



HAL
open science

Cooperation between melanoma cell states promotes metastasis through heterotypic cluster formation

Nathaniel R Campbell, Anjali Rao, Miranda V Hunter, Magdalena K Sznurkowska, Luzia Briker, Maomao Zhang, Maayan Baron, Silja Heilmann, Maxime Deforet, Colin Kenny, et al.

► **To cite this version:**

Nathaniel R Campbell, Anjali Rao, Miranda V Hunter, Magdalena K Sznurkowska, Luzia Briker, et al.. Cooperation between melanoma cell states promotes metastasis through heterotypic cluster formation. *Developmental Cell*, 2021, 56 (20), pp.2808 - 2825.e10. 10.1016/j.devcel.2021.08.018 . hal-03772733

HAL Id: hal-03772733

<https://cnrs.hal.science/hal-03772733>

Submitted on 21 Sep 2022

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.

1 **Cooperation between melanoma cell states promotes metastasis through heterotypic cluster**
2 **formation**

3

4 Nathaniel R. Campbell^{1,2,3,†}, Anjali Rao⁴, Miranda Hunter³, Magda Sznurkowska⁵, Luzia Briker⁶,
5 Maomao Zhang³, Maayan Baron⁴, Silja Heilmann^{2,‡}, Maxime Deforet^{2,§}, Colin Kenny⁷, Lorenza
6 Ferretti⁶, Ting-Hsiang Huang³, Sarah Perlee³, Manik Garg⁸, Jérémie Nsengimana^{9,¶}, Massimo
7 Saini⁵, Emily Montal³, Mohita Tagore³, Julia Newton-Bishop⁹, Mark R. Middleton¹⁰, Pippa
8 Corrie¹¹, David J. Adams¹², Roy Rabbie^{11,12}, Nicola Aceto⁵, Mitchell P. Levesque⁶, Robert A.
9 Cornell⁷, Itai Yanai⁴, Joao B. Xavier^{2,*}, Richard M. White^{3,*}

10

11 ¹ Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program, New York, NY,
12 USA

13 ² Computational and Systems Biology, Memorial Sloan Kettering Cancer Center, New York,
14 NY, USA

15 ³ Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

16 ⁴ Institute for Computational Medicine, NYU School of Medicine, New York, NY, USA

17 ⁵ Cancer Metastasis Laboratory, Department of Biomedicine, University of Basel and University
18 Hospital Basel, Basel, Switzerland

19 ⁶ Department of Dermatology, University of Zurich Hospital, University of Zurich, Zurich,
20 Switzerland

21 ⁷ Department of Anatomy and Cell Biology, University of Iowa, Iowa City, IA, USA

22 ⁸ European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI),
23 Hinxton, Cambridgeshire, UK

24 ⁹ University of Leeds School of Medicine, Leeds, UK

25 ¹⁰ Oxford NIHR Biomedical Research Centre and Department of Oncology, University of
26 Oxford, Oxford, UK

27 ¹¹ Cambridge Cancer Centre, Cambridge University Hospitals NHS Foundation Trust,
28 Cambridge, UK

29 ¹² Experimental Cancer Genetics, The Wellcome Sanger Institute, Hinxton, Cambridgeshire, UK

30 † Present address: University of Illinois College of Medicine, Chicago, IL, USA

31 ‡ Present address: Novo Nordisk Foundation Center for Stem Cell Biology, University of
32 Copenhagen, 2200 Copenhagen N, Denmark

33 § Present address: Sorbonne Université, CNRS, Institut de Biologie Paris-Seine (IBPS),
34 Laboratoire Jean Perrin (LJP), F-75005, Paris, France

35 ¶ Present address: Biostatistics Research Group, Population Health Sciences Institute, Newcastle
36 University, Newcastle, UK

37 * Co-communicating co-senior authors: J.B.X. (xavierj@mskcc.org) and R.M.W.
38 (whiter@mskcc.org)

39

40 **SUMMARY**

41 Melanomas can have multiple co-existing cell states, including proliferative versus invasive
42 subpopulations that represent a “go or grow” tradeoff. Transcriptional profiling has revealed that
43 primary melanomas maintain both of these subpopulations but how they physically and
44 mechanistically interact is poorly understood. Here we used a zebrafish model of melanoma to
45 show that cells in the invasive state (INV cells) and cells in the proliferative state (PRO cells) form
46 spatially structured heterotypic clusters and cooperate in the seeding of metastasis, which

47 maintains cell state heterogeneity. We unexpectedly found that INV cells adhere tightly to each
48 other, and form clusters with a rim of PRO cells. Intravital imaging demonstrated cooperation
49 between these subpopulations, in which INV cells facilitate the spread of less metastatic PRO cells.
50 We identified the TFAP2 neural crest transcription factor as a master regulator of both clustering
51 and the PRO/INV states. In human melanomas we saw that low expression of TFAP2 is associated
52 with capacity for clustering, supporting that this mechanism is conserved in patients. Isolation of
53 clusters from patients with metastatic melanoma revealed a subset of patients with heterotypic
54 PRO-INV clusters, providing a key clinical correlate. Our data suggest a framework for the co-
55 existence of these two divergent cell populations, in which differing cell states form heterotypic
56 clusters that promote metastasis via cell-cell cooperation.

57

58 **INTRODUCTION**

59 **Cell state heterogeneity in cancer**

60 Tumors are heterogenous populations of cells that contain a variety of subpopulations differing
61 both through genetic and non-genetic mechanisms (Hinohara and Polyak, 2019). One such form
62 of heterogeneity is transcriptional. Numerous studies using bulk or single-cell transcriptomics have
63 demonstrated the existence of transcriptional subpopulations of cells, often referred to as cancer
64 cell states (Hinohara and Polyak, 2019; Hoek and Goding, 2010). The mechanisms that generate
65 the different cell states, and how those states interact with each other remains poorly understood.

66

67 **Melanomas have multiple co-existing cell states, including PRO/INV cells**

68 Melanoma has long been noted to exhibit a wide range of phenotypic properties such as
69 pigmentation and invasiveness (Houghton et al., 1987), which is related to underlying

70 transcriptional heterogeneity. Such heterogeneity was initially studied by bulk RNA
71 microarray (Bittner et al., 2000; Hoek et al., 2006; Widmer et al., 2012) and RNA-sequencing
72 technologies (Rambow et al., 2015; Verfaillie et al., 2015), but more recent single-cell RNA
73 sequencing (Tirosh et al., 2016; Wouters et al., 2020) has increased the granularity of these
74 distinctions. Increasing evidence points to at least four distinct cell states (Rambow et al., 2018;
75 Tsoi et al., 2018; Wouters et al., 2020), with the most consistently identified ones comprising a
76 proliferative (PRO) versus invasive (INV) cell state. Individual cells tend to be PRO or INV, but
77 not both (Hoek et al., 2008; Hoek et al., 2006; Rambow et al., 2015; Tirosh et al., 2016; Verfaillie
78 et al., 2015; Widmer et al., 2012)—a tradeoff reminiscent of the “grow or go”
79 hypothesis (Hatzikirou et al., 2010; Matus et al., 2015). The PRO vs. INV populations have been
80 tightly linked to the process of phenotype switching, a phenomenon in which cells can
81 bidirectionally move between these two PRO vs. INV extremes after induction by signals such as
82 Wnt5A, EDN3, hypoxia, inflammation or nutrient deprivation from the
83 microenvironment (Carreira et al., 2006; Eichhoff et al., 2010; Falletta et al., 2017; Hoek et al.,
84 2008; Hoek et al., 2006; Kim et al., 2017; Pinner et al., 2009; Weeraratna et al., 2002; Widmer et
85 al., 2012). The PRO vs. INV state is in part controlled by the melanocyte master transcription
86 factor MITF (Carreira et al., 2006; Eichhoff et al., 2010), with the PRO cells generally being
87 MITF^{HI} and INV cells being MITF^{LO}, although many other genes such as AXL have been linked
88 to these states (Tirosh et al., 2016; Verfaillie et al., 2015). Some data posit the existence of
89 biphenotypic cells, with individual cells having characteristics of both PRO and INV cells upon
90 deletion of *Smad7* (Tuncer et al., 2019). The extent to which these states phenotype switch, or
91 remain relatively fixed in their identities, has important implications in whether new therapies
92 should be targeting the plasticity itself or the states themselves.

93

94 **Functions of co-existing cell states in tumor evolution**

95 Despite evidence that these and other (Baron et al., 2020; Rambow et al., 2019; Rambow et al.,
96 2018; Tsoi et al., 2018) subpopulations exist in tumors, little is known about how these states co-
97 exist within the tumor, or whether they cooperate to promote tumorigenic phenotypes such as
98 metastasis. While some cell states (i.e. a neural crest-like population driven by RXRG or via
99 sensitivity to iron-dependent ferroptotic cell death) have been clearly linked to resistance to MAPK
100 inhibitor therapy (Rambow et al., 2018; Tsoi et al., 2018), the role of the PRO/INV populations
101 has been best studied in the context of metastasis. Analogous to an EMT-like process in epithelial
102 cancers, it was hypothesized that PRO cells could phenotype switch to a more INV state by
103 molecules such as Wnt5A (Weeraratna et al., 2002), and become more migratory and metastatic.
104 While this switching model likely explains metastases in some patients, it does not fully explain
105 why these cell state subpopulations seem to co-exist, albeit at different ratios, in nearly all patients
106 examined. Mixing PRO with INV cells (albeit from different patients) was shown to lead to
107 polyclonal metastatic seeding (Chapman et al., 2014; Rowling et al., 2020), raising the possibility
108 that different cell states, each with distinct phenotypes, might cooperate with each other to promote
109 phenotypes such as metastasis.

110

111 **Cooperation between cell states as a mechanism for metastasis**

112 Cooperation—a social behavior where one individual increases the fitness of another—is widely
113 studied in the contexts of ecology and evolution (Ågren et al., 2019; Archetti and Pienta, 2019;
114 Foster, 2011; Hauser et al., 2009; Korolev et al., 2014). Animal development is an awesome
115 display of cooperation between cells that culminates in an adult with trillions of cells differentiated

116 into tissues and organs, but sharing goals honed by natural selection: survival and reproduction of
117 the organism (Johnston, 2014). Cancer is a departure from development. It is a disease of
118 evolution (Nowell, 1976) where cancer cells interact with each other, competing for resources but
119 also possibly cooperating with each other to potentiate malignancy (Axelrod et al., 2006).
120 Nonetheless, the role of cooperation between cells in cancer remains understudied, especially *in*
121 *vivo*. In a Wnt1-driven mouse model of breast cancer both basal and luminal cell types emerge
122 during tumorigenesis; both populations are required for tumor growth, with Wnt1 produced by the
123 luminal cells supporting growth of the basal population (Cleary et al., 2014). This model
124 demonstrated how cooperation can provide a selective pressure for the maintenance of
125 heterogeneity within tumors. Along the same lines, experiments with heterotypic tumors where
126 subpopulations overexpressed factors previously implicated in tumor progression revealed a minor
127 subclone capable of driving enhanced proliferation of the entire tumor (Marusyk et al., 2014). This
128 clone acted by secreting IL-11 to stimulate vascular growth and reorganization of the extracellular
129 matrix. When this clone was combined with a clone expressing FIGF, the otherwise nonmetastatic
130 tumors gained the ability to metastasize. There is evidence that melanoma PRO and INV
131 populations from different patients can interact with one another in the formation of
132 metastasis (Chapman et al., 2014; Rowling et al., 2020); however, whether PRO and INV cells
133 from the same tumor physically interact and cooperate, and the mechanistic links between cell
134 state and cooperation, remain unknown.

135

136 Here, we show that coexisting PRO and INV cell states can form heterotypic clusters that cooperate
137 in metastasis. Circulating tumor cell clusters have long been recognized as a particularly potent
138 mechanism for metastasis (Aceto et al., 2014; Fidler, 1973; Glaves, 1983; Liotta et al., 1976; Long

139 et al., 2016; Luo et al., 2014; Mayhew and Glaves, 1984; Watanabe, 1954), and are strongly
140 associated with worse outcome. Using a transgenic zebrafish model of melanoma (Ceol et al.,
141 2011; Patton et al., 2005; White et al., 2011), we show that PRO and INV transcriptional states
142 spontaneously aggregate into spatially ordered clusters, with a rim of PRO cells surrounding a
143 dense core of INV cells. Unexpectedly, we find that the more INV cells express higher levels of
144 adhesion molecules, a finding recapitulated in human melanoma specimens. These heterotypic
145 clusters recapitulate developmental adhesive sorting, in which embryonic cells with differential
146 levels of adhesion proteins spontaneously form similar structures. Consistent with this notion, we
147 find that this cluster structure is regulated by the developmental neural crest transcription factor
148 TFAP2, which mediates the PRO vs. INV state and metastatic seeding capacity. While phenotype
149 switching is a likely mechanism of metastasis in some patients, our data provide an alternative
150 mechanism by which relatively fixed cell states physically cooperate to promote metastasis via
151 cooperative clustering of divergent cell states.

152

153 **RESULTS**

154 **Characterization of PRO/INV cell states**

155 To address the question of how PRO and INV populations interact, we utilized a zebrafish model
156 of melanoma that allows for longitudinal single cell analysis of these heterogeneous
157 subpopulations in metastasis formation (Cagan et al., 2019; Heilmann et al., 2015). From a
158 transgenic melanoma in a BRAF^{V600E};p53^{-/-} animal (Ceol et al., 2011; Kaufman et al., 2016; Patton
159 et al., 2005; White et al., 2011) we generated a low-passage zebrafish melanoma cell line,
160 ZMEL1 (Heilmann et al., 2015), and phenotyped multiple cultures to identify populations enriched
161 for either proliferative (ZMEL1-PRO) or invasive (ZMEL1-INV) phenotypes (Figure 1a).

162 Consistent with the previous characterization of PRO and INV states (Widmer et al., 2012), we
163 observed a small but consistent difference in net proliferation, and a more substantial motility
164 difference, between the ZMEL1-PRO and -INV states (Figure 1b-c, Figure S1a-b). To confirm
165 that this recapitulates human PRO and INV states, we performed RNA-sequencing analysis (RNA-
166 seq) on these two ZMEL1 populations and found a strong association between ZMEL1-INV and
167 -PRO states and published human INV and PRO gene signatures (Hoek et al., 2006; Tirosh et al.,
168 2016; Verfaillie et al., 2015; Widmer et al., 2012), respectively, with the INV signature from Hoek
169 et al. (Hoek et al., 2006) the top gene set (Figure 1d-e, Figure S1c, Supplementary Tables 1,6,7).

170

171 To compare the metastatic potential of ZMEL1-PRO and ZMEL1-INV—a far more stringent assay
172 than *in vitro* migration—we transplanted each population orthotopically into the subcutaneous
173 tissue of transparent *casper* zebrafish and followed their growth and metastasis by whole-fish
174 fluorescence microscopy (Heilmann et al., 2015) (Figure 1f-g, Figure S1d-g). Fish harboring
175 ZMEL1-INV tumors were significantly more likely to have distant metastases three days post-
176 transplant (3 dpt), particularly in the caudal region of the fish, an anatomical region relatively
177 resistant to metastasis (Heilmann et al., 2015). To investigate this difference in detail, we
178 transplanted each population intravenously in larval *casper* zebrafish where we followed the
179 seeding of metastases by confocal time-lapse microscopy. ZMEL1-INV cells extravasated more
180 effectively than ZMEL1-PRO cells within the first dpt (Figure 1h, Supplementary Video 1). To
181 quantify this difference, we tracked metastatic progression in similarly transplanted larval fish by
182 daily whole-fish imaging; ZMEL1-INV cells invaded into the caudal tissue in a significantly
183 higher proportion of fish at the experiment endpoint (4-6 dpt, Figure 1i). Since the cells were
184 injected intravenously, these findings implicate extravasation as a key step of metastatic spread

185 where INV cells are more effective than PRO cells. To characterize the growth dynamics of these
186 tumors and metastases, we co-transplanted PRO and INV cells in a 1:1 ratio and then isolated cells
187 from resultant primary tumors and metastases. While primary tumors showed similar engraftment
188 efficiency and maintained the initial 1:1 ratio of PRO and INV cells, metastases, initially seeded
189 more efficiently by INV cells, became dominated by PRO cells over time, highlighting the
190 increased *in vivo* proliferative potential of this population (Figure 1j). Given the agreement, both
191 phenotypic and transcriptomic, between these ZMEL1 populations and established PRO and INV
192 cell states, we utilized ZMEL1-PRO and ZMEL1-INV to further characterize the relationship of
193 these cell states.

194

195 **PRO/INV cells form heterotypic clusters**

196 To identify functional processes differentiating PRO and INV populations, we performed Gene
197 Ontology (GO) association analysis on our RNA-seq data. This analysis unexpectedly revealed a
198 strong association between the INV state and signatures of enhanced cell-cell adhesion, with many
199 adhesion genes upregulated (Figure 2a-b, Figure S2a-c, Supplementary Tables 6,7). This
200 association was surprising, as the classical model of cell invasion involves the loss of cell adhesion
201 and the gain of individual motility, the opposite of what we observed (Gupta et al., 2005; Li et al.,
202 2015; Padmanaban et al., 2019). To test this paradoxical finding, we utilized a three-dimensional
203 (3D) cluster formation assay in low-attachment plates, which allows cells to spontaneously
204 aggregate over 1-3 days with minimal contribution from cell proliferation. Under these conditions,
205 while the ZMEL1-PRO cells tended to stay as individual cells or small clusters, the ZMEL1-INV
206 population formed strikingly large, spherical clusters (Figure 2c-d, Figure S2d-e Supplementary
207 Video 2) in agreement with increased adhesive properties. To test whether the association between

208 invasiveness and cell clustering is a general feature of melanoma, we compared the INV signature
209 defined by Hoek et al. (Hoek et al., 2006) with that of the cell-cell adhesion genes most associated
210 with ZMEL1-INV in a panel of 56 melanoma cell lines available in the Cancer Cell Line
211 Encyclopedia (CCLE) (Ghandi et al., 2019) and 472 clinical melanoma samples from The Cancer
212 Genome Atlas (TCGA) (Cancer Genome Atlas Network, 2015). In both cohorts, the expression of
213 cell-cell adhesion genes correlated strongly with the INV cell state (Figure S2f-g). To validate this
214 finding functionally, we assayed the cluster-forming ability of a panel of nine human melanoma
215 cell lines. We observed a strong correlation between cluster formation and the INV state, consistent
216 with our zebrafish and transcriptomic findings (Figure 2e). Taken together, these results indicate
217 that melanomas that are invasive and metastatic tend to form cluster aggregates.

218

219 Individual primary patient melanomas comprise both PRO and INV subpopulations, and
220 disseminated metastases preserve that diversity (Tirosh et al., 2016), raising the question of
221 whether these subpopulations interact. Circulating tumor cell (CTC) clusters—comprised either of
222 tumor cells or tumor and microenvironmental cells—are increasingly recognized for their role in
223 promoting metastatic spread, facilitating diversity at metastatic sites (Aceto et al., 2014; Cheung
224 et al., 2016; Gkountela et al., 2019; Maddipati and Stanger, 2015; Szczerba et al., 2019). Because
225 the ZMEL1-PRO and -INV populations were isolated from a single primary tumor, we sought to
226 establish whether the two could interact in clusters. Differential labeling of the PRO vs. INV cells
227 revealed that the two cell states consistently generated co-clusters with a coherent spatial structure,
228 with ZMEL1-INV cells at the core and ZMEL1-PRO cells at the rim, reminiscent of developmental
229 cadherin sorting (Foty and Steinberg, 2005) (Figure 2f-g, Figure S2h-i, Supplementary Video 3).
230 Indeed, CRISPR/Cas9 induced deletion of *cdh1* in ZMEL1-INV partially phenocopied ZMEL1-

231 PRO, both decreasing the cluster size relative to INV clusters and causing spatial sorting of mixed
232 clusters (Figure 2h, Figure S3a-c). Deletion of *cdh1* alone was insufficient, however, to induce
233 changes in the metastatic rate of ZMEL1-INV (Figure S3d-e), suggesting that a broader set of
234 adhesion and invasion genes, and not only *cdh1*, underlies the observed phenotypes. This
235 stereotyped spatial organization of ZMEL1-PRO and -INV clusters in an assay that models the
236 behavior of CTC clusters (Cheung et al., 2016; Gkountela et al., 2019) motivated us to investigate
237 whether interaction between these two populations would play a role *in vivo*.

238

239 **PRO-INV heterotypic clusters cooperate in metastasis**

240 To assay PRO-INV interactions *in vivo* during metastatic dissemination, we transplanted a 1:1
241 mixture of the ZMEL1-PRO and -INV populations intravenously as single cells in zebrafish larvae
242 and followed them by confocal time-lapse microscopy (Figure 3a, Supplementary Video 4). We
243 observed that these transplanted single cell populations—as confirmed by microscopic inspection
244 and *in vitro* cluster formation assays (Figure S2d)—spontaneously formed intravascular tumor cell
245 clusters comprised of cells from one or both cell states, consistent both with previous intravital
246 imaging (Liu et al., 2018) and with the detection of CTC clusters heterogenous for the melanoma
247 marker S100 in the blood of patients (Khoja et al., 2014). More notably, we observed that nearly
248 half (11 out of 24) of ZMEL1-PRO extravasation events were co-extravasations of heterotypic
249 tumor cell clusters with ZMEL1-INV (Figure S4a). We detected a pattern of collective motility
250 suggesting that cells from the same heterotypic cluster extravasated collectively, with ZMEL1-
251 INV cells behaving as leader cells and ZMEL1-PRO as followers. These detailed observations
252 suggest that the PRO and INV states known to coexist in primary tumors can form heterotypic
253 clusters and interact in the seeding of metastases.

254

255 To test the consequences of PRO-INV interaction in a more physiological setting, we next assessed
256 their interaction after orthotopic transplantation in adult zebrafish. We transplanted primary tumors
257 of each population alone and as a 1:1 mixture and followed their growth and metastasis by whole-
258 fish fluorescence microscopy. In the group with mixed primary tumors, we observed a significantly
259 higher number of fish with polyclonal metastasis than would be expected based on the metastatic
260 rate of each subpopulation alone if they did not interact (Figure 3b-c, Figure S4b-c). Strikingly,
261 we also observed that the less metastatic ZMEL1-PRO population had an increased rate of caudal
262 metastases in mixed tumors compared to when it was transplanted alone (Figure 3d, Figure S4d-
263 h), showing that this population benefited from cell-cell interaction with the INV cells. Moreover,
264 the more metastatic ZMEL1-INV population did not become less metastatic (Figure 3e), meaning
265 that they did not pay a significant cost for giving this benefit to ZMEL1-PRO. This type of
266 interaction, where one individual (INV) increases the fitness of another (PRO), is formally defined
267 as cooperation (Foster, 2011) (see Figure S4n for a schematic representation of social interactions
268 including cooperation). To further characterize the benefit to the ZMEL1-PRO population, we
269 performed transplants at various PRO:INV mixing ratios (1:4, 4:1, and 9:1) consisting of
270 tdTomato-expressing ZMEL1-PRO cells mixed with EGFP-expressing ZMEL1 cells (either PRO
271 or INV) and then quantified the metastases (Figure S4i-j). This confirmed that when ZMEL1-INV
272 cells comprise at least half of the primary tumor, the PRO subpopulation has an increased rate of
273 metastasis, providing context to which patients may exhibit such metastatic interaction. We
274 observed a similar cooperative interaction *in vitro* in dual-color Boyden Chamber migration assays
275 (Figure 3f-g), confirming that ZMEL1-PRO invades better when mixed with -INV cells
276 independently of the microenvironment. Experiments with conditioned media further suggested

277 this interaction is mediated by direct cell-cell contact (Figure S4k-l) and not via soluble factors,
278 and there was no evidence for ZMEL1-INV clusters protecting ZMEL1-PRO from apoptosis
279 (Figure S4m). The *in vivo* cooperative benefit was only evident early in metastatic dissemination
280 (3 dpt vs 7 dpt, Figure S4g-h), indicating that this cooperation is particularly beneficial when both
281 primary tumors and the number of disseminating tumor cells are small. Taken together, these data
282 show that the formation of heterotypic clusters enables the collective extravasation of PRO and
283 INV, facilitating cooperation that preserves cell state diversity in early metastatic lesions (Figure
284 S4n) (Foster, 2011; Hauser et al., 2009).

285

286 **TFAP2 mediates the PRO/INV state and clustering**

287 Although several molecular mechanisms have been shown to regulate the PRO and INV state in
288 melanoma (including MITF, AXL, WNT5A and BRN2 and their up- and downstream regulatory
289 networks (Cheng et al., 2015; Falletta et al., 2017; Fane et al., 2019; Hoek et al., 2008; Hoek et al.,
290 2006; Pinner et al., 2009; Rambow et al., 2015; Rambow et al., 2019; Rambow et al., 2018;
291 Shakhova et al., 2015; Tirosh et al., 2016; Verfaillie et al., 2015; Weeraratna et al., 2002; Widmer
292 et al., 2012)), there is no known connection between these programs and the formation of tumor
293 cell clusters. To identify the mechanism regulating clustering in the INV population, we performed
294 motif analysis on 1 kilobase regions associated with genes differentially expressed between
295 ZMEL1-PRO and -INV cells (Figure 4a, Supplementary Table 3). The top motif whose target
296 genes were enriched in the PRO cells putatively binds the NFIC and TFAP2A transcription factors.
297 One of the TFAP2 family members itself, *tfap2e*, was also one of the most differentially expressed
298 genes between the PRO and INV cells, with its expression being over 100-fold higher in the PRO
299 versus INV cells (Figure S5a, Supplementary Table 1). The TFAP2 family of transcription factors

300 plays essential roles in neural crest and melanocyte cell fate during development (de Croze et al.,
301 2011; Hoffman et al., 2007; Kaufman et al., 2016; Li and Cornell, 2007; Luo et al., 2002; Seberg
302 et al., 2017a; Seberg et al., 2017b; Van Otterloo et al., 2010) and has been implicated as part of a
303 regulatory network promoting the PRO state (Hoek et al., 2006; Rambow et al., 2015; Tirosh et
304 al., 2016; Verfaillie et al., 2015). This raised the hypothesis that TFAP2 was acting as a master
305 regulator of the clustering phenotype observed in the INV population. To test this, we performed
306 RNA-seq of ZMEL1 cells in 3D (clustered) culture, and asked which genes were differentially
307 expressed in 3D compared to 2D (non-clustered) conditions (Figure 4b, Figure S5b,
308 Supplementary Tables 2,4). In both PRO and INV, we again found enrichment of a motif that
309 binds TFAP2—specifically, TFAP2E—when looking at up- and downregulated genes together.
310 This is consistent with the known redundancy of *tfap2a* and *tfap2e* in zebrafish (Van Otterloo et
311 al., 2010), and highly suggestive of a role for TFAP2 in mediating clustering.

