739 metastasis. *Cell Rep* **6**, 868–877 (2014).
740 43. Moshkin, Y. M. *et al.* Histone chaperones ASF1 and NAP1 differentially modulate
741 removal of active histone marks by LID-RPD3 complexes during NOTCH silencing. *Mol Cell*
742 **35**, 782–793 (2009).
743 44. Varier, R. A. *et al.* Recruitment of the Mammalian Histone-modifying EMSY Complex
744 to Target Genes Is Regulated by ZNF131. *J Biol Chem* **291**, 7313–7324 (2016).
745 45. Nishibuchi, G. *et al.* Physical and functional interactions between the histone H3K4
746 demethylase KDM5A and the nucleosome remodeling and deacetylase (NuRD) complex. *J*
747 *Biol Chem* **289**, 28956–28970 (2014).

748 46. Ohguchi, H. *et al.* Lysine Demethylase 5A Is Required for MYC-Driven Transcription in
749 Multiple Myeloma. *Blood Cancer Discovery* **2**, 370–387 (2021).
750 47. Nishibuchi, G. *et al.* Physical and functional interactions between the histone H3K4
751 demethylase KDM5A and the nucleosome remodeling and deacetylase (NuRD) complex. *The*
752 *Journal of Biological Chemistry* **289**, 28956–28970 (2014).
753 48. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with
754 TurboID. *Nature Biotechnology* **36**, 880–887 (2018).
755 49. Shokri, L. *et al.* A Comprehensive Drosophila melanogaster Transcription Factor
756 Interactome. *Cell reports* **27**, 955-970.e7 (2019).
757 50. Uezu, A. *et al.* Identification of an elaborate complex mediating postsynaptic inhibition.
758 *Science* **353**, 1123–1129 (2016).
759 51. Zhang, B., Zhang, Y. & Liu, J.-L. Highly effective proximate labeling in Drosophila. *G3*
760 *(Bethesda, Md.)* **11**, (2021).
761 52. Szczesniak, L. M., Bonzerato, C. G. & Wojcikiewicz, R. J. H. Identification of the bok
762 interactome using proximity labeling. *Frontiers in cell and developmental biology* **9**, 689951
763 (2021).
764 53. Mannix, K. M., Starble, R. M., Kaufman, R. S. & Cooley, L. Proximity labeling reveals
765 novel interactomes in live Drosophila tissue. *Development* **146**, (2019).
766 54. Sanchez, A. D. *et al.* Proximity labeling reveals non-centrosomal microtubule-organizing
767 center components required for microtubule growth and localization. *Current Biology* **31**,
768 3586-3600.e11 (2021).

769 55. Chua, X. Y., Aballo, T., Elnemer, W., Tran, M. & Salomon, A. Quantitative
770 Interactomics of Lck-TurboID in Living Human T Cells Unveils T Cell Receptor Stimulation-
771 Induced Proximal Lck Interactors. *Journal of Proteome Research* **20**, 715–726 (2021).
772 56. Cho, K. F. *et al.* Proximity labeling in mammalian cells with TurboID and split-TurboID.
773 *Nature Protocols* **15**, 3971–3999 (2020).
774 57. Stockhammer, A. *et al.* When less is more – Endogenous tagging with TurboID as a tool
775 to study the native interactome of adaptor protein complexes. *bioRxiv* 2021.11.19.469212
776 (2022) doi:10.1101/2021.11.19.469212.
777 58. Drelon, C., Rogers, M. F., Belalcazar, H. M. & Secombe, J. The histone demethylase
778 KDM5 controls developmental timing in Drosophila by promoting prothoracic gland
779 endocycles. *Development* **146**, dev182568 (2019).
780 59. Artan, M. *et al.* Interactome analysis of Caenorhabditis elegans synapses by TurboID-
781 based proximity labeling. *Journal of Biological Chemistry* **297**, 101094 (2021).
782 60. Hu, S. *et al.* Identification of mutant p53-specific proteins interaction network using
783 TurboID-based proximity labeling. *Biochemical and Biophysical Research Communications*
784 **615**, 163–171 (2022).
785 61. Chevalier, B. *et al.* Differential CFTR-Interactome Proximity Labeling Procedures
786 Identify Enrichment in Multiple SLC Transporters. *Int J Mol Sci* **23**, 8937 (2022).
787 62. Uçkun, E. *et al.* In vivo Profiling of the Alk Proximitome in the Developing Drosophila
788 Brain. *Journal of Molecular Biology* **433**, 167282 (2021).
789 63. Burmistrz, M., Krakowski, K. & Krawczyk-Balska, A. RNA-Targeting CRISPR-Cas
790 Systems and Their Applications. *Int J Mol Sci* **21**, E1122 (2020).

791 64. Szklarczyk, D. *et al.* The STRING database in 2021: customizable protein-protein
792 networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic*
793 *Acids Res* **49**, D605–D612 (2021).
794 65. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of
795 biomolecular interaction networks. *Genome Research* **13**, 2498–2504 (2003).
796 66. Kaushal, A. *et al.* Essential role of Cp190 in physical and regulatory boundary formation.
797 *Science Advances* **8**, eabl8834 (2022).
798 67. Ramírez, F. *et al.* High-resolution TADs reveal DNA sequences underlying genome
799 organization in flies. *Nature Communications* **9**, 189 (2018).
800 68. Sabirov, M. *et al.* Mechanisms of CP190 Interaction with Architectural Proteins in
801 Drosophila Melanogaster. *International Journal of Molecular Sciences* **22**, (2021).
802 69. Melnikova, L. S., Molodina, V. V., Kostyuchenko, M. V., Georgiev, P. G. & Golovnin,
803 A. K. The BEAF-32 Protein Directly Interacts with Z4/putzig and Chriz/Chromator Proteins
804 in Drosophila melanogaster. *Doklady. Biochemistry and biophysics* **498**, 184–189 (2021).
805 70. Chathoth, K. T. *et al.* The role of insulators and transcription in 3D chromatin
806 organization of flies. *Genome Research* (2022) doi:10.1101/gr.275809.121.
807 71. Schwartz, Y. B. & Cavalli, G. Three-Dimensional Genome Organization and Function in
808 Drosophila. *Genetics* **205**, 5–24 (2017).
809 72. He, R. & Kidder, B. L. H3K4 demethylase KDM5B regulates global dynamics of
810 transcription elongation and alternative splicing in embryonic stem cells. *Nucleic Acids*
811 *Research* **45**, 6427–6441 (2017).
812 73. Hinohara, K. *et al.* KDM5 Histone Demethylase Activity Links Cellular Transcriptomic
813 Heterogeneity to Therapeutic Resistance. *Cancer Cell* **34**, 939-953.e9 (2018).

814 74. Liu, X., Greer, C. & Secombe, J. KDM5 interacts with Foxo to modulate cellular levels
815 of oxidative stress. *PLoS Genetics* **10**, e1004676 (2014).
816 75. Chen, X. *et al.* Structures of the human Mediator and Mediator-bound preinitiation
817 complex. *Science* **372**, eabg0635 (2021).
818 76. Zou, Z., Ohta, T., Miura, F. & Oki, S. ChIP-Atlas 2021 update: a data-mining suite for
819 exploring epigenomic landscapes by fully integrating ChIP-seq, ATAC-seq and Bisulfite-seq
820 data. *Nucleic Acids Research* **50**, W175–W182 (2022).
821 77. Hu, Y. *et al.* An integrative approach to ortholog prediction for disease-focused and other
822 functional studies. *BMC Bioinformatics* **12**, 357 (2011).
823 78. Hinohara, K. *et al.* KDM5 Histone Demethylase Activity Links Cellular Transcriptomic
824 Heterogeneity to Therapeutic Resistance. *Cancer Cell* **34**, 939-953.e9 (2018).
825 79. Spain, M. M., Caruso, J. A., Swaminathan, A. & Pile, L. A. Drosophila SIN3 isoforms
826 interact with distinct proteins and have unique biological functions. *J Biol Chem* **285**, 27457–
827 27467 (2010).
828 80. Zamurrad, S., Hatch, H. A. M., Drelon, C., Belalcazar, H. M. & Secombe, J. A
829 drosophila model of intellectual disability caused by mutations in the histone demethylase
830 KDM5. *Cell reports* **22**, 2359–2369 (2018).
831 81. Iwase, S. *et al.* A Mouse Model of X-linked Intellectual Disability Associated with
832 Impaired Removal of Histone Methylation. *Cell reports* **14**, 1000–1009 (2016).
833 82. Poeta, L., Padula, A., Lioi, M. B., van Bokhoven, H. & Miano, M. G. Analysis of a set of
834 KDM5C regulatory genes mutated in neurodevelopmental disorders identifies temporal
835 coexpression brain signatures. *Genes* **12**, (2021).

836 83. Gonçalves, T. F. *et al.* KDM5C mutational screening among males with intellectual
837 disability suggestive of X-Linked inheritance and review of the literature. *European Journal*
838 *of Medical Genetics* **57**, 138–144 (2014).
839 84. Brookes, E. *et al.* Mutations in the intellectual disability gene KDM5C reduce protein
840 stability and demethylase activity. *Human Molecular Genetics* **24**, 2861–2872 (2015).
841 85. Hatch, H. A. M. & Secombe, J. Molecular and cellular events linking variants in the
842 histone demethylase KDM5C to the intellectual disability disorder Claes-Jensen syndrome.
843 *The FEBS Journal* (2021) doi:10.1111/febs.16204.
844 86. Abrahams, B. S. *et al.* SFARI Gene 2.0: a community-driven knowledgebase for the
845 autism spectrum disorders (ASDs). *Mol Autism* **4**, 36 (2013).
846 87. Gonzalez-Mantilla, A. J., Moreno-De-Luca, A., Ledbetter, D. H. & Martin, C. L. A
847 Cross-Disorder Method to Identify Novel Candidate Genes for Developmental Brain
848 Disorders. *JAMA Psychiatry* **73**, 275–283 (2016).
849 88. Eggert, H., Gortchakov, A. & Saumweber, H. Identification of the Drosophila interband-
850 specific protein Z4 as a DNA-binding zinc-finger protein determining chromosomal structure.
851 *Journal of Cell Science* **117**, 4253–4264 (2004).
852 89. Shalaby, N. A. *et al.* Systematic discovery of genetic modulation by Jumonji histone
853 demethylases in Drosophila. *Scientific Reports* **7**, 5240 (2017).
854 90. Maharjan, M., McKowen, J. K. & Hart, C. M. Overlapping but Distinct Sequences Play
855 Roles in the Insulator and Promoter Activities of the Drosophila BEAF-Dependent scs'
856 Insulator. *Genetics* **215**, 1003–1012 (2020).
857 91. Yin, J. & Wang, G. The Mediator complex: a master coordinator of transcription and cell
858 lineage development. *Development* **141**, 977–987 (2014).

