`

Isogenic hiPSC models of Turner syndrome reveal shared roles of inactive X and Y in the human cranial neural crest network

Darcy T. Ahern [1,2] , Prakhar Bansal [1,2], Isaac V. Faustino [2], Heather R. Glatt-Deeley [2], Yuvabharath Kondaveeti [2], Erin C. Banda [2] and Stefan F. Pinter [1,2,3 *]

[1] Graduate Program in Genetics and Developmental Biology, UCONN Health, University of Connecticut, Farmington, CT, United States

[2] Department of Genetics and Genome Sciences, UCONN Health, University of Connecticut, Farmington, CT, United States

[3] Institute for Systems Genomics, University of Connecticut, Farmington, CT, United States

[*] Corresponding author: spinter@uchc.edu

`

SUMMARY

Modeling the developmental etiology of viable human aneuploidies can be challenging in rodent models, where synteny with human chromosomes is affected, or primate-specific biology is implicated. In humans, monosomy-X (45,X) causes Turner syndrome (TS), altering craniofacial, skeletal, endocrine, and cardiovascular development, which remain largely unaffected in 39,X-mice. We derived human 45,X and isogenic euploid induced pluripotent stem cells (hiPSCs) from male and female mosaics to model how human monosomy-X may impact early embryonic development. Because several neural crest (NC) derived cell types are hypothesized to underpin craniofacial and cardiovascular changes in TS, we derived anterior neural crest from our hiPSCs and performed a highly-powered and extensive differential expression study. Across all three isogenic panels, 45,X neural crest cells (NCCs) show impaired acquisition of the PAX7/SOX10 double-positive NC state relative to euploid 46,XY controls. Monosomy-X NCCs also share similarly disrupted expression of NC-specific genes relative to their isogenic euploid control regardless of whether the latter carry an inactive X or Y. Gene-set enrichment analyses indicate monosomy-X NCCs increase cholesterol biosynthesis genes while dampening ribosomal protein and nuclear-encoded mitochondrial genes. These metabolic pathways are also over-represented in gene modules that are preserved in monogenic conditions involving neurocristopathy. Importantly, these gene modules are also significantly enriched in 28% of all TS-associated terms of the human phenotype ontology, and point to specific sex-linked genes that are expressed from two copies in euploid males and females alike, which may qualify as candidate haploinsufficient drivers of TS phenotypes in NC-derived lineages. Our study demonstrates that isogenic hiPSC-derived NCC panels representing monosomy-X can serve as a powerful model of early NC development in TS, and provides euploid and X-monosomic transcriptomic datasets relevant to exploration of TS and NC biology.

`

59 **INTRODUCTION**

60 The absence of the second sex chromosome (monosomy-X) represents the only viable human
61 monosomy in humans, resulting in Turner syndrome (TS, 1:2000 live births). Although fewer than
62 0.3% of conceptuses with monosomy-X (45,X karyotype) complete development [1,2], it is
63 considered viable, because most sex-linked genes are expressed from a single copy in placental
64 mammals: the Y lost most of its ancestral genes originally shared with the proto-X, except for
65 dosage-sensitive regulators [3], while the X became subject to dosage-compensation in the form
66 of X chromosome inactivation (XCI) in XX females [4]. However, many genes retained on the Y have
67 a homologous X-linked copy that escapees XCI. In addition to these X/Y-pairs ("gametologs"),
68 genes in the recombining pseudoautosomal region (PAR) shared by X and Y also escape XCI in XX
69 females [4].

70 The compound haploinsufficiency of these gametolog and PAR genes is hypothesized to
71 underpin both cardinal short stature and premature ovarian failure in persons with TS, alongside
72 variably penetrant (40-80%) phenotypes [5], like lymphedema of hands and feet, neurocognitive
73 changes, as well as renal and cardiac malformations. Joining skeletal and craniofacial changes,
74 bicuspid aortic valve and aortic defects in TS implicate cell types derived from neural crest cells
75 (NCCs). These transitory multipotent progenitors arise from the dorsal side of the closing neural
76 tube and migrate ventrally to contribute to the developing face, skin, heart, adrenal glands, joints
77 and peripheral nervous system [6]. Because all of these organ systems are variably impacted in
78 subsets of TS patients [7], the neural crest is thought to be particularly vulnerable to monosomy-X.