312

313 We next sought to test whether TFAP2 plays a functional role in melanoma cluster formation and
314 metastasis. We performed CRISPR/Cas9 deletion of *tfap2a* and *tfap2e* in ZMEL1-PRO, which
315 typically forms poor clusters, and found a significant increase in clustering only in the context of
316 *tfap2a/e* double knockout (Figure 4c-d, Figure S5c-e). We also found that the *tfap2a/e* knockout
317 compared with a non-targeting control had a small but reproducible decrease in cell proliferation,
318 along with an increase in the persistence of migration (Figure S5f-i), consistent with the phenotype
319 differences between the INV and PRO populations. We next wanted to determine whether this
320 phenotypic switch mediated by TFAP2 translated to an *in vivo* effect on metastasis. We
321 orthotopically transplanted control or *tfap2a/e* knockout cells into adult *casper* fish and measured
322 both primary tumor growth and metastatic dissemination. The *tfap2a/e* knockout cells formed

323 primary tumors that grew significantly slower than controls (Figure 4e), which was expected from
324 their slower *in vitro* proliferation. Despite this decrease in primary tumor growth, we found similar
325 rates of overall and caudal metastasis, suggesting that loss of *tfap2a/e* induces a higher proportion
326 of cells to metastasize (Figure S5j-k). To test this idea more directly, we assessed the effect of
327 *tfap2a/e* on metastasis in a proliferation-independent assay by intravenous transplant. Time lapse
328 confocal microscopy revealed that loss of *tfap2a/e* led to metastatic extravasation in a significantly
329 higher proportion of fish (Figure 4f-g), consistent with a report that TFAP2A overexpression in
330 human cells slows metastatic spread (Huang et al., 1998). Taken together, these data suggest that
331 TFAP2 is not only a major regulator of the PRO vs. INV cell state, but that it also controls tumor
332 cell clustering and regulates metastasis via an effect on extravasation.

333

334 **TFAP2 correlates with clustering in human melanoma**

335 We next wanted to determine whether the effects of TFAP2 we observed in the zebrafish were
336 conserved in human melanoma. TFAP2A is a member of several gene expression profiles
337 describing the proliferative state (Rambow et al., 2015; Tirosh et al., 2016; Verfaillie et al., 2015),
338 including the Hoek et al. set (Hoek et al., 2006), and is critical for melanoma cell proliferation
339 (Figure S6a). Consistent with this, increased expression in primary tumors of either TFAP2A or
340 genes associated with the PRO state is associated with worse clinical outcomes in two large
341 independent clinical cohorts (Figure S6b-g), likely reflecting the known prognostic effects of
342 mitotic rate and primary tumor size in melanoma TNM staging (Gershenwald et al., 2017;
343 Thompson et al., 2011). We next examined TFAP2A expression in a panel of 56 melanoma cell
344 lines (CCLE) (Ghandi et al., 2019) and 472 clinical melanoma samples (TCGA) (Cancer Genome
345 Atlas Network, 2015), and asked how this correlated with their PRO/INV signatures defined by

346 Hoek et al. (Hoek et al., 2006). In both cohorts, we confirmed that the PRO and INV states were
347 strongly anti-correlated. Samples with higher TFAP2A expression exhibited a more PRO gene
348 signature, and conversely, samples with lower TFAP2A expression exhibited a more INV
349 signature (Figure 5a, Figure S6h). As expected based on these results, TFAP2A expression was
350 well correlated with MITF expression in these cohorts (Figure S6i-j). Further, we asked whether
351 the association between TFAP2A expression and the PRO/INV signatures was maintained at the
352 level of single cells. We analyzed available single cell RNA-seq data across a panel of 23 human
353 melanoma patients (Jerby-Arnon et al., 2018), and found a similar relationship: individual cells
354 with high TFAP2A tend to have a higher PRO score, whereas cells with low TFAP2A tend to have
355 a higher INV score (Figure S6k). Within the TCGA dataset, tumor samples collected from primary
356 sites had higher levels of TFAP2A compared to metastatic lesions despite similar expression of
357 pan-melanoma markers (Figure 5b, Figure S6l-m), in agreement with a prior report (Tellez et al.,
358 2007). Further, in the two patients for which paired primary and metastatic samples were available,
359 TFAP2A expression was lower in the metastatic lesion. A direct measurement of the relative ratio
360 of TFAP2^{HI} to TFAP2^{LO} cells in the tumors, and its correlation with patient prognosis will await
361 future longitudinal prospective single cell analysis. Next, we examined TFAP2A expression in the
362 panel of human melanoma cell lines used in Figure 2e and found that cluster-forming lines had
363 lower TFAP2A expression than non-clustering lines (Figure S6n-o). To test this association across
364 cells that better preserve the heterogeneity observed clinically in melanoma, we examined a panel
365 of four short-term human melanoma cultures (Raaijmakers et al., 2015). Cluster formation
366 correlated strongly with lower expression of TFAP2A (Figure 5c), consistent with our observation
367 that TFAP2 loss drives melanoma clustering. Collectively, our data confirm that the association

368 we had discovered in zebrafish—between TFAP2, the PRO/INV state and tumor cell clustering—
369 also occurs in human melanoma.

370

371 **TFAP2 regulates genes associated with metastasis and cell-cell adhesion**

372 To gain further insight into the mechanism by which TFAP2 regulates melanoma phenotypes, we
373 performed RNA-seq of the *tfap2a/e* knockout cells versus controls. We first validated that the
374 *tfap2a/e* knockout recapitulated the observed differences between ZMEL1-PRO and -INV by
375 performing gene set association analysis (GSAA) using the gene sets that had passed false
376 discovery cutoff (FDR < 0.05) in our ZMEL1-INV vs. -PRO RNA-seq analysis. We observed a
377 high concordance in the top dysregulated pathways—including multiple INV and GO adhesion
378 gene sets associated with TFAP2 loss—confirming that TFAP2 regulates pathways distinguishing
379 ZMEL1-PRO and -INV (Figure 5d-e, Supplementary Tables 5,6,7). Specific genes upregulated
380 upon TFAP2 loss and associated with either the INV state or adhesion include several with known
381 functions in melanoma metastasis (Figure 5f, Figure S7a; e.g. TGFBI (Lauden et al., 2014),
382 VEGFC (Streit and Detmar, 2003), CTGF (Finger et al., 2014), and CDH2 (Mrozik et al., 2018)).
383 In order to understand the mechanism by which TFAP2 regulates PRO/INV state and cell-cell
384 adhesion, we performed TFAP2A CUT&RUN (Cleavage Under Targets and Release Using
385 Nuclease) in SKMEL28 cells, the human melanoma cell line with the highest expression of PRO-
386 state genes out of those we characterized. This allowed us to understand the genes bound by
387 TFAP2A in melanoma (Rambow et al., 2015; Seberg et al., 2017b). Consistent with the known
388 roles of TFAP2 as both a transcriptional activator and repressor (Ren and Liao, 2001; Seberg et
389 al., 2017b), we observed significant enrichment for TFAP2A peaks in genes that are upregulated
390 upon *tfap2a/e* knockout in ZMEL1-PRO cells (Figure 5f asterisks, Figure S7b-d; e.g. TGFBI,

391 CDH2), suggesting it acts as a repressor of those loci. We did not observe evidence of a stress
392 response or changes in *mitfa* expression resulting from *tfap2a/e* knockout (Supplementary Table
393 7), lending further support to a model of direct regulation by TFAP2. Taken together, these data
394 highlight the direct and pleiotropic effects of TFAP2 loss on metastatic spread, further confirming
395 a role for TFAP2 in cell state and suggesting downstream mediators.

396

397 **PRO-INV heterotypic CTC clusters exist in the blood of melanoma patients**

398 To further elucidate the translational relevance of our findings and to better characterize the role
399 of PRO-INV heterotypic CTC clusters in metastatic dissemination, we isolated CTC clusters from
400 the blood of patients with metastatic melanoma (Figure 6a). We sampled the peripheral blood of
401 nine patients with advanced metastatic melanoma and performed microfluidics-based capture of
402 CTC clusters (Xu et al., 2015). We were able to isolate CTC clusters from four of these nine
403 patients (44%), from which we isolated a total of 32 CTC clusters, consistent with prior rates of
404 cluster detection (Khoja et al.; Long et al.; Luo et al.; Ruiz et al.; Sarioglu et al., 2015). We
405 characterized the cell states present in these clusters by staining with antibodies against TFAP2A
406 for PRO cells, SOX9 for INV cells, and CD45 to exclude immune cells (Figure 6b, Figure S8).
407 Overall, we found that 19% of these clusters were heterotypic, composed of a mixture of PRO and
408 INV cells (Figure 6c). As expected, this result varied across patients, with some patients having
409 no clusters and the others having a mix of homotypic and heterotypic clusters (Figure 6d,
410 Supplementary Table 10). These results provide important validation that the heterotypic clusters
411 we identified in zebrafish also occur clinically in a subset of patients with metastatic melanoma.

412

413 **Longitudinal single-cell RNA-seq reveals stability of PRO but not INV state**

414 The above data suggest a model in which PRO and INV cell clusters, regulated by TFAP2, form
415 the unit of initial metastatic seeding in certain patients. However, once seeding has occurred, it is
416 still possible that either of these cell states can undergo phenotype switching and contribute to
417 metastatic outgrowth. This possibility was suggested by our finding that metastases tend to become
418 dominated by PRO cells over time (Figure 1j). To test this more formally, we conducted a large-
419 scale longitudinal analysis of cell state at the single cell level, interrogating the effects of cell-cell
420 interaction, tumor formation, and metastasis. We performed single-cell RNA-seq of over 40,000
421 ZMEL1 cells from both the PRO and INV cell states across four different conditions: (1) *in vitro*
422 individual culture; (2) *in vitro* co-culture; (3) *in vivo* primary tumors; and (4) *in vivo* metastatic
423 lesions (Figure 7a). Strikingly, ZMEL1-PRO and -INV subpopulations were highly pure *in vitro*
424 and—despite substantial gene expression changes associated with the dramatic
425 microenvironmental pressures *in vivo*—remained largely discrete throughout all conditions
426 (Figure 7b). In order to quantify the stability of the two populations, we calculated PRO and INV
427 scores for each cell based on gene sets derived from ZMEL1 bulk RNA-seq, and trained a classifier
428 based on *in vitro* individual culture samples (Figure 7c-d). Consistent with our results from
429 conditioned media experiments (Figure S4k-l), we observed very little effect of co-culture upon
430 the transcriptomes of ZMEL1-PRO and -INV cells, with both populations remaining more than
431 99% pure. Strikingly, a fraction of INV cells in tumors, especially from the metastases, upregulated
432 PRO-state genes, increasingly occupying a PRO/INV double-positive state. This is in contrast to
433 PRO cells, which remained stable in the PRO state, and is consistent with prior *in vivo* evidence
434 that an INV-to-PRO switch is favored (Pinner et al., 2009). Further validating a role for TFAP2 as
435 a master regulator of melanoma cell state, we found that ZMEL1-INV cells that gained a PRO-
436 like gene expression program also reactivated *tfap2e* (Figure 7e, Figure S9a). Overall, these data

437 support a model of cooperation whereby clusters comprised of distinct PRO and INV
438 subpopulations promote co-metastatic seeding, and metastatic outgrowth is increasingly
439 dominated by PRO-like cells.

440

441 To interrogate signaling pathways that could be important for PRO-INV interaction *in vivo*, we
442 analyzed our single-cell data using CellPhoneDB, which allowed us to identify ligand-receptor
443 pairs that were significantly enriched between PRO and INV cells (Efremova et al., 2020) (Figure
444 S9b-c). Among the most enriched candidate receptor-ligand pairings between the PRO/INV cells
445 *in vivo* were IGF2-IGF2R and VEGFA-EFNB2. The IGF pathway in particular is especially
446 interesting, as IGF signaling has previously been shown by us and others to be an important
447 regulator of melanoma cell growth (Zhang et al., 2018). While it has traditionally been assumed
448 that IGF ligands come solely from the stroma, these data suggest that the melanoma cells
449 themselves may be one source of such ligands. In addition, a longstanding observation in the
450 melanoma literature has been the phenomenon of vasculogenic mimicry, in which tumor cells take
451 on characteristics of blood vessels, and this has been shown to be mediated in part by the VEGF
452 axis. Future experiments aimed at disrupting these pathways will be important to functionally
453 confirm their effect on metastatic cooperation.

454

455 **DISCUSSION**

456 Both individual and collective mechanisms of metastasis can occur in melanoma (Long et al.,
457 2016) and other cancers (Pearson, 2019; Reichert et al., 2018). Phenotype switching between PRO
458 and INV states has long been postulated to be a mechanism for individual seeding of metastasis in
459 melanoma (Hoek et al., 2008; Kim et al., 2017; Pinner et al., 2009; Vandamme and Berx, 2014).

460 Separately, circulating tumor cell (CTC) clusters, a mode of collective metastasis, have been
461 shown to have increased metastatic potential (Aceto et al., 2014; Cheung et al., 2016), and patients
462 with detected CTC clusters have worse clinical outcomes (Giuliano et al., 2018; Long et al., 2016).
463 Cooperation has previously been reported both in epithelial cancers (Celià-Terrassa et al., 2012;
464 Neelakantan et al., 2017; Tsuji et al., 2009) and between melanoma PRO and INV states in the
465 context of primary tumor collective cell invasion (Chapman et al., 2014) and metastatic
466 tropism (Rowling et al., 2020), but the mechanisms that explain the relationship between the
467 PRO/INV states and cooperative metastasis have remained unknown. We provide for the first time
468 a clear mechanism that explains how these two subpopulations, which coexist in the primary
469 tumor, cooperate in metastasis formation. We find that PRO and INV cells form heterotypic
470 clusters which are controlled by the neural crest transcription factor TFAP2, and provide direct
471 evidence of the presence of heterotypic PRO-INV CTC clusters in a subset of melanoma patients.
472 As far as we know, this is the first time anyone has identified clusters of PRO-INV heterotypic
473 cancer cells in melanoma patients, providing an important translational link to our mechanistic
474 work in the zebrafish. Our data on heterotypic CTC clusters in patients are consistent with the
475 hypothesis that in individual patients, either individual or collective migration may predominate.
476 This is further supported by reported melanoma CTC cluster detection rates between 2 and 55
477 percent (Khoja et al.; Long et al.; Luo et al.; Ruiz et al.; Sarioglu et al., 2015), and the observation
478 that certain melanoma patients exhibit polyclonal metastatic seeding (Rabbie et al., 2019; Sanborn
479 et al., 2015). While the phenotype switching model predicts dynamic switching of individual cells
480 between PRO and INV states (analogous to an epithelial-to-mesenchymal transition) as a
481 necessary feature of individual metastasis (Hoek et al., 2008; Kim et al., 2017; Pinner et al., 2009;
482 Vandamme and Berx, 2014), our finding that PRO and INV can cooperate while remaining as

483 distinct phenotypic populations suggests that tumors can preserve diversity during initial
484 metastatic seeding without the need for large-scale cell state switching on a rapid time scale.

485

486 Our data do not exclude the possibility that phenotype switching, or plasticity, is an operative
487 mechanism that promotes metastasis. Our single cell analysis of metastatic outgrowth
488 demonstrates that INV cells, once they arrive, can still switch to a double-positive PRO/INV state,
489 indicating that phenotype switching in the INV to PRO direction may be important after initial
490 seeding. Instead, our data indicate that dynamic switching, at least on a short time scale, is not an
491 absolute requirement. It is likely that in patients—in which tumors are extraordinarily
492 heterogeneous at both genetic and epigenetic levels—both mechanisms (fixed cell state and plastic
493 cell states) can co-exist and may both be important. Recent data suggest that within tumors, there
494 may be “high plasticity” cell states which are more prone to switching compared to other
495 cells (Marjanovic et al., 2020). It is likely that the relative balance between fixed versus plastic
496 cell states is governed by the epigenetic state of the cell, since chromatin-related proteins such as
497 JARID1 are known to be involved in determination of cell state (Harmeyer et al., 2017; Roesch et
498 al., 2010). A major unanswered question, however, is what might be the molecular driver of such
499 plasticity. Going from the PRO to INV state is linked to molecules such as WNT5A (Dissanayake
500 et al., 2007; Weeraratna et al., 2002), which drives subsequent metastatic ability, and could act in
501 both paracrine and autocrine ways. Conversely, fewer molecules driving the INV to PRO direction
502 are known. We previously demonstrated a role for EDN3 in this process (Kim et al., 2017), and
503 given the known role of endothelin signaling in melanocyte and melanoma proliferation, this is
504 likely one such factor. One important consideration is whether one direction (i.e. PRO-to-INV
505 versus INV-to-PRO) is favored over the other, which would be enabled by monitoring of cell states

506 in real time as they traverse the bloodstream. While obtaining such samples from patients would
507 be ideal, animal models might provide insights, although even in those models such experiments
508 are still technically challenging. Previous work along this line has suggested that INV cells marked
509 by BRN2 may be biased towards switching more readily to the PRO state, which would be
510 consistent with our single-cell data (Pinner et al., 2009). Defining the range of mechanisms that
511 mediate plasticity, and whether some cells are more easily switched than others, remains an
512 important area for future exploration.

513

514 Our data demonstrate that cell cluster formation driven by TFAP2 loss is a pro-metastatic feature
515 of INV cells, with pleiotropic increases in cell-cell adhesion and cell clustering enabling
516 cooperation with PRO cells. Further, we demonstrate the functional role of TFAP2 in regulating
517 cell state and clustering. TFAP2 is known to have overlapping downstream targets with MITF, the
518 best characterized driver of the PRO state, and promoters of these targets are frequently bound
519 simultaneously by TFAP2 and MITF (Seberg et al., 2017b). Despite this, they exert at least
520 partially independent functions, as evidenced by only partial rescue of *tfap2a/e* knockout with
521 overexpression of *mitfa* (Van Otterloo et al., 2010) and our own data showing that CRISPR
522 knockout of *tfap2a/e* did not alter expression of *mitfa*. How TFAP2 itself is regulated in this
523 context, however, remains an open question. DNA methylation has been linked to expression of
524 PRO/INV genes (Verfaillie et al., 2015) and to TFAP2A expression (Hallberg et al., 2014; Zeng
525 et al., 2013); however, further work is required to fully elucidate these relationships. This
526 mechanism is consistent with the recent report that breast cancer epigenetic state and CTC cluster
527 formation are tightly linked (Gkountela et al., 2019), and suggests that clusters may act to
528 potentiate an already more metastatic cell population. Given that the INV state in melanoma is

529 also associated with increased resistance to targeted therapy (Konieczkowski et al., 2014; Muller
530 et al., 2014; Verfaillie et al., 2015), pharmacologic disruption of CTC clusters could be an
531 attractive target to slow metastasis and decrease the distant spread of drug-resistant cells.

532

533 **ACKNOWLEDGEMENTS**

534 This work was supported by an NIH Research Program Grant under award number R01CA229215
535 to J.B.X. and R.M.W. N.R.C. was supported by the Kirschstein-NRSA predoctoral fellowship
536 (F30) from the NIH under award number F30CA220954, by a Medical Student Research Grant
537 from the Melanoma Research Foundation, and by a Medical Scientist Training Program grant from
538 the NIH under award number T32GM007739. This work was also supported by an NIH Research
539 Program Grant under award number R01AR062547 to R.A.C. The Leeds Melanoma cohort was
540 funded by Cancer Research UK grants C588/A19167, C8216/A6129, and C588/A10721 and NIH
541 grant R01CA83115. We thank the University Research Priority Program (URPP) of the University
542 of Zurich for access to the melanoma biobank and early passage cultures. The results published
543 here are in part based upon data generated by the TCGA Research Network:
544 <https://www.cancer.gov/tcga>.

545

546 **AUTHOR CONTRIBUTIONS**

547 N.R.C., J.B.X. and R.M.W. designed the study and wrote the manuscript, on which all authors
548 commented. N.R.C., A.R., M.H., M.Sz., M.Z., T.-H.H., S.P., M.Sa, E.M., and M.T. performed
549 experiments. C.K. and R.A.C. performed and analyzed CUT&RUN experiments. N.R.C. analyzed
550 *in vitro*, *in vivo*, and RNA-seq data. N.R.C. and M.H. analyzed human CTC cluster data. N.R.C.,
551 A.R., M.B., and I.Y. analyzed single-cell RNA-seq data. S.H. developed methods for *in vivo*

552 imaging and analysis. M.D. developed methods for *in vitro* cell tracking and image analysis. L.F.
553 and M.P.L. generated low-passage patient-derived cell lines, and L.B. and M.P.L. obtained patient
554 PBMC samples. M.G., J.N., J.N.-B., M.R.M., P.C., D.J.A. and R.R. analyzed clinical data.

555

556 **DECLARATION OF INTERESTS**

557 M.P.L. receives research funding from Roche and Novartis. R.M.W. is a paid consultant to N-of-
558 One Therapeutics, a subsidiary of Qiagen. R.M.W. is on the Scientific Advisory Board of Consano,
559 but receives no income for this. R.M.W. receives royalty payments for the use of the *casper* line
560 from Carolina Biologicals.

561 **Figure 1. PRO and INV coexist in zebrafish melanoma, with INV cells metastasizing more**
562 **frequently due to increased extravasation.**

563 **a.** Proliferative (PRO) and invasive (INV) subpopulations were identified from the ZMEL1
564 zebrafish melanoma cell line, which was originally isolated from a transgenic zebrafish and can
565 be transplanted into transparent *casper* zebrafish (adapted with permission from (Heilmann et al.,
566 2015)). **b-c.** Tracking of individual cells by time-lapse microscopy (both $p < 0.001$ by linear
567 regression, $N=4$ independent experiments). **b.** Growth curves (mean \pm SE of mean, smoothed with
568 moving window average of 5 time points) and doubling time (mean [95% CI]: 27.0 h [26.9, 27.1]
569 vs. 29.7 h [29.6, 29.9] for ZMEL1-PRO and -INV, respectively). **c.** (left) Representative
570 displacements of 500 tracks, and (right) model estimates \pm 95% CI for alpha, the slope of the log-
571 log plot of mean squared displacement vs. lag time (τ) for each ZMEL1-PRO and -INV. Larger
572 alpha indicates more persistent motion, with $\alpha=1$ for diffusive and $\alpha=2$ for projectile motion. **d.**
573 (left) The INV signature from Hoek et al. (Hoek et al., 2006) was the top gene set by Gene Set
574 Association Analysis (GSAA) of ZMEL1-INV vs. -PRO RNA-seq. (Right) Dual waterfall plot of
575 GSAA ranked by false discovery rate (FDR). Literature PRO/INV gene sets are indicated with an
576 asterisk and colored according to FDR. **e.** Heatmap of genes in Hoek INV signature that are
577 differentially expressed between ZMEL1-PRO and -INV (\log_2 fold change cutoff \pm 1.5, $p_{\text{adj}} <$
578 0.05). Human ortholog gene names are used for clarity (see Figure S1e for zebrafish gene names).
579 **f.** Segmentation of representative images of ZMEL1-PRO and -INV tumors and distant metastases
580 (e.g. to caudal region [box]) at 3 days post-transplant (3dpt). Original images shown in Figure S1e.
581 **g.** Quantification of caudal metastases seeded by ZMEL1 populations at 3 dpt (OR [95% CI]: 11.62
582 [1.43, 94.53]; $p=0.022$ by logistic regression; $N=3$ independent experiments with PRO/INV 10/10,
583 31/33, and 13/13 fish per group, respectively; $n=110$ fish total; plot shows mean \pm SD). **h.**

584 Representative images at 1 dpt from time lapse confocal microscopy of ZMEL1 cells transplanted
585 intravenously in larval zebrafish. Arrowhead indicates group of cells invading from the notochord
586 (NC) and caudal hematopoietic tissue (CHT) into the tail fin mesenchyme (TF). Images are
587 representative of n=13 fish per cell type. See Supplementary Video 1 for full time lapse. **i.**
588 Quantification of caudal tissue invasion by imaging at 4-6dpt (N=3 independent experiments with
589 PRO/INV 23/23, 21/21, and 19/23 fish per group, respectively; OR [95% CI]: 13.58 [5.56, 33.18];
590 $p < 0.001$ by logistic regression, plot shows mean \pm SD). **j.** Relative number of ZMEL1-PRO and -
591 INV cells isolated and quantified by flow cytometry from primary tumors and metastases of fish
592 transplanted with a 1:1 mixture of ZMEL1-PRO and -INV (primary tumors from n=6 fish;
593 metastases from n=4 fish; $p = 0.51$ and $p = 0.031$, respectively, by one-sample two-sided t-test with
594 Bonferroni correction).

595 **Figure 2. Cluster formation by INV state drives spatial patterning of melanoma clusters.**

596 **a.** (left) Plot of top Gene Ontology (GO) gene set by GSAA for ZMEL1-INV vs. -PRO RNA-seq.
597 (right) Dual waterfall plot of top/bottom 250 gene sets from GO analysis (for full plot see Figure
598 S2a). Adhesion GO gene sets are indicated with an asterisk and colored according to false
599 discovery rate (FDR) **b.** Heatmap of genes in adhesion GO gene sets (FDR < 0.05, n=3) that are
600 differentially expressed between ZMEL1-PRO and -INV (\log_2 fold change cutoff ± 1.5 , $p_{\text{adj}} <$
601 0.05). Human ortholog gene names are used for clarity (see Figure S2b-c for absolute expression
602 and zebrafish gene names). **c.** (top) Schematic of assay and (bottom) quantification of cluster
603 formation in low-bind plates after 2 days (N=6 independent experiments, $p < 0.001$ by two-sided t-
604 test, plot shows mean \pm SD). **d.** Representative images of clusters formed after 3 days. **e.** Human
605 melanoma cell lines ranked by cluster forming ability (left to right: low to high) with PRO/INV
606 gene expression scores from Hoek et al. (Hoek et al., 2006) (Spearman correlation and Bonferroni-
607 corrected p-values shown, scale bar 100 μm). **f.** Co-clusters of 1:1 mixture of ZMEL1-PRO and -
608 INV. **g.** (left) 3D opacity rendering and (right) central slice (slice 54 of 115) of confocal imaging
609 through co-cluster of ZMEL1-PRO and ZMEL1-INV. **h.** Co-clusters of 1:1 mixture of ZMEL1-
610 INV and ZMEL1-INV with either control (sg_scr) or *cdh1* (sg_ *cdh1*) sgRNA.