859 92. Ferreyra Solari, N. E. *et al.* The NSL Chromatin-Modifying Complex Subunit KANSL2
860 Regulates Cancer Stem-like Properties in Glioblastoma That Contribute to Tumorigenesis.
861 *Cancer Res* **76**, 5383–5394 (2016).
862 93. Mittal, P. & Roberts, C. W. M. The SWI/SNF complex in cancer - biology, biomarkers
863 and therapy. *Nat Rev Clin Oncol* **17**, 435–448 (2020).
864 94. Wanior, M., Krämer, A., Knapp, S. & Joerger, A. C. Exploiting vulnerabilities of
865 SWI/SNF chromatin remodelling complexes for cancer therapy. *Oncogene* **40**, 3637–3654
866 (2021).
867 95. Singh, M. *et al.* Histone Acetyltransferase MOF Orchestrates Outcomes at the Crossroad
868 of Oncogenesis, DNA Damage Response, Proliferation, and Stem Cell Development. *Mol*
869 *Cell Biol* **40**, e00232-20 (2020).
870 96. Cao, J. *et al.* Histone demethylase RBP2 is critical for breast cancer progression and
871 metastasis. *Cell Rep* **6**, 868–877 (2014).
872 97. Cui, J. *et al.* Cross-talk between HER2 and MED1 regulates tamoxifen resistance of
873 human breast cancer cells. *Cancer Res* **72**, 5625–5634 (2012).
874 98. Kaushal, A. *et al.* CTCF loss has limited effects on global genome architecture in
875 Drosophila despite critical regulatory functions. *Nature Communications* **12**, 1011 (2021).
876 99. Aguilan, J. T., Kulej, K. & Sidoli, S. Guide for protein fold change and p-value
877 calculation for non-experts in proteomics. *Mol Omics* **16**, 573–582 (2020).
878 100. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data
879 analysis. *Nucleic Acids Research* **44**, W160–W165 (2016).
880 101. Ramírez, F. *et al.* High-resolution TADs reveal DNA sequences underlying genome
881 organization in flies. *Nat Commun* **9**, 189 (2018).

882

Fig1

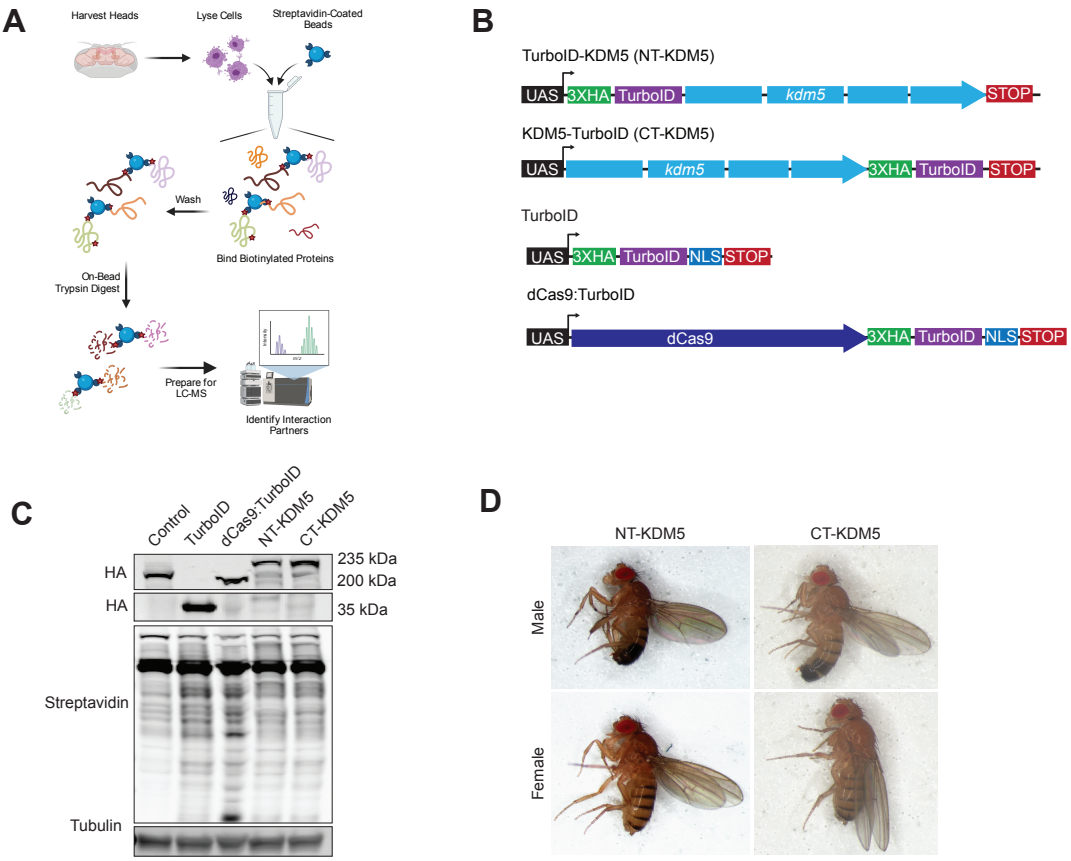

A
Harvest Heads
Lyse Cells
Streptavidin-Coated Beads
Wash
Bind Biotinylated Proteins
On-Bead Trypsin Digest
Prepare for LC-MS
Identify Interaction Partners
B
TurboID-KDM5 (NT-KDM5)
UAS
3XHA
TurboID
kdm5
STOP
KDM5-TurboID (CT-KDM5)
UAS
kdm5
3XHA
TurboID
STOP
TurboID
UAS
3XHA
TurboID
NLS
STOP
dCas9:TurboID
UAS
dCas9
3XHA
TurboID
NLS
STOP
C
Control
TurboID
dCas9:TurboID
NT-KDM5
CT-KDM5
HA
235 kDa
200 kDa
HA
35 kDa
Streptavidin
Tubulin
D
NT-KDM5
CT-KDM5
Male
Female