79 Yet, direct experimental evidence for this hypothesis has been lacking. Because many murine
80 orthologs of human PAR genes reside on autosomes [8,9], their gene dosage is maintained in X-
81 monosomic mice, which consequently fail to model most TS phenotypes [10,11]. To address this
82 problem, we derived and cytogenomically validated X-monosomic human induced pluripotent
83 stem cells (hiPSCs) alongside isogenic euploid control lines from mosaic fibroblasts. We recently
84 reported that such isogenic panels from three unrelated donors recapitulate a hypothesized
85 impairment of trophoblast cell fates due to monosomy-X [12].

86 Here, we derive anterior NCCs from each of these hiPSC panels to assess the impact of human
87 monosomy-X on the neural crest. We perform an in-depth transcriptomic and cellular analysis

`

88 comparing monosomy-X to their otherwise isogenic euploid controls, and find monosomy-X
89 hiPSCs give rise to significantly fewer PAX7/SOX10 double-positive NCCs compared to 46,XY
90 euploid controls. Differential expression reflecting monosomy-X is highly concordant across male
91 and female-derived panels, and jointly enriched in pathways associated with cholesterol
92 metabolism, NCC-relevant cell signaling and translation. Importantly, NCC gene modules that
93 reflect monosomy-X by weighted gene co-expression network analysis (WGCNA) are also
94 preserved in hiPSC-derived NCC models of monogenic conditions traditionally considered to
95 represent neurocristopathy.
96
97 **RESULTS**
98 Over half a century of detailed studies in the chick and other model systems have shed
99 light on the signaling pathways and intricate gene regulatory networks (GRNs) that govern the
100 specification and migration of NCCs [13,14]. These exquisite developmental studies have also
101 informed progressively intricate protocols for NCC derivation from embryonic stem cells and
102 hiPSCs that recapitulate terminal NCC fates as a function of their axial identity [15–17]. Due to our
103 interest in the anterior/craniofacial and vagal neural crest, which populate the developing face
104 and heart, we applied a highly robust method for NCC specification via moderate WNT and BMP
105 activation that is reported to enable reproducible derivation of anterior NCCs at ~ 70% efficiency
106 [17,18]. We applied this method to a total of 11 distinct hiPSC lines from all three of our
107 cytogenomically validated hiPSC panels [12], and quantified the dorsal neural tube marker PAX7
108 alongside the canonical neural crest marker SOX10, which jointly mark NCC specification.
109 Applying automated immunofluorescence quantification, we observe the expected rate of $PAX7^+$
110 $SOX10^+$ double-positive cells (60-70%) across all euploid lines, though female euploid lines with
111 somewhat reduced efficiency (-10%) relative to male euploid lines (Fig. 1A). While monosomy-X
112 hiPSCs can also differentiate to $PAX7^+SOX10^+$ NCCs, their overall efficiency is significantly reduced
113 by about half in both male-derived isogenic panels (Fig. 1B, male 1: $p=0.002$, male2: $p=1.1x10^{-5}$,
114 Mann-Whitney U). Interestingly, female-derived monosomy-X NCCs arise at a similar same rate
115 as their euploid counterparts, rendering this isogenic comparison non-significant ($p=0.3$) but still
116 significantly reduced relative to male euploid controls ($p\leq1.3x10^{-3}$). Specifically, we find the $PAX7^+$

`

117 rate is largely unaffected by monosomy-X across all three panels (Fig. S1), and whereas it is
118 SOX10^{+} in male1 and male2-derived monosomy-X lines that is significantly reduced (male 1:
119 p=0.001, male2: p=9.8 x10^{-6}, Mann-Whitney U).