611 **Figure 3. PRO and INV cooperate in metastasis via co-extravasation.**

612 **a.** In the first 24 hours following intravenous transplant of ZMEL1-PRO and -INV, a mixed cluster
613 of both populations (left, arrowhead) extravasated from the caudal hematopoietic tissue (CHT)
614 into the tail fin mesenchyme (TF)—with ZMEL1-INV leading (middle, arrow) and ZMEL1-PRO
615 following (right, arrow). **b.** Segmented representative image of adult zebrafish with orthotopic
616 transplant of 1:1 mixture of ZMEL1-PRO and -INV. Arrowheads indicate polyclonal metastases,
617 including to the kidney and caudal regions (left and right boxes, respectively). Original image
618 shown in Figure S4b. **c.** Number of fish co-transplanted with a 1:1 mixture of ZMEL1-PRO and -
619 INV that have no caudal metastases (None), caudal metastases comprised of exclusively PRO or
620 INV, or caudal metastases formed by co-metastasis (Co-Met) of both cell types (N=5 independent
621 experiments with 17, 15, 16, 15, and 17 fish each; 80 fish total; $p < 0.001$ by Mantel-Haenszel's test
622 for null hypothesis of no interaction). **d.** Percentage of fish with ZMEL1-PRO caudal metastasis 3
623 dpt following orthotopic transplant of ZMEL1-PRO or a 1:1 mixture of ZMEL1 PRO:INV (OR
624 [95% CI]: 3.31 [1.10, 9.96]; $p = 0.033$ by logistic regression). **e.** Proportion of fish with ZMEL1-
625 INV caudal metastasis 3 dpt following orthotopic transplant of ZMEL1-INV or a 1:1 mixture of
626 ZMEL1 PRO:INV (OR [95% CI]: 1.32 [0.61, 2.88]; $p = 0.49$ by logistic regression). For **c-e**: N=5
627 independent experiments with PRO/MIX/INV 18/17/18, 13/15/14, 15/16/15, 12/15/15, and
628 15/17/16 fish per group, respectively; 231 fish total; plots show mean \pm SD. **f.** Number of ZMEL1-
629 PRO cells in Boyden Chamber assay migrating per 20X field when alone or mixed with ZMEL1-
630 INV ($p = 0.042$ by linear regression). **g.** Number of ZMEL1-INV cells in Boyden Chamber assay
631 migrating per 20X field when alone or mixed with ZMEL1-PRO ($p = 0.91$ by linear regression).
632 For **f-g**: N=3 independent experiments for each EGFP and tdTomato labeling; plots show mean \pm
633 SD.

634 **Figure 4. TFAP2 distinguishes PRO vs. INV state and modulates clustering and metastasis.**
635 **a.** HOMER de-novo motif analysis on genes upregulated in ZMEL1-PRO vs. -INV (\log_2 fold
636 change cutoff ± 1.5 , $p_{\text{adj}} < 0.05$, $\pm 500\text{bp}$ of transcription start site [TSS]). **b.** HOMER de-novo
637 motif analysis of genes differentially expressed between ZMEL1-INV in 3D (clusters) vs. 2D (no
638 clusters) (\log_2 fold change cutoff ± 1.5 , $p_{\text{adj}} < 0.05$, $\pm 500\text{bp}$ of TSS). **c.** Cluster size after 2 days in
639 ZMEL1-PRO with CRISPR-Cas9 inactivation of *tfap2a* and *tfap2e* alone or in combination
640 (*sg_tfap2a/e*) versus control (*sg_scr*) (p-values by linear regression; N=3 independent
641 experiments). **d.** Representative images of clusters formed after 2 days from ZMEL1-PRO with
642 *sg_scr* vs. *sg_tfap2a/e*. **e.** Growth of ZMEL1-PRO orthotopic primary tumors with *sg_scr* vs.
643 *sg_tfap2a/e* (p=0.011 by linear regression; N=3 independent experiments with *sg_scr/sg_tfap2a/e*
644 24/22, 22/22, 24/24 fish per group, respectively; n=138 fish total). **f.** Representative image of
645 extravasated (arrows) and partially extravasated (arrow-head) ZMEL1-PRO cells with *sg_tfap2a/e*
646 following intravenous transplant in *casper* fish with FLK-RFP transgene labeling the vasculature.
647 **g.** Proportion of larval fish intravenously transplanted with ZMEL1-PRO with *sg_scr* or
648 *sg_tfap2a/e* with extravasated cells at 1 dpt, as quantified from confocal time lapse microscopy
649 (OR [95% CI]: 2.20 [1.05, 4.61]; p=0.038 by logistic regression; N=3 independent experiments
650 with *sg_scr/sg_tfap2a/e* 20/20, 22/23, and 22/22 fish per group, respectively; n=129 fish total).

651 **Figure 5. TFAP2 correlates with clustering in human melanoma and regulates genes**
652 **associated with metastasis and cell-cell adhesion**

653 **a.** Human melanoma cell lines in the Cancer Cell Line Encyclopedia (CCLE, n=56) plotted as
654 PRO and INV scores (Hoek et al., 2006) calculated from RNA-seq and colored according to
655 TFAP2A mRNA expression. Pearson correlation coefficients between TFAP2A and PRO/INV
656 scores are shown on axes. **b.** TFAP2A mRNA expression in The Cancer Genome Atlas (TCGA)
657 melanoma (SKCM) cohort comparing primary tumors and metastases ($p < 0.001$ by Wilcoxon rank
658 sum test with Bonferroni correction). **c.** Low-passage human melanoma cell lines ranked by
659 increased cluster forming ability (left to right) with TFAP2A expression quantified by
660 immunofluorescence (plot and top; Spearman correlation shown; scale bar 20 μm) and clustering
661 (bottom, scale bar 500 μm). **d.** GSAA was run using gene sets and GO gene sets with FDR < 0.05
662 from INV vs. PRO RNA-seq (n=39 gene sets; cyan points in Figures 1d and 2a). Bars show
663 Normalized Association Score (NAS) for CRISPR (ZMEL1-PRO sg_ *tfap2a/e* vs. sg_ *scr*) and INV
664 vs. PRO for each gene set, with black outline representing FDR < 0.05 for CRISPR experiment. **e.**
665 Plot of Hoek et al. (Hoek et al., 2006) INV signature by GSAA for ZMEL1-PRO sg_ *tfap2a/e* vs.
666 sg_ *scr* RNA-seq. **f.** Heatmap of top genes in Hoek INV and GO Adhesion gene sets that are
667 differentially expressed between ZMEL1-PRO sg_ *tfap2a/e* and sg_ *scr* (\log_2 fold change cutoff \pm
668 0.5, $p_{\text{adj}} < 0.05$). Asterisks (*) indicate genes with associated TFAP2A CUT&RUN peaks. Human
669 ortholog gene names are used for clarity (see Figure S7a for zebrafish gene names).

670 **Figure 6. PRO-INV heterotypic CTC clusters exist in the blood of melanoma patients.**
671 **a.** Schematic of experimental design. **b.** (left) IF staining for TFAP2A and SOX9 and (right)
672 nuclear quantification in a PRO-INV heterotypic melanoma CTC cluster. Scale bar is 5 μ m. **c-d.**
673 Classification of human melanoma CTC clusters based on quantification of nuclear TFAP2A and
674 SOX9 staining as (c) homotypic (PRO-only or INV-only) vs. PRO-INV heterotypic, and (d) PRO,
675 PRO-INV, and INV. For **c-d** n=32 clusters from four patients. Five additional patients were
676 analyzed with no CTC clusters identified. For **d** each stacked box represents one CTC cluster.

677 **Figure 7. Longitudinal single-cell RNA-seq reveals stability of PRO but not INV state.**

678 **a.** Schematic of experiment. Prior to flow cytometry and 10X single-cell RNA-seq, ZMEL1-PRO
679 and -INV cells were either grown *in vitro* (individual or co-culture) or isolated from zebrafish
680 orthotopically transplanted with a 1:1 mixture of the two subpopulations (primary tumors or
681 metastases). **b.** Uniform Manifold Approximation and Projection (UMAP) (McInnes et al., 2018)
682 dimensionality reduction of 40,293 ZMEL1 cells sequenced as in (a). Individual culture (IND);
683 co-culture (CO); primary tumors (PRI); metastases (MET). **c.** PRO and INV scores based on
684 ZMEL1 bulk RNA-seq are plotted for all cells in gray. ZMEL1-PRO (purple) and ZMEL1-INV
685 (green) for the indicated condition are colored. Diagonal line represents the classifier used in (d).
686 **d.** Confusion matrices comparing initial cell identity with observed cell classification based on a
687 linear classifier trained on *in vitro* individual culture samples. **e.** ZMEL1-INV cells plotted as in
688 (c) colored according to *tfap2e* mRNA expression reveal re-activation of *tfap2e* upon metastatic
689 dissemination.

690 **METHODS**

691 **Cloning**

692 Zebrafish-specific expression plasmids were generated by Gateway Cloning (Fisher) into the
693 pDestTol2pA2 backbone (Tol2kit, plasmid #394) (Kwan et al., 2007). nls-mCherry (Tol2kit,
694 plasmid #233) was cloned under the ubiquitin promoter. tdTomato was cloned under the zebrafish
695 *mitfa* promoter.

696

697 **Cell culture**

698 The establishment of the ZMEL1 zebrafish melanoma cell line from a tumor in a *mitfa*-
699 BRAF^{V600E}/*p53*^{-/-} zebrafish was described previously (Heilmann et al., 2015; Kim et al., 2017).
700 ZMEL1 was grown at 28°C in a humidified incubator in DMEM (Gibco #11965) supplemented
701 with 10% FBS (Seradigm), 1X penicillin/streptomycin/glutamine (Gibco #10378016), and 1X
702 GlutaMAX (Gibco #35050061). The ZMEL1-PRO and -INV populations were identified based
703 on phenotyping of multiple concurrent ZMEL1 cultures. ZMEL1 populations were validated by
704 RNA-seq confirming expression of expected transgenes. Human melanoma cell lines were
705 maintained in DMEM (Gibco #11965) supplemented with 10% FBS (Seradigm), 1X
706 penicillin/streptomycin/glutamine (Gibco #10378016), with the exception of HMCB, which was
707 maintained in MEM (Gibco #11095080) supplemented with 10% FBS (Seradigm), 1% Sodium
708 Pyruvate (Gibco #11360070), 1% MEM-Non-Essential Amino Acids (Gibco #11140050), 10mM
709 HEPES (Gibco #15630080), and 1X penicillin/streptomycin (Gibco #15140122). Low-passage
710 human melanoma cell lines were established and cultured as previously described (Raaijmakers et
711 al., 2015). Cells were routinely confirmed to be free from mycoplasma (Lonza Mycoalert). Human
712 cell lines were either purchased directly from ATCC or verified by STR profiling.

713

714 **Fluorescently labeled cell lines**

715 The ZMEL1 cell line constitutively expresses EGFP under the *mitfa* promoter (Heilmann et al.,
716 2015). ZMEL1 lines additionally expressing nls-mCherry under the ubiquitin promoter for cell
717 tracking experiments were generated with the Neon Transfection System (Fisher) followed by
718 FACS sorting. ZMEL1 lines expressing tdTomato under the *mitfa* promoter were generated
719 through CRISPR/Cas9 mutation of the constitutive EGFP (see CRISPR/Cas9 below) followed by
720 Neon Transfection and FACS sorting.

721

722 **CRISPR/Cas9**

723 The Alt-R CRISPR-Cas9 System (Integrated DNA Technologies) was used for CRISPR-Cas9
724 experiments following the manufacturer's protocols for use with the Neon Transfection System
725 (Fisher) and adherent cells. Successful nucleofection was confirmed by visualizing ATTO 550
726 labeled tracrRNA one day post-nucleofection. Successful loss of full-length protein expression
727 was verified by visualizing loss of EGFP expression (EGFP) or by Western blot (*cdh1*, *tfap2a*,
728 *tfap2e*). Control scramble sgRNA (sg_scr) sequence was used from Wang et al. (Wang et al.,
729 2015). sgRNA sequences are listed in Supplementary Table 8.

730

731 **Cluster formation assay**

732 ZMEL1 cells were trypsinized, centrifuged at 300g for 3 minutes, and resuspended in standard
733 culture media. A Corning Ultra-Low Attachment Surface 96-well plate (#3474) was seeded with
734 5×10^4 cells/well in a final volume of 100uL. Clusters were allowed to form over the course of 2-3
735 days in a humidified 28°C incubator, or were used for time-lapse microscopy on a Zeiss

736 AxioObserver Z1 equipped with an incubation chamber using a 5x/0.16NA objective. For human
737 melanoma cell lines, a round bottom Corning Ultra-Low Attachment Surface 96-well plate (#7007)
738 was seeded with 5×10^3 cells/well in a final volume of 100uL. Clusters were allowed to form over
739 the course of 24-48h in a humidified 37°C incubator. Cell lines were ranked according to their
740 relative abilities to form dense three-dimensional clusters.

741

742 **Cluster quantification**

743 The average cluster size of each image was quantified using a MATLAB implementation of the
744 characteristic length scale equation from Smeets et al (Smeets et al., 2016). Cluster mixing and
745 spatial sorting were quantified for each cluster using a custom MATLAB segmentation routine.
746 To quantify cluster mixing, background corrected images for EGFP and tdTomato were each
747 segmented using a low and high threshold to improve detection of small and large clusters,
748 respectively. Segmentation from each channel was merged, and the coefficient of variation (CV)
749 of cluster area was calculated for clusters filtered to have a size corresponding to 1 cell or larger.
750 The composition of each cluster (filtered to have a size corresponding to a size of approximately
751 4 cells or larger) was classified as red-only, green-only, or red-green mix. Cluster spatial sorting
752 was calculated for individual large (equivalent diameter greater than $45\mu\text{m}$) mixed red-green
753 clusters by calculating the weighted average of radial intensity profiles for each channel. The
754 difference between the weighted averages for each channel was calculated and normalized by the
755 radius of the cluster. With this dimensionless metric, small values correspond to well-mixed
756 clusters, whereas larger values correspond to a high degree of spatial segregation.

757

758 **Cluster confocal microscopy**

759 ZMEL1 clusters were fixed at room temperature for 45 minutes by adding an equal volume of 4%
760 PFA in PBS to 2-day cultures of clusters (2% final PFA concentration). Fixed clusters were washed
761 once with PBS and counterstained with Hoechst 33342 (Fisher H1399) and transferred to a 96-
762 well glass-bottom plate (Mattek, PBK96G-1.5-5-F) before imaging. Individual clusters were
763 imaged on a Leica TCS SP5-II inverted point-scanning confocal microscope with a or 40x/1.10NA
764 objective. 3D reconstruction was performed using Volocity (PerkinElmer, v6.3). Individual slices
765 were visualized with ImageJ.

766

767 **Cell tracking**

768 Time-lapse microscopy was performed on a Zeiss AxioObserver Z1 equipped with an incubation
769 chamber. A 96-well plate (Corning 353072) was seeded with 1.2×10^4 ZMEL1 cells (PRO or INV)
770 admixed such that 1/6 of the population stably expressed nls-mCherry under the Ubi promoter,
771 and allowed to adhere overnight. Cells were imaged every 5 minutes for 24 hours. Centroids of
772 nuclei were identified, and tracks generated using a MATLAB implementation of the IDL tracking
773 methods developed by John Crocker, David Grier, and Eric Weeks
774 (physics.georgetown.edu/matlab/). For each imaging location, growth was calculated based on the
775 number of nuclei present at each time point, assuming equal numbers of cells move in and out of
776 the field of view over time. For each track, mean squared displacement (MSD) was calculated as
777 previously described (Gorelik and Gautreau, 2014) over a range of lag times ($5 \leq \tau < 100$ min). The
778 log-log plot of MSD vs. τ provides information about both the diffusion coefficient (intercept) and
779 persistence (slope, α) of cells. For a cell moving randomly, $\alpha=1$; and for a cell moving along a
780 straight line, $\alpha=2$ (Gorelik and Gautreau, 2014). N=4 independent replicates were performed, each
781 consisting of 6 technical replicates per cell type. Growth rates were quantified with a linear mixed-

782 effects model using the fitlme function in MATLAB with the model, 'log(cell_number) ~ time +
783 cell_type:time + (1. replicate)'. Motility (α) was quantified as the slope of the linear model,
784 'log(MSD) ~ 1 + cell_type*log(τ)'. Growth plots represent the smoothed (moving window average
785 of 5 time points [= 25 min]) average cell number \pm SE normalized to the cell number at time zero.
786

787 **Boyden chamber migration**

788 Cell migration of PRO/PRO, INV/INV, and PRO/INV mixtures was quantified using a Boyden
789 chamber (transwell) assay. A 1:1 mixture of ZMEL1 cells labeled with EGFP and tdTomato (5×10^4
790 cells total per well) were added to each transwell insert (Corning, 353492, 3.0 μ m pore) in a 24-
791 well plate (Corning, 353504) in 500 μ L of complete media. The lower chamber was filled with
792 500 μ L of complete media and cells were allowed to migrate for 2 days. Cells were fixed for 15
793 minutes at room temperature with 4% PFA in PBS, washed once with PBS, and non-migrated cells
794 removed with cotton-tipped swabs. Migrated cells were stained with Hoechst 33342 (Fisher
795 H1399) and ≥ 9 fields/well imaged on a Zeiss AxioObserver Z1 with a 10x/0.45NA objective.
796 Nuclei were segmented using intensity thresholding of background-corrected Hoechst staining
797 followed by an intensity-based watershed step to separate adjacent objects. Cell identity was
798 established based on fluorophore expression within the mask defining each nucleus.

799

800 **Conditioned media**

801 Conditioned media was collected from confluent 10cm dishes of ZMEL1-PRO and -INV cultures
802 following 2-3 days of growth and filtered through a 0.45 μ m syringe filter (Fisher #09-720-005) to
803 remove viable cells. Filtered conditioned media or fresh complete media was mixed 1:1 with fresh
804 complete media and used for subsequent assays. For proliferation assays, ZMEL1-PRO or -INV

805 cells (1.4×10^4 cells/well) were plated in white wall 96-well plates (Corning #3610) in a 1:1 mixture
806 of fresh complete media with either ZMEL1-PRO or -INV conditioned media or fresh complete
807 media. Relative proliferation was measured by CellTiter-Glo® Luminescent Cell Viability Assay
808 (Promega) according to manufacturer's protocol two days after plating. For Boyden Chamber
809 assays, ZMEL1-PRO or -INV cells (5×10^4 cells total per well) were added to each transwell insert
810 (Corning, 353492, 3.0um pore) in a 24-well plate (Corning, 353504) in 500uL of fresh complete
811 media. The lower chamber was filled with either ZMEL1-PRO or -INV conditioned media or fresh
812 complete media and cells were allowed to migrate for 2 days. Non-migrated cells were removed
813 with cotton-tipped swabs. Migrated cells were stained with Hoechst 33342 (Fisher H1399) and 26
814 fields/well imaged on a Zeiss AxioObserver Z1 with a 10x/0.45NA objective. Centroids of nuclei
815 were identified and counted using a MATLAB implementation of the IDL tracking methods
816 developed by John Crocker, David Grier, and Eric Weeks (physics.georgetown.edu/matlab/).

817

818 **Cluster caspase-3/7 assay**

819 ZMEL1 cells were plated into cluster formation assay as described above using multiple ratios of
820 PRO:INV cells (1:0, 4:1, 1:1, 1:4, and 0:1) with two identical plates per replicate. After 2 days,
821 one plate was used for quantification of caspase-3/7 activity using Caspase-Glo 3/7 Assay System
822 (Promega) and the other for quantification of cell number using CellTiter-Glo Luminescent Cell
823 Viability Assay (Promega) each according to manufacturer's protocol. For each replicate caspase-
824 3/7 activity was normalized to cell number. The effect of various PRO:INV mixing ratios was
825 quantified by linear regression with the fraction of INV cells as a continuous variable.

826

827 **Zebrafish husbandry**

828 Zebrafish were housed in a dedicated facility maintained at 28.5°C with a light/dark cycle (14
829 hours on, 10 hours off). All anesthesia was performed using Tricaine-S (MS-222, Syndel USA,
830 Ferndale, WA) with a 4g/L, pH 7.0 stock. All procedures adhered to Memorial Sloan Kettering
831 Cancer Center IACUC protocol number 12-05-008.

832

833 **Larval transplantation**

834 Transplantation of ZMEL1 cells into 2dpf *casper* zebrafish larvae was performed as previously
835 described (Heilmann et al., 2015; Kim et al., 2017). Briefly, ZMEL1 cells were prepared by
836 trypsinization, centrifuged at 300g for 3 minutes, and resuspended at a concentration of either
837 2.5×10^7 or 5.0×10^7 cells/mL in 9:1 DPBS:H₂O (Gibco 14190-144). Cells were injected into the
838 Duct of Cuvier of 2dpf *casper* or *casper* FLK-RFP (labeling the vasculature with RFP) fish using
839 a Picoliter Microinjector (Warner Instruments, PLI-100A) with a glass capillary needle (Sutter,
840 Q100-50-10) made on a laser-based needle puller (Sutter, P-2000). For mixing studies, ZMEL1-
841 PRO and ZMEL1-INV differentially labeled with EGFP or tdTomato were mixed at a 1:1 ratio
842 prior to injection. Fish with successful transplants based on the presence of circulating cells and/or
843 cells arrested in the caudal vasculature were either used for time-lapse confocal microscopy (see
844 “Zebrafish confocal time-lapse imaging”) or individually housed and followed by daily imaging
845 on a Zeiss AxioZoom V16 fluorescence microscope.

846

847 **Adult transplantation**

848 Transplantation of ZMEL1 cells into adult *casper* zebrafish was performed as previously
849 described (Heilmann et al., 2015; Kim et al., 2017). Briefly, adult zebrafish were irradiated on two
850 sequential days with 15 Gy on a cesium irradiator (Shepherd) and were transplanted 3-4 days later.

851 On the day of transplant, ZMEL1 cells were prepared by trypsinization, washed once with DPBS,
852 and resuspended at a concentration of 1.67×10^8 cells/mL in DPBS (Gibco 14190-144). Cells were
853 injected subcutaneously, caudal to the cloaca, on the ventral side of zebrafish anesthetized with
854 Tricaine-S. Fish were imaged on days 1, 3, and 7 post-transplant on a Zeiss AxioZoom V16
855 fluorescence microscope. For 1:1 mixing studies, fish were injected with an equivalent final
856 concentration of either ZMEL1-PRO, ZMEL1-INV, or a 1:1 mixture of the two populations
857 differentially labeled with EGFP or tdTomato. A total of N=5 independent mixing experiments
858 were performed, N=3 for INV-EGFP/PRO-tdTomato and N=2 for INV-tdTomato/PRO-EGFP,
859 each with at least 12 fish per group (n=231 fish total). For variable ratio mixing studies, fish were
860 injected with an equivalent final concentration of tdTomato-labeled ZMEL1-PRO mixed at a 1:4,
861 4:1, or 9:1 ratio with EGFP-labeled ZMEL1-PRO or ZMEL1-INV. A total of N=3 independent
862 variable ratio mixing experiments were performed, each with at least 16 fish per group (n=334 fish
863 total).

864

865 **Zebrafish imaging and image quantification**

866 ***Whole-fish larval imaging***

867 Larval zebrafish transplanted as described above were anesthetized with Tricaine-S and imaged
868 on a bed of 2% agarose (KSE Scientific, BMK-A1705) in E3. Images were manually scored for
869 cells that had invaded the tail fin parenchyma at experiment endpoint.

870

871 ***Adult imaging***

872 Adult zebrafish were imaged as previously described (Heilmann et al., 2015). Briefly, on days 1,
873 3, and 7 post-transplant fish were anesthetized with Tricaine-S and imaged on a bed of 2% agarose

874 using a monochrome camera for fluorescence and brightfield, and color camera for observing
875 pigmentation. Fluorescence images were manually scored for the following pre-specified binary
876 outcomes:

- 877 • Distant metastases: tumor not adjacent or contiguous with primary tumor
- 878 • Caudal metastases: distant metastasis caudal to anal fin
- 879 • Dorsal metastases: distant metastasis on dorsum of fish, near insertion of dorsal fin

880 For each binary outcome, the population composition was also scored (PRO, INV, or both).

881 Primary tumor growth was also quantified over time for each fish using a previously described
882 custom MATLAB pipeline (Heilmann et al., 2015). An adapted version of this pipeline with an
883 adaptive threshold segmentation was utilized to allow easier visualization of representative
884 images.

885

886 ***Zebrafish larval confocal time-lapse imaging***

887 Larval zebrafish were transplanted as described above. Fish with successful intravenous
888 transplants were anesthetized with Tricaine-S and embedded in 1% low-melt agarose (Sigma
889 A9045) in E3 containing 0.28ug/mL Tricaine-S in a glass-bottom square-well 96-well plate
890 (Arrayit 96-well Microplate SuperClean, Cat M96FC, Lot 150901). Wells were filled with E3
891 containing 0.28ug/mL Tricaine-S and the plate sealed with a PCR microseal (BioRad Microseal
892 'B' Film, #MSB1001). Up to 45 larval zebrafish per experiment were imaged in parallel on a GE
893 IN Cell Analyzer 6000 every 15-20 minutes for 24-30h. A single z-stack was acquired for each
894 fish using a 10X/0.45NA objective and 8-10um z-steps. Because the ventral edge of the caudal
895 vein exists as a single plane at this developmental stage (2-3dpf) (Isogai et al., 2001), maximum

896 intensity projections were generated for each fish and manually scored for the presence of ZMEL1
897 cells that extravasate ventrally from the caudal vein and invade into the tail fin.

898

899 **Western blot**

900 Cell lysates were collected by sonication in RIPA buffer (Thermo #89901) with 1X Halt Protease
901 and Phosphatase Inhibitor Cocktail (Thermo #78441) followed by centrifugation (14,000rpm for
902 10min at 4°C) and collection of the supernatant. Protein concentration was quantified by Bradford
903 (Sigma B6916-500mL) according to manufacturer's protocol. Samples were mixed with 6X
904 reducing loading buffer (Boston BioProducts #BP-111R) and denatured at 95°C for 10 minutes.
905 Samples were run on a Mini PROTEAN TGX gel (BioRad) and transferred using Turbo Mini
906 Nitrocellulose Transfer Pack (Bio-Rad, catalog #1704158). Membranes were blocked with 5%
907 nonfat dry milk in TBST (1X TBS + 0.1% Tween 20) for 1 hour before incubation with primary
908 antibody in PBS overnight at 4°C. Membranes were washed with TBST and incubated with
909 secondary antibody in 5% nonfat dry milk for 1 hour at room temperature. Membranes were
910 washed with TBST and developed with ECL (Amersham, RPN2109) using an Amersham Imager
911 600 (GE) or chemiluminescence film. Antibodies: anti-hs_CDH1 (BD #610181, lot 8082613),
912 anti-dr_Tfap2a (LifeSpan Biosciences, #LS-C87212, log 113877), anti-dr_Tfap2e (Fisher, #PA5-
913 72631, lot UA2709682A), anti-hs_TFAP2A (Cell Signaling, #3215, clone C83E10, lot 2), anti-
914 hs_cyclophilin B (Fisher #PA1-027A, lot SD248938).

915

916 **Immunofluorescence**

917 Cells were allowed to adhere for 2 days to glass chamber slides (PEZGS0816) coated with poly-
918 D-lysine (Sigma #P-0899). Slides were fixed for 15 minutes with 4% PFA in PBS, washed 3 times

919 with PBS, and permeabilized with 0.1% Triton X-100 (Fisher #BP151-100) in PBS for 15 min.
920 Slides were blocked with 5% donkey serum (Sigma #S30-M), 1% BSA (Fisher #BP1600-100),
921 and 0.1% Triton X-100 in PBS for 1 hour, followed by incubation in primary antibody overnight
922 at 4°C. Slides were washed 3 times with PBS and incubated in secondary antibody (anti-mouse
923 [Cell Signaling, #4408S]; anti-rabbit [Cell Signaling, #8889S]) for 1.5 hours, followed by 3 washes
924 with PBS and counterstain with (1:10,000) Hoechst 33342 in PBS for 45 minutes. Slides were
925 washed with PBS and mounted in ProLong Glass Antifade Mountant (Fisher #P36984). All
926 incubations were carried out at room temperature unless otherwise noted. Antibodies: rabbit anti-
927 hs_TFAP2A (Cell Signaling, #3215, clone C83E10, lot 2), mouse anti-hs_H3 (Cell Signaling,
928 #14269, clone 1B1B2, lot 6). Slides were imaged in a minimum of 9 fields on a Zeiss
929 AxioObserver Z1 using a 20x/0.80NA objective. Nuclei were segmented using intensity
930 thresholding of background-corrected Hoechst staining followed by an intensity-based watershed
931 step to separate adjacent objects. TFAP2A expression was quantified for each nucleus as the ratio
932 of TFAP2A to Histone H3 staining intensity.