Fig 2

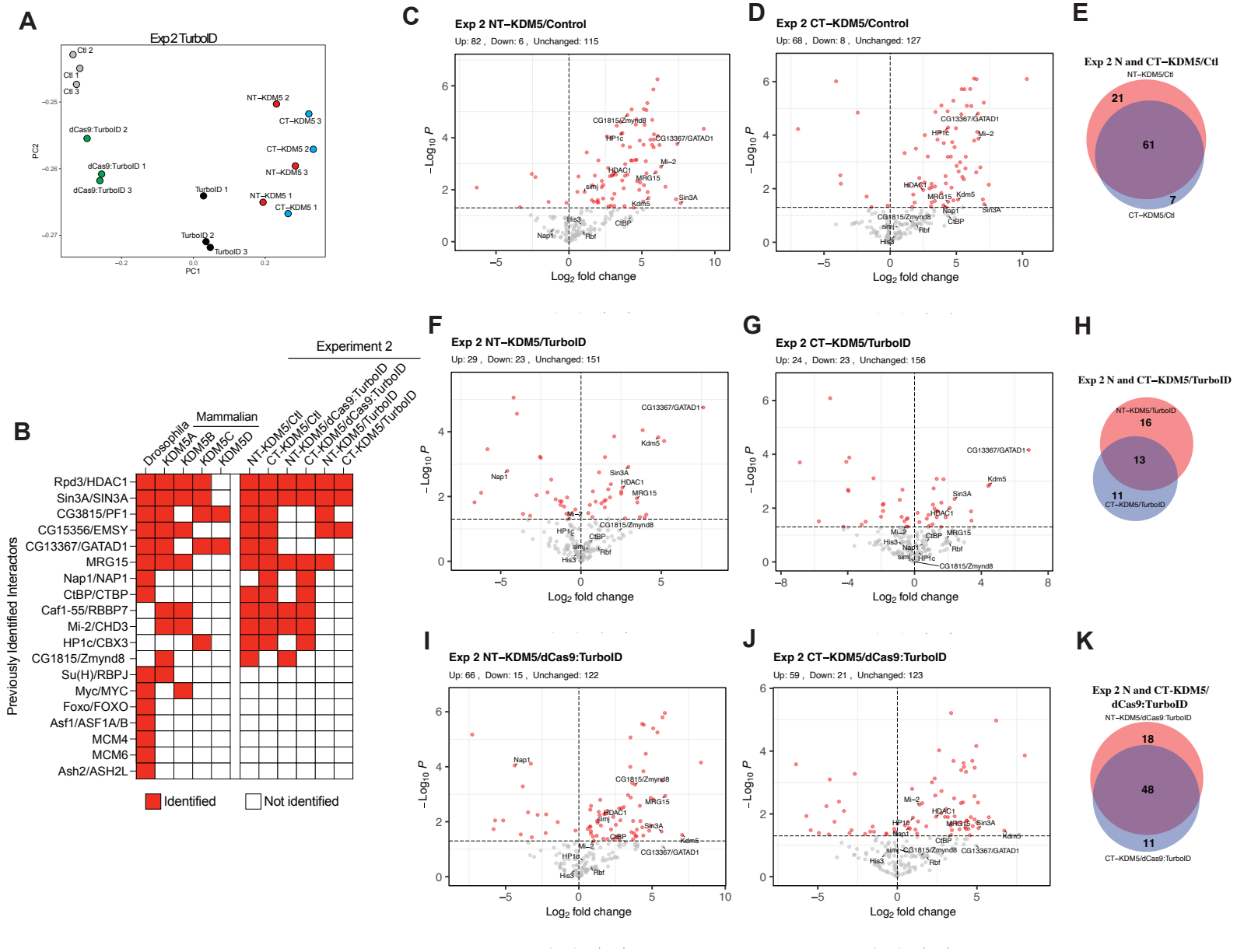

A
Exp 2 TurboID
PC1
PC2
B
Experiment 2
Mammalian
Drosophila
KDM5A
KDM5B
KDM5C
KDM5D
NT-KDM5/Ctl
CT-KDM5/Ctl
NT-KDM5/dCas9:TurboID
CT-KDM5/dCas9:TurboID
NT-KDM5/TurboID
CT-KDM5/TurboID
Previously Identified Interactors
Rpd3/HDAC1
Sin3A/SIN3A
CG3815/PF1
CG15356/EMSY
CG13367/GATAD1
MRG15
Nap1/NAP1
CtBP/CTBP
Caf1-55/RBBP7
Mi-2/CHD3
HP1c/CBX3
CG1815/Zmynd8
Su(H)/RBPJ
Myc/MYC
Foxo/FOXO
Asf1/ASF1A/B
MCM4
MCM6
Ash2/ASH2L
Identified
Not identified
C
Exp 2 NT-KDM5/Control
Up: 82 , Down: 6 , Unchanged: 115
D
Exp 2 CT-KDM5/Control
Up: 68 , Down: 8 , Unchanged: 127
E
Exp 2 N and CT-KDM5/Ctl
NT-KDM5/Ctl
CT-KDM5/Ctl
21
61
7
F
Exp 2 NT-KDM5/TurboID
Up: 29 , Down: 23 , Unchanged: 151
G
Exp 2 CT-KDM5/TurboID
Up: 24 , Down: 23 , Unchanged: 156
H
Exp 2 N and CT-KDM5/TurboID
NT-KDM5/TurboID
CT-KDM5/TurboID
16
13
11
I
Exp 2 NT-KDM5/dCas9:TurboID
Up: 66 , Down: 15 , Unchanged: 122
J
Exp 2 CT-KDM5/dCas9:TurboID
Up: 59 , Down: 21 , Unchanged: 123
K
Exp 2 N and CT-KDM5/dCas9:TurboID
NT-KDM5/dCas9:TurboID
CT-KDM5/dCas9:TurboID
18
48
11
Log2 fold change
−Log10 P

Fig 3

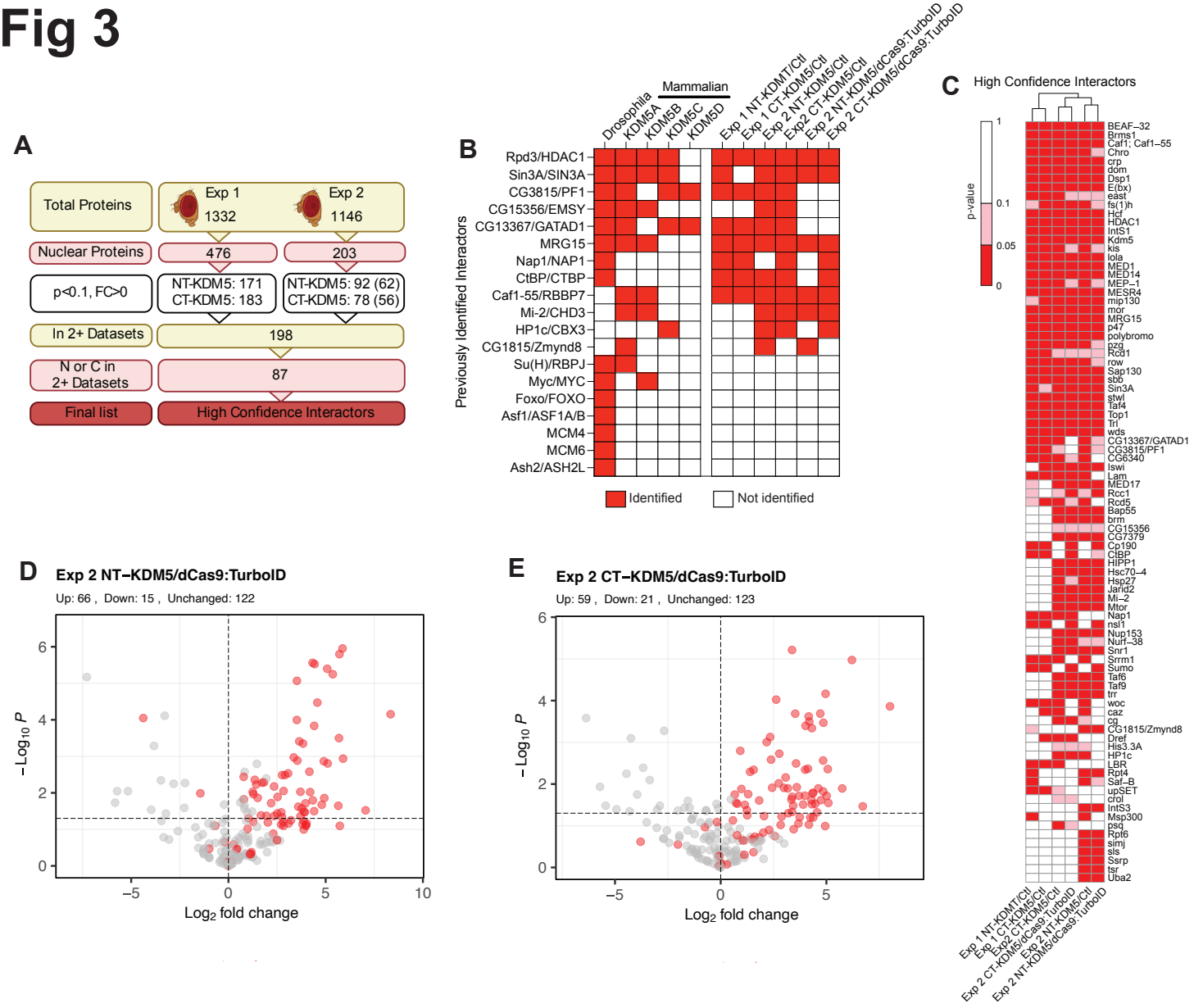
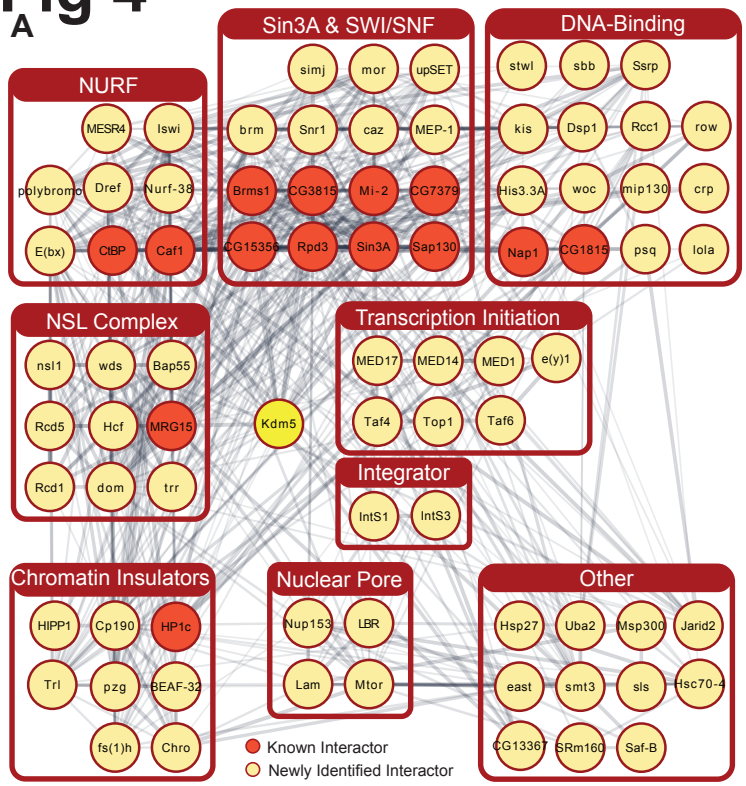
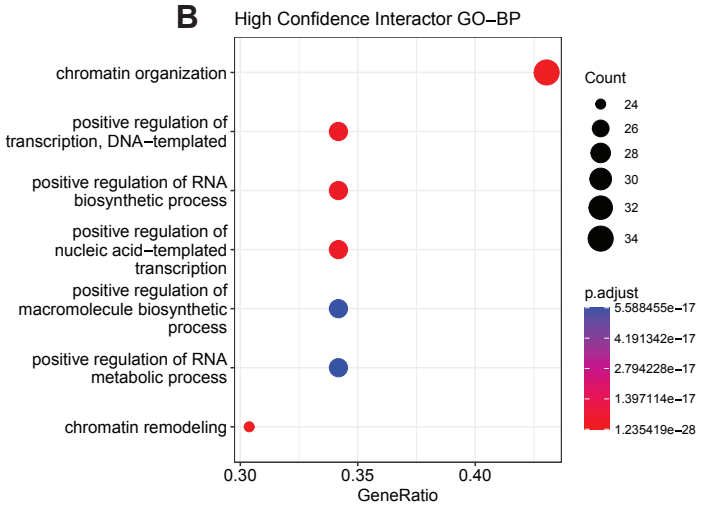
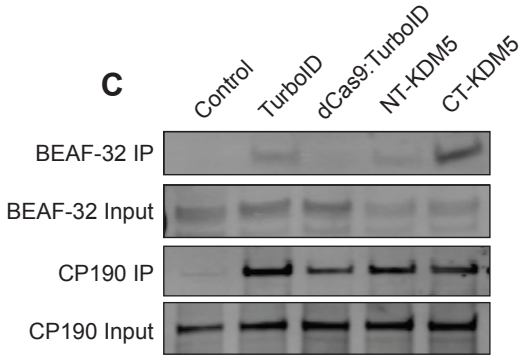

