



HAL
open science

Transient loss of Polycomb components induces an epigenetic cancer fate

Victoria Parreno, Vincent Loubiere, Bernd Schuettengruber, Maksim Erokhin, Balázs Győrffy, M. Di Stefano, L. Fritsch, Jérôme Moreaux, Darya Chetverina, A-M. Martinez, et al.

► **To cite this version:**

Victoria Parreno, Vincent Loubiere, Bernd Schuettengruber, Maksim Erokhin, Balázs Győrffy, et al..
Transient loss of Polycomb components induces an epigenetic cancer fate. 2023. hal-04252107

HAL Id: hal-04252107

<https://hal.umontpellier.fr/hal-04252107>

Preprint submitted on 20 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/366885931>

Transient loss of Polycomb components induces an epigenetic cancer fate

Preprint · January 2023

DOI: 10.1101/2023.01.04.522799

CITATIONS

0

READS

120

11 authors, including:



Victoria Parreno

The Institute of Human Genetics

2 PUBLICATIONS 34 CITATIONS

SEE PROFILE



Vincent Loubiere

Research Institute of Molecular Pathology

13 PUBLICATIONS 336 CITATIONS

SEE PROFILE



Bernd Schuettengruber

Université de Montpellier

90 PUBLICATIONS 5,515 CITATIONS

SEE PROFILE



Maksim Erokhin

Russian Academy of Sciences

65 PUBLICATIONS 553 CITATIONS

SEE PROFILE

1 Transient loss of Polycomb components induces an epigenetic cancer fate

3 Authors

4 Parreno, V.^{1*}, Loubière, V.^{1,2*}, Schuettengruber, B.¹, Erokhin, M.³, Györfy, B.^{4,5}, Di Stefano, M.¹, Fritsch, L.¹,
5 Moreaux, J.^{6,7}, Chetverina, D.³, Martinez, A-M.^{1#} & Cavalli, G.^{1#}

7 Affiliations

8
9 ¹ Institute of Human Genetics, CNRS, University of Montpellier, Montpellier, France

10 ² Research Institute of Molecular Pathology (IMP), Vienna BioCenter (VBC), Vienna, Austria.

11 ³ Institute of Gene Biology, Russian Academy of Sciences, 34/5 Vavilov Street, Moscow, 119334 Russia

12 ⁴ TTK Cancer Biomarker Research Group, Institute of Enzymology, Magyar tudosok korutja 2, 1117, Budapest,
13 Hungary

14 ⁵ Semmelweis University Department of Bioinformatics and 2nd Department of Pediatrics, Tuzolto u. 7-9, 1094,
15 Budapest, Hungary

16 ⁶ Department of Biological Hematology, CHU Montpellier, 34295 Montpellier, France

17 ⁷ UFR Medicine, University of Montpellier, 34003 Montpellier, France

18
19 * These authors contributed equally to this work

20
21 # Corresponding authors: anne-marie.martinez@igh.cnrs.fr (A.M.M.) and giacomo.cavalli@igh.cnrs.fr (G.C.)

23 Summary

24 Cell fate depends on genetic, epigenetic and environmental inputs that are interconnected, making it difficult to
25 disentangle their respective contributions to cell fate decisions¹⁻³, and epigenetic reprogramming is a major
26 contributor to tumor plasticity and adaptation⁴⁻⁶. Although cancer initiation and progression are generally
27 associated with the accumulation of somatic mutations^{7,8}, substantial epigenomic alterations underlie many
28 aspects of tumorigenesis and cancer susceptibility⁹⁻¹⁸, suggesting that genetic mechanisms alone may not be
29 sufficient to drive malignant transformations¹⁹⁻²³. However, whether purely non-genetic reprogramming
30 mechanisms are sufficient to initiate tumorigenesis irrespective of mutations is unknown. Here, we show that a
31 transient perturbation of transcriptional silencing mediated by Polycomb-Group proteins is sufficient to induce an
32 irreversible switch to a cancer cell fate in *Drosophila*. This is linked to the irreversible derepression of genes that
33 can drive tumorigenesis, including JNK and JAK-STAT signalling pathways and *zfh1*, the fly homolog of the ZEB1
34 oncogene, which we show to be a necessary driver of the cancer fate. These data show that a reversible
35 perturbation of Polycomb-Group protein levels can induce cancer in the absence of driver mutations and suggest
36 that this is achieved through epigenetic inheritance of altered cell fates.

38 Main Text

39 Over the last decades, large-scale projects greatly expanded the known repertoire of cancer-associated genetic
40 mutations affecting epigenetic factors²⁴⁻²⁸, including chromatin remodelers and modifiers controlling histone
41 marks²⁹⁻³¹, DNA methylation³², micro-RNA³³ and 3D-genome folding³⁴ that might support cancer progression³⁵,
42 corroborating the role of epigenetic aberrations in the etiology of hematological and solid malignancies³⁶⁻³⁸.
43 Indeed, epigenetic modifications are used as biomarkers and are targeted by epi-drugs in cancer therapy³⁹.
44 Tumorigenesis is therefore associated with genetic as well as epigenetic determinants⁴⁰⁻⁴²⁻⁴³.

45 The fact that several hallmarks of human cancer⁴³⁻⁴⁸ may be acquired through epigenome dysregulation suggests
46 that epigenetic alterations play causal roles in cancer⁹. For example, pancreatic cancer metastases can arise
47 without the emergence of driver DNA mutations, but coincide with major epigenetic changes^{49,50}. Recently,
48 epigenetic contributions were shown to play key roles in several cases of metastatic progression^{51,52}. In some

49 pediatric cancers, such as ependymomas, low numbers of mutations have been detected, suggesting that
50 epigenetic changes may drive tumorigenesis⁵³. Combined, these data led to the proposal that cancer is not a
51 consequence of DNA mutations only^{14,54}.

52 However, whether purely non-genetic reprogramming mechanisms are sufficient to initiate tumorigenesis
53 irrespective of mutations remains an open question. Polycomb Group (PcG) proteins are epigenetic factors
54 forming two main classes of complexes called Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2,
55 respectively), which are highly conserved from fly to human and play a critical role in cellular memory, by ensuring
56 the stable repression of major developmental genes throughout development. PcG dysregulation leads to cell fate
57 changes^{55,56}, developmental transformations and is associated with multiple types of cancers⁵⁷⁻⁵⁹, including breast
58 and prostate cancer⁶⁰. PRC2 deposits the H3K27me3 repressive mark, whereas PRC1, which contains the PH,
59 PC, PSC and the SCE subunits in flies, is responsible for H2AK118Ub deposition⁶¹⁻⁶³. Contrasting with the
60 extreme redundancy found in mammals⁵⁸, most PcG components are coded by a single gene in *Drosophila*,
61 making this system more tractable for functional studies.

62

63 **A transient epigenetic perturbation is sufficient to initiate tumors**

64 Earlier work has shown that both null mutations or constant RNAi knock-down (KD) targeting both *ph* homologs
65 (*ph-p* and *ph-d*) drive similar transcriptional defects and are sufficient to induce loss of differentiation and cell
66 overproliferation⁶⁴⁻⁶⁸. To test whether a transient epigenetic perturbation might be sufficient to trigger an
67 irreversible change in cell fate, we set up a thermosensitive *ph*-RNAi system allowing for the reversible KD of *ph*
68 in the developing larval eye imaginal disc (ED, see Figure 1a, 1b; Extended Data Fig. 1). As expected, constant
69 PH depletion induces tumors (Figure 1c,d; Extended Data Fig. 1a-d) as well as reduced adult survival (Extended
70 Data Fig. 1e). Strikingly, a transient 24h depletion of PH at the L1 stage, where the ED starts developing, is also
71 sufficient to trigger the formation of tumors, characterized by overgrowth, the loss of apico-basal polarity and
72 differentiation markers (Figure 1c-e; Extended Data Fig. 1b-d). Importantly, these tumors show normal levels of
73 the PH protein, both at day 9 (transient *ph*-KD d9) and day 11 (transient *ph*-KD d11) after egg laying (AEL) (Figure
74 1b, Extended Data Fig. 1a). EDs continue to grow significantly even after PH recovery (Figure 1e) and cannot
75 differentiate (Extended Data Fig. 1d), suggesting that the tumor state is stable and maintained independently of
76 its epigenetic trigger. Likewise, PH depletion at the L2 or early L3 stage induced tumors (Extended Data Fig. 1b-
77 d). Of note, a 24 h exposure at 29°C at the L3 stage is sufficient to deplete PH in our system (Extended Data Fig.
78 1f). Consistent with the constant need of PRC1 function to prevent tumorigenesis, a transient depletion of PSC-
79 SU(Z)2, another core subunit of PRC1 for which null mutations drive neoplastic transformation^{69,70}, was also
80 sufficient to induce tumorigenesis (Extended Data Fig. 2).

81 To identify potential driver mutations that might arise as a consequence of PH depletion, we collected eggs from
82 two independent batches of mated females and subjected them to a transient KD, a constant KD or no *ph*-KD
83 (control condition), before sequencing their genomic DNA. Given the 100% penetrance of the tumor phenotype
84 and the relatively low number of genes that can act as cancer drivers in *Drosophila*⁷¹, we reasoned that such
85 driver mutations should be repeatedly recovered in independent tumors. However, when using the first of the two
86 control batches as a reference in the search for somatic single nucleotide variants (SNV) or insertions and
87 deletions (InDel)⁷², we found that the six tumor samples showed comparable numbers of variants as well as
88 similar features distribution compared to the second control sample (Figure 1f,g). Importantly, the large majority
89 of variants were detected in a single sample, and few mutated genes were shared between samples (Figure 1h,
90 see material and methods). Only three genes were found mutated in all six tumor samples and not in control,
91 namely *spz3*, *CG34398* and *CG9525*, none of them having known functions in tumorigenesis. Finally, no
92 enrichment for any specific Gene ontology (GO) was found for genes mutated in at least two *ph*-KD samples. This
93 evidence strongly argues against the presence of recurrent driver mutations in these tumors, reminiscent of
94 previously reports suggesting that genome instability is not a prerequisite for neoplastic epithelial growth caused
95 by *ph*-KD⁶⁵.

96 In sum, these data show that the transient depletion of tumor suppressor PRC1 components is sufficient to switch
97 cells into a neoplastic state that is maintained even after normal levels are re-established. Since the same
98 genotype can generate both a normal phenotype or a tumor depending on a transient gene regulatory
99 modification, we defined these tumors as epigenetically-initiated cancers (EICs).

100

101 **EICs are characterized by activated JAK-STAT and JNK signaling**

102 To tackle the molecular mechanisms leading to the formation of EICs, we compared the transcriptomes of control
103 (no *ph*-KD), transient and constant *ph*-KD to temperature-matched controls, generated with a similar RNAi system
104 targeting the *white* gene, which is dispensable for normal eye development (see material and methods, Figure 2,
105 Extended Data Table 1). As expected, both systems are hardly distinguishable at 18°C (no RNAi), as well as upon
106 transient *w*-KD (Figure 2a; Extended Data Fig. 3a). Consistent with our previous work^{64,68}, constant *ph*-KD is
107 associated with the up-regulation of 340 genes – including canonical PcG targets such as Hox and developmental
108 Transcription Factor (TF) genes – and the down-regulation of 2,110 genes, including most key regulators of ED
109 development (Figure 2a; Extended Data Table 1; Extended Data Fig. 3b). Importantly, only a subset of these
110 genes were also differentially expressed in transient *ph*-KD at d9 AEL (256 and 812, respectively), and even less
111 at later d11 AEL (154 and 446, respectively), suggesting a progressive yet incomplete rescue of the transcriptome
112 (Figure 2a; Extended Data Fig. 3, 4a; Extended Data Table 1). Therefore, a substantial portion of the
113 transcriptome can be restored upon reinstating normal levels of PH.

114 Hierarchical clustering of differentially expressed genes identified six clusters with specific dynamics (Figure 2b;
115 Extended Data Table 2). Genes of Cluster 3, 4 and 6 are downregulated and gradually or fully return to normal
116 levels of expression (Figure 2b; Extended Data Table 2). In contrast, upregulated genes (Cluster 1, 2 and 5) show
117 3 different patterns (Figure 2b; Extended Data Table 2). To identify direct PcG target genes correlated to these
118 patterns, we performed PH ChIP-Seq, and H3K27me3 and H2AK118Ub CUT&RUNs – the canonical marks of
119 PcG-mediated transcriptional repression – in control EDs (see material and methods, Extended data 5a and
120 Extended Data Table 2). Cluster 1 genes are specifically upregulated in EICs, are not enriched for any specific
121 gene ontology (GO) and are generally not targeted by PcG (Figure 2b,c and Extended Data Table 2), suggesting
122 that their up-regulation is mostly an indirect consequence of PH depletion. In contrast, direct PcG target genes
123 are over-represented in clusters 2 and 5 (Extended Data Fig. 5b), which are enriched for GOs related to
124 developmental TFs (Figure 2b, c).

125 Cluster 5 includes many TFs that are bound by both PRC1 and PRC2 and correspond to canonical PcG targets
126 such as *en*, *eve*, *wg*, *Scr*, *Su(z)12*, *Ubx* and *wg*. Their progressive recovery of control levels of expression at d9
127 or d11 AEL precludes these genes from being required for the maintenance of EICs. On the other hand, cluster
128 2 contains genes which share the interesting feature of being irreversibly upregulated, irrespective of PH recovery.
129 Among them, we noted an over-representation of direct PcG target genes involved in JAK/STAT (*upd1*, *upd2*,
130 *upd3*, *socs36E*, *zfh1*, *chinmo*) and JNK (*Ets21C*, *puc*) signaling pathways (Figure 2d, Extended Data Table 2),
131 which were shown to play an active role in the development of tumors in various *Drosophila* cancer models^{66,69,73}.
132 In parallel, genes from cluster 2 which are not bound by PRC1 are enriched for DNA-replication related GOs
133 (Figure 2c), and likely are induced by indirect mechanisms. Cluster 2 also contains the Insulin-like peptide 8 (*Ilp8*)
134 gene responsible for developmental delay of tumor-bearing larvae⁶³, whose irreversible induction is a strong
135 signature of abnormal JNK signaling activation (Extended Data Tables 1 and 2).

136 Irreversibly upregulated genes of cluster 2 contain several key drivers of tumorigenesis in *Drosophila* conserved
137 in mammals. Among these, we noticed the presence of *zfh1*, a TF that is both a direct PcG target (Extended Data
138 Fig. 5a), is activated by multiple signaling pathways, including JAK/STAT and Notch signaling and which can
139 sustain self-renewal and tumor growth⁷⁴⁻⁷⁶. Furthermore, its mammalian homolog ZEB1 can induce epithelial
140 to mesenchymal transition⁷⁷. Consistent with its transcriptional upregulation, ZFH1 protein is increased in both
141 constant and transient tumors (Extended Data Fig. 4b,c). Importantly, *zfh1* RNAi reduces the growth of *ph*-
142 dependent tumors and partially restores cell polarity and photoreceptor differentiation (Figure 2e-h), indicating
143 that it is a *bona fide* driver of the tumor phenotype.

144 Together, these results demonstrate that EICs are driven by a restricted set of irreversibly upregulated genes,
145 including JNK and JAK/STAT signaling components, rather than by the vast pleiotropic dysregulation of cancer
146 genes that is observed upon constant PH depletion. Furthermore, we sought to investigate why this subset of
147 genes remains irreversibly upregulated after restoration of normal PH levels.

148

149 **Irreversible transcriptional changes drive *ph*-dependent tumorigenesis**

150 To identify the features that discriminate reversible genes from irreversible ones, we focused on cluster 2
151 (irreversible) and cluster 5 (reversible) genes that were either bound by PH or enriched for H3K27me3 or

152 H2AK118Ub (the canonical marks of PcG-mediated transcriptional repression, Extended Data Table 2). As
153 expected, reversible genes (n= 68) show no expression changes after transient *ph*-KD, whereas irreversible
154 genes (n= 61) remain upregulated (Extended Data Table 2, Figure 3a). PRC1-target genes are transcribed at
155 similarly low levels in both clusters in the absence of *ph*-KD, and induced at comparable levels upon constant *ph*-
156 KD; ruling out that higher transcriptional levels are the reason for irreversible genes being unable to recover
157 normal transcription after transient *ph*-KD (Figure 3a,b).

158 We then explored the possibility that chromatin might not be correctly re-established at irreversible genes. We
159 performed ChIP-Seq and CUT&RUN in constant and transient (d11) *ph*-KD EICs. PH ChIP-Seq showed a loss of
160 PH binding for both reversible and irreversible gene sets upon constant *ph*-KD, as expected, while PH binding
161 was remarkably recovered after transient *ph*-KD (Figure 3c, d; Extended Data Fig. 5h). CUT&RUNs against
162 H3K27me3 indicates a stark decrease upon constant *ph*-KD and a reciprocal increase in H3K27Ac – its activating
163 counterpart – both at reversible and irreversible genes (Figure 3 d-h and Extended Data Fig. 5c) and a similar yet
164 weaker trend was observed for H2AK118Ub (Extended Data Fig. 5e-f). Interestingly, both clusters show
165 comparable H3K27me3 levels after transient *ph*-KD, with a concomitant decrease in H3K27ac (Figure 3d,g,h).
166 Inspection of individual loci showed that qualitative and quantitative recovery is generally observed at both
167 reversible and irreversible genes (Figure 3d-h and Extended Data Fig. 5c,e,f,h). Taken together, these results
168 indicate a surprising uncoupling between the impact of transient *ph*-KD on transcription and chromatin, whereby
169 irreversible transcriptional changes drive tumorigenesis despite the re-establishment of an essentially normal
170 chromatin landscape.

171

172 **Transcriptional repressor binding distinguishes reversible and irreversible genes**

173 To test whether differential levels of PRC1 binding might explain the repression of reversible genes upon transient
174 *ph*-KD, we focused on reversible (n= 32) and irreversible (n= 26) genes whose promoters are bound by PH (see
175 material and methods, Extended Data Table 2). Despite showing similar transcriptional levels (Figure 4a), PH
176 ChIP-Seq in control EDs (no *ph*-KD) indicates that PH binding and H2AK118Ub levels are significantly stronger
177 at the TSSs of reversible genes compared to irreversible ones (Figure 4b, Extended Data 5g). A similar trend is
178 observed for H3K27me3, although not significant (Extended Data 5d). This was further confirmed using available
179 ChIP-Seq data from WT EDs^{64,68} for the two other tumor-suppressor PRC1 members, PC and PSC, and the PRC2
180 member SU(Z)12, suggesting that the fully assembled PRC1 binds together with PRC2 to these promoter regions
181 (Figure 4c,d). We therefore analyzed the binding profiles of sequence-specific DNA binding factors that are known
182 to recruit PcG proteins – namely Pleiohomeotic (PHO), Sp1-like factor for pairing sensitive-silencing (SPPS)⁷⁸,
183 Combgap (CG)⁷⁹ and Trithorax-like (TRL)⁸⁰ – and found that all 4 show similar trends, with increased binding at
184 the promoters of reversible genes (Figure 4e).

185 Finally, DNA sequence analysis shows that consensus motifs for the TRL, and ADF⁸¹ recruiters of PcG
186 components are enriched at reversible genes (Figure 4f)^{82,83}. TSSs of reversible genes are significantly enriched
187 for consensus of several transcriptional repressors, such as E(SPL)bHLH repressors including hairy (*h*), as well
188 as HKB, BRK, DPN, KLU, TTK, DAR1, KLF15 and GRH (Figure 4d).

189 Together, these results indicate a tight equilibrium, whereby less potent recruitment of PRC1 coupled with reduced
190 binding of independent transcriptional repressors make irreversible genes more sensitive to transient *ph*-KD, while
191 the upregulation of reversible genes requires constant *ph*-KD.