120 We next performed a deep and highly-powered RNA-seq study on a subset of these
121 samples, with a minimum of 4 replicates per cell line, grouped by a combination of donor and
122 karyotype ("condition") comprising 7-12 replicates each. This enables our analysis to assess
123 monosomy-X driven changes within each isogenic context (ie. donor). Indeed, each paired set of
124 monosomy-X and euploid controls segregate from each other along the first and second principal
125 components (Fig. 2A), while hierarchical clustering groups all male-derived monosomy-X samples
126 together (Fig. S2A). We next compared X-monosomic and euploid NCC expression levels of
127 marker gene sets identified in two prior human NCC differentiation studies [15,18], orthologs of
128 chick neural crest GRNs [19], and genes differentially expressed (DEGs) in the murine neural crest
129 [20], as well as from a 3D *in vitro* model of the folding human neural tube [21]. Indeed, gene sets from
130 these five studies were largely sufficient to segregate euploid and X-monosomic samples (Fig.
131 S2B), and DEGs identified in the human 3D neural tube model were the most highly expressed,
132 while early anterior neural crest (eANC), murine Hox-negative NC markers, and migratory NC
133 markers identified in the chick were roughly tied. We also confirmed the axial identity of our
134 NCCs by the uniformly low expression of posterior HOX genes (Fig. S2C). Assessing median-
135 normalized gene set levels, we observe significantly lower levels of early anterior neural crest
136 (eANC) and higher levels of late ANC (lANC) markers in euploid over X-monosomic lines in all
137 three isogenic panels (Fig. 2B), and respectively higher expression of NC over neuroectodermal
138 genes (NE) in male euploid over X-monosomic NCCs (Fig. S2D). Likewise, genes expressed in p75^{+}
139 hESC-derived NCCs are also significantly higher in euploid over X-monosomic hiPSC-derived NCCs,
140 which is also true for orthologs of GRNs identified in the migratory chick neural crest (Fig. 2C,D).

141 To delineate which groups of X/Y-linked genes may be haploinsufficient in NCCs, we
142 plotted PAR and gametolog pair genes alongside X-specific genes that escaped XCI in female
143 euploid NCCs, as reflected in allelic variant counts of the phased female X (Fig. S2E). Largely driven
144 by PAR genes, monosomy-X NCCs segregate from their isogenic euploid controls, which
145 expectedly split female 46,XX and male 46,XY samples. DEGs escaping XCI in euploid female NCCs

`

146 (lesser allele frequency, LAF≥0.1, binomial p≤0.05) included 7 PAR1 genes, 5 X-linked gametologs
147 and 16 previously-reported escapee genes, alongside 8 genes that may reflect novel NCC-specific
148 escapees or have reactivated in a subset of cells despite robust bulk expression of *XIST* (Fig. S2F).
149 We next correlated the PAX7$^+$ SOX10$^+$ fraction with the median expression of PAR, Pair, other
150 escapee genes and their total ("All") in each sample (Fig. 2E). Including the NCC marker gene sets
151 in this analysis, we observe that the PAX7$^+$ SOX10$^+$ fraction best correlates with levels of PAR and
152 Pair genes, which likewise correlate with the late ANC, p75$^+$ and chick migratory GRN sets.
153 Because NCC undergo an epithelial-to-mesenchymal transition that modulates cell cycle
154 progression [22,23], more mature NCC markers are expectedly anti-correlated with cell cycle and
155 earlier NCC markers. In sum, these data suggest monosomy-X hinders or delays the maturation
156 of early to late neural crest, in keeping with their reduction of SOX10$^+$ positive cells.