933

934 **RNA-seq**

935 ***Samples***

936 For 2D culture, three replicate cultures at 70-80% confluence for each ZMEL1-PRO and -INV
937 were utilized. For 3D culture, three replicate cultures for each ZMEL1-PRO and -INV grown in
938 Corning Ultra-Low Attachment Surface 6-well plates (Corning #3471) for 48 hours were utilized.
939 One individual replicate of RNA-seq of ZMEL1-PRO grown in 3D culture was excluded due to
940 low RIN score, poor SeQC metrics and poor clustering with other replicate samples. For *tfap2a/e*

941 CRISPR, three independent batches of ZMEL1-PRO cells nucleofected with either *sg_scr* or
942 *sg_tfap2a/e* and grown in 2D conditions were utilized at 8-16 days post nucleofection.

943

944 ***Sequencing and analysis***

945 Total RNA was extracted with the RNeasy Plus Mini kit with QiaShredder (Qiagen). Purified RNA
946 was delivered to GENEWIZ (South Plainfield, NJ) for mRNA preparation with the TruSeq RNA
947 V2 kit (Illumina) and 100bp (2D) or 150bp (3D and CRISPR) paired-end sequencing on the
948 Illumina HiSeq2500. After quality control with FASTQC (Babraham Bioinformatics) and
949 trimming with TRIMMOMATIC (Bolger et al., 2014) when necessary, reads were aligned to
950 GRCz10 (Ensembl version 81) using STAR (Dobin et al., 2013), with quality control via
951 SeQC (DeLuca et al., 2012). Differential expression was calculated with DESeq2 (Love et al.,
952 2014) using the output of the --quantMode GeneCounts feature of STAR. The rlog function was
953 used to generate log₂ transformed normalized counts. Pathway and Gene Ontology (GO) analysis
954 were performed with GSAA using the following parameters: gsametric Weighted_KS, demetric
955 Signal2Noise, permute gene_set, rnd_type no_balance, scoring_scheme weighted, norm
956 MeanDiv (Xiong et al., 2012; Xiong et al., 2014). A full list of gene sets used for GSAA can be
957 found in Supplementary Table 6. Ortholog mapping between zebrafish and human was performed
958 with DIOPT (Hu et al., 2011) (Supplementary Table 9). Only orthologs with a DIOPT score greater
959 than 6 were used for GSAA and heatmap generation. In cases of more than one zebrafish ortholog
960 of a given human gene, the zebrafish gene with the highest average expression was selected. De-
961 novo motif analysis was performed with the HOMER (Heinz et al., 2010) function findMotifs.pl,
962 using the zebrafish genome (GRCz10) and searching for motifs of lengths 8, 10, and 16 within ±
963 500bp of the TSS of differentially expressed genes. Motifs were annotated using JASPAR (Khan

964 et al., 2017). A Cell-Cell Adhesion gene set was defined from the Core Enrichment genes from
965 comparing ZMEL1-INV vs. -PRO to the gene set, *GO Cell Cell Adhesion Via Plasma Membrane*
966 *Adhesion Molecules*.

967

968 **TFAP2A CUT&RUN**

969 ***Sample preparation***

970 Anti-TFAP2A Cleavage Under Targets and Release Using Nuclease (CUT&RUN) sequencing
971 was performed in wild-type and *TFAP2A;TFAP2C* double-mutant SKMEL28 cell lines as
972 described (Skene and Henikoff, 2017) with minor modifications. Cells in log-phase culture
973 (approximately 80% confluent) were harvested by cell scraping (Corning), centrifuged at 600g
974 (Eppendorf, centrifuge 5424) and washed twice in calcium-free wash-buffer (20 mM HEPES,
975 pH7.5, 150 mM NaCl, 0.5 mM spermidine and protease inhibitor cocktail, cOmplete Mini, EDTA-
976 free Roche). Pre-activated Concanavalin A-coated magnetic beads (Bangs Laboratories, Inc) were
977 added to cell suspensions (2×10^5 cells) and incubated at 4°C for 15 mins. Antibody buffer (wash-
978 buffer with 2mM EDTA and 0.03% digitonin) containing anti-TFAP2A (abcam, ab108311) or
979 Rabbit IgG (Millipore, 12-370) was added and cells were incubated overnight at 4°C. The next
980 day, cells were washed in dig-wash buffer (wash buffer containing 0.025% digitonin) and pA-
981 MNase was added at a concentration of 500 $\mu\text{g}/\text{mL}$ (pA-MNase generously received from Dr.
982 Steve Henikoff). The pA-MNase reactions were quenched with 2X Stop buffer (340mM NaCl,
983 20mM EDTA, 4mM EGTA, 0.05% Digitonin, 100 $\mu\text{g}/\text{mL}$ RNase A, 50 $\mu\text{g}/\text{mL}$ Glycogen and 2
984 pg/mL sonicated yeast spike-in control). Released DNA fragments were Phosphatase K (1 $\mu\text{L}/\text{mL}$,
985 Thermo Fisher Scientific) treated for 1 hr at 50°C and purified by phenol/chloroform-extracted and

986 ethanol-precipitated. Fragment sizes analysed using an 2100 Bioanalyzer (Agilent). All
987 CUT&RUN experiments were performed in duplicate.

988

989 *Library preparation and data analysis*

990 CUT&RUN libraries were prepared using the KAPA Hyper Prep Kit (Roche). Quality control
991 post-library amplification was conducted using the 2100 Bioanalyzer (Agilent) for fragment
992 analysis. Libraries were pooled to equimolar concentrations and sequenced with paired-end 100
993 bp reads on an Illumina HiSeq X platform. Paired-end FastQ files were processed through FastQC
994 (Babraham Bioinformatics) for quality control. Reads were trimmed using Trim Galore Version
995 0.6.3 (Developed by Felix Krueger at the Babraham Institute) and Bowtie2 version
996 2.1.0 (Langmead and Salzberg, 2012) was used to map the reads against the hg19 genome
997 assembly. The mapping parameters and peak calling analysis was performed as previously
998 described (Meers et al., 2019). Called peaks were annotated with ChIPseeker v1.18.0 (Yu et al.,
999 2015), with Distal Intergenic peaks excluded from downstream analysis. P-values for overlap with
1000 differential expression gene sets were calculated by comparing against overlap with randomly
1001 selected gene sets (n=10,000 iterations).

1002

1003 **Isolation of human CTC clusters**

1004 Peripheral blood samples were collected from 9 patients with Stage IV metastatic melanoma within
1005 6 months of death and enriched for CTCs on a Parsortix Cell Separation Cassette (GEN3D6.5,
1006 ANGLE) (Xu et al., 2015). In-cassette staining was performed following fixation with 4%
1007 paraformaldehyde and permeabilization with 0.2% Triton-100X and 4% donkey serum (DS) in
1008 PBS. Samples were incubated for 1 hour with primary antibodies—rabbit anti-Sox9 (HPA001758,

1009 Sigma, 1:50), mouse anti-TFAP2A (sc-12726, Santa Cruz Biotechnology, 1:50) and rat anti-CD45
1010 (MA5-17687, ThermoFisher, 1:500)—followed by 1 hour incubation with secondary antibodies—
1011 donkey anti-mouse AF488 (A21202, ThermoFisher, 1:800), donkey anti-rabbit AF647 (A31573,
1012 ThermoFisher, 1:800), and donkey anti-rat DyLight 555 (SA5-10027, ThermoFisher, 1:1000)—
1013 and subsequent DAPI staining. Images of CTC clusters were acquired with Leica DMI 4000/6000
1014 using 500ms exposure for AF488 and AF647, and 300ms exposure for DyLight 555 acquisitions.
1015 Presumptive CTC clusters were manually identified as two or more adjacent TFAP2A-positive
1016 cells. Nuclei were segmented using intensity thresholding of background-corrected DAPI staining
1017 followed by a distance-based watershed step and manual curation to separate adjacent objects, with
1018 subsequent exclusion of CD45-positive cells. Nuclear TFAP2A and SOX9 intensity were
1019 quantified for each cell from background-corrected images and each cell classified as PRO vs.
1020 INV based on the relative expression of each marker.

1021

1022 **Single-cell RNA-seq (scRNA-seq)**

1023 ***Sample preparation***

1024 *In vitro* samples were cultured under standard conditions with either ZMEL1-PRO (tdTomato) and
1025 -INV (EGFP) separately (individual culture) or mixed together in a 1:1 ratio for 11 days (co-
1026 culture). Cells were trypsinized, resuspended in DMEM supplemented with 2% FBS, and flow
1027 sorted using a SY3200 (Sony) for DAPI-negative pure ZMEL1-PRO and -INV populations prior
1028 to droplet-based scRNA-seq. For *in vivo* samples, adult *casper* zebrafish were transplanted with a
1029 1:1 mixture of ZMEL1-PRO (tdTomato) and -INV (EGFP) cells as described above (Adult
1030 transplantation). Tumors were allowed to grow for 6 days (primary tumors) or 13 days
1031 (metastases). At experimental timepoint, tumors were surgically excised and minced with a fresh

1032 scalpel (primary tumors from n=6 fish; metastases from n=4 fish). Each sample was placed in a
1033 15mL tube containing 3mL of 0.9X DPBS with 0.16 mg/mL of Liberase TL (Sigma
1034 #5401020001), incubated at room temperature for 15 minutes followed by trituration using a wide-
1035 bore P1000 (Fisher #2069G), and incubated for an additional 15 minutes. 500 μ L FBS was added
1036 and each sample was triturated again and passed through a 70 μ m cell strainer (Corning #352350).
1037 Samples were centrifuged 500g for 5 minutes, resuspended in DMEM supplemented with 2% FBS,
1038 and flow sorted using a SY3200 (Sony) for DAPI-negative pure ZMEL1-PRO and -INV
1039 populations prior to droplet-based scRNA-seq.

1040

1041 ***Droplet-based scRNA-seq***

1042 For droplet-based scRNA-seq, experiments were performed using the 10X Genomics Chromium
1043 platform, with the Chromium Single Cell 3' Library & Gel Bead Kit v3.1 (1000128) and
1044 Chromium Single Cell 3' Chip G (1000127). ~8,000 cells per condition were centrifuged 300g for
1045 5 minutes and resuspended in DMEM supplemented with 10% FBS and loaded to each channel
1046 for GEM generation and barcoded single cell libraries were generated according to manufacturer's
1047 instructions. Libraries were diluted to 2nM and 75bp paired end sequencing was performed using
1048 the Illumina NextSeq 500. Between 150-200 million paired reads were generated for each library.

1049

1050 ***scRNA-seq processing and analysis***

1051 Raw sequencing data were processed using the CellRanger 3.1.0 pipeline developed by 10X
1052 Genomics. First, a custom zebrafish genome was generated based on GRCz10 with the addition
1053 of exogenous transgenes EGFP, tdTomato, and human BRAF-V600E using the command
1054 *cellranger mkref*. Next, the command *cellranger count* was utilized to perform alignment, filtering,

1055 barcode counting and UMI counting of all the samples. The Seurat R package (Version 3.1.4) was
1056 used for quality control, normalization and dimensionality reduction. Low quality cells with a)
1057 features less than 200 or greater than 5000, b) total counts less than 5000, or c) mitochondrial
1058 content greater than 15%, were discarded from the analysis. After filtering, normalization was
1059 performed using Seurat's SCTransform procedure with default parameters to perform a regularized
1060 negative binomial regression based on the 3,000 most variable genes. Uniform Manifold
1061 Approximation and Projection (UMAP) (McInnes et al., 2018) dimensionality reduction was
1062 performed using default parameters. The log transformed normalized count data were extracted
1063 from Seurat and used for downstream analysis in MATLAB. Plots of PRO vs. INV scores were
1064 generated as described in Tirosh et al. (Tirosh et al., 2016) using PRO and INV gene sets defined
1065 from the top 250 most differentially expressed genes in each population by ZMEL1 bulk RNA-
1066 seq. Briefly, log-transformed data were mean-centered. For each gene set a control gene set was
1067 defined to control for variations in sequencing depth and library complexity by randomly selecting
1068 100 genes from the same expression bin (n=25 bins), such that a 50 gene set would have a control
1069 gene set of 5,000 genes. The score for each sample was defined as the mean expression of the gene
1070 set minus the mean expression of the respective control gene set. A binary classifier for PRO versus
1071 INV state was defined by logistic regression on *in vitro* individual culture samples and used to
1072 classify cells from all conditions. Candidate ligand-receptor pairings for *in vivo* samples were
1073 identified using CellPhoneDB (Efremova et al., 2020).

1074

1075 **Analysis of publicly available RNA-seq data**

1076 ***Cancer Cell Line Encyclopedia (CCLE)***

1077 RNA-seq expression data (RSEM genes TPM, version 20180929) were downloaded from the
1078 Broad CCLE (<https://portals.broadinstitute.org/ccle>). Plots of PRO vs. INV scores and their
1079 correlations with TFAP2A mRNA expression were generated as for ZMEL1 scRNA-seq data
1080 above using log-transformed [$\log_2(\text{TPM}+1)$] and mean-centered data with PRO and INV gene sets
1081 from Hoek et al. (Hoek et al., 2006).

1082

1083 *The Cancer Genome Atlas (TCGA)*

1084 Skin Cutaneous Melanoma (SKCM) mRNA expression (v2 RSEM genes normalized, version
1085 2016_01_28, 472 samples from 469 patients) was downloaded from the Broad GDAC Firehose
1086 (<http://firebrowse.org/>). PRO and INV gene expression scores and their correlations with TFAP2A
1087 mRNA expression were calculated and plotted as described for CCLE data above. Log-
1088 transformed normalized expression of TFAP2A and the pan-melanoma markers (Gaynor et al.,
1089 1981; Xiong et al., 2019) S100A1 and S100B were compared between samples from primary
1090 tumors and metastatic sites (Wilcoxon rank sum test with Bonferroni correction).

1091

1092 *Human melanoma single-cell RNA-seq*

1093 Processed data for human melanoma single-cell RNA-seq (Tirosh et al., 2016) were downloaded
1094 from NCBI GEO (GSE72056). PRO and INV gene expression scores and their correlations with
1095 TFAP2A mRNA expression were calculated and plotted as described for CCLE data above.

1096

1097 **DepMap**

1098 TFAP2A was queried through the Broad Institute Dependency Map (DepMap) portal
1099 (<https://depmap.org/portal/>) using the CRISPR (Avana) Public 19Q3 dataset.

1100

1101 **External validation using the AVAST-M melanoma cohort**

1102 Bulk RNA-seq data from 194 primary melanoma patients were extracted from the phase III
1103 adjuvant AVAST-M melanoma cohort (Corrie et al., 2018; Garg et al., 2020). Variance stabilizing
1104 transformation (VST) was applied to the raw counts using the *varianceStabilizingTransformation*
1105 function from the package DESeq2 (Love et al., 2014) (v1.22.2).

1106

1107 ***TFAP2A survival analysis***

1108 VST normalized expression data were used as a continuous variable in a multivariate Cox
1109 regression model, using the *coxph* function of the survival package (Therneau, 2020; Therneau
1110 and Grambsch, 2000) (v2.42-3) in R (v3.5.1). Progression-free survival was calculated as the time
1111 from diagnosis to the last follow-up or death/progression to metastatic disease, whichever occurred
1112 first. The following clinical covariates were considered in the multivariate Cox regression model;
1113 age at diagnosis, gender, stage (AJCC 7th edition), ECOG (Eastern Cooperative Oncology Group
1114 Performance Status), NClass (regional lymph involvement) and treatment (bevacizumab or
1115 placebo) (Garg et al., 2020).

1116

1117 ***PRO/INV survival analysis***

1118 VST expression data corresponding to the genes listed in the PRO and INV gene sets were
1119 extracted (Hoek et al., 2006; Tirosh et al., 2016; Verfaillie et al., 2015; Widmer et al., 2012). For
1120 each sample, the expression values of these genes were standardized to have zero mean and unit
1121 standard deviation. The mean of these standardized expression values was computed to obtain a
1122 vector score corresponding to the PRO and INV expression score vectors. These PRO and INV

1123 expression score vectors were divided into “high” and “low” expression groups using the median
1124 cut-off. Cox regression models were then fitted by means of the coxph function of the survival
1125 package (Therneau, 2020; Therneau and Grambsch, 2000) (v2.42-3) in R (v3.5.1). The hazard
1126 ratio (HR) (95% CI) and p-values corresponding to the “high” expression score group were
1127 reported in both univariate and multivariate analyses.

1128

1129 **External validation using the Leeds Melanoma Cohort**

1130 Primary tumor expression of TFAP2A as well as PRO and INV signatures were tested for
1131 association with melanoma specific survival and relapse-free survival by Cox proportional hazards
1132 in a large population-based cohort (n=703, accession number EGAS00001002922) (Nsengimana
1133 et al., 2018). Signature scores were created by averaging z-transformed gene expressions and
1134 dichotomized by median split.

1135

1136 **Statistical analysis**

1137 Statistical analysis and figure generation were performed in MATLAB (Mathworks, R2016a).
1138 RNA-seq analysis was performed in R (R Foundation for Statistical Computing, 3.4.0). Image
1139 processing and analysis was performed using MATLAB, Zen (Zeiss), ImageJ (NIH), and Volocity
1140 (PerkinElmer, v6.3). The Leeds Melanoma Cohort was analyzed in STATA v14 (StataCorp,
1141 Texas, USA). Unless otherwise noted, bar plots represent mean \pm standard deviation (SD) of
1142 independent experiments, and dots represent means of independent experiments. Abbreviations
1143 for p-values are as follows: * p<0.05, ** p<0.01, *** p<0.001.

1144

1145 **Data and reagent availability**

1146 All RNA-seq data generated in this study are available via the NCBI GEO repository
1147 (GSE151679), with bulk RNA-seq counts and differential expression tables in the Supporting
1148 Information. TFAP2A CUT&RUN data are available via the NCBI GEO repository (GSE153020).
1149 Cell lines generated in this work are available upon request. Raw data are available upon request.

1150

1151 **Code availability**

1152 A MATLAB-based image analysis pipeline for quantifying melanoma in zebrafish was previously
1153 published (Heilmann et al., 2015). Additional scripts are available upon request.

1154

1155 **Figure S1. (Related to Figure 1)**

1156 **a.** Number of ZMEL1 cells migrating in Boyden Chamber assay ($p < 0.001$ by linear regression,
1157 $N = 3$ independent experiments for each of 2 fluorophores). **b.** Log-log plot of mean squared
1158 displacement (MSD) vs. lag time (τ) over the range of $5 \leq \tau < 100$ minutes with model fits
1159 overlaid ($N = 4$ independent experiments, see Methods for details). The slope (α) provides
1160 quantification of the persistence of motility, where a cell moving randomly will have $\alpha = 1$, and a
1161 cell moving along a straight line will have $\alpha = 2$ (Gorelik and Gautreau, 2014). Black line with $\alpha = 1$
1162 is shown for comparison. **c.** Heatmap of genes in the Hoek INV signature that are differentially
1163 expressed in ZMEL1-INV vs. -PRO (\log_2 fold change cutoff ± 1.5 , $p_{\text{adj}} < 0.05$). As in Figure 1e, but
1164 with zebrafish gene names. **d.** Quantification of overall distant metastases seeded by ZMEL1
1165 populations at 3 dpt (OR [95% CI]: 4.49 [1.94, 10.43]; $p < 0.001$ by logistic regression; $N = 3$
1166 independent experiments with PRO/INV 10/10, 31/33, and 13/13 fish per group, respectively;
1167 $n = 110$ fish total; plot shows mean \pm SD). **e.** Representative images of ZMEL1-PRO and -INV
1168 tumors and distant metastases (e.g. to caudal region [box]) at 3 days post-transplant (3dpt). **f-g.**
1169 Quantification of (f) overall metastasis and (g) caudal metastasis seeded by ZMEL1 populations
1170 at 3dpt stratified by small vs. large primary tumor size at 3dpt. ($p = 0.71$ for overall metastasis and
1171 $p = 0.69$ for caudal metastasis for comparison of small vs. large primary tumors by logistic
1172 regression; $N = 3$ independent experiments with PRO/INV 10/10, 31/33, and 13/13 fish per group,
1173 respectively; $n = 110$ fish total; plot shows mean \pm SD).

1174 **Figure S2. (Related to Figure 2)**

1175 **a.** Dual waterfall plot of all gene sets from GO analysis. Adhesion GO gene sets are indicated with
1176 an asterisk and colored according to false discovery rate (FDR). **b-c.** Heatmap of genes in adhesion
1177 GO gene sets (FDR < 0.05, n=3) that are differentially expressed between ZMEL1-PRO and -INV
1178 (\log_2 fold change cutoff ± 1.5 , $p_{\text{adj}} < 0.05$). As in Figure 2b, but with (b) absolute expression data,
1179 and (c) zebrafish gene names. **d.** Representative images of early stages of ZMEL1-INV cluster
1180 formation in Figure 2d. Times are calculated from initial plating of assay. **e.** Quantification of
1181 coefficient of variation ($CV = \sigma/\mu$, where σ is the population standard deviation and μ the
1182 population mean) of cluster area of individual PRO and INV clusters at 2 days ($p=0.026$ by two-
1183 tailed t-test, N=3 independent experiments). **f-g.** Human melanoma samples from (f) The Cancer
1184 Cell Line Encyclopedia (CCLE, n=56) and (g) The Cancer Genome Atlas (TCGA) melanoma
1185 (SKCM, n=472) cohort are plotted as INV (Hoek et al., 2006) versus Cell-Cell Adhesion scores
1186 calculated from RNA-seq. Pearson correlation coefficient between scores is shown. **h.** (left) 3D
1187 opacity rendering and (right) central slice of confocal imaging through co-cluster of ZMEL1-PRO
1188 and ZMEL1-INV. **i.** Quantification of (top) mixing and (bottom) sorting of heterotypic ZMEL1
1189 clusters as in Figure 2f. Red (tdTomato) and green (EGFP) labeled ZMEL1 cells were mixed
1190 PRO:PRO, PRO:INV, and INV:INV at indicated ratios. (top) Quantification of all clusters
1191 revealed that nearly all clusters mix, regardless of cell type. (bottom) Spatial sorting was
1192 significantly enriched in PRO:INV clusters compared to PRO:PRO and INV:INV controls
1193 ($p<0.001$, $p<0.001$, and $p=0.038$ for 1:4, 1:1, and 4:1, respectively, by one-way ANOVA on mean
1194 of each replicate versus respective PRO:PRO and INV:INV controls [greater of two p-values
1195 reported]).

1196 **Figure S3. (Related to Figure 2)**

1197 **a.** Cluster size after 2 days in ZMEL1-INV with CRISPR deletion of *cdh1* (*sg_cdh1*) versus control
1198 (*sg_scr*) ($p=0.0016$, two-tailed paired t-test, $N=5$ independent experiments). **b.** (top)
1199 Representative images of effect of *sg_cdh1* versus *sg_scr* on cluster size at 3 days. (bottom)
1200 Representative images of mixing effects of *sg_cdh1* versus *sg_scr*. Control (*sg_scr*) ZMEL1-INV
1201 cells mixed with ZMEL1-PRO show clear spatial sorting, whereas *sg_cdh1* ZMEL1-INV cells
1202 mixed with ZMEL1-PRO demonstrate decreased sorting. **c.** Western blot confirmed two different
1203 *sg_cdh1*'s decreased Cdh1 protein expression in ZMEL1-INV to a level comparable to ZMEL1-
1204 PRO. *sg_cdh1* (1) was utilized for all phenotypic experiments. **d-e.** ZMEL1-INV orthotopic
1205 primary tumors with *sg_cdh1* do not seed (d) distant metastases and (e) caudal metastases in a
1206 significantly different percentage of zebrafish than with *sg_scr* control ($p=0.56$ and $p=0.44$,
1207 respectively, at 7 dpt by logistic regression; $N=3$ independent experiments with *sg_scr/sg_cdh1*
1208 16/15, 21/20, 19/19 fish per group, respectively; $n=110$ fish total).

1209 **Figure S4. (Related to Figure 3)**

1210 **a.** Number of observed extravasation events of ZMEL1-PRO either alone or in the form of co-
1211 extravasation with ZMEL1-INV following intravenous transplant ($p=0.18$ by two-tailed paired t-
1212 test, $N=4$ independent experiments with 14, 15, 22, and 22 fish each; $n=73$ fish total). **b.** Adult
1213 zebrafish with orthotopic transplants of 1:1 mixtures of ZMEL1-PRO and -INV seed polyclonal
1214 metastases (arrowheads), including to the kidney and caudal regions (left and right boxes,
1215 respectively). **c.** Number of fish co-transplanted with a 1:1 mixture of ZMEL1-PRO and -INV that
1216 have no caudal metastases (None), caudal metastases comprised of exclusively PRO or INV, or
1217 caudal metastases formed by co-metastasis (Co-Met) of both cell types ($N=5$ independent
1218 experiments with 17, 15, 16, 15, and 17 fish each; 80 fish total; $p<0.001$ by Mantel-Haenszel's test
1219 for null hypothesis of no interaction; as in Figure 3c for each independent experiment). **d.** ZMEL1-
1220 PRO and **e.** ZMEL1-INV showed similar levels of overall metastasis at 3 dpt in mixed tumors
1221 compared to when each was transplanted alone ($p=0.83$ and $p=0.13$, respectively, by logistic
1222 regression). **f.** Proportion of fish with caudal ZMEL1-PRO only, ZMEL1-INV only, or PRO/INV
1223 co-metastasis (co-met) at 3dpt (PRO vs. INV: $p=0.0054$; "PRO only" in PRO vs. 1:1 mix: $p=0.033$;
1224 "INV only" in INV vs. 1:1 mix: $p=0.49$; all by logistic regression). Alternate presentation of data
1225 in Figure 3d-e. **g-h.** (g) ZMEL1-PRO and (h) ZMEL1-INV showed similar levels of caudal
1226 metastasis at 7 dpt in mixed tumors compared to when each was transplanted alone ($p=0.54$ and
1227 $p=0.63$, respectively, by logistic regression). For **b-h:** $N=5$ independent experiments with
1228 PRO/MIX/INV 18/17/18, 13/15/14, 15/16/15, 12/15/15, and 15/17/16 fish per group, respectively;
1229 231 fish total; plots show mean \pm SD). **i-j.** Adult zebrafish were orthotopically transplanted with
1230 an equivalent final concentration of tdTomato-labeled ZMEL1-PRO mixed at a 1:4, 4:1, or 9:1
1231 ratio with EGFP-labeled ZMEL1-PRO or ZMEL1-INV. tdTomato-labeled ZMEL1-PRO (i)

1232 overall metastases and (j) caudal metastases were quantified 3dpt to measure the impact of
1233 cooperation with ZMEL1-INV. ZMEL1-PRO metastasized more when co-transplanted with
1234 ZMEL1-INV than with ZMEL1-PRO in Red:Green 1:4 transplants (OR [95% CI]: 2.78 [1.09,
1235 7.10]; p=0.032 by logistic regression). A total of N=3 independent variable ratio mixing
1236 experiments were performed, each with at least 16 fish per group (n=108/113/113 fish per
1237 replicate; n=334 fish total). **k-l.** Proliferation (k) and Boyden chamber migration (l) were
1238 quantified in ZMEL1-PRO and -INV cells with or without conditioned media (CM) from ZMEL1-
1239 PRO and -INV cells (p-values by linear regression; N=3 independent experiments for proliferation;
1240 N=4 independent experiments for migration). **m.** Caspase-3/7 activity normalized to cell number
1241 for clusters of indicated ratio of ZMEL1-PRO and -INV cells at 2 days. P-value for trend of
1242 normalized caspase-3/7 activity with increasing INV composition by linear regression (adjusted
1243 R-squared 0.91). **n.** The interaction between INV (donor) and PRO (recipient) cells can be
1244 schematically represented as a donor-recipient interaction (Hauser et al., 2009) falling into a
1245 regime of cooperation.