192

193 **EICs are autonomous, immortal tumors that become more aggressive over time**

194 Most EICs-bearing larvae die after d11 AEL, preventing the study of tumor development over time. To circumvent
195 this limitation, allografts of imaginal disc tissue into the abdomen of adult *Drosophila* hosts are commonly used to
196 assess the tumorigenic potential of a tissue, and we previously showed that *ph* mutant EDs continuously grow
197 until they eventually kill the host⁶⁶. To be able to track transplanted EICs, we developed a variant of our
198 themosensitive system that constitutively expresses GFP in the eye, while a UAS-RFP cassette can be used as
199 a reporter of ongoing *ph*-KD (Extended Data Fig. 6a, f-i). Importantly, this system induces EICs with similar
200 penetrance, morphological and transcriptional defects compared to the previous one (Extended Data Fig. 6b-e;
201 Extended Data Table 3). We then performed a series of allografts using this line (Figure 5a). Of note, host flies

202 were always kept at a restrictive temperature after transplant (18°C), precluding that further overgrowth would
203 result from ongoing *ph*-RNAi.

204 Constant *ph*-KD tumors were able to expand and invade a fraction of injected host flies at restrictive temperature
205 (Figure 5b). Transient *ph*-KD EICs behaved similarly, indicating that their overgrowth results from an autonomous,
206 stably acquired state (Figure 5b). To quantitatively measure tumor growth over time, we set up an arborescent
207 allograft scheme, allowing us to trace the tumor of origin (Extended Data Fig. 7a). Tumors derived from both,
208 constant or transient PH depletion, maintained their ability to expand in host flies over 10 rounds of transplantation
209 (about 3 months). Tumor growth penetrance – defined as the percentage of hosts flies bearing GFP positive cells
210 20 days after transplantation – did not decrease but rather increased over generations of transplantation (Figure
211 5b), as the survival of host flies kept on decreasing (Figure 5c). Median lifespan of the hosts after the first
212 transplant with constant or transient EICs was of 20 and 21,5 days, respectively (Figure 5d), compared to 31,5
213 days for control tissues. In contrast, hosts transplanted with the fifth generation of allografts survived 18 (constant
214 *ph*-KD) and 11 days (transient *ph*-KD, d11), and only 10 days for tenth generation allografts (Figure 5e).
215 Furthermore, tumors produced metastases in regions and organs located far away from the injection site, with an
216 increasing penetrance (Extended Data Fig. 7b, c).

217 These results suggest that, rather than being a milder form of neoplastic transformation, EICs tumorigenic
218 potential is autonomously maintained, increases over time, and can propagate months after *ph*-RNAi has been
219 removed.

220

221 Discussion

222 Despite great interest in epigenetic determinants of disease states, it is difficult to discriminate between genetic,
223 environmental and cell-intrinsic epigenetic contributions, which might cooperate to induce tumorigenesis⁵². Our
224 model system shows that upon transient depletion of PRC1 subunits, cells derail from their normal fate and
225 undergo neoplastic transformation (Extended Data Fig. 8). This transformation coincides with the irreversible
226 activation of genes including key JNK and JAK/STAT pathway members, which in turn might sustain cell growth,
227 proliferation, loss of cell polarity, cell migration and cytokine activity even in the presence of normal PRC1 function.
228 One main difference between these irreversibly activated genes and reversible PcG target genes is a more potent
229 PcG binding on the latter ones, consistent with higher levels of PcG recruiters. We posit that, even if PRC1 is
230 wiped out from these genes upon depletion, stronger PcG binding, perhaps concomitant with the function of
231 additional repressive TFs, facilitates the restoration of a repressive chromatin environment once PRC1 levels are
232 re-established. In contrast, weaker PRC1 binding at irreversible genes might not suffice to fully restore silencing,
233 resulting in a self-sustaining aberrant cell state that stimulates tumor progression (Extended Data Fig. 8). Of note,
234 even if some of these genes are rather strongly transcribed, the levels of H3K27Ac on their promoters and gene
235 bodies are distinctly lower compared to a random set of equally transcribed genes (not shown). It will be interesting
236 to analyze the basis for this difference and its significance.

237 Although the system described here has not been analyzed before, previous work has shown that self-sustaining
238 alternate cell states can be triggered by transient perturbations in immortalized breast cells⁸⁴ or other cultured
239 cells^{85,86}, including neural progenitor cells subjected to transient inhibition of the PRC2 complex⁸⁷. PRC2
240 impairment in mouse striatal neurons induces progressive neurodegeneration by triggering a self-sustaining
241 transcription derailment program over time⁸⁸. Furthermore, KO or transient chemical inhibition of PRC2 also led
242 cells to enter a quasi mesenchymal state that depends on ZEB1, the mouse homolog of fly *Zfh1*, is highly
243 metastatic and associates with poor patient survival⁷⁷. Therefore, epigenetic events might play a major role at
244 early stages of oncogenesis or during tumour progression in some mammalian cancers^{54,89,90}. Our survey of a
245 large database of different types of solid cancers (Extended Data Fig. 9) as well as of data from several Multiple
246 Myeloma patient cohorts (Extended Data Fig. 10) indicates that low expression levels of genes encoding canonical
247 PRC1 subunits is associated with several tumours and with poor patient prognosis, consistent with a putative
248 suppressive role for PRC1 in some tumor types. Future work might address the role of epigenetic perturbations
249 in these tumors and in other physiological processes.

250

251 References

252

- 253 1 Cavalli, G. & Heard, E. Advances in epigenetics link genetics to the environment and disease. *Nature*
254 **571**, 489-499 (2019). <https://doi.org/10.1038/s41586-019-1411-0>
- 255 2 Waddington, C. H. The epigenotype. 1942. *Int J Epidemiol* **41**, 10-13 (2012).
256 <https://doi.org/10.1093/ije/dyr184>
- 257 3 Fitz-James, M. H. & Cavalli, G. Molecular mechanisms of transgenerational epigenetic inheritance. *Nat*
258 *Rev Genet* **23**, 325-341 (2022). <https://doi.org/10.1038/s41576-021-00438-5>
- 259 4 Shaffer, S. M. *et al.* Rare cell variability and drug-induced reprogramming as a mode of cancer drug
260 resistance. *Nature* **546**, 431-435 (2017). <https://doi.org/10.1038/nature22794>
- 261 5 Hata, A. N. *et al.* Tumor cells can follow distinct evolutionary paths to become resistant to epidermal
262 growth factor receptor inhibition. *Nat Med* **22**, 262-269 (2016). <https://doi.org/10.1038/nm.4040>
- 263 6 Nam, A. S., Chaligne, R. & Landau, D. A. Integrating genetic and non-genetic determinants of cancer
264 evolution by single-cell multi-omics. *Nat Rev Genet* **22**, 3-18 (2021). [https://doi.org/10.1038/s41576-020-](https://doi.org/10.1038/s41576-020-0265-5)
265 [0265-5](https://doi.org/10.1038/s41576-020-0265-5)
- 266 7 McGranahan, N. & Swanton, C. Biological and therapeutic impact of intratumor heterogeneity in cancer
267 evolution. *Cancer Cell* **27**, 15-26 (2015). <https://doi.org/10.1016/j.ccell.2014.12.001>
- 268 8 Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546-1558 (2013).
269 <https://doi.org/10.1126/science.1235122>
- 270 9 Flavahan, W. A., Gaskell, E. & Bernstein, B. E. Epigenetic plasticity and the hallmarks of cancer. *Science*
271 **357** (2017). <https://doi.org/10.1126/science.aal2380>
- 272 10 Brock, A., Chang, H. & Huang, S. Non-genetic heterogeneity--a mutation-independent driving force for
273 the somatic evolution of tumours. *Nat Rev Genet* **10**, 336-342 (2009). <https://doi.org/10.1038/nrg2556>
- 274 11 Timp, W. & Feinberg, A. P. Cancer as a dysregulated epigenome allowing cellular growth advantage at
275 the expense of the host. *Nat Rev Cancer* **13**, 497-510 (2013). <https://doi.org/10.1038/nrc3486>
- 276 12 Rosenquist, R., Esteller, M. & Plass, C. Introduction: Epigenetics in cancer. *Semin Cancer Biol* **51**, iv-v
277 (2018). <https://doi.org/10.1016/j.semcancer.2018.07.002>
- 278 13 Timsah, Z. *et al.* Grb2 depletion under non-stimulated conditions inhibits PTEN, promotes Akt-induced
279 tumor formation and contributes to poor prognosis in ovarian cancer. *Oncogene* **35**, 2186-2196 (2016).
280 <https://doi.org/10.1038/onc.2015.279>
- 281 14 Chatterjee, A., Rodger, E. J. & Eccles, M. R. Epigenetic drivers of tumourigenesis and cancer metastasis.
282 *Semin Cancer Biol* **51**, 149-159 (2018). <https://doi.org/10.1016/j.semcancer.2017.08.004>
- 283 15 Turajlic, S., Sottoriva, A., Graham, T. & Swanton, C. Resolving genetic heterogeneity in cancer. *Nat Rev*
284 *Genet* **20**, 404-416 (2019). <https://doi.org/10.1038/s41576-019-0114-6>
- 285 16 Marine, J. C., Dawson, S. J. & Dawson, M. A. Non-genetic mechanisms of therapeutic resistance in
286 cancer. *Nat Rev Cancer* **20**, 743-756 (2020). <https://doi.org/10.1038/s41568-020-00302-4>
- 287 17 Maes, K. *et al.* Epigenetic Modifiers: Anti-Neoplastic Drugs With Immunomodulating Potential. *Front*
288 *Immunol* **12**, 652160 (2021). <https://doi.org/10.3389/fimmu.2021.652160>
- 289 18 Lesch, B. J. *et al.* Intergenerational epigenetic inheritance of cancer susceptibility in mammals. *Elife* **8**
290 (2019). <https://doi.org/10.7554/eLife.39380>
- 291 19 Martincorena, I. *et al.* Somatic mutant clones colonize the human esophagus with age. *Science* **362**,
292 911-917 (2018). <https://doi.org/10.1126/science.aau3879>
- 293 20 Yizhak, K. *et al.* RNA sequence analysis reveals macroscopic somatic clonal expansion across normal
294 tissues. *Science* **364** (2019). <https://doi.org/10.1126/science.aaw0726>
- 295 21 Yokoyama, A. *et al.* Age-related remodelling of oesophageal epithelia by mutated cancer drivers. *Nature*
296 **565**, 312-317 (2019). <https://doi.org/10.1038/s41586-018-0811-x>
- 297 22 Yoshida, K. *et al.* Tobacco smoking and somatic mutations in human bronchial epithelium. *Nature* **578**,
298 266-272 (2020). <https://doi.org/10.1038/s41586-020-1961-1>
- 299 23 Teixeira, V. H. *et al.* Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-
300 invasive lung cancer lesions. *Nat Med* **25**, 517-525 (2019). <https://doi.org/10.1038/s41591-018-0323-0>
- 301 24 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-
302 74 (2012). <https://doi.org/10.1038/nature11247>
- 303 25 Hutter, C. & Zenklusen, J. C. The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. *Cell*
304 **173**, 283-285 (2018). <https://doi.org/10.1016/j.cell.2018.03.042>
- 305 26 International Cancer Genome, C. *et al.* International network of cancer genome projects. *Nature* **464**,
306 993-998 (2010). <https://doi.org/10.1038/nature08987>
- 307 27 Bernstein, B. E. *et al.* The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* **28**, 1045-
308 1048 (2010). <https://doi.org/10.1038/nbt1010-1045>
- 309 28 Stunnenberg, H. G., International Human Epigenome, C. & Hirst, M. The International Human
310 Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. *Cell* **167**, 1897 (2016).
311 <https://doi.org/10.1016/j.cell.2016.12.002>
- 312 29 Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by Polycomb and COMPASS families.
313 *Science* **352**, aad9780 (2016). <https://doi.org/10.1126/science.aad9780>
- 314 30 Butera, A., Melino, G. & Amelio, I. Epigenetic "Drivers" of Cancer. *J Mol Biol* **433**, 167094 (2021).
315 <https://doi.org/10.1016/j.jmb.2021.167094>

- 316 31 Farlik, M. *et al.* DNA Methylation Dynamics of Human Hematopoietic Stem Cell Differentiation. *Cell Stem*
317 *Cell* **19**, 808-822 (2016). <https://doi.org/10.1016/j.stem.2016.10.019>
- 318 32 Muller, D. & Gyorffy, B. DNA methylation-based diagnostic, prognostic, and predictive biomarkers in
319 colorectal cancer. *Biochim Biophys Acta Rev Cancer* **1877**, 188722 (2022).
320 <https://doi.org/10.1016/j.bbcan.2022.188722>
- 321 33 Pon, J. R. & Marra, M. A. Driver and passenger mutations in cancer. *Annu Rev Pathol* **10**, 25-50 (2015).
322 <https://doi.org/10.1146/annurev-pathol-012414-040312>
- 323 34 Kloetgen, A., Thandapani, P., Tsirigos, A. & Aifantis, I. 3D Chromosomal Landscapes in Hematopoiesis
324 and Immunity. *Trends Immunol* **40**, 809-824 (2019). <https://doi.org/10.1016/j.it.2019.07.003>
- 325 35 Feinberg, A. P., Koldobskiy, M. A. & Gondor, A. Epigenetic modulators, modifiers and mediators in
326 cancer aetiology and progression. *Nat Rev Genet* **17**, 284-299 (2016).
327 <https://doi.org/10.1038/nrg.2016.13>
- 328 36 Schuyler, R. P. *et al.* Distinct Trends of DNA Methylation Patterning in the Innate and Adaptive Immune
329 Systems. *Cell Rep* **17**, 2101-2111 (2016). <https://doi.org/10.1016/j.celrep.2016.10.054>
- 330 37 Cancer Genome Atlas Research, N. *et al.* Genomic and epigenomic landscapes of adult de novo acute
331 myeloid leukemia. *N Engl J Med* **368**, 2059-2074 (2013). <https://doi.org/10.1056/NEJMoa1301689>
- 332 38 Spencer, D. H. *et al.* CpG Island Hypermethylation Mediated by DNMT3A Is a Consequence of AML
333 Progression. *Cell* **168**, 801-816 e813 (2017). <https://doi.org/10.1016/j.cell.2017.01.021>
- 334 39 Bates, S. E. Epigenetic Therapies for Cancer. *N Engl J Med* **383**, 650-663 (2020).
335 <https://doi.org/10.1056/NEJMra1805035>
- 336 40 You, J. S. & Jones, P. A. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell* **22**,
337 9-20 (2012). <https://doi.org/10.1016/j.ccr.2012.06.008>
- 338 41 Feinberg, A. P. & Tycko, B. The history of cancer epigenetics. *Nat Rev Cancer* **4**, 143-153 (2004).
339 <https://doi.org/10.1038/nrc1279>
- 340 42 Baylin, S. B. & Jones, P. A. Epigenetic Determinants of Cancer. *Cold Spring Harb Perspect Biol* **8** (2016).
341 <https://doi.org/10.1101/cshperspect.a019505>
- 342 43 Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov* **12**, 31-46 (2022).
343 <https://doi.org/10.1158/2159-8290.CD-21-1059>
- 344 44 Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57-70 (2000).
345 [https://doi.org/10.1016/s0092-8674\(00\)81683-9](https://doi.org/10.1016/s0092-8674(00)81683-9)
- 346 45 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
347 <https://doi.org/10.1016/j.cell.2011.02.013>
- 348 46 Alonso-Curbelo, D. *et al.* A gene-environment-induced epigenetic program initiates tumorigenesis.
349 *Nature* **590**, 642-648 (2021). <https://doi.org/10.1038/s41586-020-03147-x>
- 350 47 Feinberg, A. P. A genetic approach to cancer epigenetics. *Cold Spring Harb Symp Quant Biol* **70**, 335-
351 341 (2005). <https://doi.org/10.1101/sqb.2005.70.027>
- 352 48 Ramos, P. *et al.* Small cell carcinoma of the ovary, hypercalcemic type, displays frequent inactivating
353 germline and somatic mutations in SMARCA4. *Nat Genet* **46**, 427-429 (2014).
354 <https://doi.org/10.1038/ng.2928>
- 355 49 Makohon-Moore, A. P. *et al.* Limited heterogeneity of known driver gene mutations among the
356 metastases of individual patients with pancreatic cancer. *Nat Genet* **49**, 358-366 (2017).
357 <https://doi.org/10.1038/ng.3764>
- 358 50 McDonald, O. G. *et al.* Epigenomic reprogramming during pancreatic cancer progression links anabolic
359 glucose metabolism to distant metastasis. *Nat Genet* **49**, 367-376 (2017).
360 <https://doi.org/10.1038/ng.3753>
- 361 51 Fennell, K. A. *et al.* Non-genetic determinants of malignant clonal fitness at single-cell resolution. *Nature*
362 **601**, 125-131 (2022). <https://doi.org/10.1038/s41586-021-04206-7>
- 363 52 Pascual, G. *et al.* Dietary palmitic acid promotes a prometastatic memory via Schwann cells. *Nature* **599**,
364 485-490 (2021). <https://doi.org/10.1038/s41586-021-04075-0>
- 365 53 Mack, S. C. *et al.* Epigenomic alterations define lethal CIMP-positive ependymomas of infancy. *Nature*
366 **506**, 445-450 (2014). <https://doi.org/10.1038/nature13108>
- 367 54 Feinberg, A. P. The Key Role of Epigenetics in Human Disease Prevention and Mitigation. *N Engl J Med*
368 **378**, 1323-1334 (2018). <https://doi.org/10.1056/NEJMra1402513>
- 369 55 Chan, H. L. & Morey, L. Emerging Roles for Polycomb-Group Proteins in Stem Cells and Cancer. *Trends*
370 *Biochem Sci* **44**, 688-700 (2019). <https://doi.org/10.1016/j.tibs.2019.04.005>
- 371 56 Sparmann, A. & van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nat*
372 *Rev Cancer* **6**, 846-856 (2006).
- 373 57 Koppens, M. & van Lohuizen, M. Context-dependent actions of Polycomb repressors in cancer.
374 *Oncogene* **35**, 1341-1352 (2016). <https://doi.org/10.1038/onc.2015.195>
- 375 58 Schuettengruber, B., Bourbon, H. M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and
376 Trithorax: 70 Years and Counting. *Cell* **171**, 34-57 (2017). <https://doi.org/10.1016/j.cell.2017.08.002>
- 377 59 Parreno, V., Martinez, A. M. & Cavalli, G. Mechanisms of Polycomb group protein function in cancer.
378 *Cell Res* **32**, 231-253 (2022). <https://doi.org/10.1038/s41422-021-00606-6>