157 To determine whether global differential expression points to a common monosomy-X
158 signature across isogenic panels, we called and compared DEGs common to any two or all three
159 panels. Altogether, over 30% of all DEGs were shared in all pairwise comparisons ($p = 1.1x10^{-13}$,
160 $p = 1.2x10^{-21}$ and $p = 3.7x10^{-137}$, hypergeometric test), with the two male-derived monosomy-X
161 panels reflecting the most significant overlap (Fig. 3A). We also assessed the direction of gene
162 expression changes in these overlapping DEG sets, which perfectly segregate all monosomy-X
163 samples from their isogenic euploid controls. Indeed, DEGs shared by all three ($p = 3.7x10^{-124}$,
164 sign test), as well as any two isogenic panels ($p = 7.4x10^{-195}$, $5.1x10^{-11}$, $8.0x10^{-47}$) change
165 concordantly in highly significant fashion (Fig. 3A). This level of overlap and concordant change
166 is remarkable given that monosomy-X samples ('mXO1/2' & 'fXO') were only assessed relative to
167 euploid controls in their own isogenic context.

168 To identify cellular pathways and developmental processes commonly impacted by the
169 lack of X or Y, we performed gene-set enrichment analysis (GSEA) for all three monosomy-X NCC
170 panels as ordered by their averaged Wald statistic ('ave.XO'). We also compared female-to-male
171 euploid ('fmE1/2'), as well karyotypically-identical male1-to-male2 monosomy-X ('m1m2O')
172 NCCs. Mirroring the global overlap of DEGs (Fig. 3A), significantly enriched gene-sets revealed a
173 highly concordant pattern across all three isogenic panels (Fig. 3B): across Wikipathway
174 (p.adj<0.02), KEGG and Hallmark (Fig. S3, Table S1) gene term collections, monosomy-X NCCs

upregulate genes associated with cell cycle, cholesterol and lipid biosynthesis, and down-regulate genes linked to translation and oxidative phosphorylation. While the impact of monosomy-X in the female NCC panel is generally milder than the two male panels, the concordance in GSEA terms mirrors the highly significant overlap in DEGs, and both hyperlipidemia and hypercholesteremia are seen in patients with TS [24]. Likewise, in the human phenotype ontology (HPO), monosomy-X NCCs show dysregulated gene sets relating to craniofacial, joint and cognitive development that are also frequently observed in patients with Turner Syndrome (Table S1), including 'Down-slated palpebral fissures', 'joint hypermobility, 'hyperactivity' (all p.adj<0.0001). Indeed, the enrichment of HPO terms relating to Turner Syndrome ("ORPHA881") in monosomy-X NCC vs. control comparisons is particularly striking (Fig. 3C), and recovers 28/114 (24.6%) of all HPO terms linked to TS (Table S1), none of which were significant in any of the control comparisons. In sum, these GSEA results underscore that differential expression in X-monosomic NCCs relative to euploid controls recovers gene terms of known TS phenotypes that relate to the neural crest, as well as beyond, and implicates cellular and developmental pathways impacted by monosomy-X.

We next performed weighted gene co-expression network analysis (WGCNA) to elucidate the relationships between monosomy-X and altered pathways, and prioritize individual X/Y-linked genes as potential dosage-sensitive drivers. This analysis assigned genes to 29 modules, 22 of which are driven by contrasting expression of euploid and X-monosomic expression, as evidenced by a loss of module preservation in networks lacking euploid samples (Fig. S4A). Eight modules (groupA: 1,6,7,13,19,20,26,27) are significantly correlated ($R \geq 0.3$, $p \leq 0.05$) with the $PAX7^{+}SOX10^{+}$ rate, and six modules (groupB: 3,5,8,9,11,15) are significantly anti-correlated ($R \leq -0.3$, $p \leq 0.05$, Fig. 4A). Consistently, the $PAX7^{+}SOX10^{+}$ correlated groupA modules associate with mature NCC markers, whereas $PAX7^{+}SOX10^{+}$ anti-correlated groupB modules reflect an earlier, NE-biased cycling cell signature. The $PAX7^{+}SOX10^{+}$ positively correlated groupA was enriched in translation and mitochondrial function terms, while the $PAX7^{+}SOX10^{+}$ anti-correlated groupB was over-represented for cholesterol biosynthesis, cell cycle-related terms, and several signaling pathways, including Notch, focal adhesion/PI3AKT/mTOR and TGFbeta signaling (Fig. 4A, S4B, Table S2). To further contextualize the relationship between these modules and biological