A
Total Proteins
Exp 1
1332
Exp 2
1146
Nuclear Proteins
476
203
p<0.1, FC>0
NT-KDM5: 171
CT-KDM5: 183
NT-KDM5: 92 (62)
CT-KDM5: 78 (56)
In 2+ Datasets
198
N or C in 2+ Datasets
87
Final list
High Confidence Interactors
B
Drosophila
Mammalian
KDM5A
KDM5B
KDM5C
KDM5D
Exp 1 NT-KDM5/Ctl
Exp 1 CT-KDM5/Ctl
Exp 2 NT-KDM5/Ctl
Exp2 CT-KDM5/Ctl
Exp 2 NT-KDM5/dCas9:TurboID
Exp 2 CT-KDM5/dCas9:TurboID
Previously Identified Interactors
Rpd3/HDAC1
Sin3A/SIN3A
CG3815/PF1
CG15356/EMSY
CG13367/GATAD1
MRG15
Nap1/NAP1
CtBP/CTBP
Caf1-55/RBBP7
Mi-2/CHD3
HP1c/CBX3
CG1815/Zmynd8
Su(H)/RBPJ
Myc/MYC
Foxo/FOXO
Asf1/ASF1A/B
MCM4
MCM6
Ash2/ASH2L
Identified
Not identified
C
High Confidence Interactors
p-value
1
0.1
0.05
0
BEAF–32
Brms1
Caf1; Caf1–55
Chro
crp
dom
Dsp1
E(bx)
east
fs(1)h
Hcf
HDAC1
IntS1
Kdm5
kis
lola
MED1
MED14
MEP–1
MESR4
mip130
mor
MRG15
p47
polybromo
pzg
Rcd1
row
Sap130
sbb
Sin3A
stwl
Taf4
Top1
Trl
wds
CG13367/GATAD1
CG3815/PF1
CG6340
Iswi
Lam
MED17
Rcc1
Rcd5
Bap55
brm
CG15356
CG7379
Cp190
CtBP
HIPP1
Hsc70–4
Hsp27
Jarid2
Mi–2
Mtor
Nap1
nsl1
Nup153
Nup1–38
Snr1
Srrm1
Sumo
Taf6
Taf9
trr
woc
caz
cg
CG1815/Zmynd8
Dref
His3.3A
HP1c
LBR
Rpt4
Saf–B
upSET
crol
IntS3
Msp300
psq
Rpt6
simj
sls
Ssrp
tsr
Uba2
Exp 1 NT-KDM5/Ctl
Exp 1 CT-KDM5/Ctl
Exp2 CT-KDM5/Ctl
Exp 2 CT-KDM5/dCas9:TurboID
Exp 2 NT-KDM5/Ctl
Exp 2 NT-KDM5/dCas9:TurboID
D
Exp 2 NT–KDM5/dCas9:TurboID
Up: 66 , Down: 15 , Unchanged: 122
–Log10 P
Log2 fold change
E
Exp 2 CT–KDM5/dCas9:TurboID
Up: 59 , Down: 21 , Unchanged: 123
–Log10 P
Log2 fold change

Fig 4


A
NURF
MESR4
Iswi
polybromo
Dref
Nurf-38
E(bx)
CtBP
Caf1
Sin3A & SWI/SNF
simj
mor
upSET
brm
Snr1
caz
MEP-1
Brms1
CG3815
Mi-2
CG7379
CG15356
Rpd3
Sin3A
Sap130
DNA-Binding
stwl
sbb
Ssrp
kis
Dsp1
Rcc1
row
His3.3A
woc
mip130
crp
Nap1
CG1815
psq
lola
NSL Complex
nsl1
wds
Bap55
Rcd5
Hcf
MRG15
Rcd1
dom
trr
Kdm5
Transcription Initiation
MED17
MED14
MED1
e(y)1
Taf4
Top1
Taf6
Integrator
IntS1
IntS3
Chromatin Insulators
HIPP1
Cp190
HP1c
Trl
pzg
BEAF-32
fs(1)h
Chro
Nuclear Pore
Nup153
LBR
Lam
Mtor
Other
Hsp27
Uba2
Msp300
Jarid2
east
smt3
sls
Hsc70-4
CG13367
SRm160
Saf-B
Known Interactor
Newly Identified Interactor


B
High Confidence Interactor GO–BP
chromatin organization
positive regulation of transcription, DNA–templated
positive regulation of RNA biosynthetic process
positive regulation of nucleic acid–templated transcription
positive regulation of macromolecule biosynthetic process
positive regulation of RNA metabolic process
chromatin remodeling
0.30
0.35
0.40
GeneRatio
Count
24
26
28
30
32
34
p.adjust
5.588455e−17
4.191342e−17
2.794228e−17
1.397114e−17
1.235419e−28


C
Control
TurboID
dCas9:TurboID
NT-KDM5
CT-KDM5
BEAF-32 IP
BEAF-32 Input
CP190 IP
CP190 Input

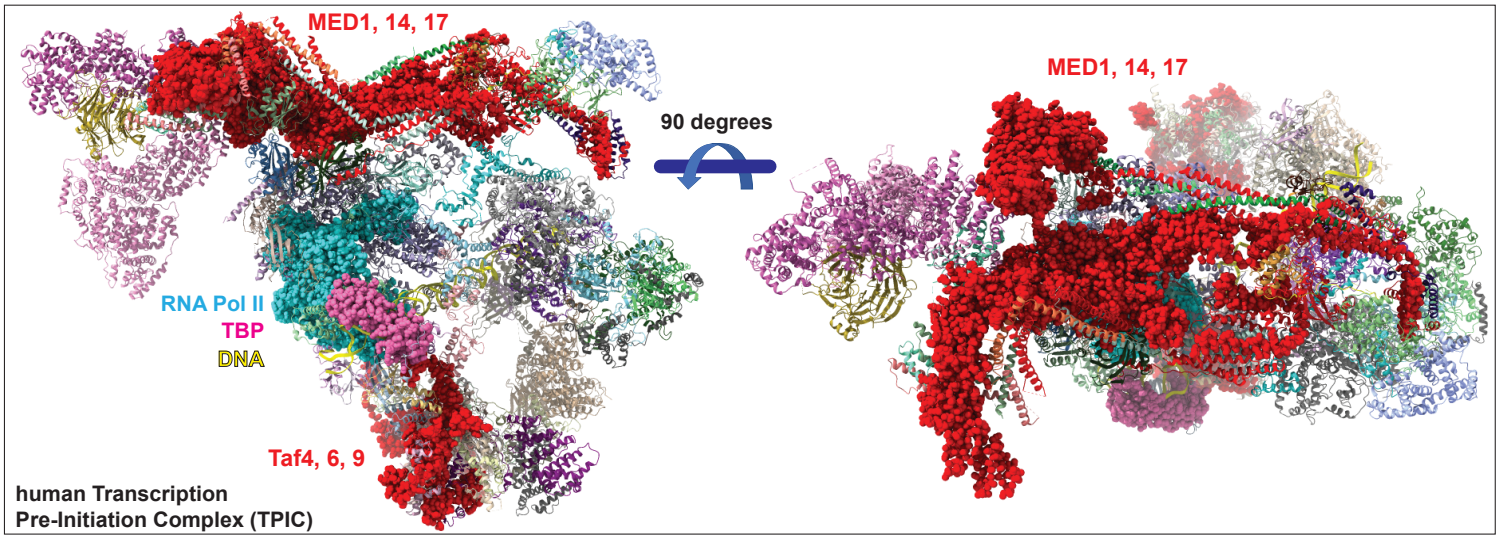

D
MED1, 14, 17
90 degrees
MED1, 14, 17
RNA Pol II
TBP
DNA
Taf4, 6, 9
human Transcription
Pre-Initiation Complex (TPIC)