- 379 60 Schuettengruber, B., Martinez, A. M., Iovino, N. & Cavalli, G. Trithorax group proteins: switching genes
380 on and keeping them active. *Nat Rev Mol Cell Biol* **12**, 799-814 (2011). <https://doi.org/10.1038/nrm3230>
- 381 61 Gonzalez, C. Drosophila melanogaster: a model and a tool to investigate malignancy and identify new
382 therapeutics. *Nat Rev Cancer* **13**, 172-183 (2013). <https://doi.org/10.1038/nrc3461>
- 383 62 Mohr, S. E. & Perrimon, N. Drosophila melanogaster: a simple system for understanding complexity. *Dis*
384 *Model Mech* **12** (2019). <https://doi.org/10.1242/dmm.041871>
- 385 63 Bilder, D., Ong, K., Hsi, T. C., Adiga, K. & Kim, J. Tumour-host interactions through the lens of Drosophila.
386 *Nat Rev Cancer* **21**, 687-700 (2021). <https://doi.org/10.1038/s41568-021-00387-5>
- 387 64 Loubiere, V. *et al.* Coordinate redeployment of PRC1 proteins suppresses tumor formation during
388 Drosophila development. *Nat Genet* **48**, 1436-1442 (2016). <https://doi.org/10.1038/ng.3671>
- 389 65 Sievers, C., Comoglio, F., Seimiya, M., Merdes, G. & Paro, R. A deterministic analysis of genome integrity
390 during neoplastic growth in Drosophila. *PLoS One* **9**, e87090 (2014).
391 <https://doi.org/10.1371/journal.pone.0087090>
- 392 66 Martinez, A. M. *et al.* Polyhomeotic has a tumor suppressor activity mediated by repression of Notch
393 signaling. *Nat Genet* **41**, 1076-1082 (2009). <https://doi.org/10.1038/ng.414>
- 394 67 Beuchle, D., Struhl, G. & Muller, J. Polycomb group proteins and heritable silencing of Drosophila Hox
395 genes. *Development* **128**, 993-1004 (2001). <https://doi.org/10.1242/dev.128.6.993>
- 396 68 Loubiere, V., Papadopoulos, G. L., Szabo, Q., Martinez, A. M. & Cavalli, G. Widespread activation of
397 developmental gene expression characterized by PRC1-dependent chromatin looping. *Sci Adv* **6**,
398 eaax4001 (2020). <https://doi.org/10.1126/sciadv.aax4001>
- 399 69 Classen, A. K., Bunker, B. D., Harvey, K. F., Vaccari, T. & Bilder, D. A tumor suppressor activity of
400 Drosophila Polycomb genes mediated by JAK-STAT signaling. *Nat Genet* **41**, 1150-1155 (2009).
401 <https://doi.org/10.1038/ng.445>
- 402 70 Bunker, B. D., Nellimoottil, T. T., Boileau, R. M., Classen, A. K. & Bilder, D. The transcriptional response
403 to tumorigenic polarity loss in Drosophila. *Elife* **4** (2015). <https://doi.org/10.7554/eLife.03189>
- 404 71 Hamaratoglu, F. & Atkins, M. Rounding up the Usual Suspects: Assessing Yorkie, AP-1, and Stat
405 Coactivation in Tumorigenesis. *Int J Mol Sci* **21** (2020). <https://doi.org/10.3390/ijms21134580>
- 406 72 Rossi, F., Attolini, C. S., Mosquera, J. L. & Gonzalez, C. Drosophila Larval Brain Neoplasms Present
407 Tumour-Type Dependent Genome Instability. *G3 (Bethesda)* **8**, 1205-1214 (2018).
408 <https://doi.org/10.1534/g3.117.300489>
- 409 73 Beira, J. V., Torres, J. & Paro, R. Signalling crosstalk during early tumorigenesis in the absence of
410 Polycomb silencing. *PLoS Genet* **14**, e1007187 (2018). <https://doi.org/10.1371/journal.pgen.1007187>
- 411 74 Leatherman, J. L. & Dinardo, S. Zfh-1 controls somatic stem cell self-renewal in the Drosophila testis and
412 nonautonomously influences germline stem cell self-renewal. *Cell Stem Cell* **3**, 44-54 (2008).
413 <https://doi.org/10.1016/j.stem.2008.05.001>
- 414 75 Boukhatmi, H., Martins, T., Pillidge, Z., Kamenova, T. & Bray, S. Notch Mediates Inter-tissue
415 Communication to Promote Tumorigenesis. *Curr Biol* **30**, 1809-1820 e1804 (2020).
416 <https://doi.org/10.1016/j.cub.2020.02.088>
- 417 76 Enomoto, M., Takemoto, D. & Igaki, T. Interaction between Ras and Src clones causes interdependent
418 tumor malignancy via Notch signaling in Drosophila. *Dev Cell* **56**, 2223-2236 e2225 (2021).
419 <https://doi.org/10.1016/j.devcel.2021.07.002>
- 420 77 Zhang, Y. *et al.* Genome-wide CRISPR screen identifies PRC2 and KMT2D-COMPASS as regulators of
421 distinct EMT trajectories that contribute differentially to metastasis. *Nat Cell Biol* **24**, 554-564 (2022).
422 <https://doi.org/10.1038/s41556-022-00877-0>
- 423 78 Brown, J. L., Grau, D. J., DeVido, S. K. & Kassis, J. A. An Sp1/KLF binding site is important for the
424 activity of a Polycomb group response element from the Drosophila engrailed gene. *Nucleic Acids Res*
425 **33**, 5181-5189 (2005). <https://doi.org/10.1093/nar/gki827>
- 426 79 Ray, P. *et al.* Combgap contributes to recruitment of Polycomb group proteins in Drosophila. *Proc Natl*
427 *Acad Sci U S A* **113**, 3826-3831 (2016). <https://doi.org/10.1073/pnas.1520926113>
- 428 80 Oh, H. *et al.* Genome-wide association of Yorkie with chromatin and chromatin-remodeling complexes.
429 *Cell Rep* **3**, 309-318 (2013). <https://doi.org/10.1016/j.celrep.2013.01.008>
- 430 81 Orsi, G. A. *et al.* High-resolution mapping defines the cooperative architecture of Polycomb response
431 elements. *Genome Res* **24**, 809-820 (2014). <https://doi.org/10.1101/gr.163642.113>
- 432 82 Kassis, J. A. & Brown, J. L. Polycomb group response elements in Drosophila and vertebrates. *Adv*
433 *Genet* **81**, 83-118 (2013). <https://doi.org/10.1016/B978-0-12-407677-8.00003-8>
- 434 83 Schuettengruber, B. *et al.* Cooperativity, specificity, and evolutionary stability of Polycomb targeting in
435 Drosophila. *Cell Rep* **9**, 219-233 (2014). <https://doi.org/10.1016/j.celrep.2014.08.072>
- 436 84 Iliopoulos, D., Hirsch, H. A. & Struhl, K. An epigenetic switch involving NF-kappaB, Lin28, Let-7
437 MicroRNA, and IL6 links inflammation to cell transformation. *Cell* **139**, 693-706 (2009).
438 <https://doi.org/10.1016/j.cell.2009.10.014>
- 439 85 Mayran, A. *et al.* Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate.
440 *Nat Genet* **50**, 259-269 (2018). <https://doi.org/10.1038/s41588-017-0035-2>
- 441 86 Reizel, Y. *et al.* FoxA-dependent demethylation of DNA initiates epigenetic memory of cellular identity.
442 *Dev Cell* **56**, 602-612 e604 (2021). <https://doi.org/10.1016/j.devcel.2021.02.005>

- 443 87 Holoch, D. *et al.* A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in
444 mammals. *Nat Genet* **53**, 1686-1697 (2021). <https://doi.org/10.1038/s41588-021-00964-2>
445 88 von Schimmelmann, M. *et al.* Polycomb repressive complex 2 (PRC2) silences genes responsible for
446 neurodegeneration. *Nat Neurosci* **19**, 1321-1330 (2016). <https://doi.org/10.1038/nn.4360>
447 89 Jaffe, L. F. Epigenetic theories of cancer initiation. *Adv Cancer Res* **90**, 209-230 (2003).
448 [https://doi.org/10.1016/s0065-230x\(03\)90007-8](https://doi.org/10.1016/s0065-230x(03)90007-8)
449 90 Holliday, R. A new theory of carcinogenesis. *Br J Cancer* **40**, 513-522 (1979).
450 <https://doi.org/10.1038/bjc.1979.216>
451
452

453 **Figures**
454

Figure 1

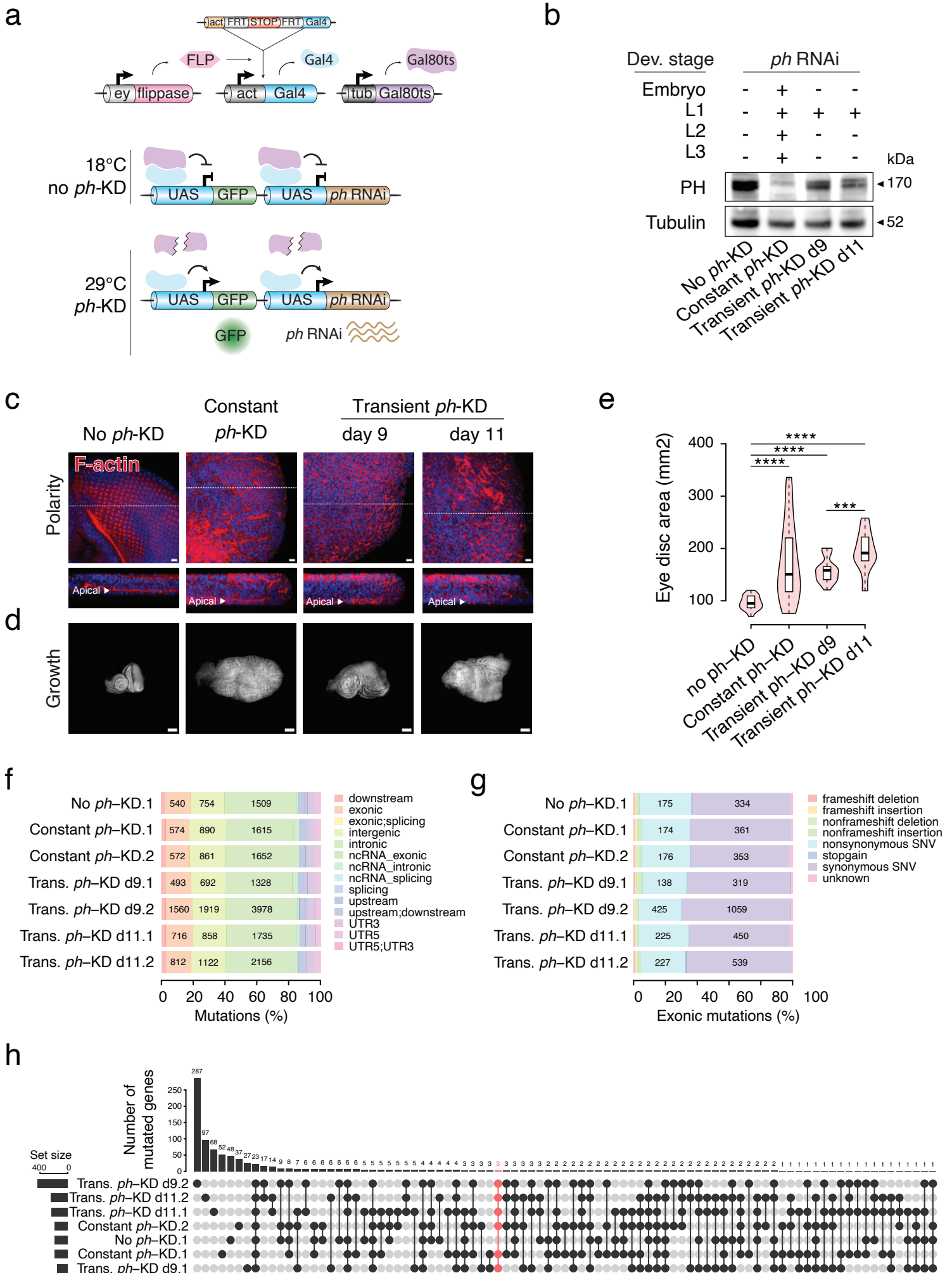


Figure 1: Transient PRC1 depletion is sufficient to initiate tumors

a- Scheme depicting the system for conditional depletion of PH. The expression of *ph*-RNAi, as well as that of the GFP marker, is under the control of UAS sequences. Cells that express *ey-FLP* (in pink) catalyze FLP-out of a transcriptional stop (in orange) in developing EDs, allowing the expression of *act-Gal4* (in light blue). The *tub-Gal80^{ts}* (in purple) encodes a ubiquitously expressed, temperature-sensitive Gal4 repressor. At restrictive temperature (29°C), but not at 18°C, Gal80^{ts} is inactivated. Gal4 can therefore bind to UAS sequences, allowing the expression of *ph*-RNAi and GFP that is used as a read-out of *ph*-KD. **b-** Western Blot analysis of PH expression in EDs of L3 larvae subjected to no *ph*-KD (control), constant or transient *ph*-KD at L1 stage. **c-** Confocal microscopy analysis of F-actin stained with rhodamine-phalloidin (red) showing a well-organized monolayer-stratified epithelium with apical F-actin seen in planar (top) and xy cross-sections (bottom) in no *ph*-KD (control), while polarity is disrupted in constant or transient *ph*-KD EDs. DNA is stained with DAPI (blue). **d-e-** Comparative measurement of ED size quantified as an overall DAPI staining area (d) showing ED sizes (e) under no *ph*-KD (control), constant or transient *ph*-KD conditions (wilcoxon test, ***pval<1e-3, ****pval<1e-5). **f-g-** Feature distribution of all (f) or exonic (g) somatic SNPs and InDels from no *ph*-KD (control), constant or transient *ph*-KD tissues (2 biological replicates). Numbers are shown for highly represented features (>10%). **h-** Mutated gene overlaps between conditions (including frameshift, nonsynonymous and stop gain mutations). Scale bars: 10 μm (c), 100 μm (d).

Figure 2

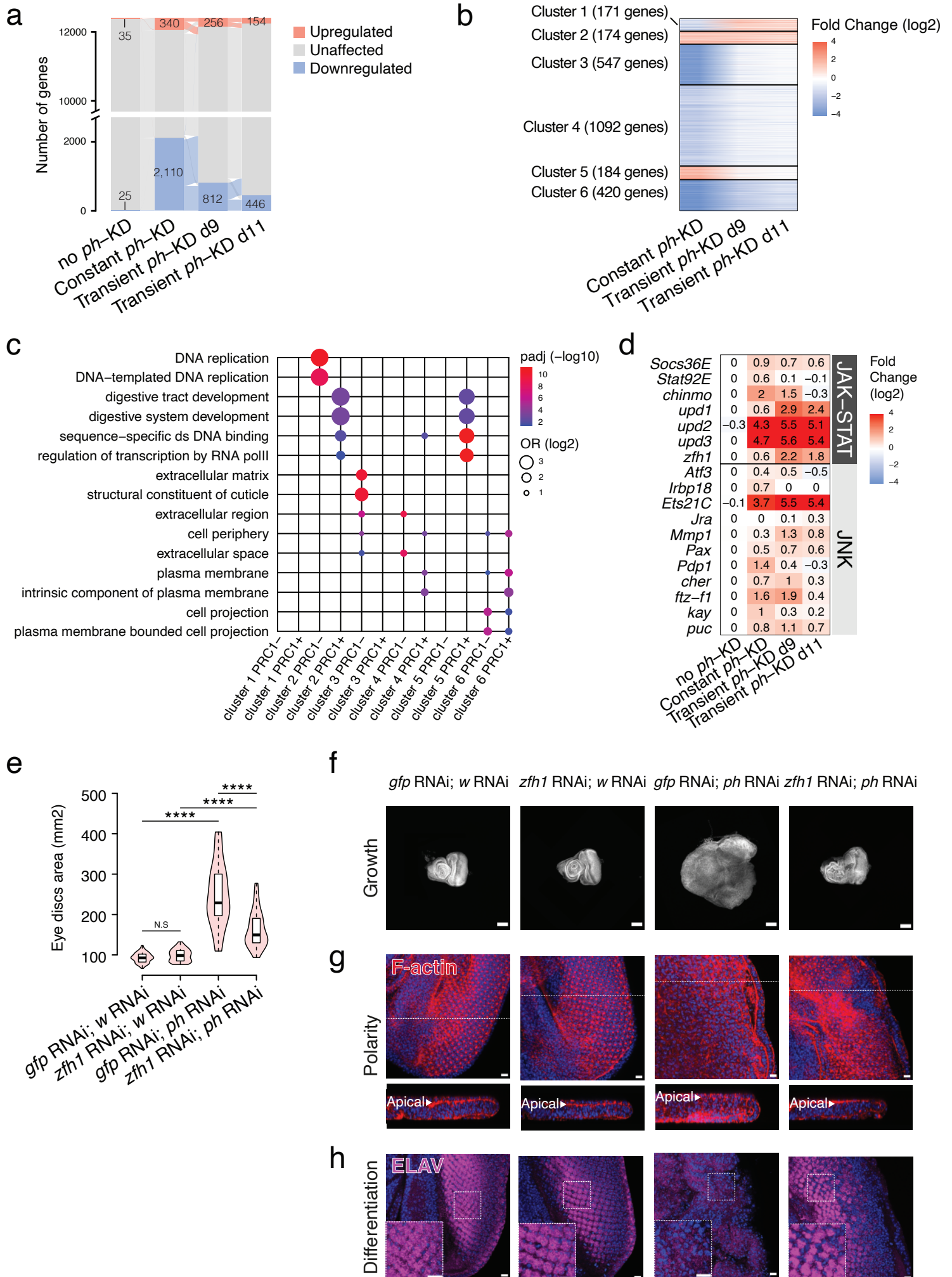


Figure 2: EICs maintain a subset of the dysregulated tumoral transcriptome and their maintenance requires the ZFH1 TF.

a- Alluvial plot displaying the dynamics of transcriptional states between no *ph*-KD (control), constant and transient *ph*-KD. Transitions between upregulated (orange), unaffected (grey) and downregulated (blue) states are indicated by thin lines of the same respective colors. **b-** Clustering of genes differentially expressed in constant and/or transient *ph*-KD, but not in the control no *ph*-KD. **c-** Gene ontologies significantly associated with each gene cluster, further stratified for the presence or absence of PH peaks located in the gene body or up to 2.5 kb upstream to the TSS. **d-** Transcriptional fold changes in genes involved in JAK-STAT (top) and JNK (bottom) signaling pathways upon PH depletion. **e-** Comparative size of *ph*-RNAi EDs in the presence or absence of *zfh1*-RNAi and the corresponding control RNAi (targeting the *white* and *gfp* genes, respectively). **f-h** stainings of *ph*-RNAi EDs in the presence or absence of *zfh1*-RNAi and the corresponding control RNAs (directed against *white* and *gfp*, respectively), marking the DAPI area (growth, **f**), F-actin (polarity, **g**) and neuronal marker ELAV (differentiation, **h**). Scale bars: 100 μ m (**f**), 10 μ m (**g,h**).