`

204 processes, we turned to a semantic similarity map of significantly-enriched GO terms
205 (p.adjust≤0.01, hypergeometric distribution) in modules that also significantly co/anti-correlated
206 with the PAX7^{+}SOX10^{+} rate (Fig. 4B, -/+ labeled modules). Terminal neural-crest related terms
207 ("cranial facial parasympathetic nerve", "endocardial cushion", "anterior/posterior pattern
208 specification", "angiogenesis") were enriched in groupB modules, whereas metabolic functions
209 that connect cell cycle and splicing to translation and mitochondrial function were represented
210 by both group A&B modules. Overall, these enriched metabolic terms mirrored the GSEA results
211 (Fig. 3B,C, Table S1).

212 If metabolic processes reflected in these modules play important roles in neural crest
213 development, we would expect them to be detectable in tissues populated by the neural crest.
214 We tested this hypothesis in transcriptomes collected from the developing heart [25] and face [26,27],
215 spanning Carnegie stages 13-23. Both tissues ('CS_Heart' with 11/28 preserved modules, and
216 'CS_Face' with 12/28) show moderate to high module preservation (Z-score ≥ 5), 8 of which also
217 co/anti-correlate significantly with the PAX7^{+}SOX10^{+} rate in our hiPSC-derived NCCs (Fig. 4A).
218 Likewise, we assessed module preservation in hiPSC-derived NCCs of monogenic conditions
219 considered canonical neurocristopathies: 15 modules were preserved in Waardenburg [28]
220 (*SOX10*), 12 in Pierre-Robin[29] (*SOX9*), and 11 each in Familial Dysautonomia[30] (*IKBKAP*) and
221 Bohring-Opitz[31] (*ASXL1*) syndromes, most of which also overlapped with each other. In contrast,
222 only respectively 6 and 4 modules were preserved in hiPSC-derived NCC models of Floating-
223 Harbor[32] (*SRCAP*) and Branchio-Oculofacial[33] (*TFAP2A*) syndromes, while no module met a Z-
224 score threshold of 5 standard deviations in a pancreatic tumor stroma control dataset. In sum,
225 our analysis indicates that monosomy-X sensitive modules that reflect metabolic disturbances
226 also co-vary in relevant tissues with developmental time and *in vitro* NCC models of several
227 neurocristopathies.

228 Interestingly, three groupA modules (13,20,27) were exclusively preserved in our hiPSC-
229 derived monosomy-X trophoblast model, and correlated with escapee levels (PAR, Pair &
230 'otherEsc'), as well as each other (Fig. 4A, S4C) [12], raising the question in which modules these
231 genes were over-represented. Indeed, module 20 was strongly enriched in PAR genes (Fig. S4D)
232 but lacked any other significant MSigDB or GO annotations (Table S2). Likewise, most gametolog

`

233 pairs were dispersed across modules 4 & 20, or remained unassigned. One notable exception was
234 the only ribosomal protein to be encoded by two genes, as *RPS4X* was assigned to the translation-
235 enriched module 6, which was generally down-regulated in X-monosomic lines (Fig. 3B, Fig. 4A,
236 Table S1). Of note, *RPS4X* was also significantly reduced in male1-derived 45,X NCCs lacking
237 *RPS4Y1* expression.