Fig5

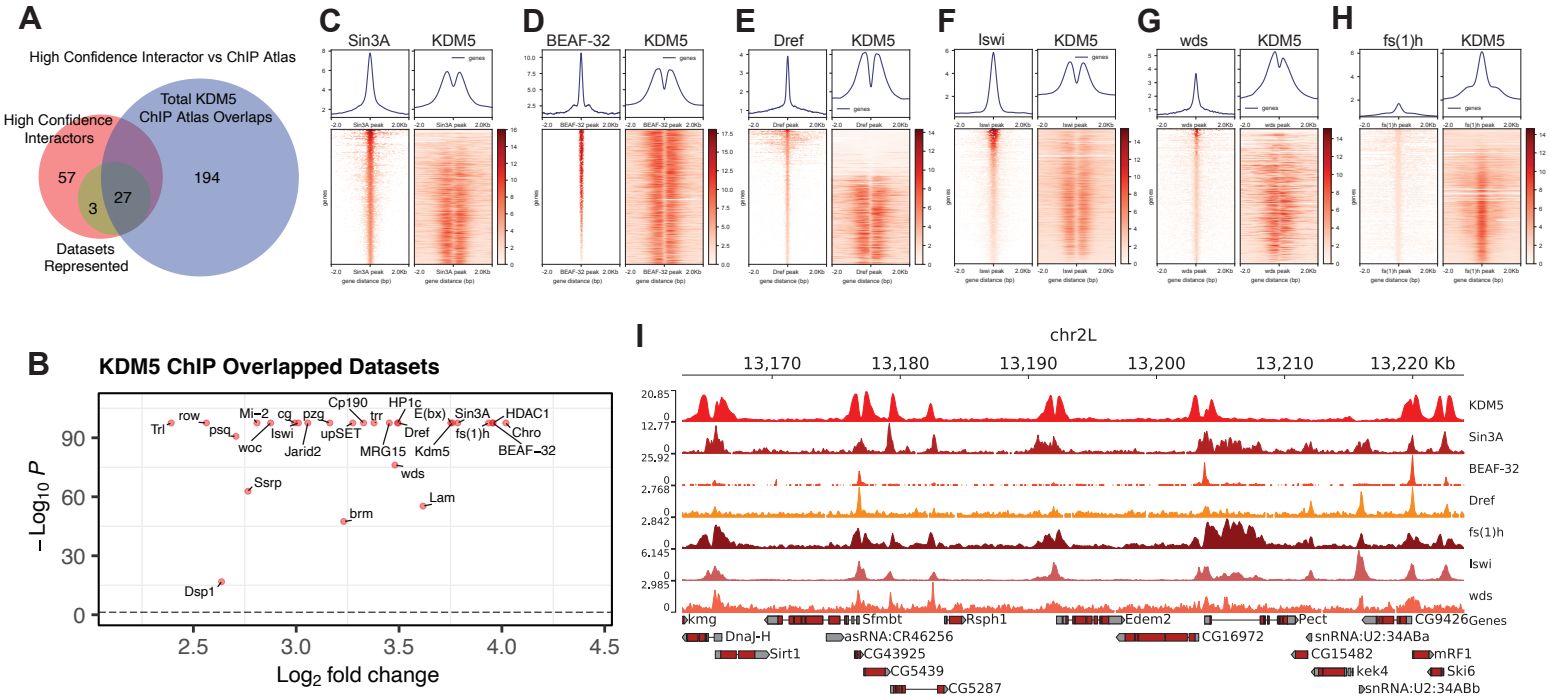

A
High Confidence Interactor vs ChIP Atlas
High Confidence Interactors
Total KDM5 ChIP Atlas Overlaps
57
3
27
194
Datasets Represented
C
Sin3A
KDM5
D
BEAF-32
KDM5
E
Dref
KDM5
F
lswi
KDM5
G
wds
KDM5
H
fs(1)h
KDM5
B
KDM5 ChIP Overlapped Datasets
Trl
row
psq
Mi-2
cg
pzg
Cp190
trr
HP1c
E(bx)
Sin3A
HDAC1
lswi
upSET
Dref
fs(1)h
Chro
woc
Jarid2
MRG15
Kdm5
BEAF-32
wds
Ssrp
Lam
brm
Dsp1
$-\mathrm{Log}_{10}\,P$
Log_2 fold change
I
chr2L
13,170
13,180
13,190
13,200
13,210
13,220 Kb
KDM5
Sin3A
BEAF-32
Dref
fs(1)h
lswi
wds
Genes
kmg
Sfmbt
Rsph1
Edem2
Pect
CG9426
DnaJ-H
asRNA:CR46256
CG16972
snRNA:U2:34ABa
Sirt1
CG43925
CG15482
mRF1
CG5439
kek4
Ski6
CG5287
snRNA:U2:34ABb

Fig6

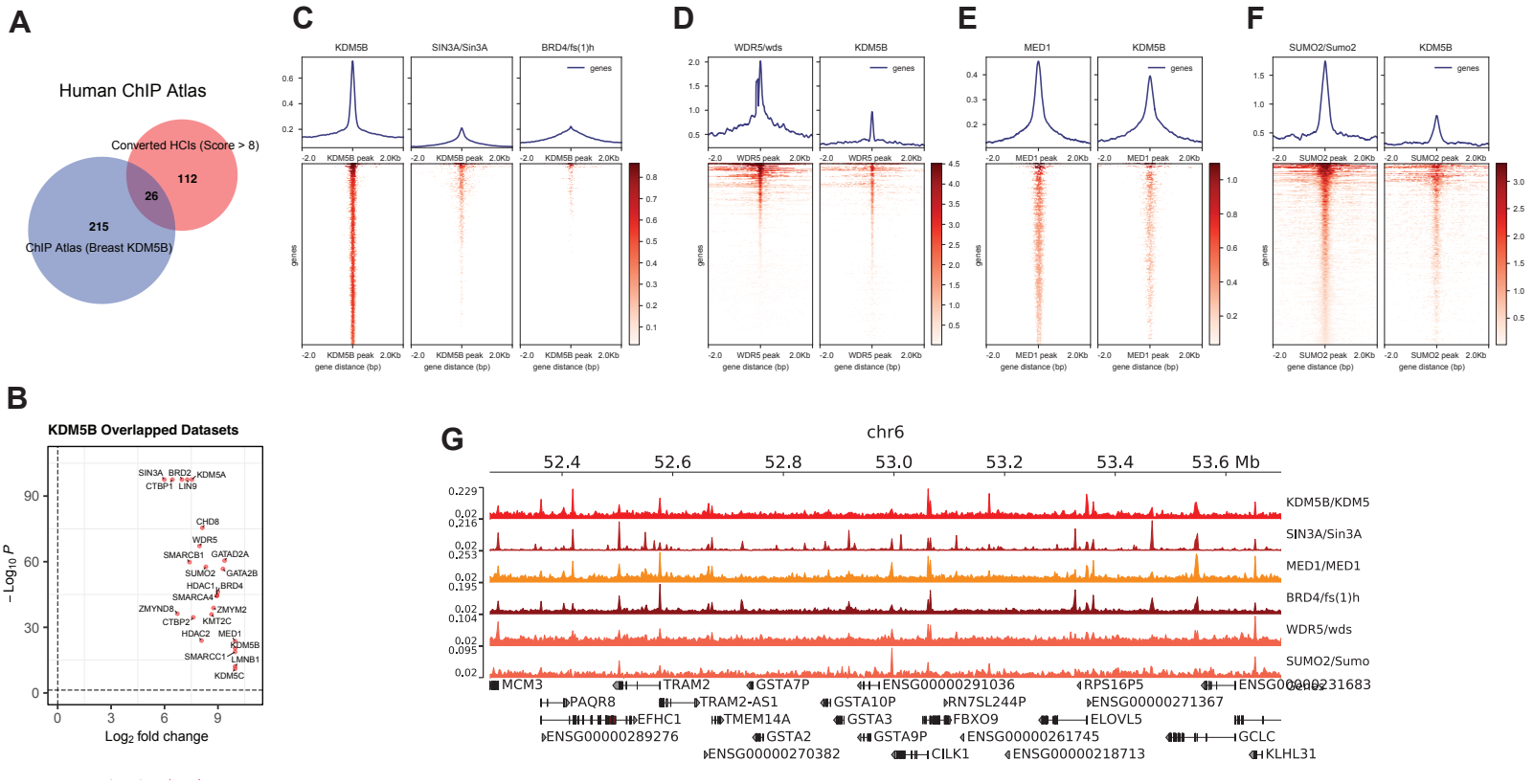
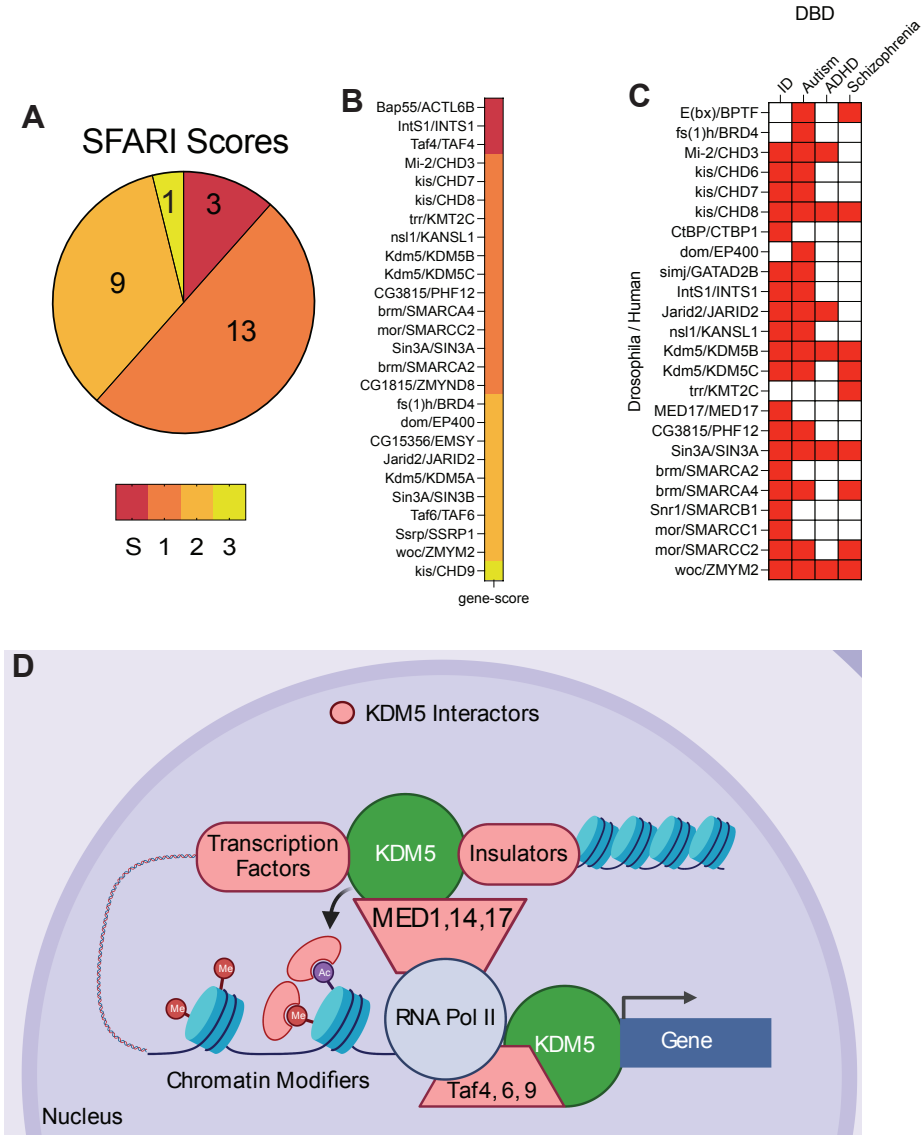
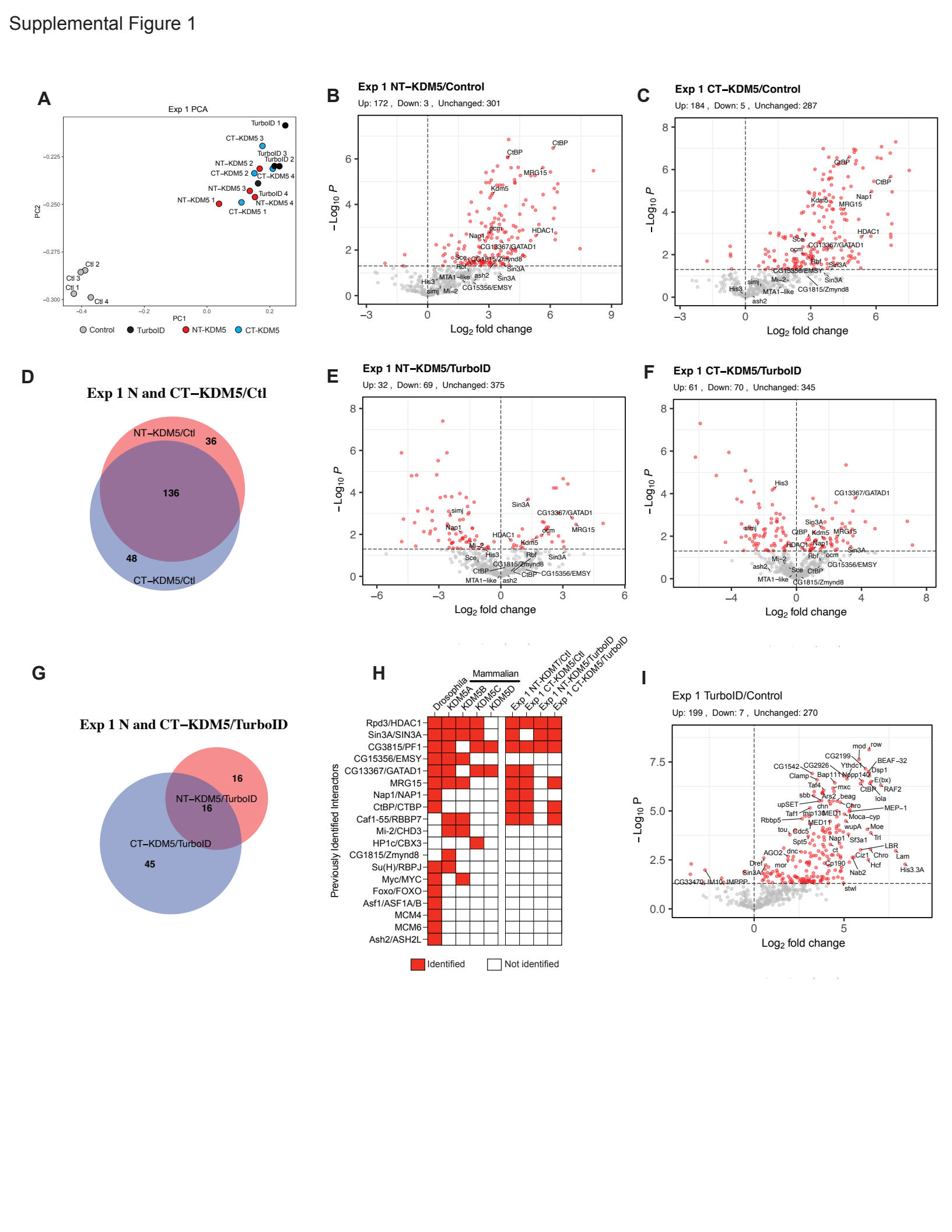