Figure 3

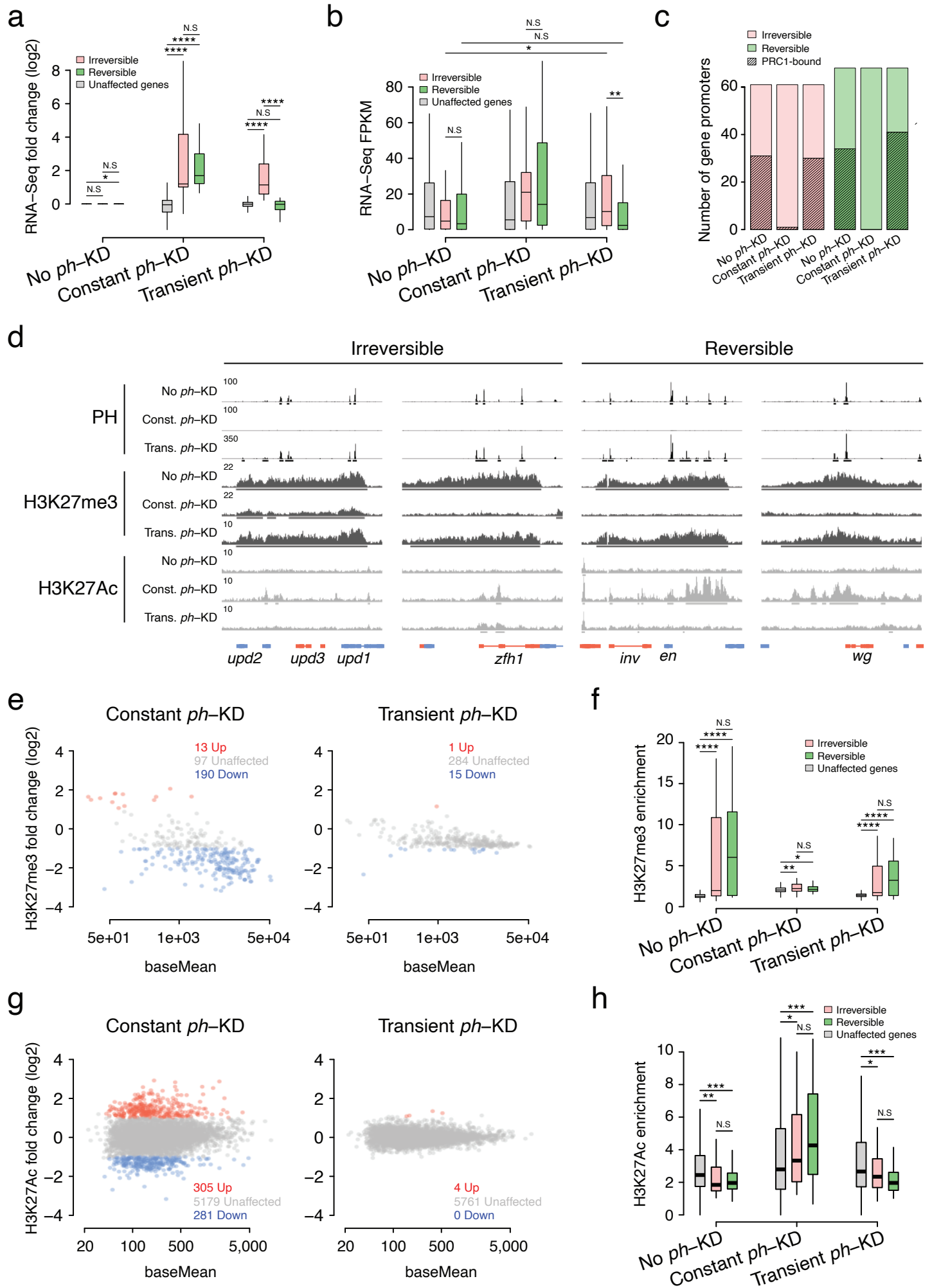


Figure 3: Stable transcriptional activation of irreversible genes despite the presence of Polycomb chromatin features.

a-, b- Fold Changes (**a**) and FPKMs (**b**) of irreversible (pink) and reversible (green) genes. All unaffected genes are used as reference (wilcoxon test, N.S- not significant, * $pval < 5e-2$, ** $pval < 1e-2$, *** $pval < 1e-3$, **** $pval < 1e-5$). **c-** Number of irreversible (pink) and reversible (green) gene promoters that are bound by PRC1 (black shading lines) in no *ph*-KD (control), constant or transient *ph*-KD. **d-** Screenshot of ChIP-Seq profiles for PH, H3K27me3 and H3K27Ac CUT&RUNs at representative genes activated irreversibly (left) or reversibly (right) under the indicated conditions. **e-** Differential analysis of H3K27me3 domains that are unaffected (grey) or exhibit decreased (blue) or increased (orange) enrichment upon constant (left) or transient (right) PH depletion. **f-** Quantification of H3K27me3 at irreversible (pink) and reversible (green) gene bodies, using all unaffected genes as reference (grey). Wilcoxon test, N.S- not significant, * $pval < 5e-2$, ** $pval < 1e-2$, *** $pval < 1e-3$, **** $pval < 1e-5$. **g-** Differential analysis of H3K27Ac peaks that are unaffected (grey) or exhibit decreased (blue) or increased (orange) enrichment upon constant (left) or transient (right) PH depletion. **h-** Quantification of H3K27Ac at irreversible (pink) and reversible (green) gene promoters, using all unaffected genes as reference (grey). Wilcoxon test, N.S- not significant, * $pval < 5e-2$, ** $pval < 1e-2$, *** $pval < 1e-3$, **** $pval < 1e-5$.

Figure 4

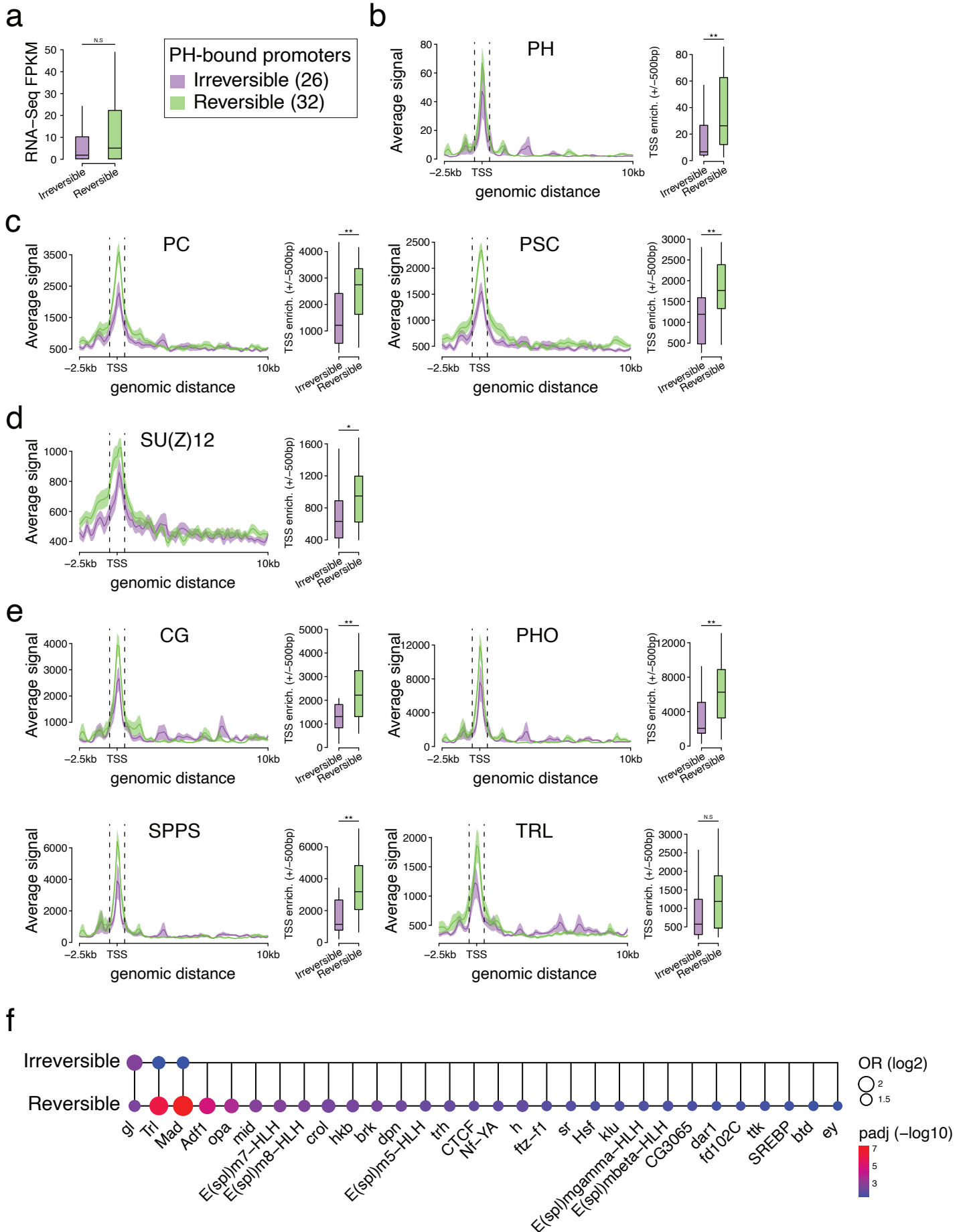


Figure 4: Irreversible genes are identified by lower levels of binding of Polycomb components and PcG recruiters

a- Boxplots of RNA levels in FPKMs of irreversible (purple) and reversible (green) PH-bound genes (wilcoxon test, N.S- not significant). **b-e-** Metaplots of average CHIP-Seqs for PRC1 (PH, PC, PSC) and PRC2 (SU(Z)12) subunits, as well as PcG recruiter proteins (CG, PHO, SPPS, TRL) to irreversible (purple) and reversible (green) genes. Quantifications are shown on the right (wilcoxon test, N.S- not significant, *pval<5e-2, **pval<1e-2, ***pval<1e-3, ****pval<1e-5). **f-** Comparative enrichment of DNA consensus binding motifs at irreversible and reversible gene promoters (-200bp/+50bp TSSs), irrespective of PH binding. All unaffected gene promoters were used as a control set.

Figure 5

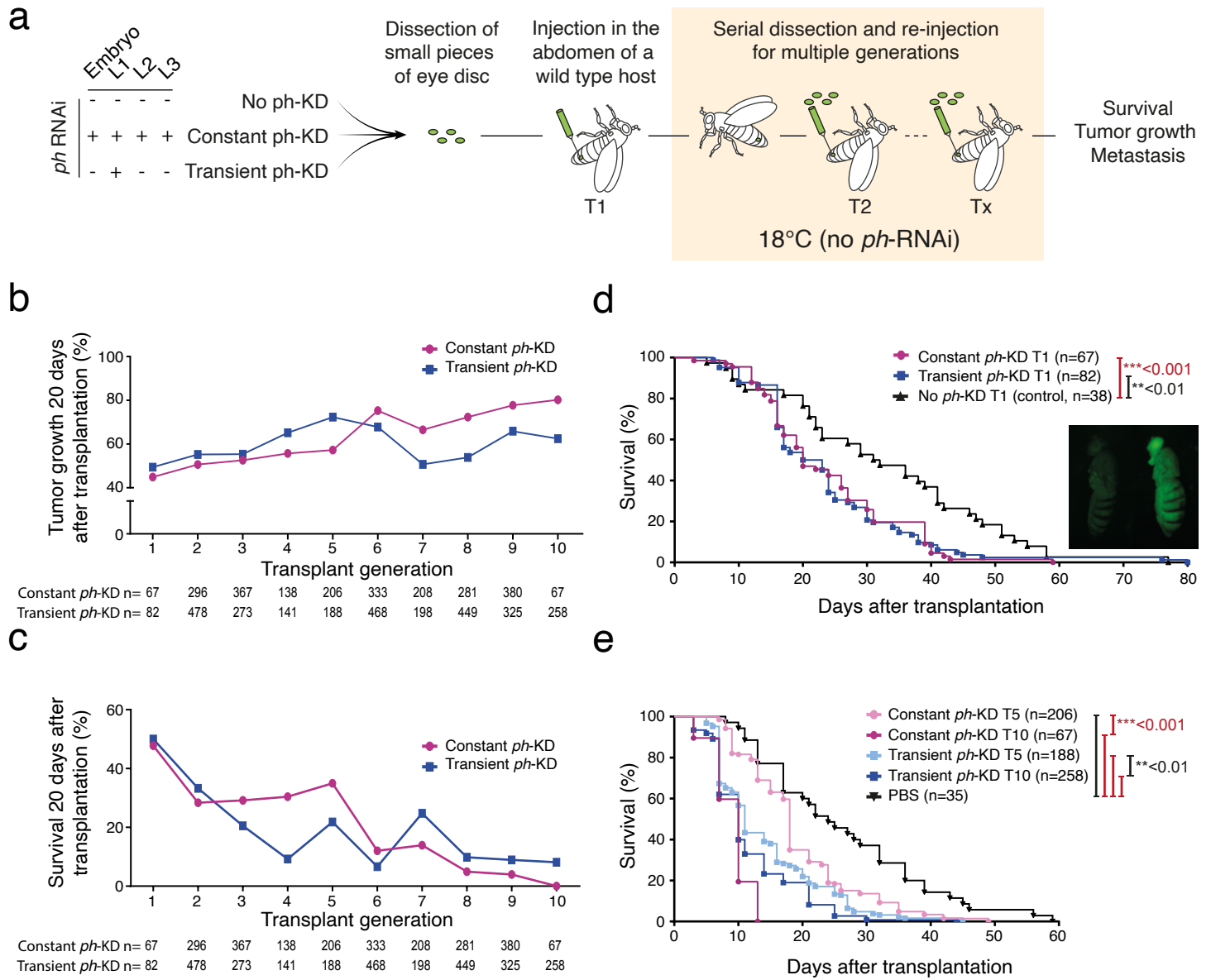


Figure 5: EICs generate immortal malignancies.

a- Schematic overview of experimental allograft workflow. GFP positive Flies of the same genotype are subjected to different temperature conditions at L1 stage, namely 18°C as control, constant PH depletion at 29°C during larval stages, or transient, 24H PH depletion at 29°C. EDs are dissected from donor larvae and allografted repeatedly into the abdomen of host flies until T10 (\approx 3 months). All allograft experiments are performed at 18°C to avoid *ph*-RNAi expression. **b-, c-** Tumor growth measured as a percentage of flies showing tumoral progression 20 days after transplantation (**b**) and survival of host flies 20 days after allograft of constant (purple) or transient (blue) *ph*-KD tumors during 10 rounds of transplantation (**c**). **d-, e-** Cumulative survival rate as percentage of flies calculated for 80 days after the first transplantation (**d**) and comparative cumulative percentage of flies that develop tumors calculated for 60 days after the fifth and the tenth transplantations of constant (purple and pink respectively) or transient (pale and dark blue respectively) *ph*-KD transplants. Significant differences were assessed using Log-rank test. n indicates the number of host flies analyzed.

455 **Methods**

456 **Fly strains**

457 Fly lines were obtained from Bloomington Drosophila Stock Center (BL) and Vienna BioCenter Core Facilities
458 (VDRC). All genotypes are indicated using *Drosophila* genetics annotation, i.e. a “;” indicates a different
459 chromosome. Genotypes are indicated starting from items located in the sex chromosomes on the left, items on
460 autosome 2 in the center and on autosome 3 on the right. A blank space is left when no specific items are present
461 in one chromosome. For heterozygous conditions, each allele is indicated and a “/” separates the two alleles on
462 homologous chromosomes.

463 The following fly strains were used for immunostaining, RNA-seq and CUT&RUN experiment:

464 y[1] w[*] P{ry[+t7.2]=ey-FLP.N}2 P{w[+mC]=GAL4-Act5C(FRT.CD2).P}D ; ; M{w[+mC]=UAS-NowGFP-NLS}ZH-
465 86Fb (BL#64095, driver line expressing Gal4 in the developing eye disc. Expresses GFP under the control of
466 Gal4).

467 ; P{w[+mC]=tubP-GAL80[ts]}20 ; UAS-phRNAi (derived from BL#7019 and VDRC#50028; used for *ph*-KD).

468 ; P{w[+mC]=tubP-GAL80[ts]}20 ; UAS-wRNAi (derived from BL#7019 and BL#33623, used for *white*-KD and the
469 generation of temperature-matched control transcriptomes).

470 ; UAS psc-su(z)2 RNAi ; P{w[+mC]=tubP-GAL80[ts]}ncd[GAL80ts-7] (derived from BL#38261, VDRC#100096
471 and BL#7018, used for *Psc/Su(z)2*-KD).

472 The following fly lines were used for the *Zfh1* rescue experiment:

473 y[1] w[*] P{ry[+t7.2]=ey-FLP.N}2 P{w[+mC]=GAL4-Act5C(FRT.CD2).P}D ; ; M{w[+mC]=UAS-NowGFP-NLS}ZH-
474 86Fb (BL#64095, driver line expressing Gal4 in the developing eye disc. Expresses GFP under the control of
475 Gal4).

476 ; UAS zfh1 RNAi (VDRC#103205, used for *zfh1*-KD).

477 ; UAS GFP RNAi (BL#9391, control RNAi).

478 ; ; UAS-wRNAi (BL#33623, control RNAi).

479 ; ; UAS-phRNAi (VDRC#50028, used for *ph*-KD).

480 The following lines were used for the allograft assay and gDNA sequencing:

481 eyFLP, UbiP63E(FRT.STOP)Stinger ; ; Act5C-Gal4(FRT.CD2), UAS-RFP /TM6BTb (derived from BL#5580,
482 BL#32249 and BL#30558; driver line expressing Gal4 and GFP in the developing eye disc. Expresses RFP under
483 the control of Gal4).

484 ; P{w[+mC]=tubP-GAL80[ts]}20 ; UAS-phRNAi (derived from BL#7019 and VDRC#50028; used for *ph*-KD).

485 w[*] ; P{w[+mC]=His2Av-mRFP1}III.1 (BL#23650).

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

489 **Experimental crosses**

490 Fly lines were maintained and amplified at 25°C on standard food medium. For all crosses, egg laying was
491 performed for 4h at 18°C, before applying the following temperature shifts:

- 492 - No *ph*-KD: 18°C, dissection 10 days After Egg Laying (AEL).
- 493 - Constant *ph* -KD: 29°C, dissection 4 days AEL.
- 494 - Larval depletion: 48 hours at 18°C, 29°C until dissection 5 days AEL.
- 495 - Transient *ph* -KD: 18°C for 48h, 29°C for 24h, 18°C until dissection at day 9 (d9) or day 11 (d11) AEL.
- 496 - Transient *ph* -KD L2 stage: 18°C for 96 hours, 29°C for 24 hours, 18°C until dissection at day 8 AEL.
- 497 - Transient *ph* -KD L3 early stage: 18°C for 120 hours, 29°C for 24 hours, 18°C until dissection 8 days
498 AEL.
- 499 - Transient *ph* -KD L3 late stage: 18°C for 168 hours, 29°C for 24 hours, 18°C until dissection 8 days
500 AEL.
- 501 - Transient *Psc*-KD: 18°C for 48 hours, 29°C for 48 hours, 18°C until dissection 8 days AEL.

502 **Immunostaining procedures**

503 Eye-antennal imaginal discs from third instar female larvae were dissected at RT in 1x PBS and fixed in 4%
504 Formaldehyde for 20 min. Tissues were permeabilized for 1h in 1x PBS + 0.5% Triton X-100 on a rotating wheel.
505 Permeabilized tissues were blocked 1 hour in 3% BSA PBTr (1x PBS + 0.1% Triton X-100), and incubated O/N
506 on a rotating wheel at 4°C with primary antibodies diluted in PBTr + 1% BSA. The following primary antibodies
507 were used: goat anti-PH¹ (1:500), mouse anti-ELAV (1:1000, DSHB, 9F8A9), mouse anti-ABD-B (1:1000, DSHS,
508 1A2E9), chicken anti-GFP (1:500, Life Technologies, D27C4), rabbit anti-zfh1² (1:2000). Then, sample were
509 washed in PBTr before adding secondary antibodies or rhodamine phalloidin in PBTr and 2 hours at RT on

510 rotating wheel. The following secondary antibodies were used: donkey anti-goat Alexa Fluor 555 (1:1000,
511 Invitrogen, A-21432), donkey anti-mouse Alexa Fluor 647 (1:1000, Invitrogen, A-31571), donkey anti-chicken
512 (1:1000, Clinisciences, 703-546-155), donkey anti-rabbit Alexa Fluor 555 (1:1000, Invitrogen, A-31572). F-actin
513 was stained by adding rhodamine phalloidin Alexa Fluor 555 (1:1000, Invitrogen, R415) or Alexa Fluor 488
514 (1:1000, Invitrogen, A12379). Tissues were washed in PBTr. DAPI staining were done at a final concentration of
515 1 µg/mL during 15 minutes. Discs were washed in PBTr. Discs were mounted in Vectashield medium (eurobio
516 scientific, catalog no. H-1000-10) and microscopy acquisition was performed on Leica SP8-UV confocal
517 microscope. Eye disc area was measured by defining the limits of the tissue using Fiji on at least 30 discs.