1246 **Figure S5. (Related to Figure 4)**

1247 **a.** Boxplots of *tfap2a-e* expression from RNA-seq of ZMEL1-PRO and -INV. **b.** HOMER de-novo
1248 motif analysis of genes differentially expressed between ZMEL1-PRO in 3D (clusters) vs. 2D (no
1249 clusters) (\log_2 fold change cutoff ± 1.5 , $p_{\text{adj}} < 0.05$, $\pm 500\text{bp}$ of transcription start site [TSS]). **c.**
1250 Cluster size in ZMEL1-PRO with CRISPR-Cas9 inactivation of *tfap2a* or *tfap2e* alone and in
1251 combination (p-values by linear regression; N=3 independent experiments). sgRNAs highlighted
1252 in purple (sg_scr) and orange (sg_*tfap2a/e* 1/3) were used for further experiments (Figure 4 and
1253 Figure S5,7). **d-e.** Western blot confirmation of CRISPR inactivation of (d) *tfap2a* and (e) *tfap2e*
1254 with each sgRNA or combination of sgRNAs. **f-i.** Tracking of individual cells by time-lapse
1255 microscopy (N=3 independent experiments). **f.** ZMEL1-PRO with sg_*tfap2a/e* has slowed growth
1256 versus sg_scr ($p < 0.001$ by linear regression, model estimates $\pm 95\%$ CI shown). **g.** Representative
1257 displacements of 500 tracks per sgRNA. **h.** Model estimates $\pm 95\%$ CI for alpha, the slope of the
1258 log-log plot of mean squared displacement vs. lag time (tau) for each ZMEL1-PRO sg_*tfap2a/e*
1259 and sg_scr ($p < 0.001$ by linear regression). Larger alpha indicates more persistent motion. **i.** Log-
1260 log plot of mean squared displacement (MSD) vs. lag time (tau) over the range of $5 \leq \text{tau} < 100$
1261 minutes with model fits overlaid (see Methods for details). The slope (α) provides quantification
1262 of the persistence of motility, where a cell moving randomly will have $\alpha=1$, and a cell moving
1263 along a straight line will have $\alpha=2$ (Gorelik and Gautreau, 2014). Black line with $\alpha=1$ is shown for
1264 comparison. **j-k.** ZMEL1-PRO orthotopic primary tumors with sg_*tfap2a/e* do not seed (j) distant
1265 metastases and (k) caudal metastases in a significantly different proportion of zebrafish than with
1266 sg_scr control ($p=0.44$ and $p=0.90$, respectively, at 7 dpt by logistic regression; N=3 independent
1267 experiments with sg_scr/sg_*tfap2a/e* 24/22, 22/22, 24/24 fish per group, respectively; n=138 fish
1268 total).

1269 **Figure S6. (Related to Figure 5)**

1270 **a.** Analysis of TFAP2A in the Dependency Map (DepMap, CRISPR (Avana) Public 19Q3) dataset
1271 reveals a dependence of melanoma proliferation on TFAP2A. **b-c.** High TFAP2A mRNA
1272 expression in primary tumors predicts worse (b) melanoma specific survival in patients in the
1273 Leeds Melanoma Cohort (HR [95% CI]: 1.6 [1.2, 2.1], $p=0.001$ upper vs. lower half by Cox
1274 proportional hazards and (c) progression free survival in patients in the AVAST-M Melanoma
1275 Cohort (multivariate Cox regression model). **d-g.** Primary tumors with high PRO or low INV
1276 expression are associated with worse outcomes in patients in (d-e) the Leeds Melanoma Cohort
1277 and (f-g) the AVAST-M Melanoma Cohort. **h.** Human melanoma samples from The Cancer
1278 Genome Atlas (TCGA) melanoma (SKCM, $n=472$) cohort are plotted as PRO and INV
1279 scores (Hoek et al., 2006) calculated from RNA-seq and colored according to TFAP2A mRNA
1280 expression. Pearson correlation coefficients between TFAP2A and PRO/INV scores are shown on
1281 axes. **i-j.** Expression of TFAP2A and MITF in human melanoma samples from (i) CCLE ($n=56$)
1282 and (j) TCGA melanoma (SKCM, $n=472$) cohort are plotted with Pearson correlation coefficient.
1283 **k.** Individual human melanoma cells are plotted as PRO and INV scores (Hoek et al., 2006)
1284 calculated from single-cell RNA-seq and colored according to TFAP2A mRNA expression (re-
1285 analyzed from Jerby-Arnon et al. (Jerby-Arnon et al., 2018)). Pearson correlation coefficients
1286 between TFAP2A and PRO/INV scores are shown on axes. **l-m.** (l) S100A1 and (m) S100B
1287 mRNA expression in The Cancer Genome Atlas (TCGA) melanoma (SKCM) cohort comparing
1288 primary tumors and metastases (p -values by Wilcoxon rank sum test with Bonferroni correction).
1289 **n.** Human melanoma cell lines ranked by cluster forming ability (left to right: low to high)
1290 demonstrate negative correlation between TFAP2A mRNA expression by RNA-seq and clustering

1291 (Spearman correlation shown). **o.** Western blot analysis of TFAP2A expression in panels of (left)
1292 low-passage human melanoma cell lines and (right) human melanoma cell lines.

1293 **Figure S7. (Related to Figure 5)**

1294 **a.** Heatmap of top genes in Hoek INV and GO Adhesion gene sets that are differentially expressed
1295 between ZMEL1-PRO *sg_tfap2a/e* and *sg_scr* (\log_2 fold change cutoff ± 0.5 , $p_{\text{adj}} < 0.05$). As in
1296 Figure 5f, but with zebrafish gene names. **b.** Distribution of TFAP2A CUT&RUN peaks as
1297 annotated by ChIPSeeker. **c-d.** Overlap of TFAP2A CUT&RUN peaks with genes upregulated in
1298 ZMEL1-PRO following CRISPR/Cas9 with (c) *sg_scr* ($p=0.7$ by bootstrapping) and (d)
1299 *sg_tfap2a/e* ($p<0.001$ by bootstrapping).

1300 **Figure S8. (Related to Figure 6)**

1301 **a.** Nuclear quantification of TFAP2A and SOX9 in melanoma CTC clusters. n=32 clusters from
1302 four patients. Five additional patients were analyzed with no CTC clusters identified. Clusters
1303 were defined as two or more TFAP2A^{pos}; CD45^{neg} cells. DAPI staining was used to generate
1304 nuclear masks for quantification. Scale bar is 5µm.

1305 **Figure S9. (Related to Figure 7)**

1306 **a.** Single-cell expression of *tfap2e* in ZMEL1-PRO and -INV cells. Individual culture (IND); co-
1307 culture (CO); primary tumors (PRI); metastases (MET). **b-c.** CellPhoneDB results indicating
1308 ligand-receptor pairs that are significantly enriched for the indicated cell-cell interactions in (b)
1309 primary tumors and (c) metastases comprised of both ZMEL1-PRO and -INV. Ligand-receptor
1310 pairings are listed in order of cell pairings on x-axis.

1311 **Supplementary Table 1**

1312 ZMEL1-INV vs. -PRO RNA-seq results. Filtered differential expression (DE; absolute \log_2 fold
1313 change ≥ 1.5 , $p_{\text{adj}} < 0.05$); all DE, log normalized counts, and raw counts.

1314

1315 **Supplementary Table 2**

1316 ZMEL1 3D vs. 2D culture RNA-seq results for ZMEL1-PRO and ZMEL1-INV. Filtered
1317 differential expression (DE; absolute \log_2 fold change ≥ 1.5 , $p_{\text{adj}} < 0.05$); all DE, log normalized
1318 counts, and raw counts.

1319

1320 **Supplementary Table 3**

1321 Annotation of top HOMER motif for 1000bp region (transcription start site ± 500 bp) associated
1322 with genes upregulated in ZMEL1-PRO. Annotation presented for genes differentially expressed
1323 in ZMEL1-PRO and -INV.

1324

1325 **Supplementary Table 4**

1326 Annotation of top HOMER motif for 1000bp region (transcription start site ± 500 bp) associated
1327 with genes differentially expressed in 3D vs. 2D culture for each ZMEL1-PRO and ZMEL1-INV.
1328 Annotations presented for top motif for each ZMEL1-PRO and ZMEL1-INV.

1329

1330 **Supplementary Table 5**

1331 ZMEL1-PRO *sg_tfp2a/e* vs. *sg_scr* RNA-seq results. Filtered differential expression (DE;
1332 absolute \log_2 fold change ≥ 0.5 , $p_{\text{adj}} < 0.05$); all DE, log normalized counts, and raw counts.

1333

1334 **Supplementary Table 6**

1335 List of all gene sets used for GSAA, including sources and full gene lists.

1336

1337 **Supplementary Table 7**

1338 Results from GSAA analyses. Results table and details for gene sets with false discovery rate

1339 (FDR) below 0.05 (up to 6 per analysis).

1340

1341 **Supplementary Table 8**

1342 Sequences for sgRNAs used for CRISPR/Cas9 experiments.

1343

1344 **Supplementary Table 9**

1345 Full zebrafish to human ortholog mapping table. Orthologs with DIOPT Score < 2 were excluded.

1346 A DIOPT Score greater than 6 was considered sufficient for use in GSAA.

1347

1348 **Supplementary Table 10**

1349 Melanoma patient CTC cluster statistics.

1350

1351 **Supplementary Video 1 (Related to Fig. 1h)**

1352 Time lapse confocal microscopy of ZMEL1 cells transplanted intravenously in larval zebrafish.

1353 Arrowhead indicates group of cells invading from the notochord (NC) and caudal hematopoietic

1354 tissue (CHT) into the tail fin mesenchyme (TF).

1355

1356 **Supplementary Video 2 (Related to Fig. 2d)**

1357 Time lapse microscopy of ZMEL1 cluster formation in low-bind plates over 3 days.

1358

1359 **Supplementary Video 3 (Related to Fig. 2f)**

1360 Time lapse microscopy of cluster formation of 1:1 mixture of ZMEL1-PRO and -INV in low-bind
1361 plates over 3 days.

1362

1363 **Supplementary Video 4 (Related to Fig. 3a)**

1364 Time lapse confocal microscopy following intravenous transplant of ZMEL1-PRO and -INV. A
1365 mixed cluster of both ZMEL1-PRO and -INV populations (11h:00m, arrowhead) extravasated
1366 from the caudal hematopoietic tissue (CHT) into the tail fin mesenchyme (TF)—with ZMEL1-
1367 INV leading (19h:00m, arrow) and ZMEL1-PRO following (23h:40m, arrow).

1368

1369 **Graphical Abstract.**

1370 Individual melanoma tumors are comprised of proliferative (PRO) and invasive (INV)
1371 subpopulations that coexist, each with a set of associated phenotypes. TFAP2 acts as a master
1372 regulator of the PRO vs. INV states and clustering, positively regulating melanoma proliferation
1373 while negatively regulating both motility/extravasation and clustering. The interaction of these two
1374 populations in clusters leads to cooperation in the seeding of metastasis, promoting the formation
1375 of heterogenous metastases via collective invasion.

1376 **REFERENCES**

- 1377 Aceto, N., Bardia, A., Miyamoto, D.T., Donaldson, M.C., Wittner, B.S., Spencer, J.A., Yu, M.,
1378 Pely, A., Engstrom, A., Zhu, H., *et al.* (2014). Circulating tumor cell clusters are oligoclonal
1379 precursors of breast cancer metastasis. *Cell* *158*, 1110-1122.
- 1380 Ågren, J.A., Davies, N.G., and Foster, K.R. (2019). Enforcement is central to the evolution of
1381 cooperation. *Nature Ecology & Evolution* *3*, 1018-1029.
- 1382 Archetti, M., and Pienta, K.J. (2019). Cooperation among cancer cells: applying game theory to
1383 cancer. *Nat Rev Cancer* *19*, 110-117.
- 1384 Axelrod, R., Axelrod, D.E., and Pienta, K.J. (2006). Evolution of cooperation among tumor cells.
1385 *Proc Natl Acad Sci U S A* *103*, 13474-13479.
- 1386 Baron, M., Tagore, M., Hunter, M.V., Kim, I.S., Moncada, R., Yan, Y., Campbell, N.R., White,
1387 R.M., and Yanai, I. (2020). The Stress-Like Cancer Cell State Is a Consistent Component of
1388 Tumorigenesis. *Cell Systems* *11*, 536-546.
- 1389 Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R.,
1390 Yakhini, Z., Ben-Dor, A., *et al.* (2000). Molecular classification of cutaneous malignant
1391 melanoma by gene expression profiling. *Nature* *406*, 536-540.
- 1392 Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina
1393 sequence data. *Bioinformatics* *30*, 2114-2120.
- 1394 Cagan, R.L., Zon, L.I., and White, R.M. (2019). Modeling Cancer with Flies and Fish. *Dev Cell*
1395 *49*, 317-324.

1396 Cancer Genome Atlas Network (2015). Genomic Classification of Cutaneous Melanoma. *Cell*
1397 *161*, 1681-1696.

1398 Carreira, S., Goodall, J., Denat, L., Rodriguez, M., Nuciforo, P., Hoek, K.S., Testori, A., Larue,
1399 L., and Goding, C.R. (2006). Mitf regulation of *Dial* controls melanoma proliferation and
1400 invasiveness. *Genes Dev* *20*, 3426-3439.

1401 Celià-Terrassa, T., Meca-Cortés, Ó., Mateo, F., Martínez de Paz, A., Rubio, N., Arnal-Estapé,
1402 A., Ell, B.J., Bermudo, R., Díaz, A., Guerra-Rebollo, M., *et al.* (2012). Epithelial-mesenchymal
1403 transition can suppress major attributes of human epithelial tumor-initiating cells. *The Journal of*
1404 *Clinical Investigation* *122*, 1849-1868.

1405 Ceol, C.J., Houvras, Y., Jane-Valbuena, J., Bilodeau, S., Orlando, D.A., Battisti, V., Fritsch, L.,
1406 Lin, W.M., Hollmann, T.J., Ferré, F., *et al.* (2011). The histone methyltransferase SETDB1 is
1407 recurrently amplified in melanoma and accelerates its onset. *Nature* *471*, 513-517.

1408 Chapman, A., Fernandez del Ama, L., Ferguson, J., Kamarashev, J., Wellbrock, C., and
1409 Hurlstone, A. (2014). Heterogeneous tumor subpopulations cooperate to drive invasion. *Cell Rep*
1410 *8*, 688-695.

1411 Cheng, P.F., Shakhova, O., Widmer, D.S., Eichhoff, O.M., Zingg, D., Frommel, S.C., Belloni,
1412 B., Raaijmakers, M.I.G., Goldinger, S.M., Santoro, R., *et al.* (2015). Methylation-dependent
1413 SOX9 expression mediates invasion in human melanoma cells and is a negative prognostic factor
1414 in advanced melanoma. *Genome Biol* *16*, 42.

1415 Cheung, K.J., Padmanaban, V., Silvestri, V., Schipper, K., Cohen, J.D., Fairchild, A.N., Gorin,
1416 M.A., Verdone, J.E., Pienta, K.J., Bader, J.S., *et al.* (2016). Polyclonal breast cancer metastases

1417 arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc Natl Acad*
1418 *Sci U S A* *113*, E854-863.

1419 Cleary, A.S., Leonard, T.L., Gestl, S.A., and Gunther, E.J. (2014). Tumour cell heterogeneity
1420 maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* *508*, 113-117.

1421 Corrie, P.G., Marshall, A., Nathan, P.D., Lorigan, P., Gore, M., Tahir, S., Faust, G., Kelly, C.G.,
1422 Marples, M., Danson, S.J., *et al.* (2018). Adjuvant bevacizumab for melanoma patients at high
1423 risk of recurrence: survival analysis of the AVAST-M trial. *Ann Oncol* *29*, 1843-1852.

1424 de Croze, N., Maczkowiak, F., and Monsoro-Burq, A.H. (2011). Reiterative AP2a activity
1425 controls sequential steps in the neural crest gene regulatory network. *Proc Natl Acad Sci U S A*
1426 *108*, 155-160.

1427 DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.D., Williams, C., Reich, M.,
1428 Winckler, W., and Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and
1429 process optimization. *Bioinformatics* *28*, 1530-1532.

1430 Dissanayake, S.K., Wade, M., Johnson, C.E., O'Connell, M.P., Leotlela, P.D., French, A.D.,
1431 Shah, K.V., Hewitt, K.J., Rosenthal, D.T., Indig, F.E., *et al.* (2007). The Wnt5A/protein kinase C
1432 pathway mediates motility in melanoma cells via the inhibition of metastasis suppressors and
1433 initiation of an epithelial to mesenchymal transition. *J Biol Chem* *282*, 17259-17271.

1434 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson,
1435 M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* *29*,
1436 15-21.

1437 Efremova, M., Vento-Tormo, M., Teichmann, S.A., and Vento-Tormo, R. (2020). CellPhoneDB:
1438 inferring cell–cell communication from combined expression of multi-subunit ligand–receptor
1439 complexes. *Nat Protoc* 15, 1484-1506.

1440 Eichhoff, O.M., Zipser, M.C., Xu, M., Weeraratna, A.T., Mihic, D., Dummer, R., and Hoek, K.S.
1441 (2010). The immunohistochemistry of invasive and proliferative phenotype switching in
1442 melanoma: a case report. *Melanoma Res* 20, 349-355.

1443 Falletta, P., Sanchez-Del-Campo, L., Chauhan, J., Effern, M., Kenyon, A., Kershaw, C.J.,
1444 Siddaway, R., Lisle, R., Freter, R., Daniels, M.J., *et al.* (2017). Translation reprogramming is an
1445 evolutionarily conserved driver of phenotypic plasticity and therapeutic resistance in melanoma.
1446 *Genes Dev* 31, 18-33.

1447 Fane, M.E., Chhabra, Y., Smith, A.G., and Sturm, R.A. (2019). BRN2, a POUerful driver of
1448 melanoma phenotype switching and metastasis. *Pigment Cell Melanoma Res* 32, 9-24.

1449 Fidler, I.J. (1973). The relationship of embolic homogeneity, number, size and viability to the
1450 incidence of experimental metastasis. *Eur J Cancer* 9, 223-227.

1451 Finger, E.C., Cheng, C.F., Williams, T.R., Rankin, E.B., Bedogni, B., Tachiki, L., Spong, S.,
1452 Giaccia, A.J., and Powell, M.B. (2014). CTGF is a therapeutic target for metastatic melanoma.
1453 *Oncogene* 33, 1093-1100.

1454 Foster, K.R. (2011). The sociobiology of molecular systems. *Nature Reviews Genetics* 12, 193-
1455 203.

1456 Foty, R.A., and Steinberg, M.S. (2005). The differential adhesion hypothesis: a direct evaluation.
1457 *Dev Biol* 278, 255-263.

1458 Garg, M., Couturier, D.-L., Nsengimana, J., Fonseca, N.A., Wongchenko, M., Yan, Y., Lauss,
1459 M., Jönsson, G.B., Newton-Bishop, J., Parkinson, C., *et al.* (2020). Tumour gene expression
1460 signature in primary melanoma predicts long-term outcomes: A prospective multicentre study.
1461 *bioRxiv*, 2020.2002.2024.961771.

1462 Gaynor, R., Herschman, H.R., Irie, R., Jones, P., Morton, D., and Cochran, A. (1981). S100
1463 protein: a marker for human malignant melanomas? *Lancet* 1, 869-871.

1464 Gershenwald, J., Scolyer, R., Hess, K., and *et al.* (2017). Melanoma of the Skin. In *AJCC Cancer*
1465 *Staging Manual: 8th Edition*, M.B. Amin, ed. (Chicago: American Joint Committee on Cancer),
1466 p. 563.

1467 Ghandi, M., Huang, F.W., Jané-Valbuena, J., Kryukov, G.V., Lo, C.C., McDonald, E.R.,
1468 Barretina, J., Gelfand, E.T., Bielski, C.M., Li, H., *et al.* (2019). Next-generation characterization
1469 of the Cancer Cell Line Encyclopedia. *Nature* 569, 503-508.

1470 Giuliano, M., Shaikh, A., Lo, H.C., Arpino, G., De Placido, S., Zhang, X.H., Cristofanilli, M.,
1471 Schiff, R., and Trivedi, M.V. (2018). Perspective on Circulating Tumor Cell Clusters: Why It
1472 Takes a Village to Metastasize. *Cancer Res* 78, 845-852.

1473 Gkoutela, S., Castro-Giner, F., Szczerba, B.M., Vetter, M., Landin, J., Scherrer, R., Krol, I.,
1474 Scheidmann, M.C., Beisel, C., Stirnimann, C.U., *et al.* (2019). Circulating Tumor Cell Clustering
1475 Shapes DNA Methylation to Enable Metastasis Seeding. *Cell* 176, 98-112 e114.

1476 Glaves, D. (1983). Correlation between circulating cancer cells and incidence of metastases. *Br J*
1477 *Cancer* 48, 665-673.

1478 Gorelik, R., and Gautreau, A. (2014). Quantitative and unbiased analysis of directional
1479 persistence in cell migration. *Nat Protocols* 9, 1931-1943.

1480 Gupta, P.B., Kuperwasser, C., Brunet, J.P., Ramaswamy, S., Kuo, W.L., Gray, J.W., Naber, S.P.,
1481 and Weinberg, R.A. (2005). The melanocyte differentiation program predisposes to metastasis
1482 after neoplastic transformation. *Nat Genet* 37, 1047-1054.

1483 Hallberg, A.R., Vorrink, S.U., Hudachek, D.R., Cramer-Morales, K., Milhem, M.M., Cornell,
1484 R.A., and Domann, F.E. (2014). Aberrant CpG methylation of the TFAP2A gene constitutes a
1485 mechanism for loss of TFAP2A expression in human metastatic melanoma. *Epigenetics* 9, 1641-
1486 1647.

1487 Harmeyer, K.M., Facompre, N.D., Herlyn, M., and Basu, D. (2017). JARID1 Histone
1488 Demethylases: Emerging Targets in Cancer. *Trends Cancer* 3, 713-725.

1489 Hatzikirou, H., Basanta, D., Simon, M., Schaller, K., and Deutsch, A. (2010). ‘Go or Grow’: the
1490 key to the emergence of invasion in tumour progression? *Mathematical Medicine and Biology: A*
1491 *Journal of the IMA* 29, 49-65.

1492 Hauser, M., McAuliffe, K., and Blake, P.R. (2009). Evolving the ingredients for reciprocity and
1493 spite. *Philos Trans R Soc Lond B Biol Sci* 364, 3255-3266.

1494 Heilmann, S., Ratnakumar, K., Langdon, E., Kansler, E., Kim, I., Campbell, N.R., Perry, E.,
1495 McMahon, A., Kaufman, C., van Rooijen, E., *et al.* (2015). A quantitative system for studying
1496 metastasis using transparent zebrafish. *Cancer Res* 75, 4272-4282.

1497 Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C.,
1498 Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription
1499 factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38,
1500 576-589.

1501 Hinohara, K., and Polyak, K. (2019). Intratumoral Heterogeneity: More Than Just Mutations.
1502 *Trends Cell Biol* 29, 569-579.

1503 Hoek, K.S., Eichhoff, O.M., Schlegel, N.C., Dobbeling, U., Kobert, N., Schaerer, L., Hemmi, S.,
1504 and Dummer, R. (2008). In vivo switching of human melanoma cells between proliferative and
1505 invasive states. *Cancer Res* 68, 650-656.

1506 Hoek, K.S., and Goding, C.R. (2010). Cancer stem cells versus phenotype-switching in
1507 melanoma. *Pigment Cell Melanoma Res* 23, 746-759.

1508 Hoek, K.S., Schlegel, N.C., Brafford, P., Sucker, A., Ugurel, S., Kumar, R., Weber, B.L.,
1509 Nathanson, K.L., Phillips, D.J., Herlyn, M., *et al.* (2006). Metastatic potential of melanomas
1510 defined by specific gene expression profiles with no BRAF signature. *Pigment Cell Res* 19, 290-
1511 302.

1512 Hoffman, T.L., Javier, A.L., Campeau, S.A., Knight, R.D., and Schilling, T.F. (2007). Tfp2
1513 transcription factors in zebrafish neural crest development and ectodermal evolution. *J Exp Zool*
1514 *B Mol Dev Evol* 308, 679-691.

1515 Houghton, A.N., Real, F.X., Davis, L.J., Cordon-Cardo, C., and Old, L.J. (1987). Phenotypic
1516 heterogeneity of melanoma. Relation to the differentiation program of melanoma cells. *J Exp*
1517 *Med* *165*, 812-829.

1518 Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., and Mohr, S.E.
1519 (2011). An integrative approach to ortholog prediction for disease-focused and other functional
1520 studies. *BMC Bioinformatics* *12*, 357.

1521 Huang, S., Jean, D., Luca, M., Tainsky, M.A., and Bar-Eli, M. (1998). Loss of AP-2 results in
1522 downregulation of c-KIT and enhancement of melanoma tumorigenicity and metastasis. *EMBO J*
1523 *17*, 4358-4369.

1524 Isogai, S., Horiguchi, M., and Weinstein, B.M. (2001). The vascular anatomy of the developing
1525 zebrafish: an atlas of embryonic and early larval development. *Dev Biol* *230*, 278-301.

1526 Jerby-Arnon, L., Shah, P., Cuoco, M.S., Rodman, C., Su, M.J., Melms, J.C., Leeson, R.,
1527 Kanodia, A., Mei, S., Lin, J.R., *et al.* (2018). A Cancer Cell Program Promotes T Cell Exclusion
1528 and Resistance to Checkpoint Blockade. *Cell* *175*, 984-997 e924.

1529 Johnston, L.A. (2014). Socializing with MYC: cell competition in development and as a model
1530 for premalignant cancer. *Cold Spring Harb Perspect Med* *4*, a014274.