A
Human ChIP Atlas
Converted HCIs (Score > 8)
112
26
215
ChIP Atlas (Breast KDM5B)
C
KDM5B
SIN3A/Sin3A
BRD4/fs(1)h
genes
KDM5B peak
gene distance (bp)
D
WDR5/wds
KDM5B
WDR5 peak
E
MED1
MED1 peak
F
SUMO2/Sumo2
SUMO2 peak
B
KDM5B Overlapped Datasets
SIN3A
BRD2
KDM5A
CTBP1
LIN9
CHD8
WDR5
SMARCB1
GATAD2A
SUMO2
GATA2B
HDAC1
BRD4
SMARCA4
ZMYND8
ZMYM2
CTBP2
KMT2C
HDAC2
MED1
KDM5B
SMARCC1
LMNB1
KDM5C
$-Log_{10}$ P
Log_2 fold change
G
chr6
52.4
52.6
52.8
53.0
53.2
53.4
53.6 Mb
KDM5B/KDM5
SIN3A/Sin3A
MED1/MED1
BRD4/fs(1)h
WDR5/wds
SUMO2/Sumo
Genes
MCM3
TRAM2
GSTA7P
ENSG00000291036
RPS16P5
PAQR8
TRAM2-AS1
GSTA10P
RN7SL244P
ENSG00000271367
EFHC1
TMEM14A
GSTA3
FBXO9
ELOVL5
ENSG00000289276
GSTA2
GSTA9P
ENSG00000261745
GCLC
ENSG00000270382
CILK1
ENSG00000218713
KLHL31

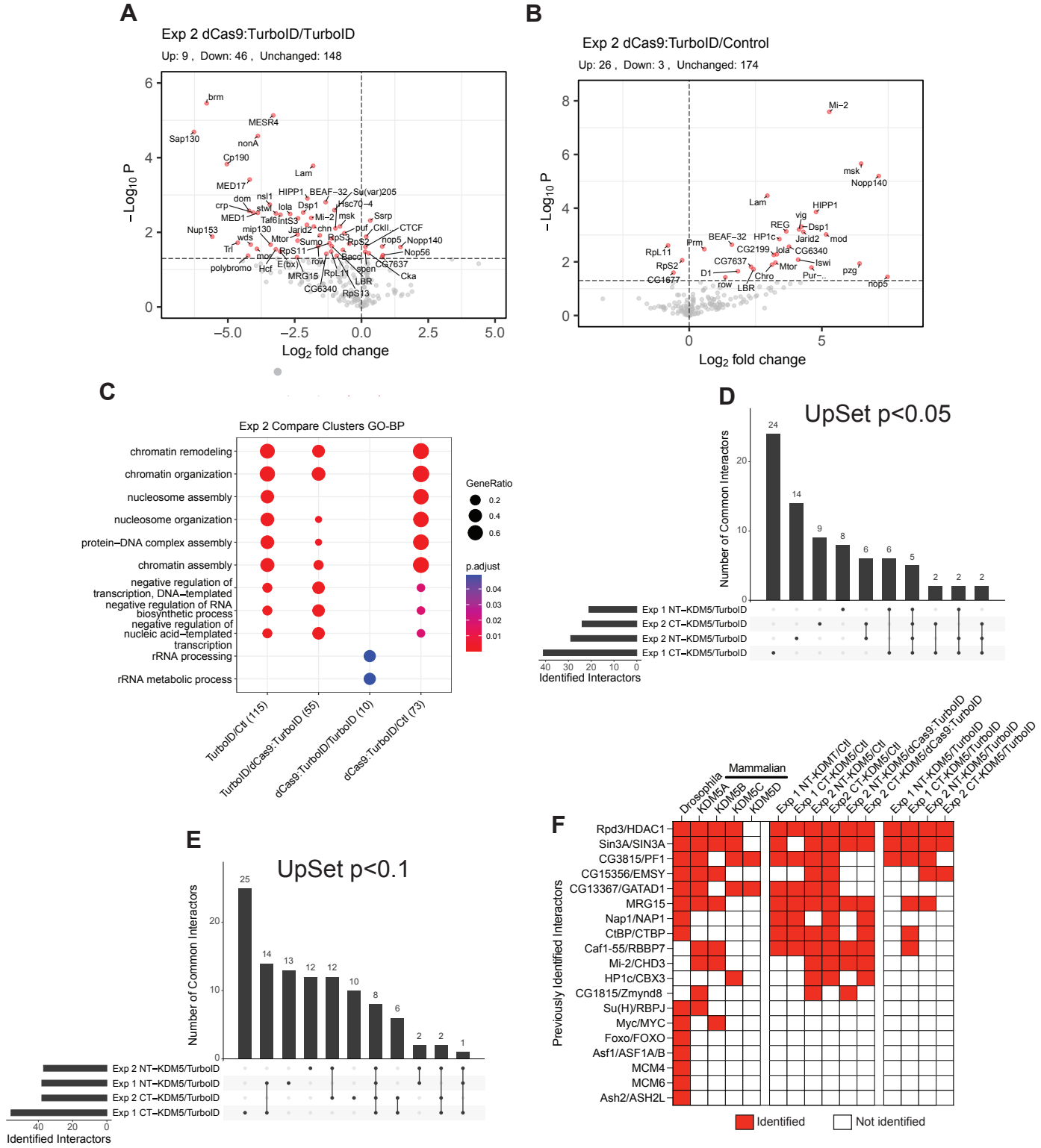
Fig7


A
SFARI Scores
1
3
9
13
S 1 2 3
B
Bap55/ACTL6B
IntS1/INTS1
Taf4/TAF4
Mi-2/CHD3
kis/CHD7
kis/CHD8
trr/KMT2C
nsl1/KANSL1
Kdm5/KDM5B
Kdm5/KDM5C
CG3815/PHF12
brm/SMARCA4
mor/SMARCC2
Sin3A/SIN3A
brm/SMARCA2
CG1815/ZMYND8
fs(1)h/BRD4
dom/EP400
CG15356/EMSY
Jarid2/JARID2
Kdm5/KDM5A
Sin3A/SIN3B
Taf6/TAF6
Ssrp/SSRP1
woc/ZMYM2
kis/CHD9
gene-score
C
DBD
ID
Autism
ADHD
Schizophrenia
E(bx)/BPTF
fs(1)h/BRD4
Mi-2/CHD3
kis/CHD6
kis/CHD7
kis/CHD8
CtBP/CTBP1
dom/EP400
simj/GATAD2B
IntS1/INTS1
Jarid2/JARID2
nsl1/KANSL1
Kdm5/KDM5B
Kdm5/KDM5C
trr/KMT2C
MED17/MED17
CG3815/PHF12
Sin3A/SIN3A
brm/SMARCA2
brm/SMARCA4
Snr1/SMARCB1
mor/SMARCC1
mor/SMARCC2
woc/ZMYM2
Drosophila / Human
D
KDM5 Interactors
Transcription Factors
KDM5
Insulators
MED1,14,17
Me
Ac
RNA Pol II
KDM5
Gene
Chromatin Modifiers
Taf4, 6, 9
Nucleus