518 **RNA-seq**

519 Third instar female larvae were dissected in Schneider medium on ice. Total RNA was extracted using TRIzol
520 reagent. RNA purification was performed using RNA Clean & Concentrator kit (Zymo Research, #R1015). Finally,
521 poly-A RNAs selection, library preparation and illumina sequencing (20M paired-end reads, 150nt) were
522 performed by Novogene (<https://en.novogene.com/>). All experiments were performed in triplicate.

523 **Genomic DNA sequencing**

524 Genomic DNA was isolated using QIAamp DNA Micro Kit (Qiagen) following manufacturer's instructions. For
525 each biological replicate around 70 eye discs from female wandering larvae were dissected. Genomic DNAs
526 were processed for library preparation by Novogene (<https://en.novogene.com/>). Briefly, genomic DNA was
527 fragmented to an average size of ~350bp and then processed for DNA library preparation by following
528 manufacturer (Illumina) paired-end protocols. Sequencing was performed using Illumina Novaseq 6000 platform
529 to generate 150bp paired-ends reads with a coverage of at least 10x for 99% of the genome.

530 **Western blot**

531 About 150 eye-antennal imaginal discs were dissected in Schneider medium on ice per replicate. In order to
532 collect enough material, eye-antennal discs were dissected in batches, snap frozen in liquid nitrogen and stored
533 at -80°C. Discs were homogenized with a Tenbroeck directly in RIPA lysis buffer (50 mM Tris pH 7.5, 150 mM
534 NaCl, 1% NP40, 0.5% Na-deoxycholate, 0.1% SDS, 2X protease inhibitor) and incubated on ice for 10min. If
535 necessary, a second round of mechanical dissociation was performed. Samples were centrifuged 10 min at
536 10,000 x g at 4°C and supernatant was transferred to a fresh tube. Proteins were quantified using BCA protein
537 assay and 10 µg were used per gel lane, before 20 min of migration at 200V in MES 20X migration buffer and
538 1h transfer (1A). Membranes were blocked for 1 hour in PBS + 0.2% Tween + 10% milk powder at RT, incubated
539 O/N with primary antibodies in PBS + 0.2% tween at 4°C on shaker and washed in PBS+0.2% Tween. The
540 following primary antibodies were used: rabbit anti-PH (1:200), rabbit anti-zfh1² (1:2000), mouse anti-beta tubulin
541 (1:5000, DSHB, AA12.1). HRP-conjugated secondary antibodies were incubated with the membrane for 2 hours
542 at RT. The following secondary antibodies were used: goat anti-rabbit (1:15 000, Sigma, A0545), Rabbit anti-
543 mouse (1:15 000, Sigma A9044). Membranes were washed in PBS+0.2% Tween and revealed using Super
544 Signal West Dura kit (Pierce) and Chemidoc Bio-rad.

545 **Allograft assay**

546 Allografts were performed as previously described (Rossi and Gonzalez 2015). In brief, EDs were dissected in
547 PBS from 3rd instar female larvae, cut into pieces and injected into the abdomen of adult female hosts genotype
548 (BL#23650). Flies were monitored every two days and tumors were dissected and re-injected when the host
549 abdomen was fully GFP.

550 **ChIP-seq experiments**

551 ChIP-Seq on 3rd instar imaginal eye discs were essentially performed as described earlier³, with minor
552 modifications. 400 eye discs were used per replicate. When necessary, several batches of dissection/collection
553 were snap frozen in liquid nitrogen and stored at -80°C to collect enough material. Chromatin was sonicated
554 using a Bioruptor Pico (Diagenode) for 10 min (30sec in, 30sec off). PH antibodies were diluted 1:100 for IP.
555 After decrosslinking, DNA was purified using MicroChIP DiaPure columns from Diagenode. DNA libraries for
556 sequencing were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina. Sequencing (paired-
557 end sequencing 150bp, approx. 4Gb/sample) was performed by Novogene (<https://en.novogene.com/>).

558 **CUT&RUN experiments**

565 CUT&RUN experiments were performed as described by Kami Ahmad in protocols.io
566 (<https://dx.doi.org/10.17504/protocols.io.umfeu3n>) with minor modifications. 50 eye discs were dissected in
567 Schneider medium, centrifuged for 3 min at 700g and washed twice with wash+ buffer before addition of
568 Concanavalin A-coated beads. MNase digestion (pAG-MNase Enzyme from Cell Signaling) was performed for
569 30 min on ice. After ProteinaseK digestion, DNA was recovered using SPRIselect beads and eluted in 50ul TE.
570 DNA libraries for sequencing were prepared using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina.
571 Sequencing (paired-end sequencing 150bp, approx. 2Gb/sample) was performed by Novogene
572 (<https://en.novogene.com/>).

573

574 **Bioinformatic analyses**

575 All in-house bioinformatic analyses were performed in R version 3.6.3 (URL <https://www.R-project.org/>).
576 Computations on genomic coordinate files and downstream computations were conducted using the data.table
577 R package (data.table: Extension of 'data.frame'. <https://r-datatable.com>, <https://Rdatatable.gitlab.io/data.table>,
578 <https://github.com/Rdatatable/data.table>, v1.14.2).

579

580 **gDNA processing and mapping of somatic variants**

581 gDNA variants calling was performed by Novogene (<https://en.novogene.com/>). Briefly, base calling was
582 performed using Illumina pipeline CASAVA v1.8.2, and subjected to quality control using fastp with the following
583 parameters: -g -q 5 -u 50 -n 15 -l 150 --min_trim_length 10 --overlap_diff_limit 1--overlap_diff_percent_limit 10.
584 Then, sequencing reads were aligned to the dm6 version of the *Drosophila* genome using Burrows-Wheeler
585 Aligner (BWA) with default parameters and duplicate reads were removed using samtools and PICARD
586 (<http://picard.sourceforge.net>). Raw SNP/InDel sets were called using GATK with the following parameters: --
587 gcpHMM 10 -stand_emit_conf 10 -stand_call_conf 30. Then, SNPs were filtered using the following criteria: SNP:
588 QD<2, FS>60, MQ<30, HaplotypeScore>13, MappingQualityRankSum<-12.5, ReadPosRankSum<-8; For
589 INDEL variants, following criteria were used: QD<2, FS>200, ReadPosRankSum<-20. UCSC known genes were
590 used for gene and region annotations. Finally, variants were compared to the first batch of control, no *ph*-KD ED
591 in the search for bona fide de novo mutations using the MuTect2 module of the GATK package. Only SNP and
592 INDEL variants that passed Mutect2 filtering (FILTER="PASS") were considered for downstream analyses.

593 **RNA-seq processing and differential analysis**

594 After initial quality checks of newly generated data using fastqc
595 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), paired-end reads were aligned to a custom index
596 consisting of the dm6 version of the *Drosophila* genome together with GFP, EGFP and mRFP1 sequences, using
597 the align function from the Rsubread R package ⁴ (v2.0.1) with the following parameters: maxMismatches = 6,
598 unique = TRUE. Then, aligned reads were counted for each *D. melanogaster* transcript (dmel_r6.36 annotation)
599 using the featureCounts function from the Rsubread R package (v2.0.1, isPairedEnd = TRUE) and differential
600 expression analysis was performed using the DESeq2 R package ⁵ (v1.26.0, design = ~replicate+condition). For
601 each condition, *ph*-RNAi were compared to temperature-matched *w*-RNAi controls.
602 Then, genes that were not affected at 18°C (padj>=0.05) but differentially expressed in at least one of the other
603 *ph*-RNAi conditions (padj<0.05 and |log2Fold2FoldChange| >1) were considered for clustering. Log2FoldChange
604 values were clipped at 5th and 95th percentile, and clustered using the supersom function from the kohonen R
605 package ⁶ (v3.0.10). Given that D9 and D11 transient *ph*-KD yielded substantially similar transcriptomes, a two-
606 layer self-organizing map was trained (layer 1: Constant *ph*-KD; layer2: D9 and D11 transient *ph*-KD) with similar
607 weights for the 2 layers, using a 3x2 grid (topology = hexagonal, toroidal = TRUE).

608

609 **CUT&RUN and ChIP-Seq processing and analyses**

610 After initial quality checks of newly generated data using fastqc, paired-end reads were aligned to the dm6 version
611 of the *Drosophila* genome using bowtie 2 ⁷ (v2.3.5.1) with the following parameters: --local --very-sensitive-local
612 --no-unal --no-mixed --no-discordant --phred33 -l 10 -X 700, and low mapping quality were discarded using
613 samtools ⁸ (-q 30, v1.10, using htlib 1.10.2-3).

614 PH, H3K27me3, H3K27Ac and H2AK118Ub peaks/domains were called for each replicate separately and on
615 merged reads – using IgG as input control – using macs2 ⁹ (v2.2.7.1) with the following parameters: --keep-dup
616 1 -g dm -f BAMPE -B --SPMR. Only the peaks that were detected in both replicates (enrichment>0 AND
617 qvalue<0.05) and using merged reads (enrichment>2 AND qvalue<0.01) were retained for further analyses, after
618 being merged with a minimum gap size of 250 bp for narrow peaks (PH and H3K27Ac) and 2.5kb for broad marks
619 (H3K27me3 and H2AK118Ub). macs2 bedgraph files were used for visualization purposes.

620 PH, H3K27me3 and H2AK118Ub bound genes were defined as overlapping with at least one confident peak
621 (see previous section) anywhere in the body of the gene and up to 2.5kb upstream of its TSS.
622 For differential analysis of H3K27me3, H3K27Ac and H2AK118Ub, peaks were merge across all conditions
623 (maximum gap of 250 bp for H3K27Ac, 2.5kb for H3K27me3 and H2AK118Ub), overlapping reads were counted
624 using the featureCounts function from the Rsubread R package (v2.0.1, isPairedEnd = TRUE) and differential
625 analysis was performed using the DESeq2 R package (v1.26.0, size factors= total number of aligned reads,
626 design= ~replicate+condition).

627 628 **GO enrichment**

629 Gene ontologies were carried out using the AnnotationDbi R package
630 (<https://bioconductor.org/packages/AnnotationDbi>, v1.48.0). Briefly, genes of interest were compared to a
631 universe set of genes consisting of all genes that passed DESeq2 initial filters, using one-sided Fisher's exact
632 test (alternative = "greater"). Obtained p-values were corrected for multiple testing using Benjamini-Hochberg's
633 method.

634 635 **Motif enrichment**

636 Non-redundant TF motifs were obtained from de Almeida et al. ¹⁰, and only the motifs that are associated to a
637 known Drosophila TF were considered. Motif were counted at the promoters of reversible and irreversible genes
638 as well as unaffected genes (TSS -200/+50bp) using the matchMotifs function from the motifmatch R package
639 (Schep A (2022). motifmatchr: Fast Motif Matching in R. R package version 1.18.0) with the following parameters:
640 p.cutoff = 5e-4, bg = "genome", genome= "dm6". Then, both sets were compared using two-sided Fisher's exact
641 test and p-values were corrected for multiple testing using Benjamini-Hochberg's method. In the case where a
642 single TF would be associated to several motifs, only the most significant value was reported.

643 644 **Analysis of Human solid tumors**

645 The differential gene expression analysis was carried out by employing Mann-Whitney test using the
646 TNMplot database which contains transcriptome-level RNA-seq data for different tumor samples from The
647 Cancer Genome Atlas (TCGA) and The Genotype-Tissue Expression (GTEx) repositories
648 (<https://pubmed.ncbi.nlm.nih.gov/33807717/>).

649 The survival analysis was carried out using the Pan-Cancer (Bladder, Lung adenocarcinoma, and Rectum
650 adenocarcinoma) or gene-array (Breast, Ovarian, and Prostate) (<https://pubmed.ncbi.nlm.nih.gov/34527184/>
651 and <https://pubmed.ncbi.nlm.nih.gov/34309564/>) datasets of the online tool www.kmplot.com (accessed on 22
652 December 2022). The Pan-Cancer dataset is based on TCGA data generated using the Illumina HiSeq 2000
653 platform with survival information derived from the published sources
654 (<https://pubmed.ncbi.nlm.nih.gov/33723286/>). The gene-array samples were obtained using Affymetrix
655 HGU133A and HGU133plus2 gene chips. The samples were MAS5 normalized and the mean expression in
656 each sample was scaled to 1000. The most reliable probe sets to represent single genes were identified using
657 JetSet (<https://dash.harvard.edu/handle/1/10295333>).

658 In the survival analysis, each cutoff value between the lower and upper quartiles of expression was analyzed
659 by Cox proportional hazards regression and false-discovery rate was computed to correct for multiple
660 hypothesis testing. Then, the best performing cutoff was used in the final analysis. Kaplan-Meier survival plots
661 were generated to visualize the survival differences, and hazard rates with 95% confidence intervals were
662 computed to numerically assess the difference between the two cohorts. The statistical analysis was
663 performed in the R statistical environment (www.r-project.org). The analysis results for single genes can be
664 validated using the platforms at www.kmplot.com and www.tnmplot.com.

665 666 **Analysis of Multiple Myeloma patient cohorts**

667 For gene expression profiling data from patients with Multiple Myeloma, we used six cohorts that included
668 Affymetrix gene expression data (HGU133plus2) of purified MM cells from the TT2 (Gene Expression Omnibus,
669 accession number GSE2658), TT3 (accession number E-TABM-1138 accession number GSE4583) and Hovon
670 (accession number GSE19784) cohorts (345, 158 and 282 newly-diagnosed MM patients treated by high-dose
671 melphalan and autologous hematopoietic stem cell transplantation); the Mulligan cohort (188 patients at relapse
672 treated by proteasome inhibitor in monotherapy); the Mtp cohort non eligible to HDT ¹¹ (63 newly-diagnosed MM
673 patients non eligible to high-dose melphalan and autologous hematopoietic stem cell transplantation) and the
674 Mtp Dara cohort ¹¹ (51 patients at relapse treated by anti-CD38 monoclonal antibody (Daratumumab)). Gene
675 expression data were normalized with the MAS5 algorithm and processing of the data was performed using the

676 webtool genomicscape (<http://www.genomicscape.com>), as done previously^{12,13}, using the R environment
677 (www.r-project.org). The prognostic value of PHC1, PHC2, PHC3, CBX2, CBX7 et BMI1 gene expression was
678 investigated using the Maxstat R function and Kaplan Meier survival curves as previously described¹⁴. The
679 differential gene expression analysis between normal bone marrow plasma cells from healthy donors and MM
680 cells from patients was carried out by employing Mann-Whitney test. The prognostic value of *PHC1*, *PHC2*,
681 *PHC3*, *CBX2*, *CBX7* and *BMI1* genes was combined using our previously published methodology¹⁴ (sum of the
682 Cox b-coefficients of each of the 6 genes, weighted by ± 1 if the patient MMC signal for a given gene is above or
683 below the probe set Maxstat value of the gene). Clustering was performed using the Morpheus software
684 (<https://software.broadinstitute.org/morpheus>) and violin plots using GraphPad Prism software
685 (<http://www.graphpad.com/scientific-software/prism/>).

686
687

Method references

- 688 1 Grimaud, C. *et al.* RNAi Components Are Required for Nuclear Clustering of Polycomb Group Response
689 Elements. *Cell* **124**, 957-971 (2006).
- 690 2 Flaherty, M. S. *et al.* chinmo is a functional effector of the JAK/STAT pathway that regulates eye
691 development, tumor formation, and stem cell self-renewal in Drosophila. *Dev Cell* **18**, 556-568 (2010).
692 <https://doi.org/10.1016/j.devcel.2010.02.006>
- 693 3 Loubiere, V. *et al.* Coordinate redeployment of PRC1 proteins suppresses tumor formation during
694 Drosophila development. *Nat Genet* **48**, 1436-1442 (2016). <https://doi.org/10.1038/ng.3671>
- 695 4 Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for
696 alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* **47**, e47 (2019).
697 <https://doi.org/10.1093/nar/gkz114>
- 698 5 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq
699 data with DESeq2. *Genome Biol* **15**, 550 (2014). <https://doi.org/10.1186/s13059-014-0550-8>
- 700 6 Wehrens, R. & Kruisselbrink, J. Flexible Self-Organizing Maps in kohonen 3.0. *Journal of Statistical*
701 *Software* **87**, 1-18 (2018). <https://doi.org/10.18637/jss.v087.i07>
- 702 7 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359
703 (2012). <https://doi.org/10.1038/nmeth.1923>
- 704 8 Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10** (2021).
705 <https://doi.org/10.1093/gigascience/giab008>
- 706 9 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
707 <https://doi.org/10.1186/gb-2008-9-9-r137>
- 708 10 de Almeida, B. P., Reiter, F., Pagani, M. & Stark, A. DeepSTARR predicts enhancer activity from DNA
709 sequence and enables the de novo design of synthetic enhancers. *Nat Genet* **54**, 613-624 (2022).
710 <https://doi.org/10.1038/s41588-022-01048-5>
- 711 11 Ovejero, S. *et al.* The BLM helicase is a new therapeutic target in multiple myeloma involved in replication
712 stress survival and drug resistance. *Front Immunol* **13**, 983181 (2022).
713 <https://doi.org/10.3389/fimmu.2022.983181>
- 714 12 Kassambara, A. & Moreaux, J. Analysis of Global Gene Expression Profiles. *Methods Mol Biol* **1792**,
715 157-166 (2018). https://doi.org/10.1007/978-1-4939-7865-6_11
- 716 13 Kassambara, A. *et al.* GenomicScape: an easy-to-use web tool for gene expression data analysis.
717 Application to investigate the molecular events in the differentiation of B cells into plasma cells. *PLoS*
718 *Comput Biol* **11**, e1004077 (2015). <https://doi.org/10.1371/journal.pcbi.1004077>
- 719 14 Alaterre, E. *et al.* Comprehensive characterization of the epigenetic landscape in Multiple Myeloma.
720 *Theranostics* **12**, 1715-1729 (2022). <https://doi.org/10.7150/thno.54453>

721 Acknowledgements

722 We thank Montpellier Resources Imagerie facility as well as the *Drosophila* facility (both affiliated to BioCampus
723 University of Montpellier, CNRS, INSERM, Montpellier, France). V.P. was supported by the EpiGenMed cluster
724 of Excellence funding (PIA of the French Ministry of Higher Education and Research) and by la Ligue Nationale
725 Contre le Cancer. V.L. was supported by the EpiGenMed cluster of Excellence funding (PIA of the French Ministry
726 of Higher Education and Research). A-M.M was supported by a grant of the Fondation ARC (contract N. 216574,
727 acronym "Epicancer). B.S.. was supported by INSERM. G.C. was supported by CNRS. Research in the G.C.
728 laboratory was supported by grants from the European Research Council (Advanced Grant 3DEpi), the European
729 CHROMDESIGN ITN project (Marie Skłodowska-Curie grant agreement No 813327), the European E-RARE
730 NEURO DISEASES grant "IMPACT", by the Agence Nationale de la Recherche (PLASMADIFF3D, grant N. ANR-
731 18-CE15-0010), by the Fondation pour la Recherche Médicale (DEI20151234396), by the MSD Avenir Foundation
732 ((Project GENE-IGH), and by the French National Cancer Institute (INCa, PIT-MM grant N. INCA-PLBIO18-362).

733
734

Author contributions

735 V.L, V.P., A-M.M., and G.C. initiated and led the project. V.P, L.F. and V.L. performed genetic experiments. V.P.
736 performed immunostaining, molecular biology and genomic experiments. V.L. and M.DS. performed
737 computational analysis of genomic datasets. B.S. performed CHIP-seq and CUT&RUN experiments. M.E., B.G
738 and D.C. performed computational analysis of different tumor types. J.M. performed computational analysis of
739 Multiple Myeloma samples. V.L, V.P., A-M.M and G.C. wrote the manuscript. All the authors discussed the data
740 and reviewed the manuscript.