1531 Kaufman, C.K., Mosimann, C., Fan, Z.P., Yang, S., Thomas, A.J., Ablain, J., Tan, J.L., Fogley,
1532 R.D., van Rooijen, E., Hagedorn, E.J., *et al.* (2016). A zebrafish melanoma model reveals
1533 emergence of neural crest identity during melanoma initiation. *Science* *351*, aad2197.

1534 Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R.,
1535 Bessy, A., Chèneby, J., Kulkarni, S.R., Tan, G., *et al.* (2017). JASPAR 2018: update of the open-
1536 access database of transcription factor binding profiles and its web framework. *Nucleic Acids*
1537 *Res* 46, D260-D266.

1538 Khoja, L., Shenjere, P., Hodgson, C., Hodgetts, J., Clack, G., Hughes, A., Lorigan, P., and Dive,
1539 C. (2014). Prevalence and heterogeneity of circulating tumour cells in metastatic cutaneous
1540 melanoma. *Melanoma Res* 24, 40-46.

1541 Kim, I.S., Heilmann, S., Kansler, E.R., Zhang, Y., Zimmer, M., Ratnakumar, K., Bowman, R.L.,
1542 Simon-Vermot, T., Fennell, M., Garippa, R., *et al.* (2017). Microenvironment-derived factors
1543 driving metastatic plasticity in melanoma. *Nat Commun* 8, 14343.

1544 Konieczkowski, D.J., Johannessen, C.M., Abudayyeh, O., Kim, J.W., Cooper, Z.A., Piris, A.,
1545 Frederick, D.T., Barzily-Rokni, M., Straussman, R., Haq, R., *et al.* (2014). A melanoma cell state
1546 distinction influences sensitivity to MAPK pathway inhibitors. *Cancer Discov* 4, 816-827.

1547 Korolev, K.S., Xavier, J.B., and Gore, J. (2014). Turning ecology and evolution against cancer.
1548 *Nat Rev Cancer* 14, 371-380.

1549 Kwan, K.M., Fujimoto, E., Grabher, C., Mangum, B.D., Hardy, M.E., Campbell, D.S., Parant,
1550 J.M., Yost, H.J., Kanki, J.P., and Chien, C.B. (2007). The Tol2kit: a multisite gateway-based
1551 construction kit for Tol2 transposon transgenesis constructs. *Dev Dyn* 236, 3088-3099.

1552 Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat*
1553 *Methods* 9, 357-359.

1554 Lauden, L., Siewiera, J., Boukouaci, W., Ramgolam, K., Mourah, S., Lebbe, C., Charron, D.,
1555 Aoudjit, F., Jabrane-Ferrat, N., and Al-Daccak, R. (2014). TGF-beta-induced (TGFB1) protein in
1556 melanoma: a signature of high metastatic potential. *J Invest Dermatol* *134*, 1675-1685.

1557 Li, F.Z., Dhillon, A.S., Anderson, R.L., McArthur, G., and Ferrao, P.T. (2015). Phenotype
1558 switching in melanoma: implications for progression and therapy. *Front Oncol* *5*, 31.

1559 Li, W., and Cornell, R.A. (2007). Redundant activities of Tfp2a and Tfp2c are required for
1560 neural crest induction and development of other non-neural ectoderm derivatives in zebrafish
1561 embryos. *Dev Biol* *304*, 338-354.

1562 Liotta, L.A., Saidel, M.G., and Kleinerman, J. (1976). The significance of hematogenous tumor
1563 cell clumps in the metastatic process. *Cancer Res* *36*, 889-894.

1564 Liu, X., Taftaf, R., Kawaguchi, M., Chang, Y.F., Chen, W., Entenberg, D., Zhang, Y., Gerratana,
1565 L., Huang, S., Patel, D.B., *et al.* (2018). Homophilic CD44 interactions mediate tumor cell
1566 aggregation and polyclonal metastasis in patient-derived breast cancer models. *Cancer Discov* *9*,
1567 96-113.

1568 Long, E., Ilie, M., Bence, C., Butori, C., Selva, E., Lalvee, S., Bonnetaud, C., Poissonnet, G.,
1569 Lacour, J.P., Bahadoran, P., *et al.* (2016). High expression of TRF2, SOX10, and CD10 in
1570 circulating tumor microemboli detected in metastatic melanoma patients. A potential impact for
1571 the assessment of disease aggressiveness. *Cancer Med* *5*, 1022-1030.

1572 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
1573 dispersion for RNA-seq data with DESeq2. *Genome Biol* *15*, 550.

1574 Luo, T., Matsuo-Takasaki, M., Thomas, M.L., Weeks, D.L., and Sargent, T.D. (2002).
1575 Transcription factor AP-2 is an essential and direct regulator of epidermal development in
1576 *Xenopus*. *Dev Biol* 245, 136-144.

1577 Luo, X., Mitra, D., Sullivan, R.J., Wittner, B.S., Kimura, A.M., Pan, S., Hoang, M.P., Brannigan,
1578 B.W., Lawrence, D.P., Flaherty, K.T., *et al.* (2014). Isolation and molecular characterization of
1579 circulating melanoma cells. *Cell Rep* 7, 645-653.

1580 Maddipati, R., and Stanger, B.Z. (2015). Pancreatic Cancer Metastases Harbor Evidence of
1581 Polyclonality. *Cancer Discov* 5, 1086-1097.

1582 Marjanovic, N.D., Hofree, M., Chan, J.E., Canner, D., Wu, K., Trakala, M., Hartmann, G.G.,
1583 Smith, O.C., Kim, J.Y., Evans, K.V., *et al.* (2020). Emergence of a High-Plasticity Cell State
1584 during Lung Cancer Evolution. *Cancer Cell* 38, 229-246.e213.

1585 Marusyk, A., Tabassum, D.P., Altrock, P.M., Almendro, V., Michor, F., and Polyak, K. (2014).
1586 Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* 514,
1587 54-58.

1588 Matus, D.Q., Lohmer, L.L., Kelley, L.C., Schindler, A.J., Kohrman, A.Q., Barkoulas, M., Zhang,
1589 W., Chi, Q., and Sherwood, D.R. (2015). Invasive Cell Fate Requires G1 Cell-Cycle Arrest and
1590 Histone Deacetylase-Mediated Changes in Gene Expression. *Dev Cell* 35, 162-174.

1591 Mayhew, E., and Glaves, D. (1984). Quantitation of tumorigenic disseminating and arrested
1592 cancer cells. *Br J Cancer* 50, 159-166.

1593 McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and
1594 Projection for Dimension Reduction. arXiv, eprint 1802.03426.

1595 Meers, M.P., Tenenbaum, D., and Henikoff, S. (2019). Peak calling by Sparse Enrichment
1596 Analysis for CUT&RUN chromatin profiling. *Epigenetics Chromatin* 12, 42.

1597 Mrozik, K.M., Blaschuk, O.W., Cheong, C.M., Zannettino, A.C.W., and Vandyke, K. (2018). N-
1598 cadherin in cancer metastasis, its emerging role in haematological malignancies and potential as
1599 a therapeutic target in cancer. *BMC Cancer* 18, 939.

1600 Muller, J., Krijgsman, O., Tsoi, J., Robert, L., Hugo, W., Song, C., Kong, X., Possik, P.A.,
1601 Cornelissen-Steijger, P.D., Geukes Foppen, M.H., *et al.* (2014). Low MITF/AXL ratio predicts
1602 early resistance to multiple targeted drugs in melanoma. *Nat Commun* 5, 5712.

1603 Neelakantan, D., Zhou, H., Oliphant, M.U.J., Zhang, X., Simon, L.M., Henke, D.M., Shaw,
1604 C.A., Wu, M.F., Hilsenbeck, S.G., White, L.D., *et al.* (2017). EMT cells increase breast cancer
1605 metastasis via paracrine GLI activation in neighbouring tumour cells. *Nat Commun* 8, 15773.

1606 Nowell, P.C. (1976). The clonal evolution of tumor cell populations. *Science* 194, 23-28.

1607 Nsengimana, J., Laye, J., Filia, A., O'Shea, S., Muralidhar, S., Pozniak, J., Droop, A., Chan, M.,
1608 Walker, C., Parkinson, L., *et al.* (2018). beta-Catenin-mediated immune evasion pathway
1609 frequently operates in primary cutaneous melanomas. *J Clin Invest* 128, 2048-2063.

1610 Padmanaban, V., Krol, I., Suhail, Y., Szczerba, B.M., Aceto, N., Bader, J.S., and Ewald, A.J.
1611 (2019). E-cadherin is required for metastasis in multiple models of breast cancer. *Nature* 573,
1612 439-444.

1613 Patton, E.E., Widlund, H.R., Kutok, J.L., Kopani, K.R., Amatruda, J.F., Murphey, R.D.,
1614 Berghmans, S., Mayhall, E.A., Traver, D., Fletcher, C.D., *et al.* (2005). BRAF mutations are
1615 sufficient to promote nevi formation and cooperate with p53 in the genesis of melanoma. *Current*
1616 *biology* : CB 15, 249-254.

1617 Pearson, G.W. (2019). Control of Invasion by Epithelial-to-Mesenchymal Transition Programs
1618 during Metastasis. *J Clin Med* 8.

1619 Pinner, S., Jordan, P., Sharrock, K., Bazley, L., Collinson, L., Marais, R., Bonvin, E., Goding,
1620 C., and Sahai, E. (2009). Intravital imaging reveals transient changes in pigment production and
1621 Brn2 expression during metastatic melanoma dissemination. *Cancer Res* 69, 7969-7977.

1622 Raaijmakers, M.I.G., Widmer, D.S., Maudrich, M., Koch, T., Langer, A., Flace, A., Schnyder,
1623 C., Dummer, R., and Levesque, M.P. (2015). A new live-cell biobank workflow efficiently
1624 recovers heterogeneous melanoma cells from native biopsies. *Exp Dermatol* 24, 377-380.

1625 Rabbie, R., Ansari-Pour, N., Cast, O., Lau, D., Scott, F., Welsh, S.J., Parkinson, C., Khoja, L.,
1626 Moore, L., Tullett, M., *et al.* (2019). Multi-site clonality analyses uncovers pervasive subclonal
1627 heterogeneity and branching evolution across melanoma metastases. *bioRxiv*, 848390.

1628 Rambow, F., Job, B., Petit, V., Gesbert, F., Delmas, V., Seberg, H., Meurice, G., Van Otterloo,
1629 E., Dessen, P., Robert, C., *et al.* (2015). New Functional Signatures for Understanding
1630 Melanoma Biology from Tumor Cell Lineage-Specific Analysis. *Cell Reports* 13, 840-853.

1631 Rambow, F., Marine, J.C., and Goding, C.R. (2019). Melanoma plasticity and phenotypic
1632 diversity: therapeutic barriers and opportunities. *Genes Dev* 33, 1295-1318.

1633 Rambow, F., Rogiers, A., Marin-Bejar, O., Aibar, S., Femel, J., Dewaele, M., Karras, P., Brown,
1634 D., Chang, Y.H., Debiec-Rychter, M., *et al.* (2018). Toward Minimal Residual Disease-Directed
1635 Therapy in Melanoma. *Cell* *174*, 843-855 e819.

1636 Reichert, M., Bakir, B., Moreira, L., Pitarresi, J.R., Feldmann, K., Simon, L., Suzuki, K.,
1637 Maddipati, R., Rhim, A.D., Schlitter, A.M., *et al.* (2018). Regulation of Epithelial Plasticity
1638 Determines Metastatic Organotropism in Pancreatic Cancer. *Dev Cell* *45*, 696-711 e698.

1639 Ren, Y., and Liao, W.S. (2001). Transcription factor AP-2 functions as a repressor that
1640 contributes to the liver-specific expression of serum amyloid A1 gene. *J Biol Chem* *276*, 17770-
1641 17778.

1642 Roesch, A., Fukunaga-Kalabis, M., Schmidt, E.C., Zabierowski, S.E., Brafford, P.A., Vultur, A.,
1643 Basu, D., Gimotty, P., Vogt, T., and Herlyn, M. (2010). A temporarily distinct subpopulation of
1644 slow-cycling melanoma cells is required for continuous tumor growth. *Cell* *141*, 583-594.

1645 Rowling, E.J., Miskolczi, Z., Nagaraju, R., Wilcock, D.J., Wang, P., Telfer, B., Li, Y., Lasheras-
1646 Otero, I., Redondo-Munoz, M., Sharrocks, A.D., *et al.* (2020). Cooperative behaviour and
1647 phenotype plasticity evolve during melanoma progression. *Pigment Cell Melanoma Res.*

1648 Ruiz, C., Li, J., Luttgen, M.S., Kolatkar, A., Kendall, J.T., Flores, E., Topp, Z., Samlowski,
1649 W.E., McClay, E., Bethel, K., *et al.* (2015). Limited genomic heterogeneity of circulating
1650 melanoma cells in advanced stage patients. *Phys Biol* *12*, 016008.

1651 Sanborn, J.Z., Chung, J., Purdom, E., Wang, N.J., Kakavand, H., Wilmott, J.S., Butler, T.,
1652 Thompson, J.F., Mann, G.J., Haydu, L.E., *et al.* (2015). Phylogenetic analyses of melanoma

1653 reveal complex patterns of metastatic dissemination. *Proc Natl Acad Sci U S A* *112*, 10995-
1654 11000.

1655 Sarioglu, A.F., Aceto, N., Kojic, N., Donaldson, M.C., Zeinali, M., Hamza, B., Engstrom, A.,
1656 Zhu, H., Sundaresan, T.K., Miyamoto, D.T., *et al.* (2015). A microfluidic device for label-free,
1657 physical capture of circulating tumor cell clusters. *Nat Methods* *12*, 685-691.

1658 Seberg, H.E., Van Otterloo, E., and Cornell, R.A. (2017a). Beyond MITF: Multiple transcription
1659 factors directly regulate the cellular phenotype in melanocytes and melanoma. *Pigment Cell*
1660 *Melanoma Res* *30*, 454-466.

1661 Seberg, H.E., Van Otterloo, E., Loftus, S.K., Liu, H., Bonde, G., Sompallae, R., Gildea, D.E.,
1662 Santana, J.F., Manak, J.R., Pavan, W.J., *et al.* (2017b). TFAP2 paralogs regulate melanocyte
1663 differentiation in parallel with MITF. *PLoS Genet* *13*, e1006636.

1664 Shakhova, O., Cheng, P., Mishra, P.J., Zingg, D., Schaefer, S.M., Debbache, J., Häusel, J.,
1665 Matter, C., Guo, T., Davis, S., *et al.* (2015). Antagonistic Cross-Regulation between Sox9 and
1666 Sox10 Controls an Anti-tumorigenic Program in Melanoma. *PLoS Genet* *11*, e1004877.

1667 Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution
1668 mapping of DNA binding sites. *Elife* *6*.

1669 Smeets, B., Alert, R., Pešek, J., Pagonabarraga, I., Ramon, H., and Vincent, R. (2016). Emergent
1670 structures and dynamics of cell colonies by contact inhibition of locomotion. *Proceedings of the*
1671 *National Academy of Sciences* *113*, 14621-14626.

1672 Streit, M., and Detmar, M. (2003). Angiogenesis, lymphangiogenesis, and melanoma metastasis.
1673 *Oncogene* 22, 3172-3179.

1674 Szczerba, B.M., Castro-Giner, F., Vetter, M., Krol, I., Gkoutela, S., Landin, J., Scheidmann,
1675 M.C., Donato, C., Scherrer, R., Singer, J., *et al.* (2019). Neutrophils escort circulating tumour
1676 cells to enable cell cycle progression. *Nature* 566, 553-557.

1677 Tellez, C.S., Davis, D.W., Prieto, V.G., Gershenwald, J.E., Johnson, M.M., McCarty, M.F., and
1678 Bar-Eli, M. (2007). Quantitative analysis of melanocytic tissue array reveals inverse correlation
1679 between activator protein-2alpha and protease-activated receptor-1 expression during melanoma
1680 progression. *J Invest Dermatol* 127, 387-393.

1681 Therneau, T.M. (2020). *A Package for Survival Analysis in R*.

1682 Therneau, T.M., and Grambsch, P.M. (2000). *Modeling Survival Data: Extending the Cox Model*
1683 (New York: Springer).

1684 Thompson, J.F., Soong, S.J., Balch, C.M., Gershenwald, J.E., Ding, S., Coit, D.G., Flaherty,
1685 K.T., Gimotty, P.A., Johnson, T., Johnson, M.M., *et al.* (2011). Prognostic significance of
1686 mitotic rate in localized primary cutaneous melanoma: an analysis of patients in the multi-
1687 institutional American Joint Committee on Cancer melanoma staging database. *J Clin Oncol* 29,
1688 2199-2205.

1689 Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A.,
1690 Rodman, C., Lian, C., Murphy, G., *et al.* (2016). Dissecting the multicellular ecosystem of
1691 metastatic melanoma by single-cell RNA-seq. *Science* 352, 189-196.

1692 Tsoi, J., Robert, L., Paraiso, K., Galvan, C., Sheu, K.M., Lay, J., Wong, D.J.L., Atefi, M.,
1693 Shirazi, R., Wang, X., *et al.* (2018). Multi-stage Differentiation Defines Melanoma Subtypes
1694 with Differential Vulnerability to Drug-Induced Iron-Dependent Oxidative Stress. *Cancer Cell*
1695 *33*, 890-904 e895.

1696 Tsuji, T., Ibaragi, S., and Hu, G.F. (2009). Epithelial-mesenchymal transition and cell
1697 cooperativity in metastasis. *Cancer Res* *69*, 7135-7139.

1698 Tuncer, E., Calçada, R.R., Zingg, D., Varum, S., Cheng, P., Freiburger, S.N., Deng, C.-X.,
1699 Kleiter, I., Levesque, M.P., Dummer, R., *et al.* (2019). SMAD signaling promotes melanoma
1700 metastasis independently of phenotype switching. *The Journal of Clinical Investigation* *129*,
1701 2702-2716.

1702 Van Otterloo, E., Li, W., Bonde, G., Day, K.M., Hsu, M.Y., and Cornell, R.A. (2010).
1703 Differentiation of zebrafish melanophores depends on transcription factors AP2 alpha and AP2
1704 epsilon. *PLoS Genet* *6*, e1001122.

1705 Vandamme, N., and Berx, G. (2014). Melanoma cells revive an embryonic transcriptional
1706 network to dictate phenotypic heterogeneity. *Front Oncol* *4*, 352.

1707 Verfaillie, A., Imrichova, H., Atak, Z.K., Dewaele, M., Rambow, F., Hulselmans, G.,
1708 Christiaens, V., Svetlichnyy, D., Luciani, F., Van den Mooter, L., *et al.* (2015). Decoding the
1709 regulatory landscape of melanoma reveals TEADS as regulators of the invasive cell state. *Nat*
1710 *Commun* *6*, 6683.

1711 Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., and
1712 Sabatini, D.M. (2015). Identification and characterization of essential genes in the human
1713 genome. *Science* 350, 1096-1101.

1714 Watanabe, S. (1954). The metastasizability of tumor cells. *Cancer* 7, 215-223.

1715 Weeraratna, A.T., Jiang, Y., Hostetter, G., Rosenblatt, K., Duray, P., Bittner, M., and Trent, J.M.
1716 (2002). Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma.
1717 *Cancer Cell* 1, 279-288.

1718 White, R.M., Cech, J., Ratanasirintrawoot, S., Lin, C.Y., Rahl, P.B., Burke, C.J., Langdon, E.,
1719 Tomlinson, M.L., Mosher, J., Kaufman, C., *et al.* (2011). DHODH modulates transcriptional
1720 elongation in the neural crest and melanoma. *Nature* 471, 518-522.

1721 Widmer, D.S., Cheng, P.F., Eichhoff, O.M., Belloni, B.C., Zipser, M.C., Schlegel, N.C.,
1722 Javelaud, D., Mauviel, A., Dummer, R., and Hoek, K.S. (2012). Systematic classification of
1723 melanoma cells by phenotype-specific gene expression mapping. *Pigment Cell Melanoma Res*
1724 25, 343-353.

1725 Wouters, J., Kalender-Atak, Z., Minnoye, L., Spanier, K.I., De Waegeneer, M., Bravo González-
1726 Blas, C., Mauduit, D., Davie, K., Hulselmans, G., Najem, A., *et al.* (2020). Robust gene
1727 expression programs underlie recurrent cell states and phenotype switching in melanoma. *Nat*
1728 *Cell Biol* 22, 986-998.

1729 Xiong, Q., Ancona, N., Hauser, E.R., Mukherjee, S., and Furey, T.S. (2012). Integrating genetic
1730 and gene expression evidence into genome-wide association analysis of gene sets. *Genome Res*
1731 22, 386-397.

1732 Xiong, Q., Mukherjee, S., and Furey, T.S. (2014). GSAASeqSP: A Toolset for Gene Set
1733 Association Analysis of RNA-Seq Data. *Scientific Reports* 4, 6347.

1734 Xiong, T.-f., Pan, F.-q., and Li, D. (2019). Expression and clinical significance of S100 family
1735 genes in patients with melanoma. *Melanoma Res* 29, 23-29.

1736 Xu, L., Mao, X., Imrali, A., Syed, F., Mutsvangwa, K., Berney, D., Cathcart, P., Hines, J.,
1737 Shamash, J., and Lu, Y.J. (2015). Optimization and Evaluation of a Novel Size Based
1738 Circulating Tumor Cell Isolation System. *PLoS ONE* 10, e0138032.

1739 Yu, G., Wang, L.-G., and He, Q.-Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP
1740 peak annotation, comparison and visualization. *Bioinformatics* 31, 2382-2383.

1741 Zeng, L., Jarrett, C., Brown, K., Gillespie, K.M., Holly, J.M., and Perks, C.M. (2013). Insulin-
1742 like growth factor binding protein-3 (IGFBP-3) plays a role in the anti-tumorigenic effects of 5-
1743 Aza-2'-deoxycytidine (AZA) in breast cancer cells. *Exp Cell Res* 319, 2282-2295.

1744 Zhang, Y.M., Zimmer, M.A., Guardia, T., Callahan, S.J., Mondal, C., Di Martino, J., Takagi, T.,
1745 Fennell, M., Garippa, R., Campbell, N.R., *et al.* (2018). Distant Insulin Signaling Regulates
1746 Vertebrate Pigmentation through the Sheddase Bace2. *Dev Cell* 45, 580-594.e587.