238 To prioritize candidate dosage-sensitive genes shared by X and Y, we squared their
239 correlation with the averaged expression of their respective modules and plotted this coefficient
240 (kME^2) over their individual correlation with the $PAX7^+SOX10^+$ rate (Fig. 4C). We also present
241 their percentile-ranked probability of haploinsufficiency (pHI % rank, from[34]) as a constraint
242 metric to reflect possible selection. Most of the highly pHI-ranked genes correlate poorly with
243 the $PAX7^+SOX10^+$ percentage, in contrast to most PAR genes, as well as Y-linked gametologs, that
244 correlate strongly but rank poorly in pHI [35], in the residual variation intolerance score (RVIS [36]),
245 or remain unranked (Y-linked gametologs). Yet, *ZBED1, SLC25A6, RPS4X* and *PLCXD1,* rank in the
246 top half of pHI or RVIS scores and correlate strongly with the $PAX7^+SOX10^+$ rate and their kME
247 (Fig. 4C). While homozygous loss-of-function mutations have revealed *PLCXD1* as dispensable in
248 humans [37], *ZBED1* was also the top candidate in our 45,X hiPSC-derived trophoblast analysis [12].
249 *RPS4X* was especially notable given that translation was over-represented in its assigned module
250 6, which was also enriched in the top3-scoring HPO terms associated with TS: 'Horseshoe kidney',
251 'Coarctation of the aorta' and 'Webbed neck' (Fig. 4D,E). Indeed, altogether 32/114 (28.1%) of all
252 TS-associated HPO terms were significantly enriched in 10 modules (p.adjust≤0.1,
253 hypergeometric distribution) (Fig. 4D/E, S4E), 9 of which were also preserved in monogenic
254 neurocristopathies (Fig. 4A).

255

256 **DISCUSSION**

257 Despite the relatively small number of genes shared by extant mammalian X & Y, only TS-
258 associated short stature has been linked conclusively to haploinsufficiency of one such gene to-
259 date, namely pseudoautosomal *SHOX* [38]. This example also highlights challenges in modeling TS
260 in rodents, as murine *SHOX* ortholog *Shox2* is autosomal [39] and thus remains unaffected in 39,X
261 mice, a pattern shared with all but two murine orthologs of human PAR1 genes [9]. Likewise, more

`

262 pan-mammalian gametologs pairs were retained on both X&Y in humans than mice [3], with the
263 remaining murine gametologs appearing largely haplo-sufficient in mouse cardiovascular
264 development [10,11].

265 Given the craniofacial and cardiovascular phenotypes in TS, new models are thus needed
266 to understand how monosomy-X impacts development, specifically in cells and tissues derived
267 from the neural crest. Our study applies a well-established hiPSC to NCC differentiation model
268 [17,18] to demonstrate that monosomy-X significantly reduces NCC specification relative to euploid
269 controls (Fig. 1, S1), which is also reflected in expression of various NC marker sets across three
270 independent isogenic panels (Fig. 2, S2). Differential expression indicates monosomy-X alters
271 important metabolic pathways, specifically increasing cholesterol biosynthesis genes, while
272 reducing ribosomal protein and nuclear-encoded mitochondrial genes (Fig. 3, S3, Table S1). This
273 finding is consistent with recent reports revealing metabolic reprogramming to play a key role in
274 transitioning NCCs from proliferative to migratory and terminally differentiated states[40,41]. During
275 the epithelial-to-mesenchymal transition (EMT), NCCs increase both ribosome biosynthesis [40]
276 and aerobic glycolysis [41], but our 45,X NCC transcriptomes suggest these processes are
277 decoupled, as many ribosomal proteins are down- while glycolytic genes are up-regulated (Fig.
278 3, Table S1), at the expense of oxidative phosphorylation, which has also been linked to NCC
279 specification [42]. In addition, cholesterol biosynthesis genes are significantly upregulated in 45,X
280 NCCs, which while critical for modulating NCC signaling [43] may also alter their patterning [44], and
281 is independently consistent with hypercholesteremia seen in TS patients [5,24].