741

742 Competing interest declaration

743 The authors declare no competing interests.

744

745 Data availability:

746 The datasets generated during and/or analyzed during the current study will be available from GEO.

747

748 Supplementary information

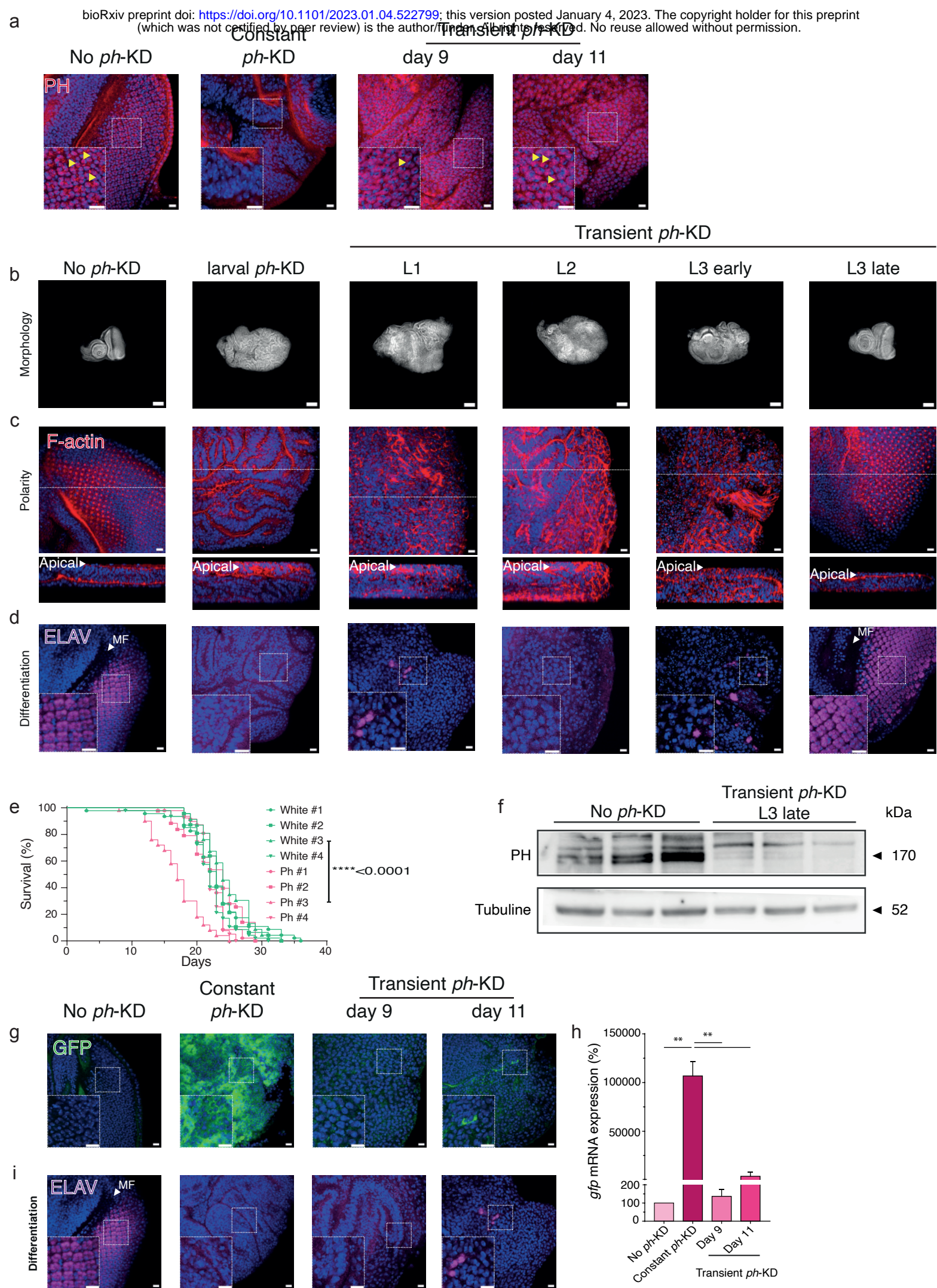
749 This manuscript includes supplementary information, namely ten Extended Data Figures and three Extended Data
750 Tables.

751 Extended Data Table 1: Differential analyses and FPKMs of no ph-KD, transient ph-KD and constant ph-KD in
752 the fly line expressing GFP marker only during depletion of PH.

753 Extended Data Table 2: Clustering of differentially expressed genes and recovery status.

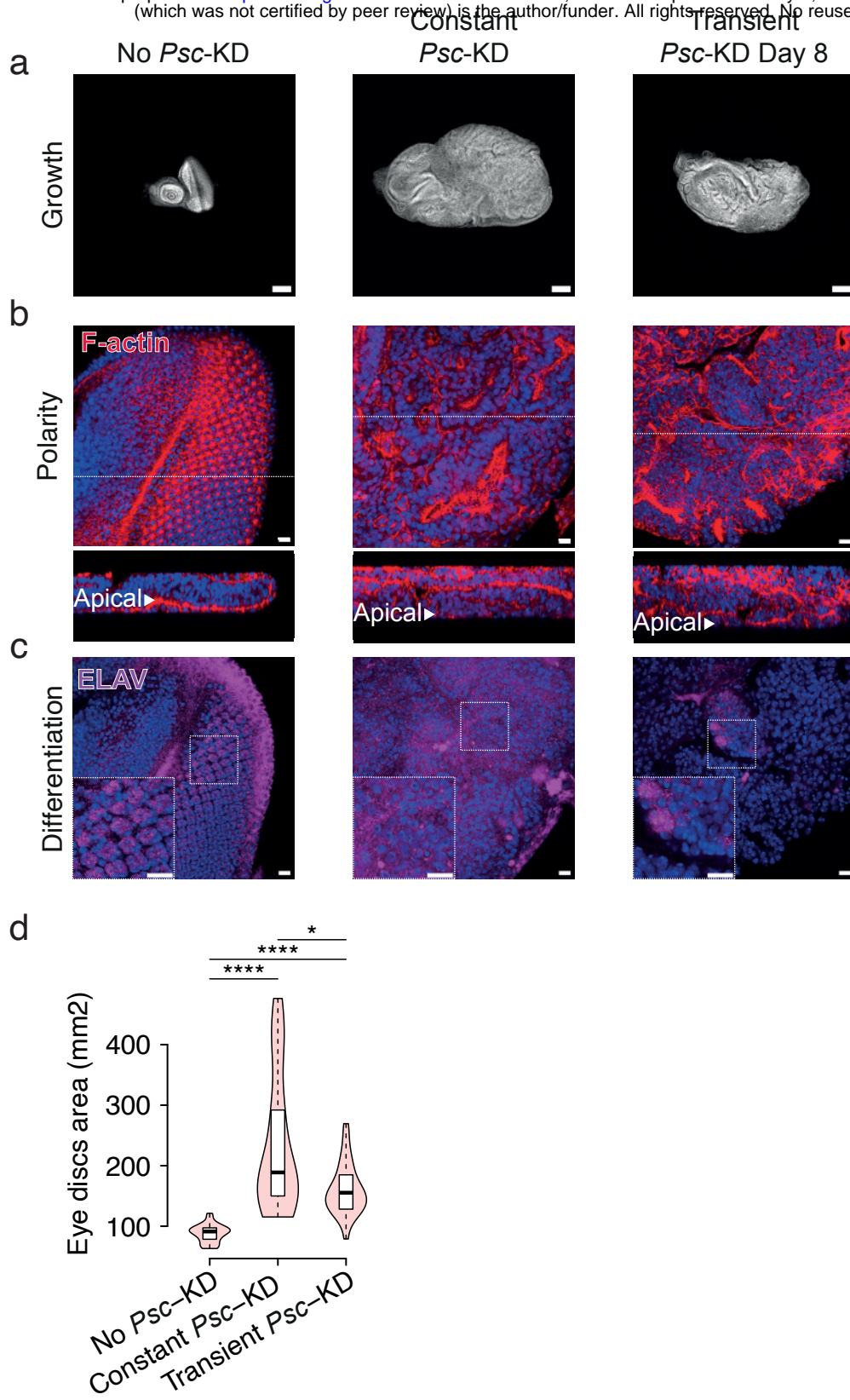
754 Extended Data Table 3 : Differential analyses and FPKMs of no ph-KD, transient ph-KD and constant ph-KD in
755 the fly line expressing constitutive GFP

756



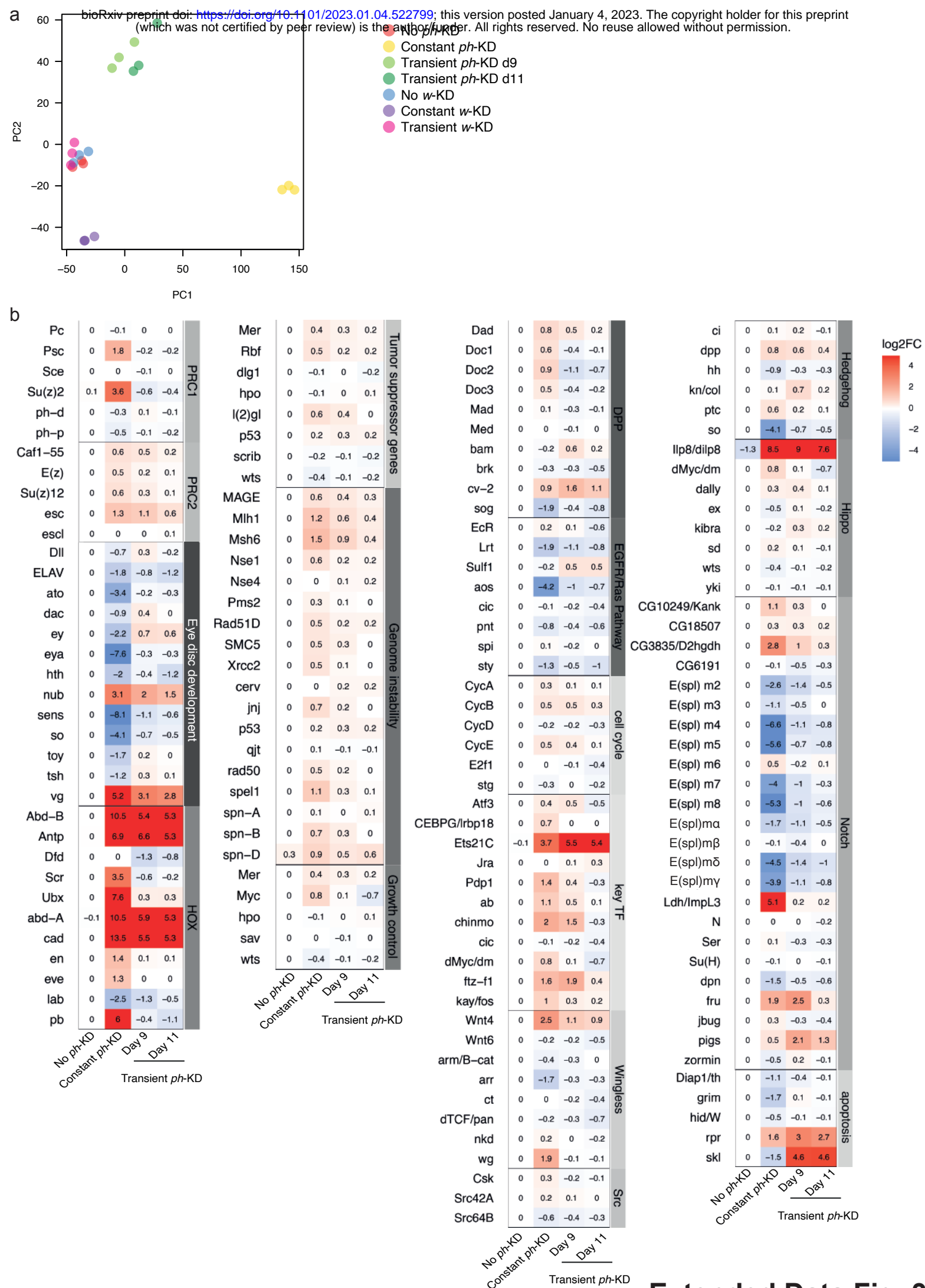
Extended Data Fig. 1: Transient *ph*-KD triggers PH depletion and generates neoplastic tumors that persist after PH recovery.

a- PH immunostaining (in red) after no *ph*-KD (control), constant or transient *ph*-KD. Tissues were counterstained with DAPI (in blue). **b-d** DAPI staining (**b**); F-actin labeling showing apico-basal polarity (in red; **c**) and neuronal marker ELAV (differentiation, in purple, **d**) staining of EDs after no *ph*-KD (control), or a transient 24H *ph*-KD during first (L1), second (L2), early third or late third (L3) larval instar respectively. **e-** Survival curves of no-*w*-KD and no-*ph*-KD flies, maintained at 29°C only at the adult stage. **f-** Western blot showing PH protein from EDs subjected to no *ph*-KD (control) or transient *ph*-KD at late third instar larval stage in biological triplicate. **g-** GFP staining (in green) after no *ph*-KD (control), constant or transient *ph*-KD. Tissues were counterstained with DAPI (in blue). **h-** RT-qPCR experiments showing *gfp* expression level after no *ph*-KD (control), constant or transient *ph*-KD. Error bars represent the standard error of the mean (SEM) for three independent experiments. ** <0.01. **i-** ELAV (differentiation, in purple) staining after no *ph*-KD (control), constant or transient *ph*-KD. Tissues were counterstained with DAPI (in blue). Scale bars: 10 μ m (a, c, d, g, i), 100 μ m (b).



Extended Data Fig. 2: Transient *Psc*-KD phenocopies transient *ph*-KD.

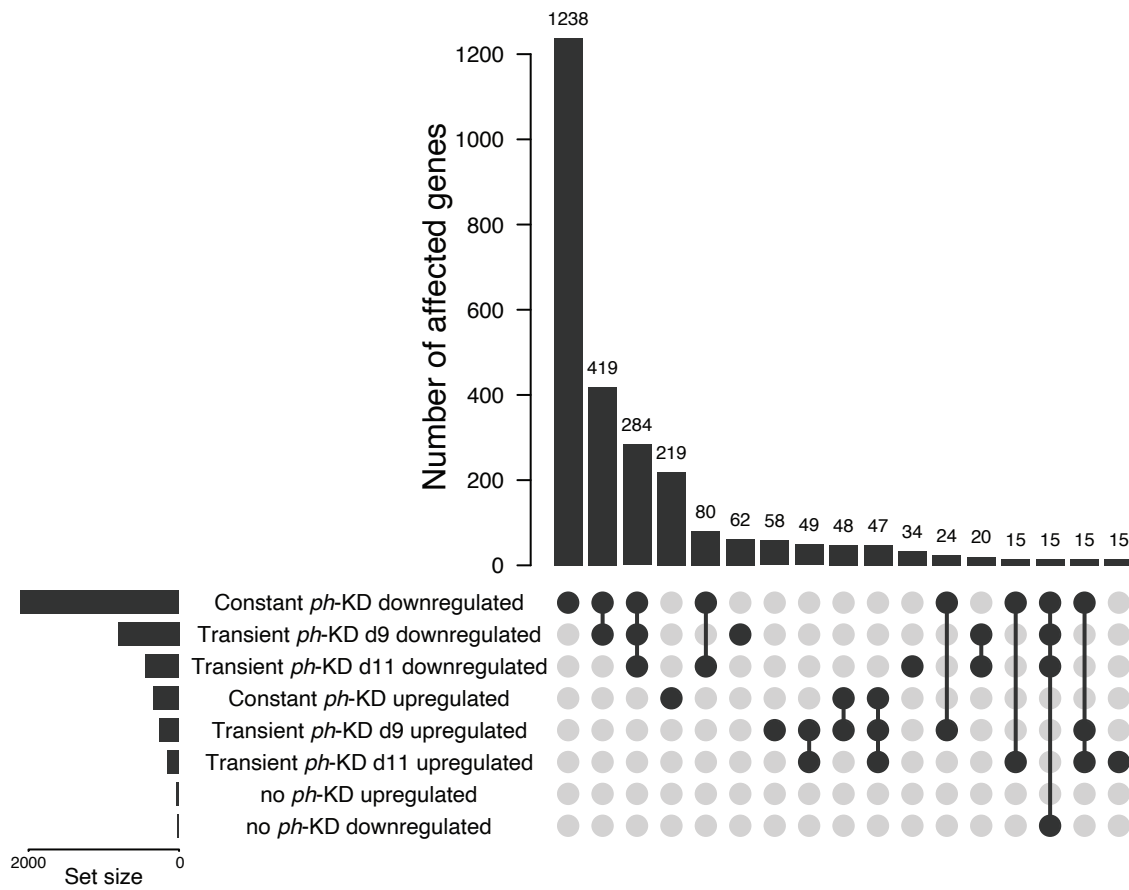
a-c DNA is stained with DAPI (blue) in EDs under the indicated conditions (**a**); F-actin staining (in red; **b**) showing apico-basal polarity and ELAV neuronal marker (differentiation, in purple, **c**) staining after no *Psc*-KD, constant or transient *Psc*-KD. **d**- Violin plots showing a comparative measure of ED size quantified as an overall DAPI staining area under no *Psc*-KD (control), constant or transient *Psc*-KD conditions (Wilcoxon test, * $p < 5e-2$, **** $p < 1e-5$).



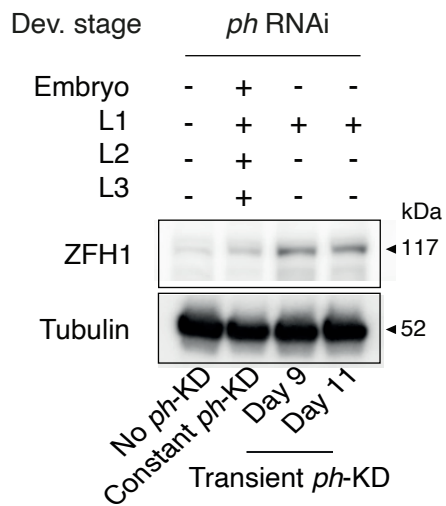
Extended Data Fig. 3: Transcriptomic analysis of the effect of a constant or transient *ph-KD*.

a- Principal Component Analysis (PCA) on the biological triplicate of gene expression data. Each dot represents one sample. PCA shows the high similarity between controls no *w-KD* and no *ph-KD*. **b-** Transcriptional fold changes upon PH depletion of genes known to be involved in tumorigenesis in *Drosophila*: PcG members, differentiation and Hox genes, tumor suppressor genes, genes involved in genome instability, cell cycle, apoptosis and growth control, members of major signaling pathways (Dpp, Decapentaplegic; EGFR, Epidermal Growth Factor Receptor; Notch; Hh, Hedgehog; Wg, Wingless; Hippo).