1747

Figure 1

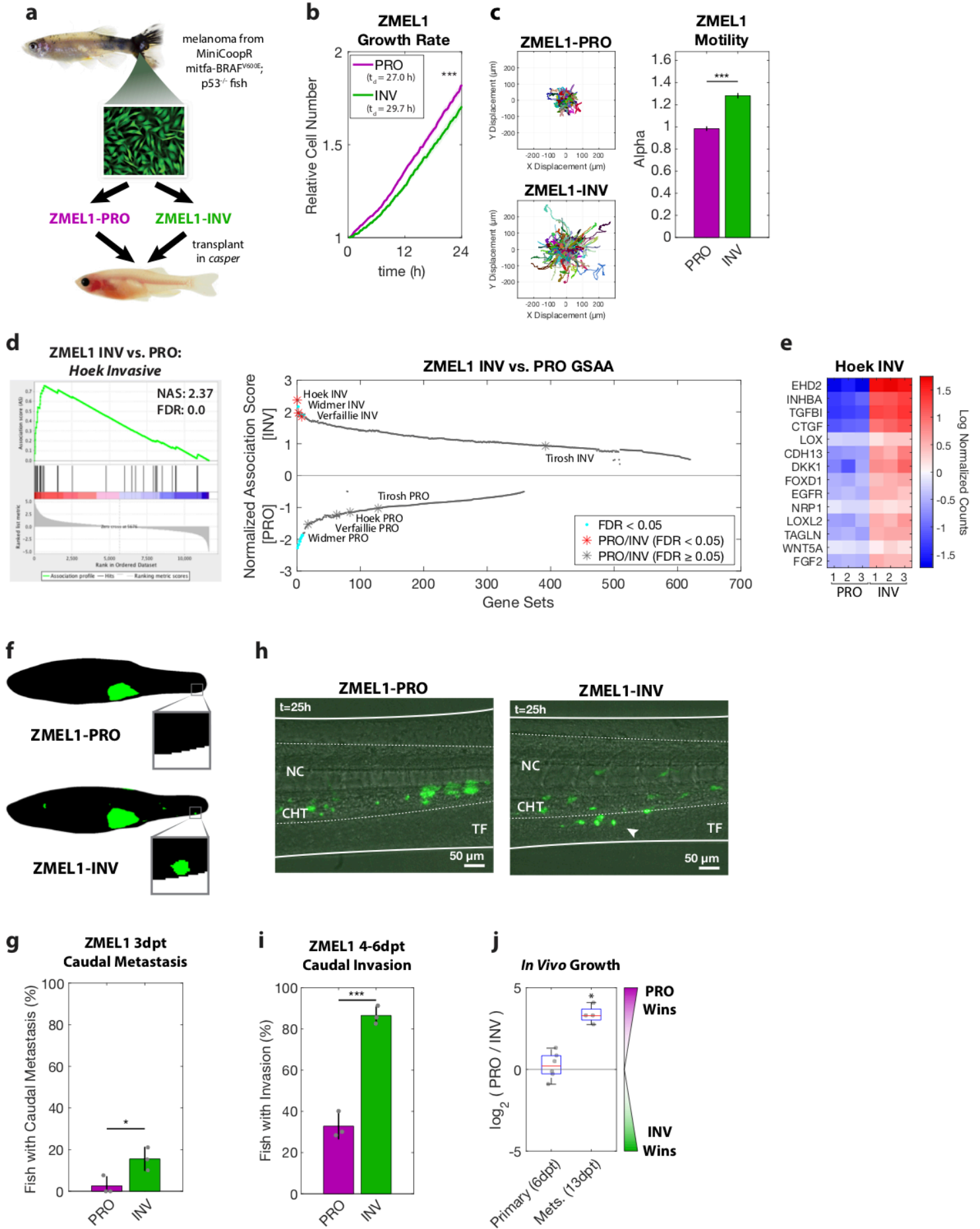


Figure 2

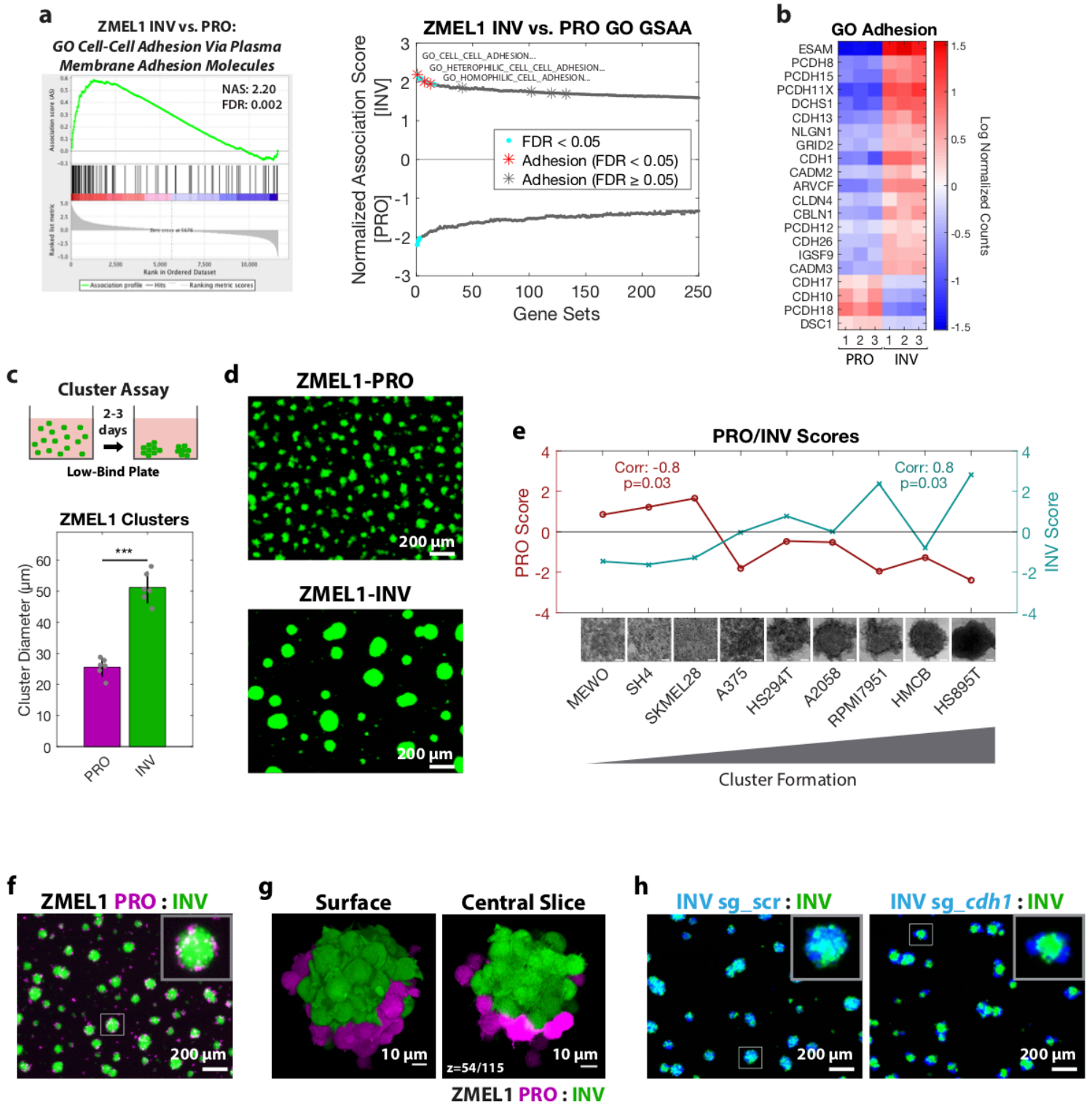


Figure 3

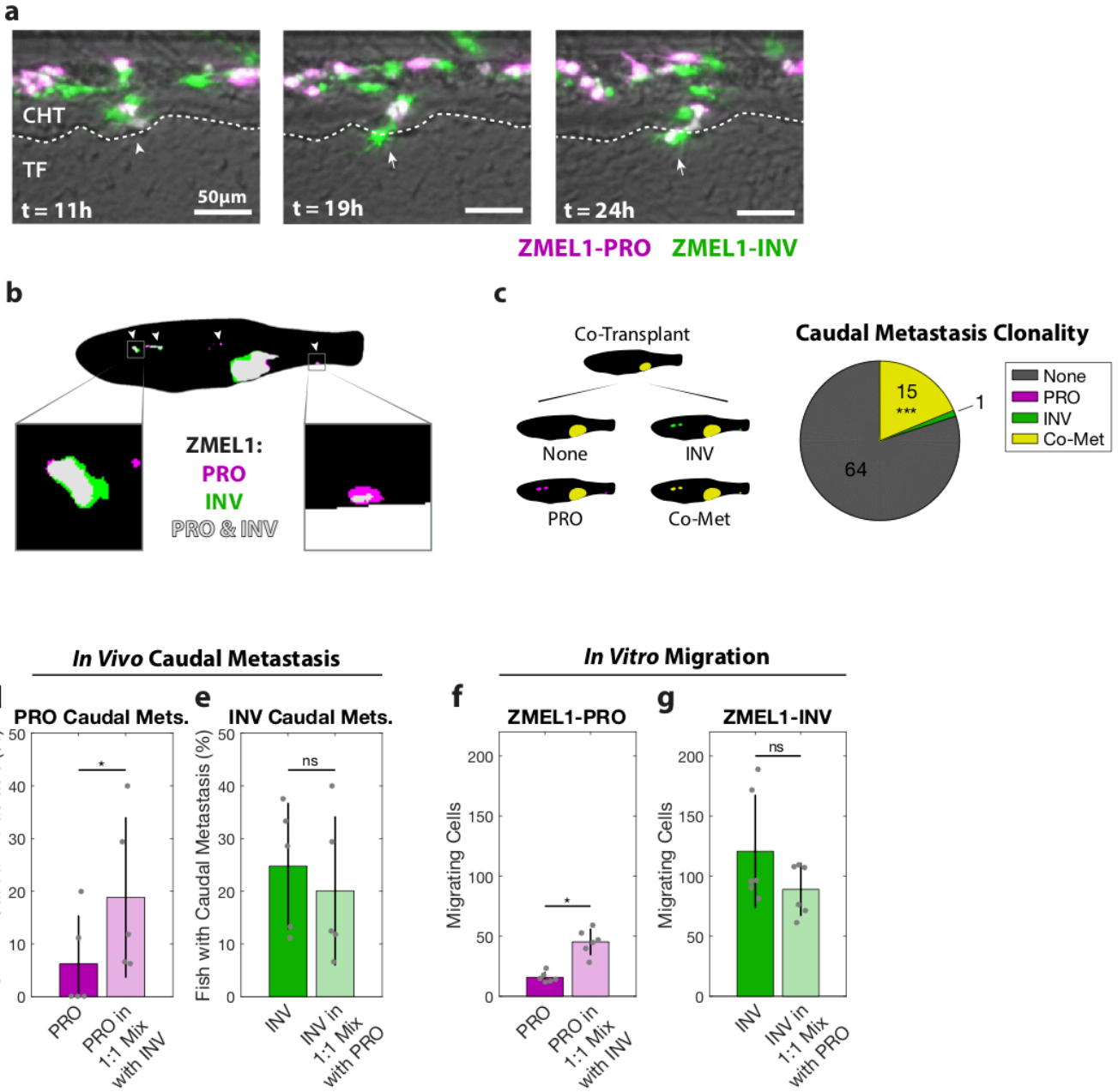


Figure 4

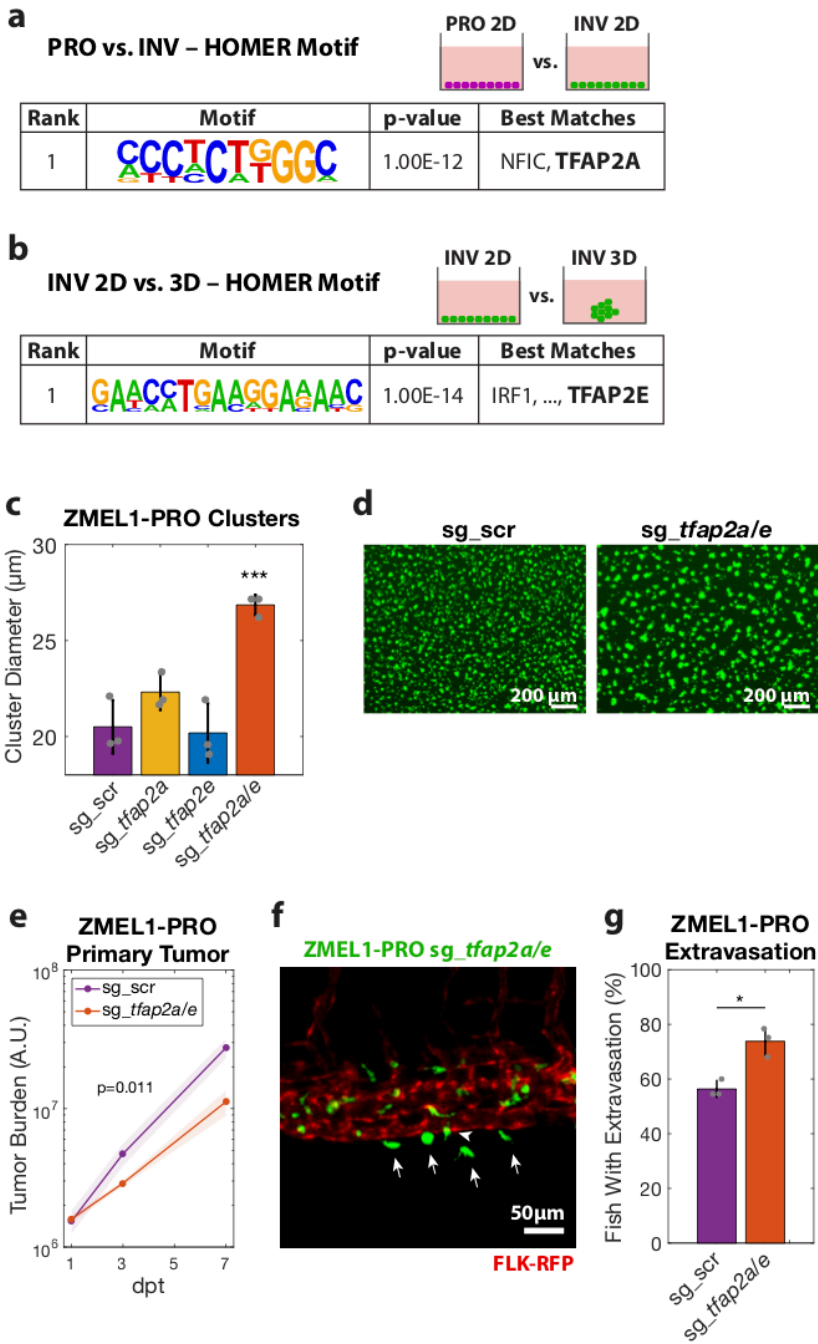


Figure 5

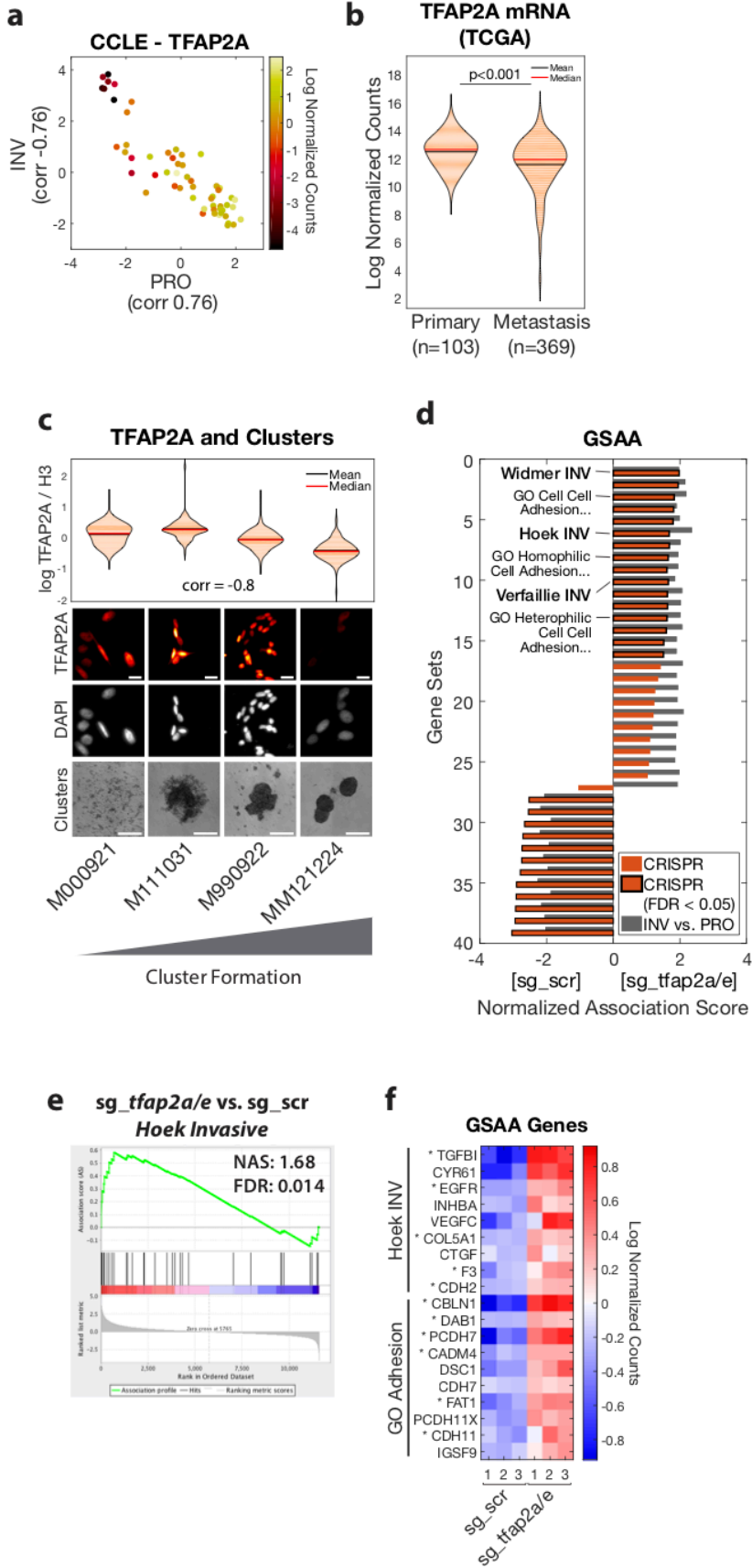
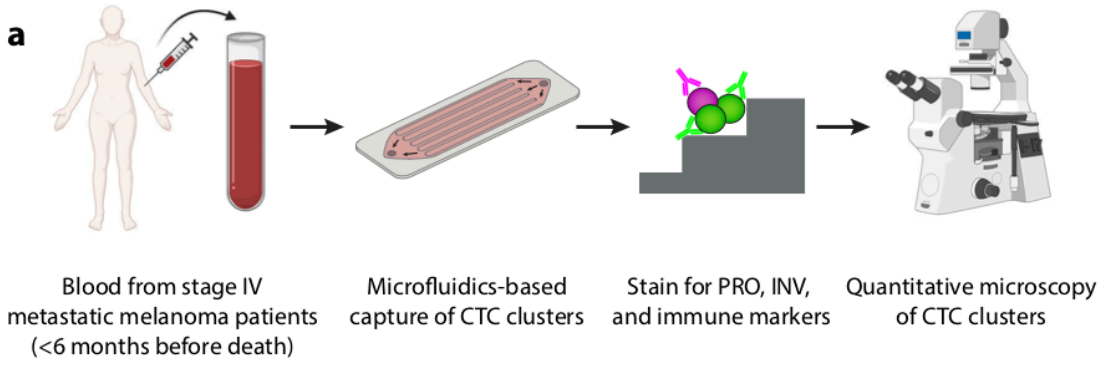
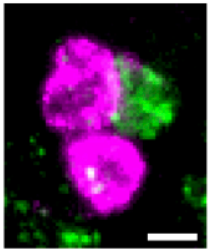


Figure 6



b IF Staining



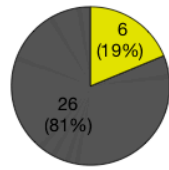
TFAP2A
SOX9

Quantification



INV PRO
-1 0 1
Log Normalized
TFAP2A / SOX9

c Human CTC Clusters



Homotypic
PRO-INV
Heterotypic

d Human CTC Clusters

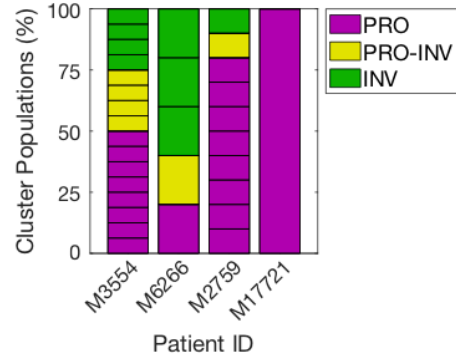
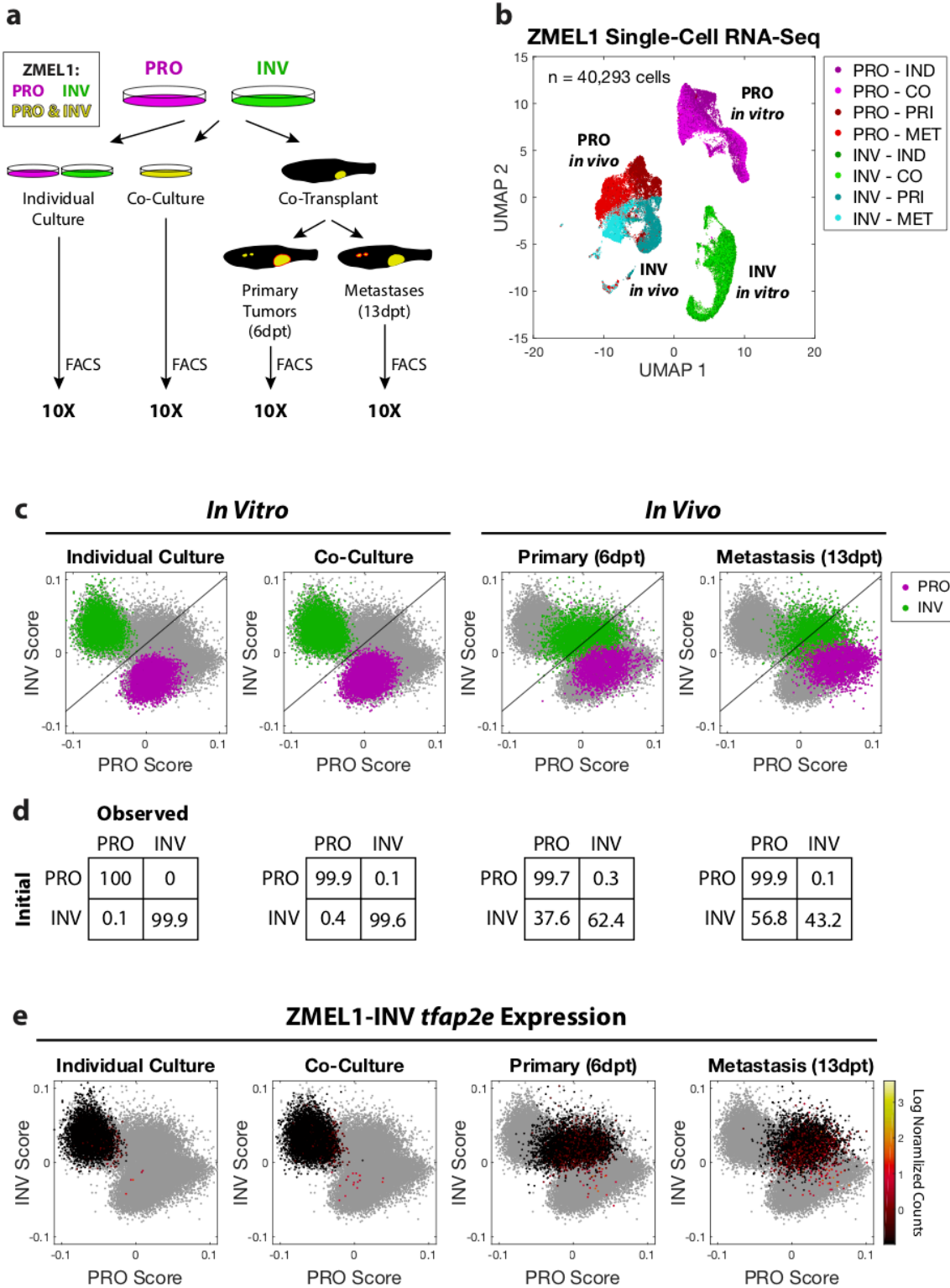


Figure 7



Graphical Abstract

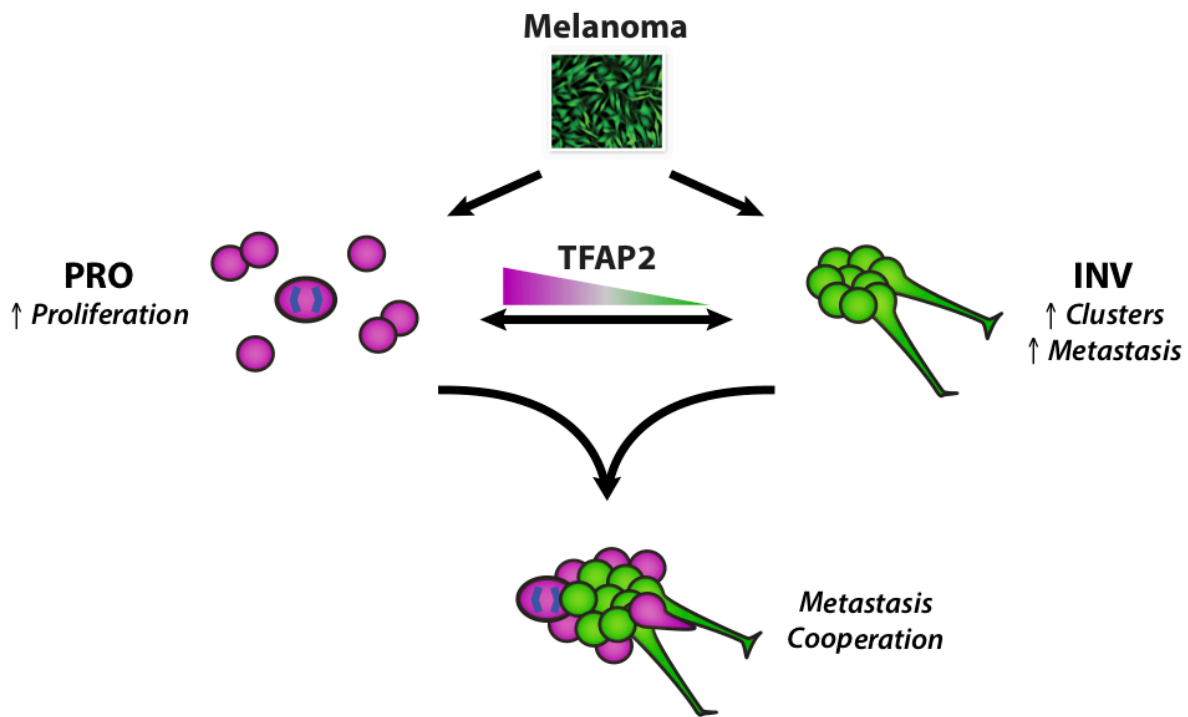
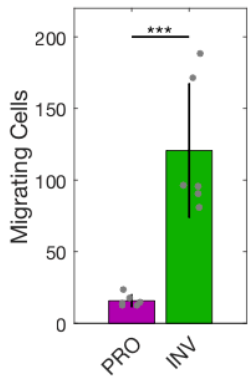
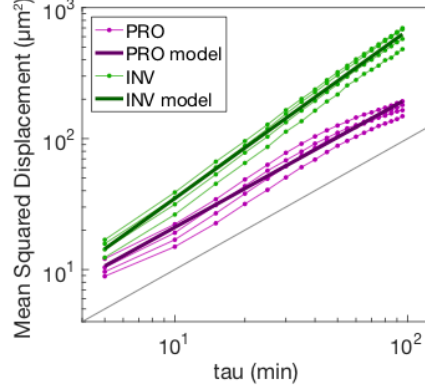