282 This link to TS phenotypes is also reflected in TS-associated HPO terms that were enriched
283 in GSEA (28/114) and WGCNA (32/114) results (Fig. 3,4). Indeed, three TS-associated HPO terms
284 were most significantly over-represented in module 6, which comprised of ribosomal proteins
285 and other translation factors that were down-regulated in 45,X NCCs. Interestingly, ribosomal
286 protein gene and gametolog *RPS4X* was also assigned to module 6, correlated with the
287 $PAX7^{+}SOX10^{+}$ rate, and ranks in the top half of pHI/RVIS scores among X-linked genes. Although
288 *RPS4X* is not implicated in cardinal TS phenotypes like short stature and gonadal dysgenesis [45,46],
289 the combined dosage *RPS4X* and *RPS4Y1* may play an important role neural crest-derived cell
290 types. To-date, escape of *RPS4X* from XCI has only been reported for primates [47], which have

`

also maintained *RPS4Y* on the Y[3]. Our study therefore raises the question whether neural-crest related TS phenotypes may be most penetrant in mammals that also maintained Y-linked *RPS4Y*, which may be resolved in comparative cardiovascular studies of mammals with sufficiently viable monosomy-X [9,48–50].

It is also plausible that the neural crest may be sensitive to the dosage of multiple PAR genes and gametologs, which poses challenges to prioritizing candidate dosage-sensitive genes by mutational constraint measures that are based on the frequency of loss-of-function variants in the population. The GenTAC (Genetically Triggered Thoracic Aortic Aneurysms and Other Cardiovascular Conditions) Registry reported that distal segmental Xp deletions minimally encompassing PAR1 and 21 additional genes can be sufficient for bicuspid aortic valve and aortic coarctation [51]. Interestingly, no such left-sided lesions were observed in 13 subjects non-mosaic for 46,X,i(Xq) iso-chromosomes, pointing to increased Xq copy-number as a potential protective factor when Xp dosage is lacking. Yet, as the iXq would be subject to XCI, such protective gene(s) would also have to escape XCI. While this group of candidate protective genes is not necessarily confined to gametologs, it should be noted that it only includes two such genes (*RPS4X, RBMX2)* [52]. In sum, monosomy-X hiPSC-derived NCCs with otherwise isogenic euploid controls, described herein provide a tractable model to resolve long-standing questions on the dosage contributions of human PAR and gametolog pairs in neural crest-derived cell types that give rise to clinically-relevant features of TS.

`

FIGURE LEGENDS

Figure 1: Monosomy-X impact on neural crest differentiation relative to isogenic euploid controls. (A) Representative immunofluorescence (IF) images of hiPSC-derived pairs of monosomy-X alongside euploid control NCCs representing three donors (Male1/2 and Female) show PAX7 (green), SOX10 (red), nuclei (Hoechst33342) and merged channels (100 um scale bar). **(B)** CellProfiler quantification of all IF images from 4-7 rounds of differentiation across all cell lines. Brackets and p-value (Mann-Whitney U test) indicate groups of cell lines compared within each and across isogenic panels. Panels separate lines from each donor, symbols and colors denote karyotype and donor.

Figure 2: Reduction of neural crest marker expression in monosomy-X relative isogenic euploid controls. (A) Principal component analysis (PCA) plot segregates RNA-seq samples by the combination of donor and karyotype, denoted "condition", with % of variance indicated on axes. **(B)** Variance-stabilized counts (vst) of early and late anterior NC (e/lANC), retinoic acid-treated NC (raANC) and neuromesodermal progenitor (NMP) markers from [18] (vertical panels) were median-normalized and averaged over euploid and 45,X (EU & XO, respectively) NC samples for each (horizontal) donor panel. Significant differences in median expression denoted by Mann-Whitney P-value. **(C,D)** as in B) for respectively p75-/+ associated markers [15], and chick NC marker sets [19]. **(E)** Pearson correlation of the matching $PAX7^{+}SOX10^{+}$ percentage with averaged marker sets from B-D, alongside cycling markers [53], NC markers from a 3D folding human neural tube model [21], and averaged pseudo-autosomal, X/Y Pair, and other escapee genes (PAR/PAIR/oESC) across all RNA-seq samples. Pairwise correlations only shown where significant ($p \leq 0.05$, Fisher transformed Pearson R).