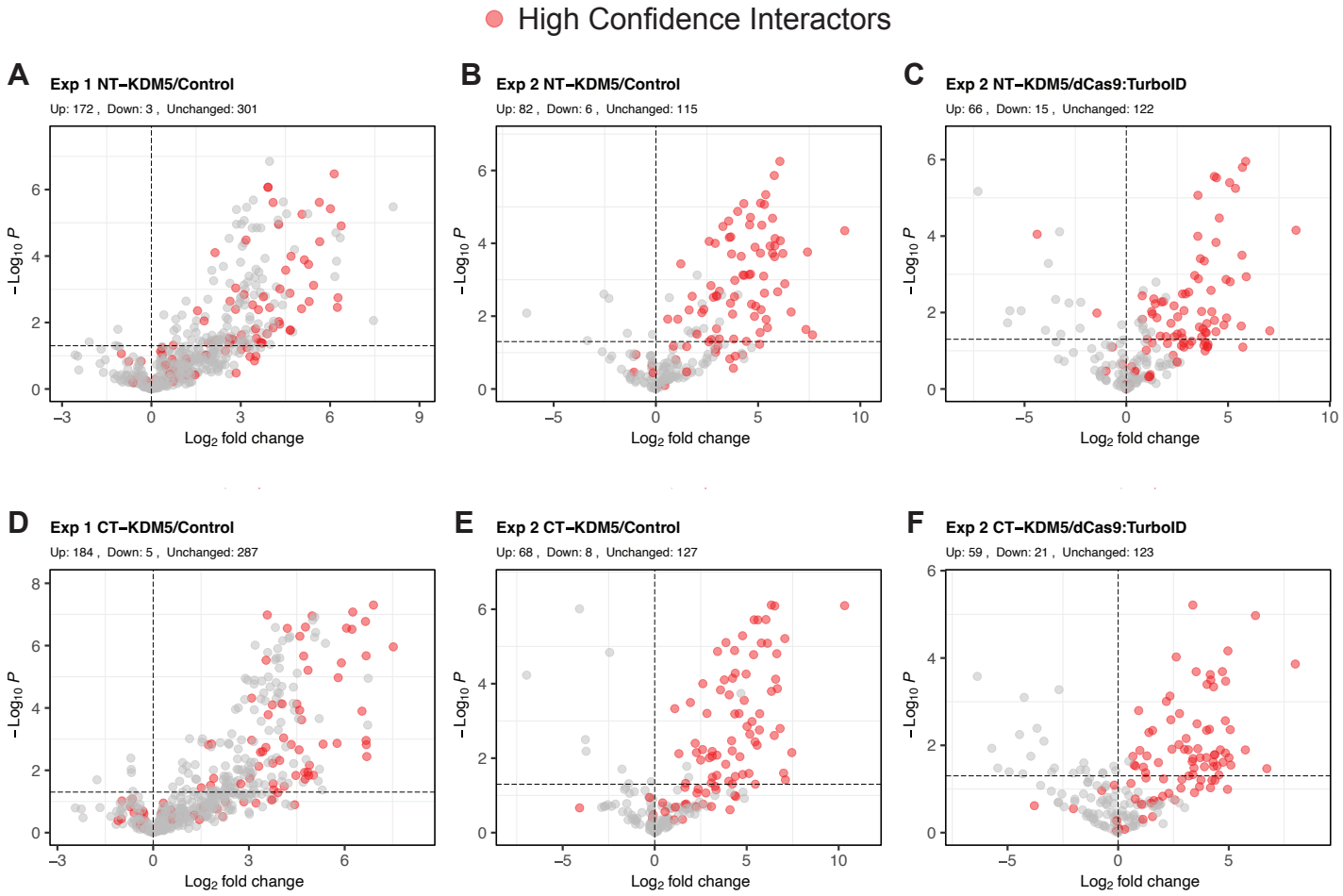
Supplemental Figure 1



Supplemental Figure 2



Supplemental Figure 3



Proximity labeling reveals a new *in vivo* network of interactors for the histone demethylase KDM5

Matanel Yheskel, Simone Sidoli, and Julie Secombe

Supplemental Figure Legenes

Figure S1: KDM5-TurboID studies using control and TurboID alone

(A) Principal components analysis of normalized nuclear protein abundances from control (Ctl), TurboID, NT-KDM5 and CT-KDM5 (experiment 1 data).

(B) Volcano plot showing data comparing NT-KDM5 to control.

(C) Volcano plot showing data comparing CT-KDM5 to control.

(D) Venn diagram showing overlap of the proteins identified using NT-KDM5 and CT-KDM5 compared to control (Ctl).

(E) Volcano plot showing data comparing NT-KDM5 to TurboID.

(F) Volcano plot showing data comparing CT-KDM5 to TurboID.

(G) Venn diagram showing overlap of the proteins identified using NT-KDM5 and CT-KDM5 compared to TurboID.

(H) Summary of known *Drosophila* KDM5 interactors, whether the interaction is conserved in mammals (mouse and/or human) and the identification of these proteins in experiment 1 with all comparisons. Three interactors identified in mammalian cells but not previously in *Drosophila* are also included (Caf1-55, Mi-2 and Zmynd8).

(I) Volcano plot showing data comparing TurboID to control.

For all volcano plots shown, known interactors are indicated with text. Red dots indicate significantly enriched proteins ($p<0.05$) and the dotted line on Y-axis indicates $p=0.05$.

Figure S2: Comparing TurboID alone to dCas9:TurboID

(A) Volcano plot showing data comparing dCas9:TurboID to TurboID alone.

(B) Volcano plot showing data comparing dCas9:TurboID to control.

(C) GO analyses of proteins enriched comparing TurboID to control (Ctl), TurboID to dCas9:TurboID, dCas9:TurboID to TurboID and dCas9:TurboID to control.

(D) UpSet plot showing the number of common interactors identified in experiments in which TurboID was used to compare datasets using $p<0.05$.

(E) UpSet plot showing the number of common interactors identified in experiments in which TurboID was used to compare datasets using $p<0.1$.

(F) Summary of all known interactors identified in all experimental comparisons from experiment 1 and experiment 2.

For all volcano plots shown, red dots indicate significantly enriched proteins ($p<0.05$) and the dotted line on Y-axis indicates $p=0.05$.

Figure S3: Summary of identification of high confidence interactors in each dataset.

(A) Volcano plot showing data comparing NT-KDM5 to control from experiment 1 with high confidence interactors indicated by red dots.

(B) Volcano plot showing data comparing NT-KDM5 to control from experiment 2 with high confidence interactors indicated by red dots.

(C) Volcano plot showing data comparing NT-KDM5 to dCas9:TurboID (experiment 2) with high confidence interactors indicated by red dots.

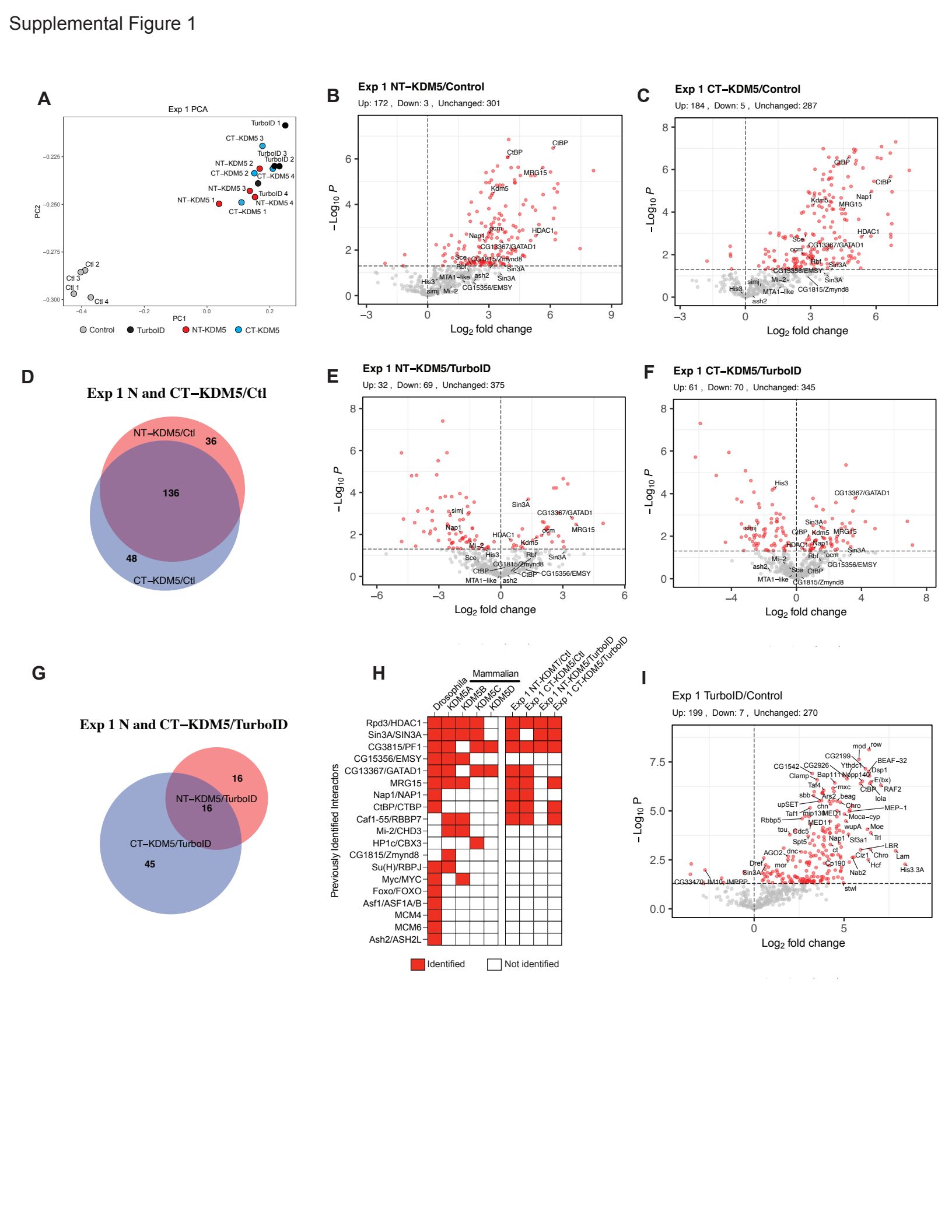
(D) Volcano plot showing data comparing CT-KDM5 to control from experiment 1 with high confidence interactors indicated by red dots.