Figure S1

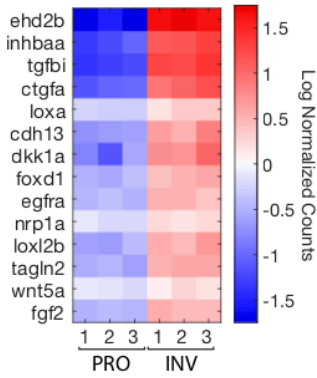
a ZMEL1
Boyden Chamber



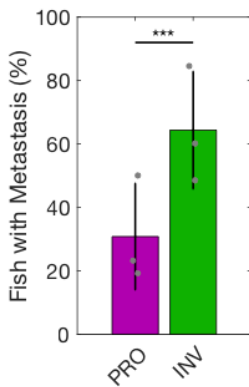
b ZMEL1 MSD vs. Tau



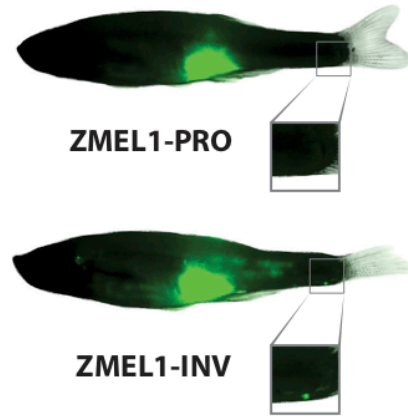
c Hoek INV



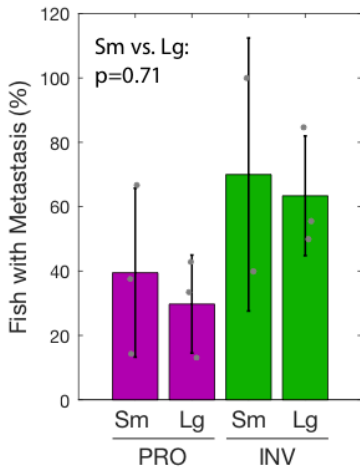
d ZMEL1 3dpt
Overall Metastasis



e



f Overall Metastasis



g Caudal Metastasis

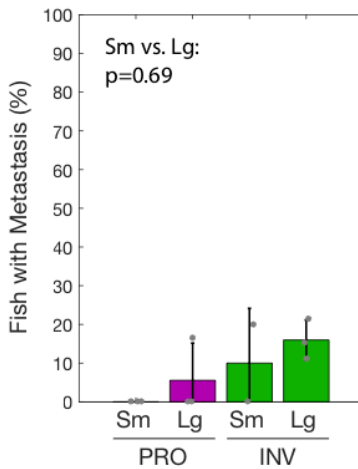


Figure S2

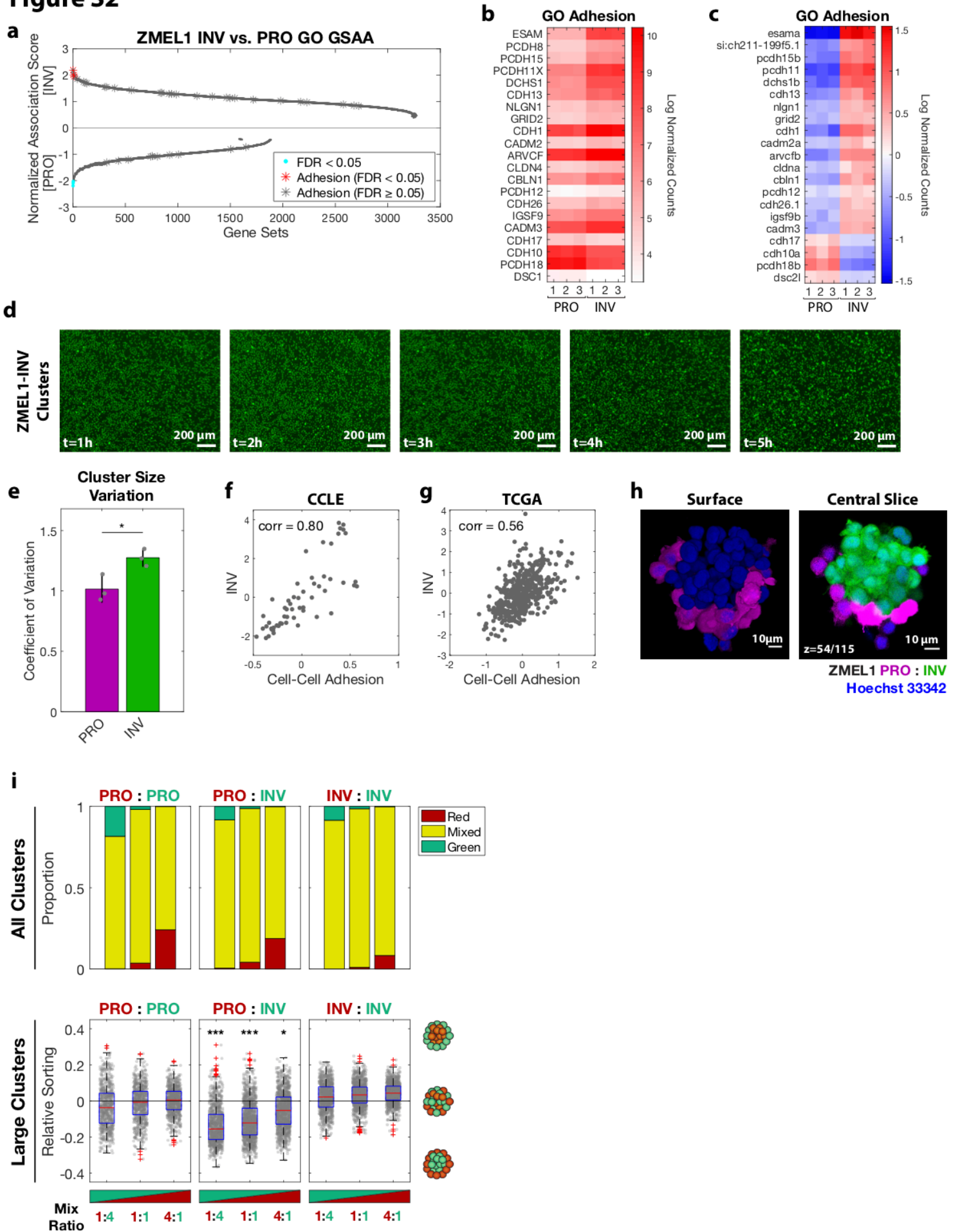


Figure S3

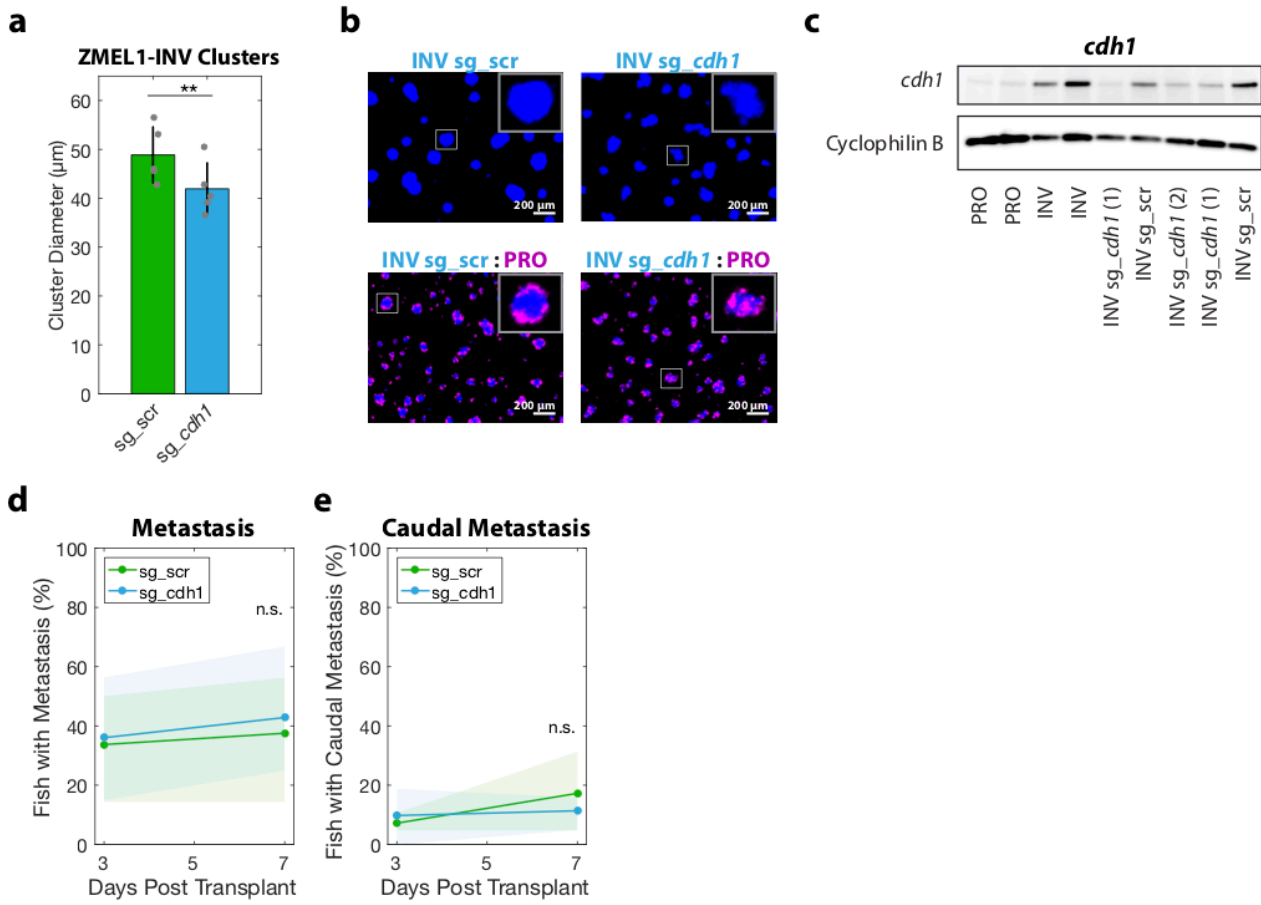


Figure S4

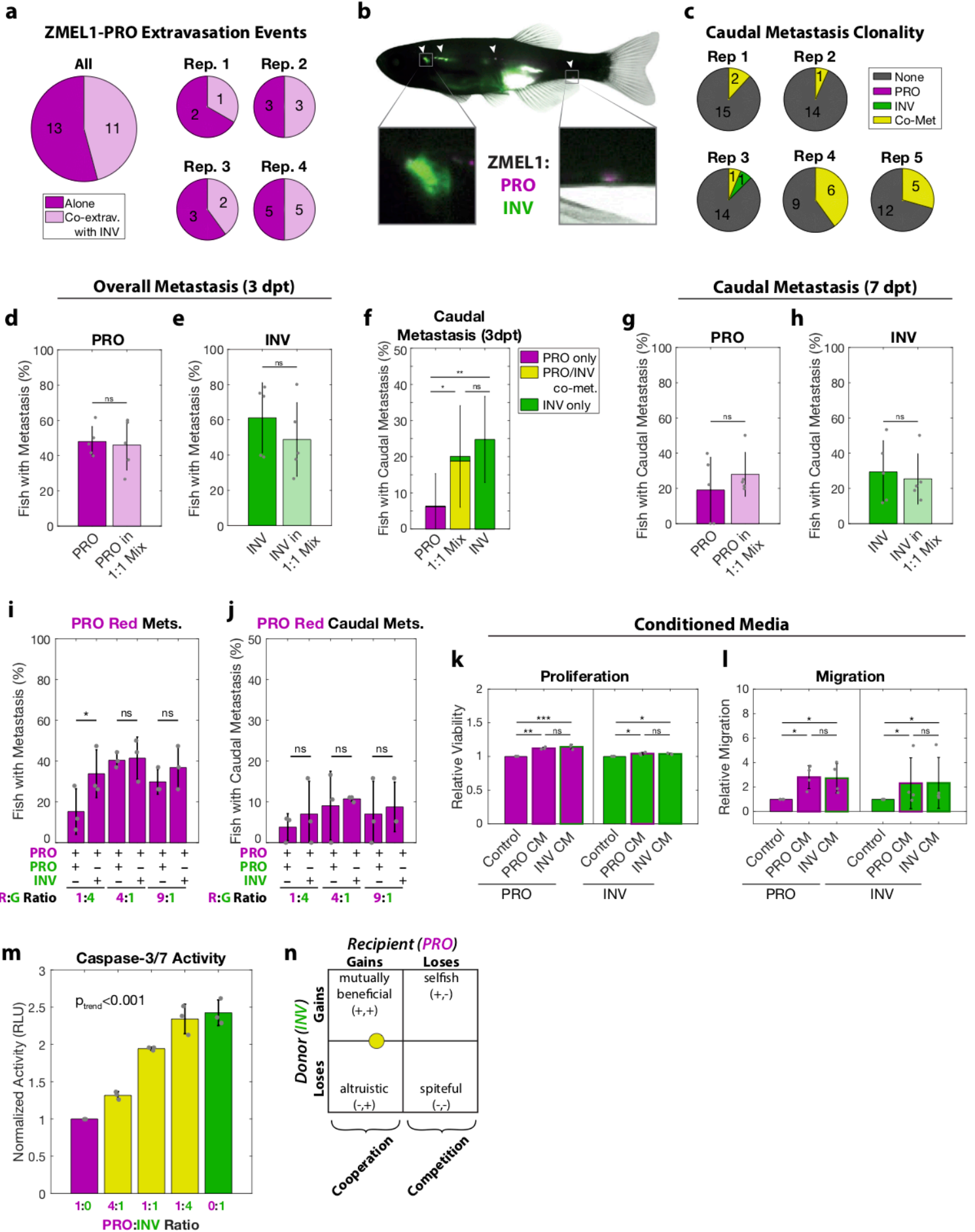


Figure S5

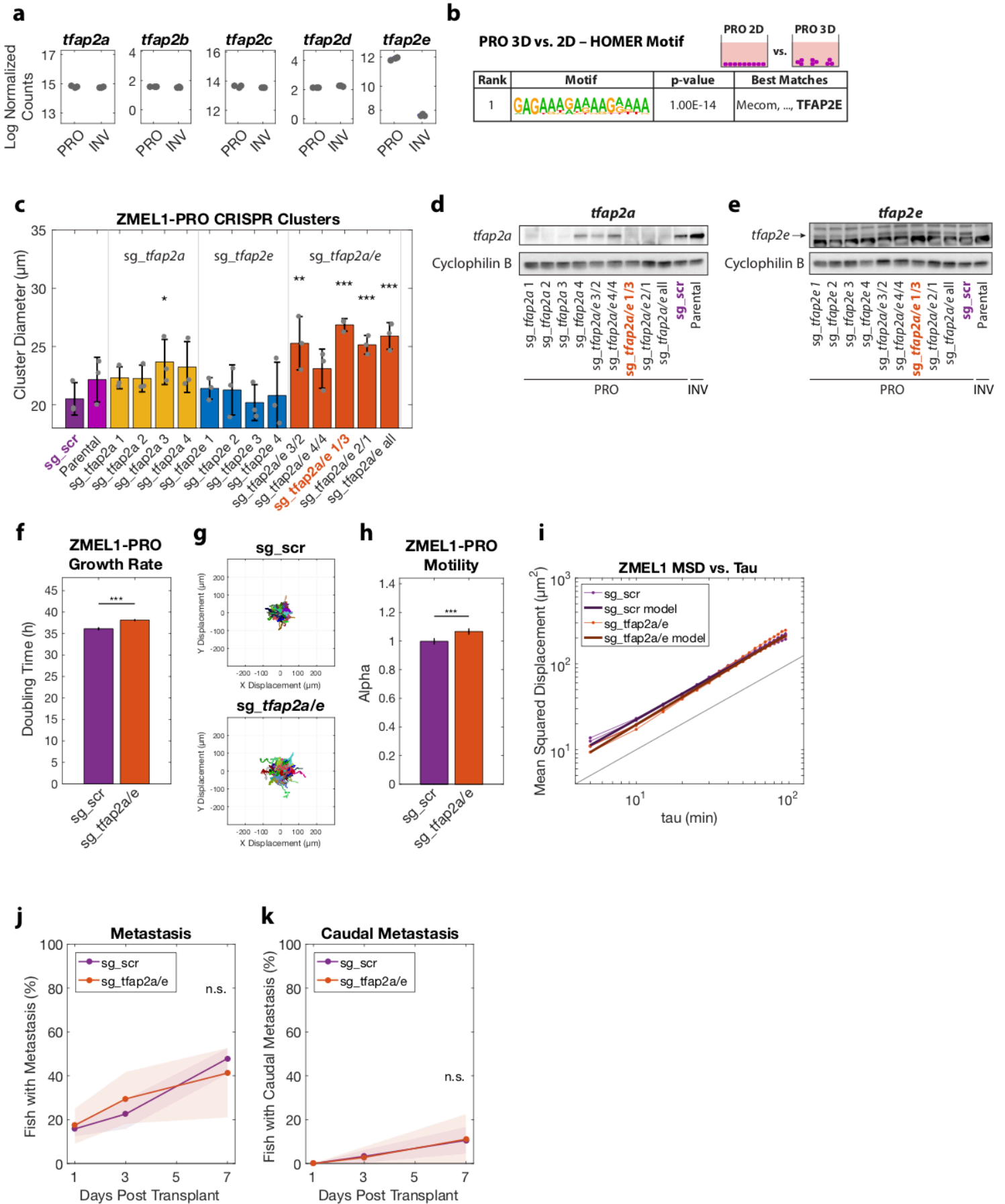


Figure S6

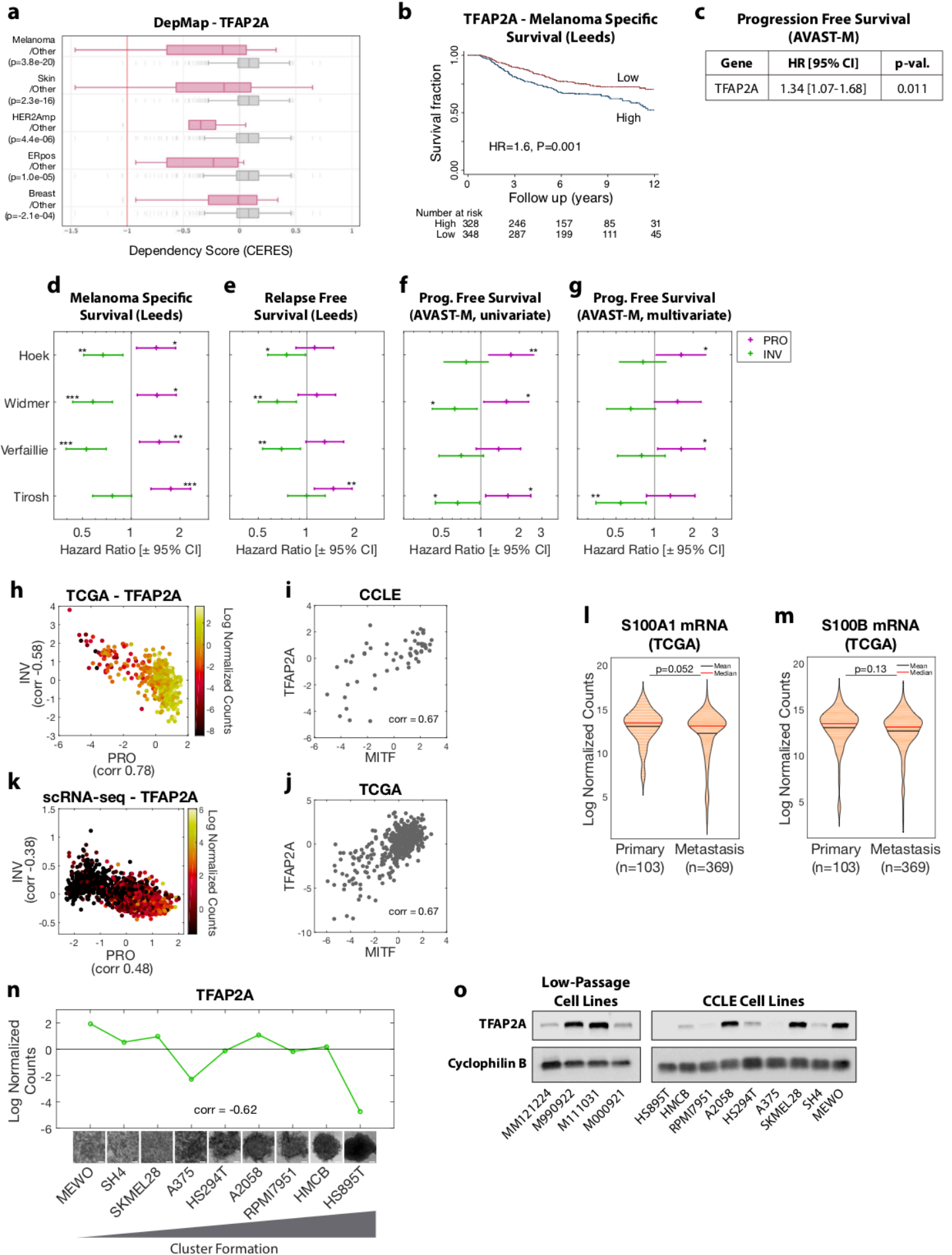


Figure S7

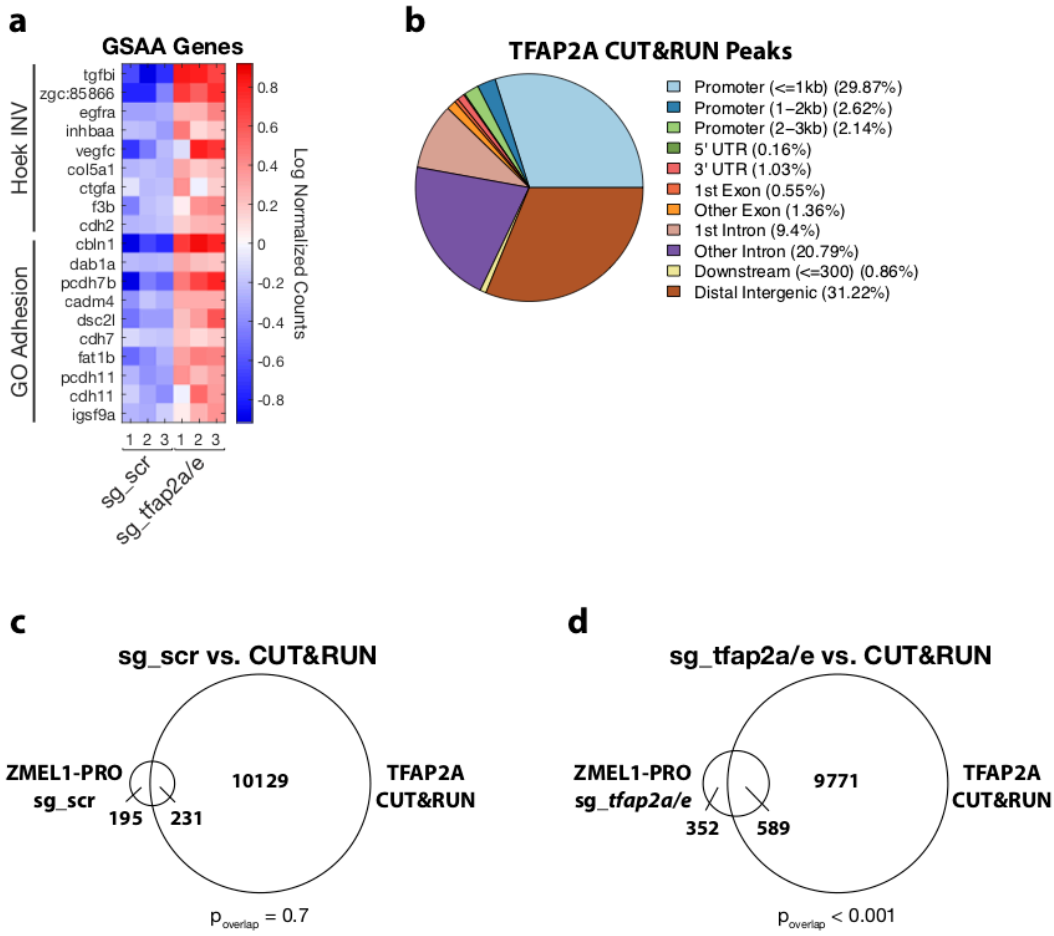


Figure S8

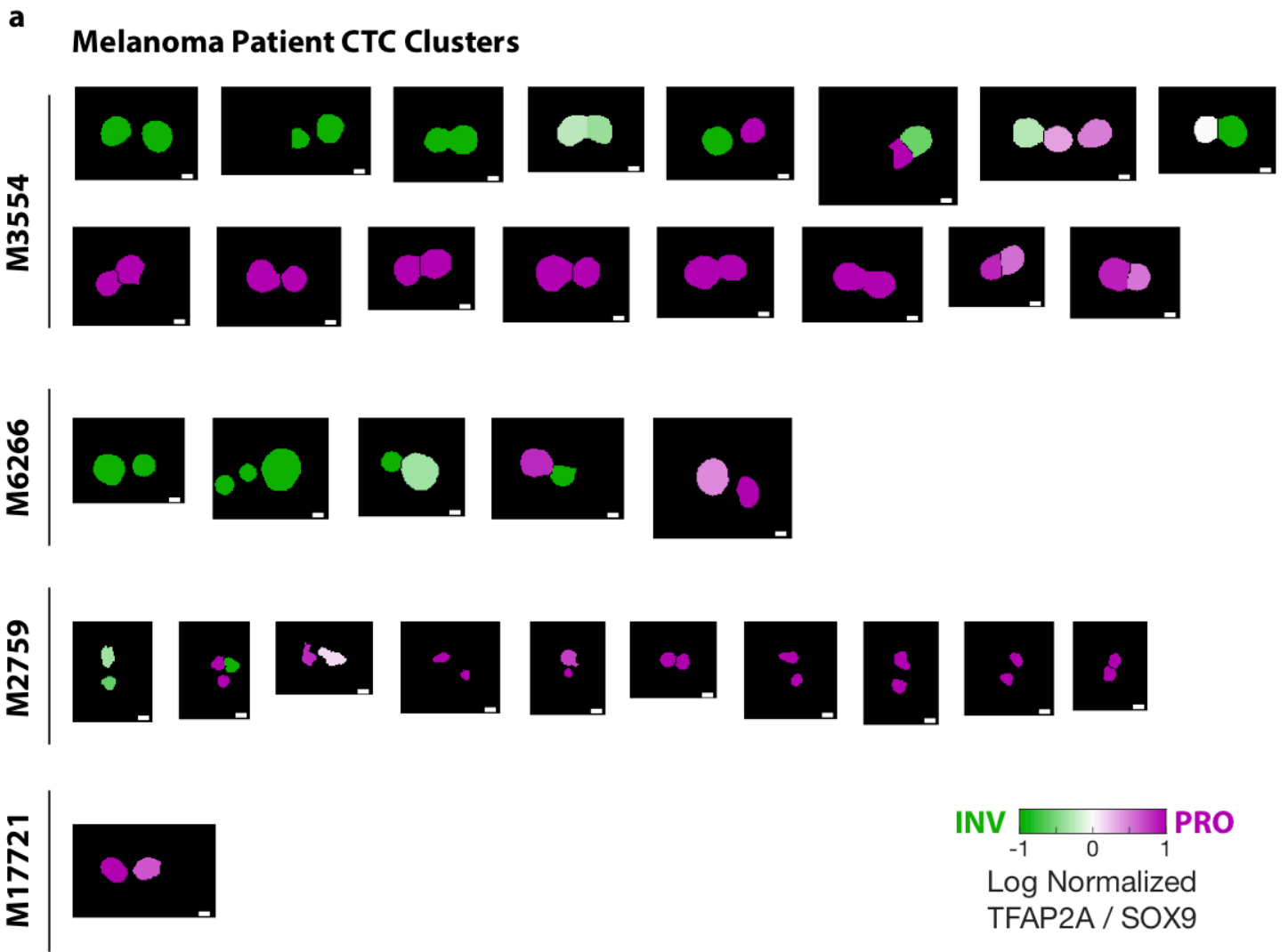
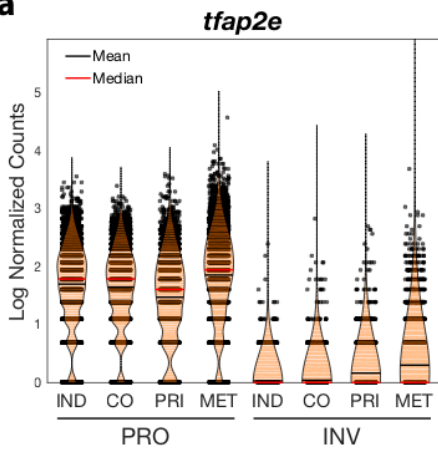
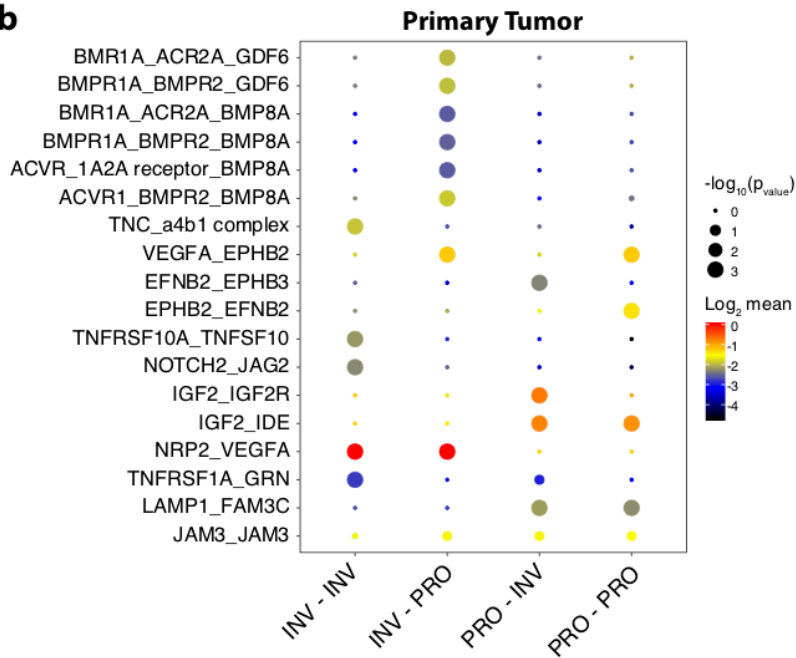


Figure S9

a



b



c