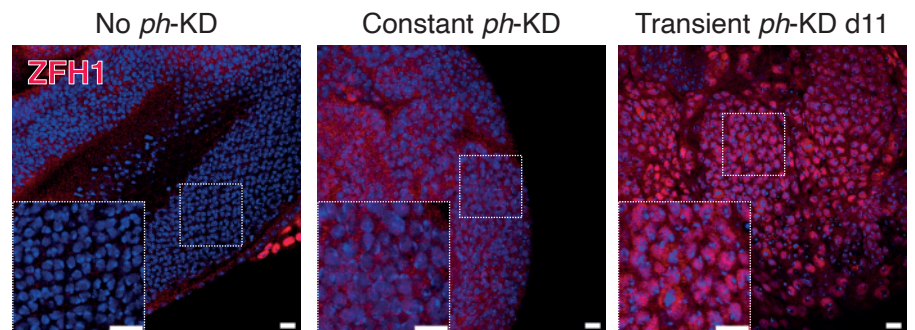
a



b



c



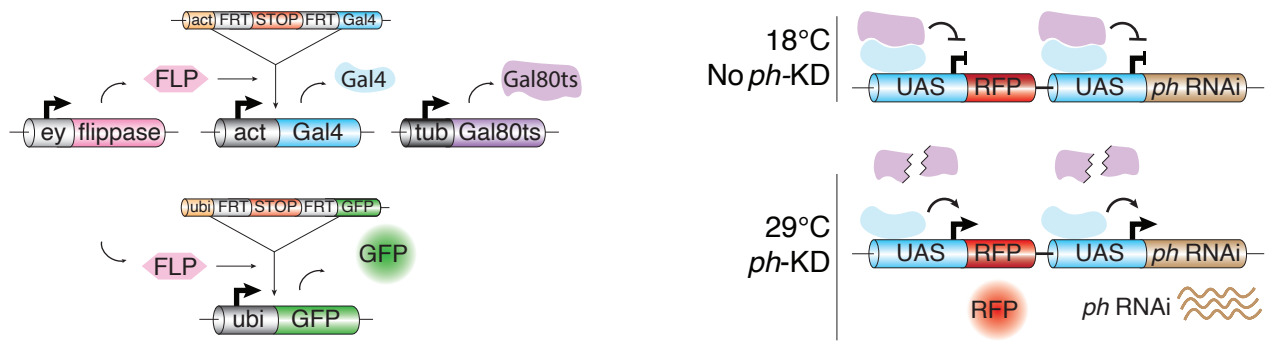
Extended Data Fig. 4: ZFH1 is included among irreversibly derepressed genes after transient PH depletion.

a- Representation of the number of overlaps for transcriptionally down-regulated or up-regulated genes between the indicated conditions. **b-** Western blot showing ZFH1 in EDs subjected to no *ph*-KD (control), constant or transient *ph*-KD. **c-** ZFH1 immunostaining (in red) after no *ph*-KD (control), constant or transient *ph*-KD. Tissues were counterstained with DAPI (in blue). Scale bars: 10 μ m.

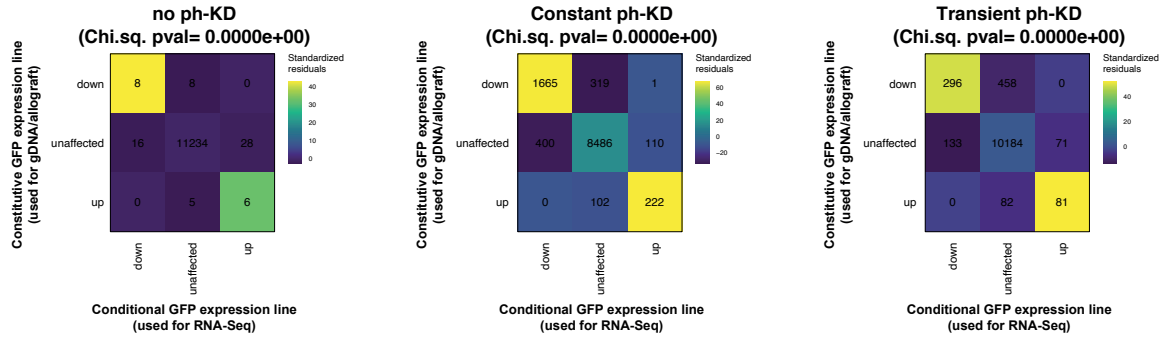
Extended Data Fig. 5: Similar recovery of chromatin landscape at reversible and irreversible genes.

a- Screenshot of PH ChIP-Seq, H2AK118Ub, H3K27me3 and H3K27Ac CUT&RUNs at the *zfh1* locus in no *ph*-KD (control) EDs. **b-** Proportion of PRC1-bound genes per cluster, using all unaffected genes as a reference. Overrepresentation of PRC1-bound genes was assessed using two tailed Fisher's exact test (N.S- not significant, * $pval < 5e-2$, ** $pval < 1e-2$, *** $pval < 1e-3$, **** $pval < 1e-5$). **c-** H3K27me3 quantification at gene bodies per cluster, further stratified for the presence or absence of PRC1 binding (in orange and grey, respectively). Wilcoxon test, N.S- not significant, * $pval < 5e-2$, ** $pval < 1e-2$, *** $pval < 1e-3$, **** $pval < 1e-5$. **d-** Average signal and quantification of H3K27me3 at the body of irreversible (pink) and reversible (green) genes (Wilcoxon test, N.S- Not significant). **e-** H2AK118Ub quantification at gene bodies per cluster, further stratified for the presence or absence of PRC1 binding (in orange and grey, respectively). Wilcoxon test, N.S- not significant, * $pval < 5e-2$, ** $pval < 1e-2$, *** $pval < 1e-3$, **** $pval < 1e-5$. **f-** Differential analysis of H2AK118Ub domains that are unaffected (grey) or exhibiting decreased (blue) or increased (orange) enrichment upon constant (left) or transient (right) PH depletion. **g-** Average signal and quantification of H2AK118Ub at the body of irreversible (pink) and reversible (green) genes (Wilcoxon test, N.S- Not significant). **h-** Quantification of PH enrichment at the promoter of irreversible (pink) and reversible (green) genes (Wilcoxon test, N.S- not significant, * $pval < 5e-2$, ** $pval < 1e-2$, *** $pval < 1e-3$, **** $pval < 1e-5$).

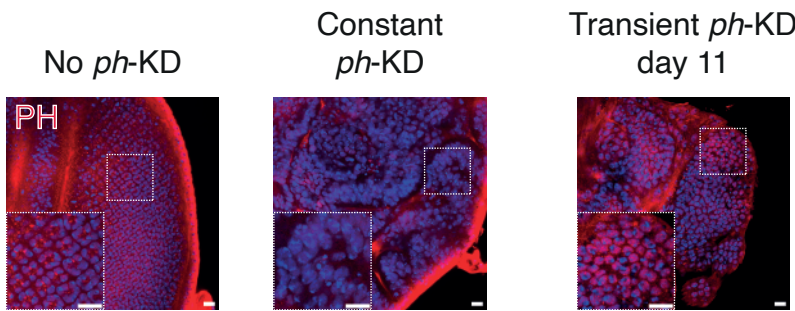
a



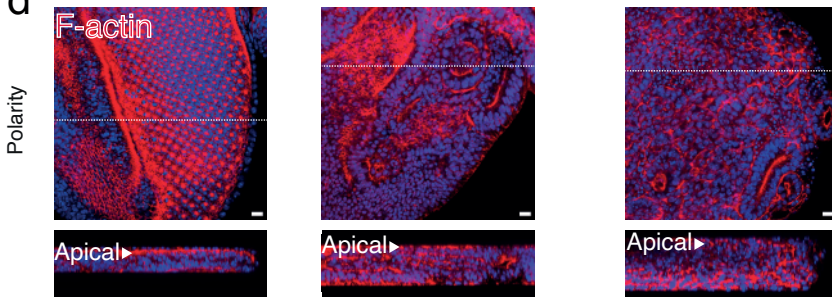
b



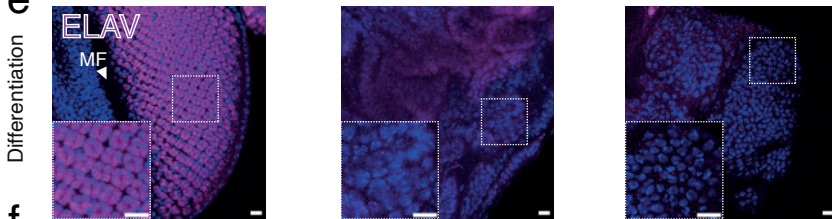
c



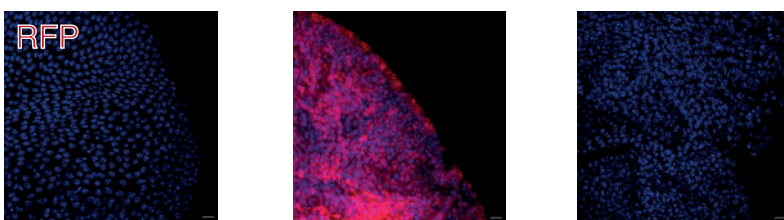
d



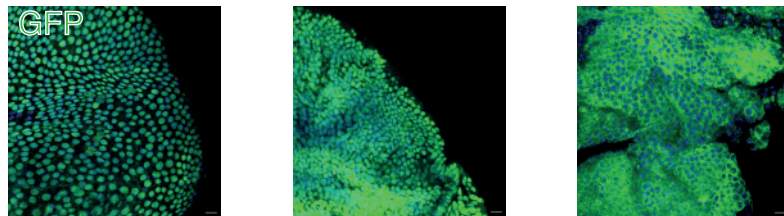
e



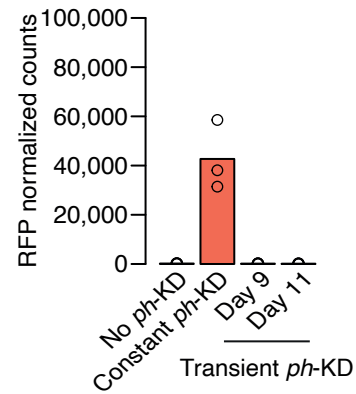
f



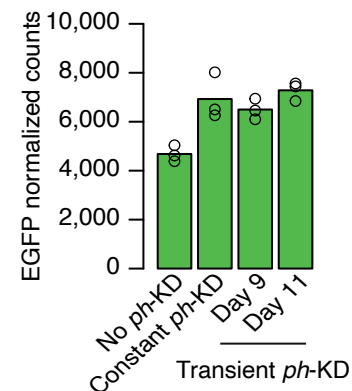
h



g



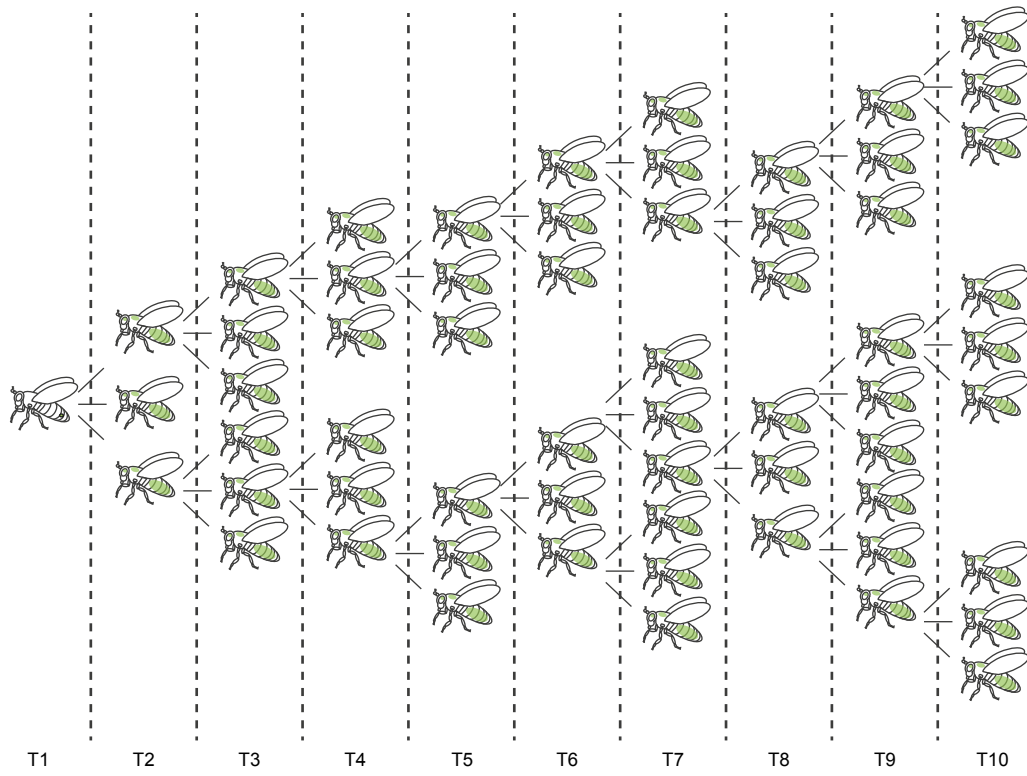
i



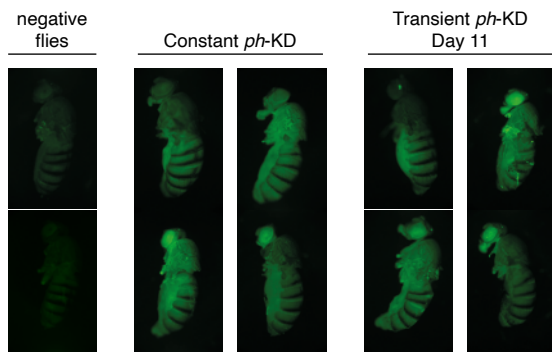
Extended Data Fig. 6: A conditional genetic tool allowing the long-term tracking of cells submitted to constant or transient PH depletion.

a- Schematic overview of the thermosensitive *ph*-RNAi genetic system used in allograft experiments. This fly line contains two fluorochromes: GFP which is ubiquitously expressed (Ubp63 promoter), in order to follow the tumor evolution through multiple transplantation cycles, while the RFP expression is a read-out of the *ph*-KD which depends, like *ph*-RNAi, on the binding of the yeast transactivator Gal4 to its UAS target sequences. At 18°C, neither *ph*-RNAi nor RFP are expressed because the Gal4 inhibitor Gal80^{ts} is active. At 29°C, the thermosensitive Gal80^{ts} is inactivated. Gal4 binds to UAS sequences allowing the expression of *ph*-RNAi and RFP. **b-** Chi-squared residuals showing the overlap between the transcriptomes of the conditional GFP line (used for RNA-Seq, Fig. 2) and the constitutive GFP line (developed for gDNA and allografts, Fig. 1 and 5). **c-** PH immunostaining (in red) after no *ph*-KD (control), constant or transient *ph*-KD. Tissues were counterstained with DAPI (in blue). **d-** F-actin staining (in red) showing apico-basal polarity after no *ph*-KD (control), constant or transient *ph*-KD. Tissues were counterstained with DAPI (in blue). **e-** ELAV staining (in purple) showing neuronal differentiation after no *ph*-KD (control), constant or transient *ph*-KD. Tissues were counterstained with DAPI (in blue). **f-** RFP immunostaining (in red) after no *ph*-KD (control), constant or transient *ph*-KD makes it possible to track the dynamics of *ph*-KD. Tissues were counterstained with DAPI (in blue). **g-** Normalized read counts of RFP mRNAs after no *ph*-KD (control), constant or transient *ph*-KD showing that transcriptional expression of RFP occurs at constant 29°C exposure but is reverted after a transient *ph*-KD. **h-** GFP staining (in green) after no *ph*-KD (control), constant or transient *ph*-KD. As expected reveal, GFP is expressed in all conditions. Tissues were counterstained with DAPI (in blue). **i-** Normalized GFP read counts after no *ph*-KD (control), constant or transient *ph*-KD showing constant expression of GFP in all conditions. Scale bars: 10 μm.

a



b

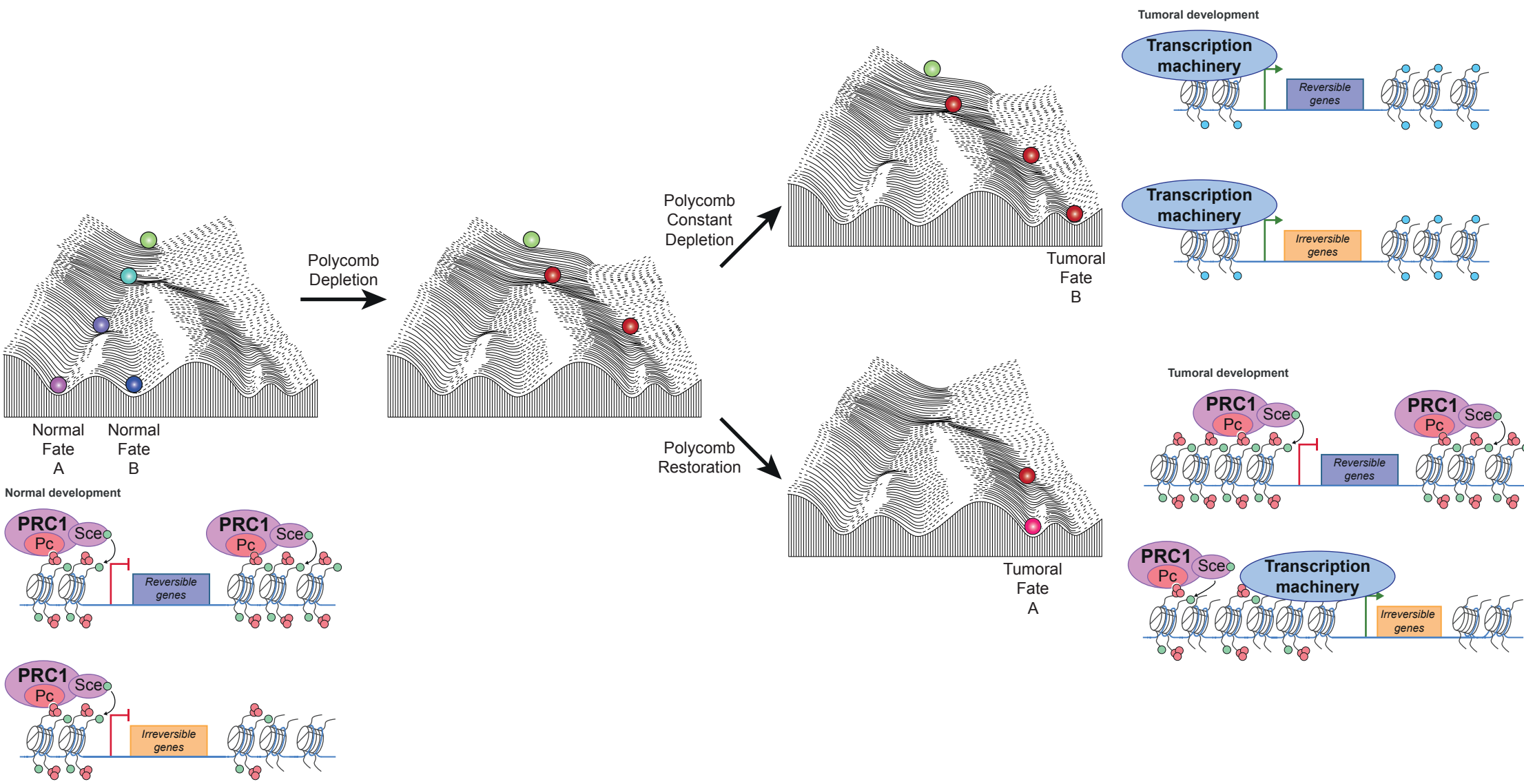


c

Number of flies (percentage)	T1					T5					T10				
	Eyes	Proboscis	Head	Thorax	Legs	Eyes	Proboscis	Head	Thorax	Legs	Eyes	Proboscis	Head	Thorax	Legs
Constant <i>ph</i> -KD	2 (2,99)	4 (5,97)	2 (2,99)	7 (10,45)	2 (2,99)	19 (9,22)	27 (13,11)	20 (9,71)	39 (18,93)	19 (9,22)	5 (7,46)	9 (38,68)	6 (8,96)	22 (32,84)	7 (10,45)
Transient <i>ph</i> -KD day 11	13 (15,85)	16 (19,51)	14 (17,07)	17 (20,73)	7 (8,54)	28 (11,70)	43 (22,87)	31 (16,49)	57 (30,32)	28 (14,89)	41 (15,89)	79 (30,62)	45 (17,44)	104 (40,31)	47 (18,22)

Extended Data Fig. 7: Comparative analysis of metastasis formation through serial transplantation of constantly PH- depleted tumors versus EICs.