(E) Volcano plot showing data comparing CT-KDM5 to control from experiment 2 with high confidence interactors indicated by red dots.

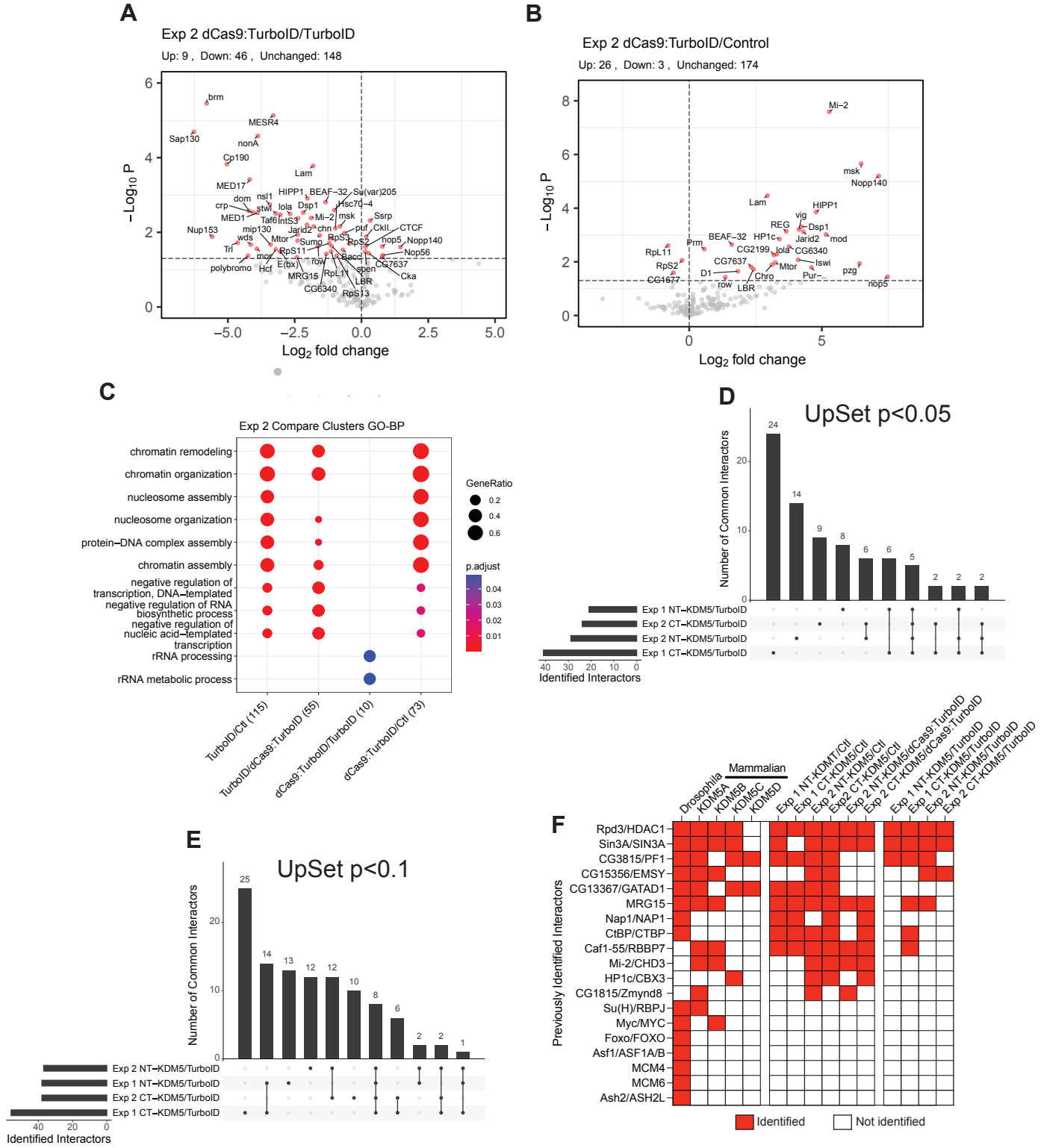
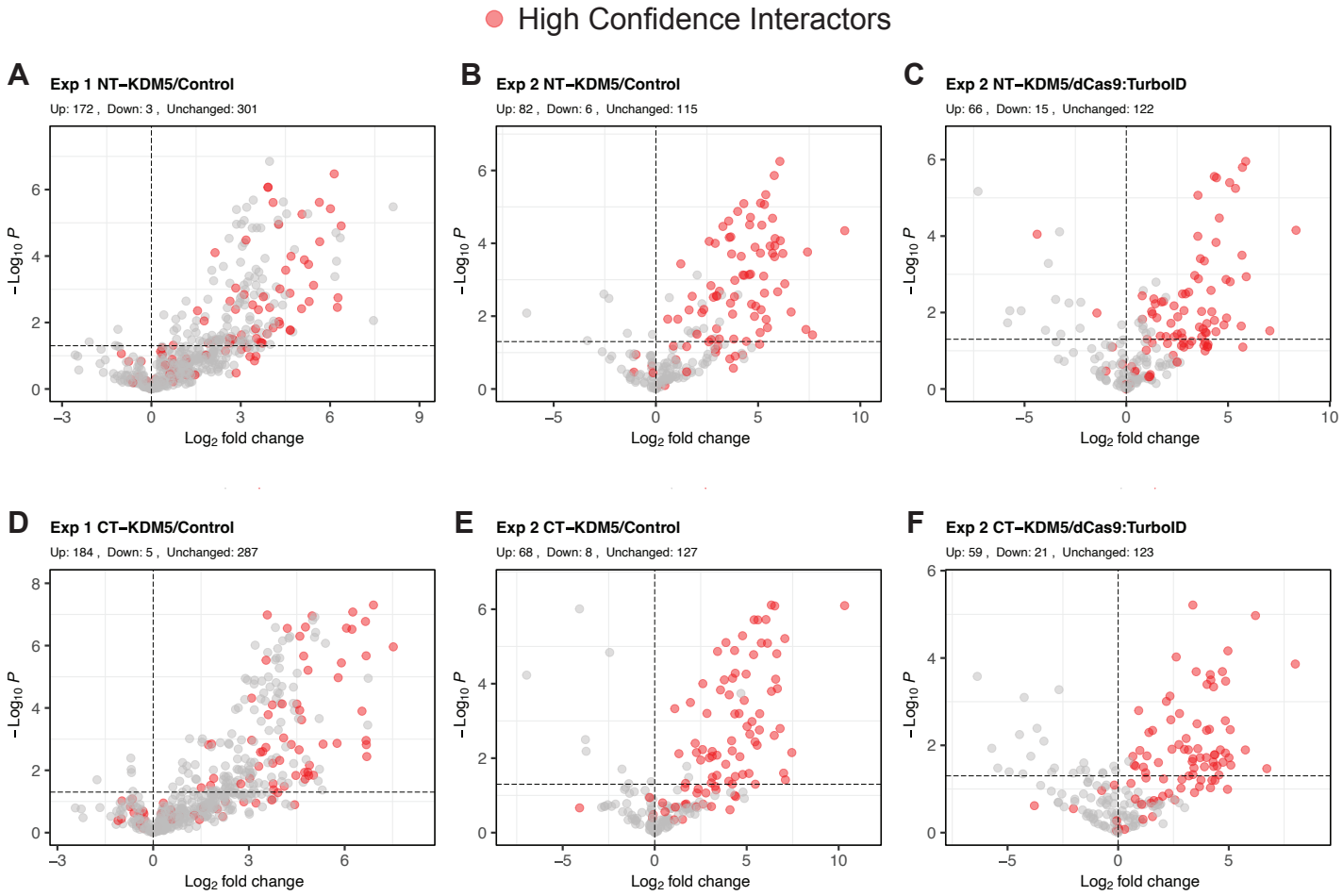
(F) Volcano plot showing data comparing CT-KDM5 to dCas9:TurboID (experiment 2) with high confidence interactors indicated by red dots.

For all volcano plots shown, known interactors are indicated with text. Red dots indicate significantly enriched proteins (p-value<0.05) and the dotted line on Y-axis indicates p=0.05.

Supplemental Figure 1



Supplemental Figure 2



Supplemental Figure 3