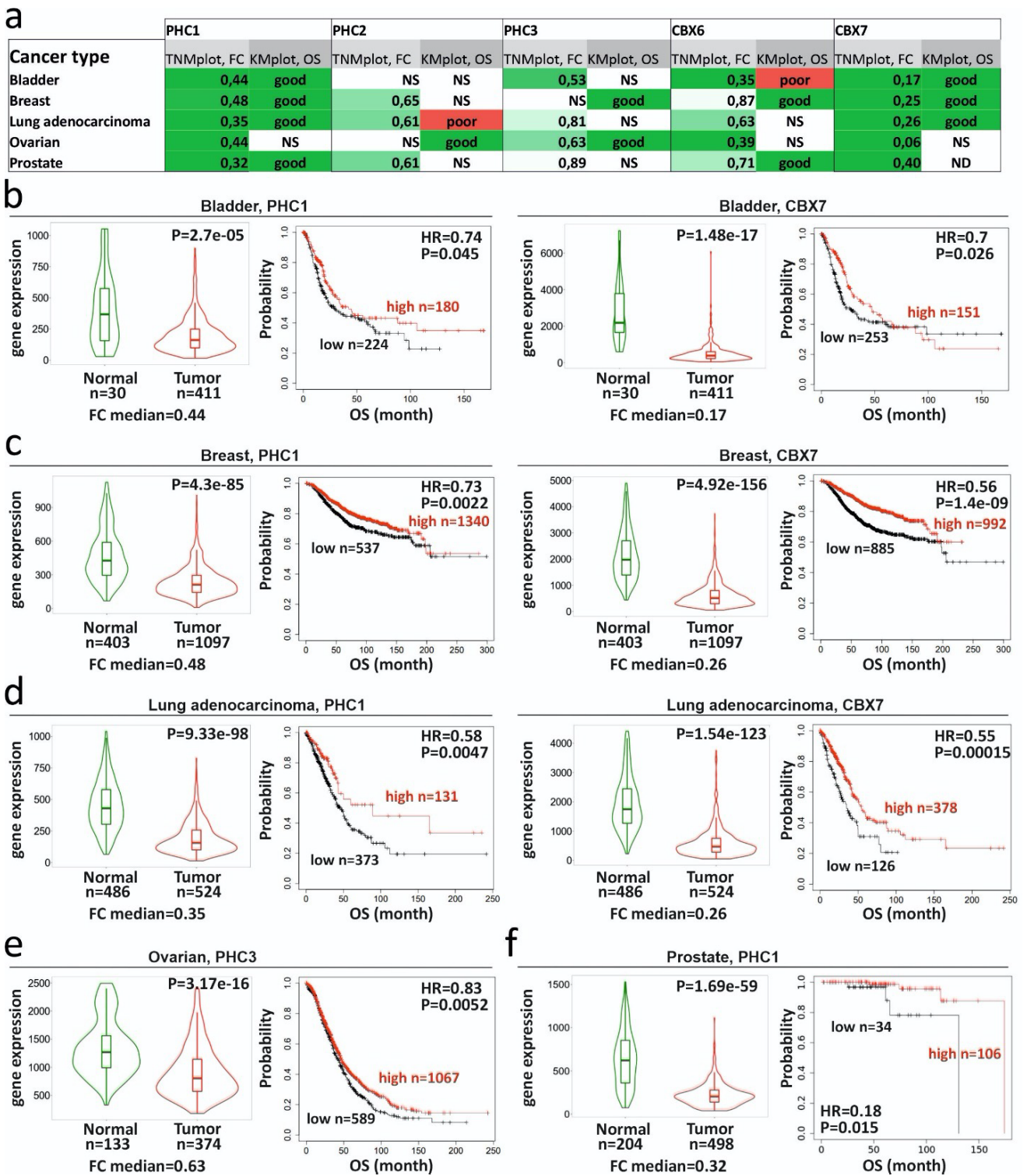
a- Tree representation of the allograft assay. A primary ED tumor derived from a constant or transient *ph*-KD is dissected from donor larvae and repeatedly allografted into the abdomen of a female host maintained at 18°C to prevent re-expression of *ph*-RNAi. Each injected fly is monitored every two days. When the host fly abdomen is completely filled with GFP positive cells, the host is dissected and tumor cells are again injected into multiple hosts. The procedure was repeated up to the T10 generation. **b-** Pictures of flies injected with grafts obtained from no *ph*-KD (control), constant or transient *ph*-KD conditions. The primary tumor can invade the abdomen and surrounding tissues. **c-** In order to score the frequency of metastases, injected flies were monitored twice a week and the appearance of metastasis in the thorax, head, proboscis, eyes and legs were noted for each generation. Values in the table represent the number and percentage (in brackets) of flies with metastases after transplantation of the 1st, 5th or 10th (T1, T5 and T10, respectively) generation of constant or transient *ph*-KD tumor transplants.



Extended Data Fig. 8

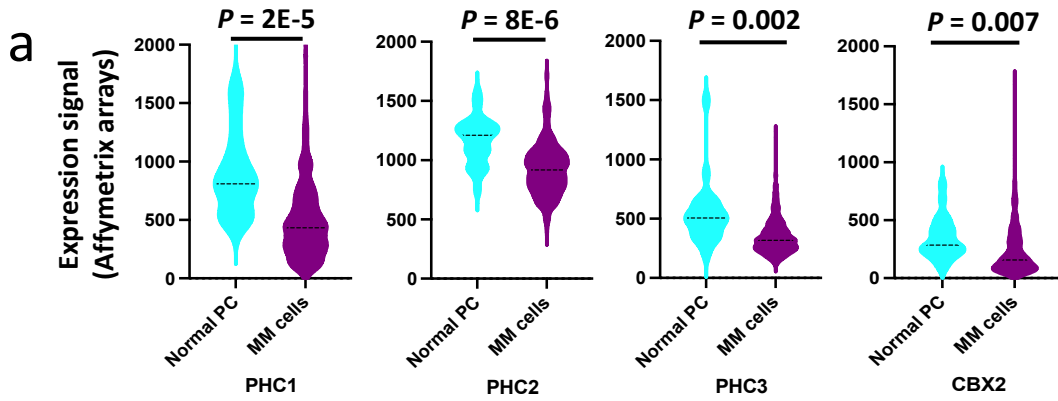
Extended Data Fig. 8: A model explaining the emergence of epigenetically induced cancers.

The model is based on the well-known Waddington landscape depicting a marble rolling down a slope with multiple choices of trajectories that depend on the hills and valleys encountered along their path. This scheme is a metaphor of the multiple possible cell fates that can arise from a single cell representing the zygote and is frequently used to signify that epigenetic inheritance contributes to the stable transmission of cell fates, once they are determined by intrinsic and extrinsic signals. In the context of this work, we posit that Polycomb components contribute to shaping the landscape and allow for multiple normal cell fates to be set up and transmitted through the developmental process. In normal development, cells (in green) on top of the hill, will go down the hill during differentiation in order to acquire normal fates. (left panel). Upon depletion of a Polycomb component, such as the PRC1 subunits PH or PSC, the landscape is modified (center panel). If depletion is stably maintained, the modified landscape forces cells to enter a pathway that is both aberrant and intrinsically stable, inducing cancer formation through loss of cell differentiation, loss of cell polarity and sustained proliferation (top right panel). If Polycomb protein levels are restored, the landscape goes back to its original shape. However, if restoration of the landscape occurs after the cells have already chosen an aberrant route (represented by the marble in the middle of the landscape), they can no longer find the healthy trajectory and will be obliged to choose among a limited set of possibilities within a diseased cell space. This ultimately can lead to the maintenance of tumoral phenotypes. In addition to the Waddington landscape panels, gene panels are added along the panels, representing a putative molecular explanation of the phenomenon described here. The chromatin and functional state of reversible and irreversible genes is shown in each condition. In the absence of Polycomb depletion, reversible genes are characterized by the presence of high levels of Polycomb (only PRC1 complexes are shown for simplicity) and by low levels or lack of transcription. Irreversible genes are also not or lowly expressed, but they contain less Polycomb in their chromatin. Removal of PRC1 subunits induces chromatin opening and derepression of both classes of genes. Upon restoration of Polycomb components however, a difference emerges. High levels of Polycomb binding might restore a silent chromatin state, whereas lower levels of Polycomb binding might be unable to restore repression at irreversible genes.



Extended Data Fig. 9: Examples of tumor suppressive role of canonical PRC1 core subunits in different cancer types.

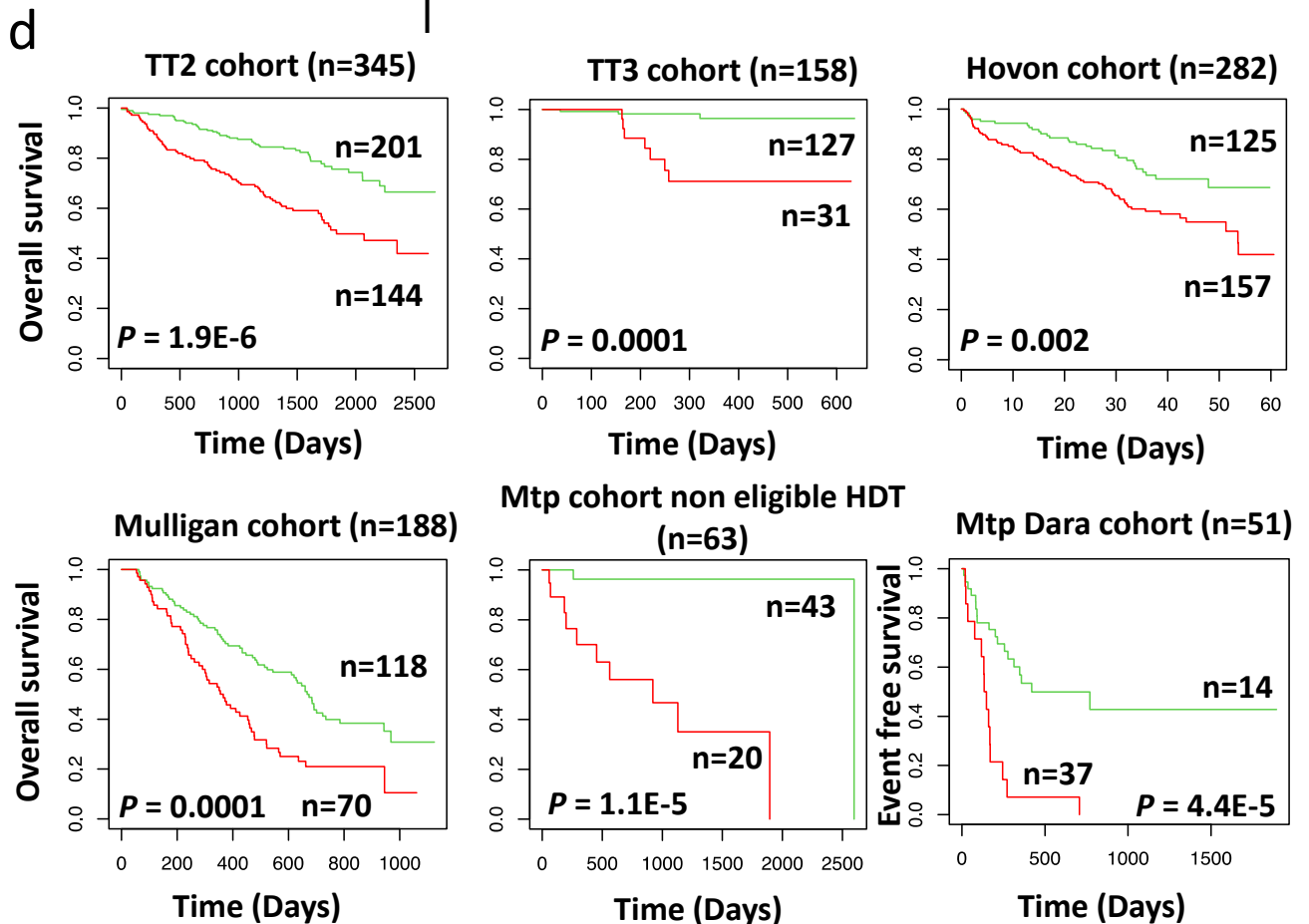
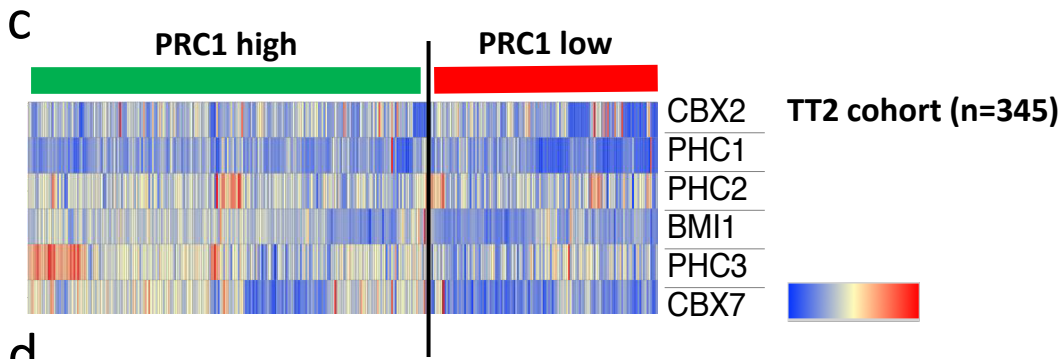
a- Clinical correlations for PRC1 in selected cancer types. Differential gene expression (TNMplot) and clinical prognosis Kaplan-Meier plot (KMplot) results are given for *PHC1*, *PHC2*, *PHC3*, *CBX6*, *CBX7* genes. TNMplot columns represent the differential gene expression analysis in tumor and matched normal tissues, which was performed using the <https://tnmplot.com/> online tool. FC median: Fold change median. Statistical significance was calculated by a Mann–Whitney U test and set at 0.01. NS – non-significant Mann–Whitney *p*-value. Green boxes indicate that gene expression is significantly lower in tumor tissues. KMplot columns show the analysis of correlation between the overall survival (OS) and the levels of gene expression. KMplot analysis was performed using the www.kmplot.com online tool. HR: hazard ratio. Statistical significance was calculated by a logrank *p*-value and was set at 0.05. NS – non-significant logrank *p*-values. Green boxes (“Good”) indicate cases in which high expression of PRC1 genes in tumors is associated with a better overall patient survival. **b-** Clinical prognosis for PHC1 (left) and CBX7 (right) in bladder cancer. For each gene the TNMplot (Violin plots, left panels) and KM plots (right panels) are shown. **c-** Clinical prognosis for PHC1 (left) and CBX7 (right) in breast cancer. For each gene the TNMplot (Violin plots, left panels) and KM plots (right panels) are shown. **d-** Clinical prognosis for PHC1 (left) and CBX7 (right) in lung adenocarcinoma. For each gene the TNMplot (Violin plots, left panels) and KM plots (right panels) are shown. **e-** Clinical prognosis for PHC3 (left) in ovarian and PHC1 (right) in prostate cancer. For each gene the TNMplot (Violin plots, left panels) and KM plots (right panels) are shown.



b

	TT2 cohort (n=345)	TT3 cohort (n=158)	Hovon cohort (n=282)	Mulligan cohort (n=188)	Montp cohort non eligible HDT (n=63)	Montp Dara cohort (n=51)
PHC1	good	good	good	NS	NS	good
PHC2	NS	NS	good	good	good	NS
PHC3	good	good	good	good	good	good
CBX2	good	good	NS	good	NS	NS
CBX7	good	good	good	good	good	good
BMI1	good	NS	good	NS	good	NS

High expression of PRC1 genes is associated with a better overall survival; $P \leq 0.05$.



Extended Data Fig. 10: A tumor suppressive role of PRC1 core subunits in Multiple Myeloma

a- *PHC1*, *PHC2*, *PHC3* and *CBX2* gene expression is significantly downregulated in malignant plasma cells (PCs) from patients with Multiple Myeloma (MM) compared to normal bone marrow PCs. Affymetrix U133 P gene expression profiles of purified bone marrow PC from 22 healthy donors and purified myeloma PCs from 345 previously untreated patients were compared using publicly available data (Gene Expression Omnibus, accession number GSE2658) from the University of Arkansas for Medical Sciences (UAMS, Little Rock, AR). Statistical difference was assayed using a Student t test. **b-** Prognostic value of the PRCA core components in MM. The prognostic value of *PHC1*, *PHC2*, *PHC3*, *CBX2*, *CBX7* et *BMI1* gene expression was analyzed in 6 independent cohorts of patients with MM using the Maxstat R function and Kaplan Meier survival curves as previously described. High expression of *PHC1*, *PHC2*, *PHC3*, *CBX2*, *CBX7* and *BMI1* was associated with significantly shorter overall survival in at least three independent cohorts of MM patients out of the six investigated (green color). The six cohorts included gene expression data of purified MM cells from the TT2, TT3 (accession number E-TABM-1138) and Hovon (accession number GSE19784) cohorts (345, 158 and 282 newly-diagnosed MM patients treated by high-dose melphalan and autologous hematopoietic stem cell transplantation); the Mulligan cohort (188 patients at relapse treated by proteasome inhibitor in monotherapy); the Mtp cohort non eligible to HDT (63 newly-diagnosed MM patients non eligible to high-dose melphalan and autologous hematopoietic stem cell transplantation) and the Mtp Dara cohort (51 patients at relapse treated by anti-CD38 monoclonal antibody (Daratumumab)). **c-** The prognostic information of *PHC1*, *PHC2*, *PHC3*, *CBX2*, *CBX7* and *BMI1* genes was combined. Patients of the TT2 cohort (n = 345) were ranked according to the increased value of the calculated score and a cluster was defined. **d-** In the TT2 cohort, a maximum difference in overall survival was obtained, using the Maxstat R package, splitting patients into high-risk for 144 patients with the lowest expression of PRC1 genes and low-risk group for the 201 patients with higher PRC1 gene expression. Using the same parameter of the TT2 training cohort, we validated the association between low expression of PRC1 genes and a poor outcome in the five other independent cohorts of patients with MM.

Extended data Table 1: Differential analyses and FPKMs of no *ph*-KD, transient *ph*-KD and constant *ph*-KD (fly line expressing a GFP marker during depletion of PH)

For each *ph*-KD/*w*-KD comparison, DESeq2 outputs are on separated sheets. “diff”column specifies whether the genes was consider as unaffected, up- or downregulated according to the criteria ($|\log_2FC| > 1$ & $p_{adj} < 0.05$). Last shee contains FPKMs for all conditions.

Extended data Table 2: Clustering of differentially expressed genes and recovery status

For each gene, its corresponding cluster is shown (Figure 2b, “Cluster”column). PH, H3K27me3 and H2AK118Ub binding are also specified and were used to define “Irreversible” and “reversible” genes (see material and methods, “Recovery” column). Finally, PH binding at the gene promoter is shown and was used to restrict “irreversible” and “reversible” gene sets (see material and methods, “PH bound promoter Recovery” column).

Extended data Table 3: Differential analyses and FPKMs of no *ph*-KD, transient *ph*-KD and constant *ph*-KD (fly line expressing GFP constitutively)

For each *ph*-KD/*w*-KD comparison, DESeq2 outputs are on separated sheets. “diff”column specifies whether the genes was consider as unaffected, up- or downregulated according to the criteria ($|\log_2FC| > 1$ & $p_{adj} < 0.05$). Last shee contains FPKMs for all conditions.