Figure 3: Concordant impact of monosomy-X on NCC transcriptomes across isogenic panels. **(A)** *Left:* Venn diagram of differentially expressed genes (DEGs) in male1, male2 and female (mXO1, mXO2) monosomy-X NCCs. Significance of pairwise overlapping DEGs in bolded p-values (hypergeometric distribution). *Right:* Differential vst heatmaps for overlapping DEG sets from (A) as denoted by arrows. Ratio and p-value (sign-test) labels denote the number of DEGs with concordant direction ("Dir.") in the triple and all pairwise overlapping DEG sets, which is also visualized in the divergent (fXO, mXO1/2) annotation panels showing the Wald statistic (DESeq2) for each gene (row dendrogram not shown). Column dendrogram segregates samples by karyotype, irrespective of donor. **(B)** Gene set enrichment analysis (GSEA) results for the WikiPathway collection, ordered the quantile-normalized mean Wald statistic across all monosomy-X conditions (ave.XO). X axis denotes the log-scaled GSEA adjusted p-value, and colors the normalized enrichment score of up- (red) and down-(blue) regulated genes associated with a given term (bubble size depicts the corresponding number of genes). Results shown for all individual isogenic monosomy-X comparisons (fXO, mXO1/2), averaged (ave.XO/ave.mXO) Wald rankings and representative control comparisons (fmE1/2: female-to-male euploid, and m1m2O: male1/2 X-monosomic samples to each other). **(C)** GSEA results for the Human Phenotype Ontology (HPO) showing significantly ($p.adjust \leq 0.1$) enriched terms associated with Turner Syndrome (Orphanet ID: 881), ordered by the quantile-normalized mean Wald statistic of male1/2 monosomy-X (ave.mXO, otherwise as in B).

`

Figure 4: Monosomy-X sensitive gene modules correlate in human development and hiPSC-derived NCC models of monogenic neurocristopathy. (A) Color labels and assigned gene totals for WGCNA modules (1-29) with corresponding: **(i)** Correlation matrix (Pearson R) of each module eigengene (average module expression) to karyotype status (euploid, 45,X) and $PAX7^{+}SOX10^{+}$ percentage, as well as averaged expression of marker gene sets (from Fig. 2), and sex-linked genes (PAR, X/Y-pair, other escapees and "All" of these classes combined). Pearson R plotted only for significant correlations (p≤0.05, Fisher transformation), with log-scaled adjusted p-values printed to the nearest integer. **(ii)** Significantly-enriched (p.adjust≤0.1, hypergeometric distribution) gene terms from the Hallmark and canonical pathways collections of the molecular signatures database (MSigDB) pointing to metabolic and developmental processes. **(iii)** Preservation statistics (integer-rounded Z-score) in transcriptomes of the developing human face [26,27] (Carnegie stages CS13-17 & 22) and heart[25] (CS13-23), alongside hiPSC-derived NCC models of monogenic neurocristopathy syndromes (Pierre-Robin/*SOX9* [29], Waardenburg/*SOX10* [28], Familial Dysautonomia/*IKBKAP* [30], Bohring-Opitz/*ASXL1* [31], Floating-Harbor/*SRCAP* [32], and Branchio-Oculofacial/*TFAP2A* [33]). Monosomy-X human iPSC-derived trophoblast-like cells (TBL [12]) and pancreatic tumor stroma (Stroma [54]) were included as positive and negative controls, respectively. **(iv)** Wald (DESeq2) statistic averaged across genes within each module. **(B)** Enrichment map of significantly enriched gene ontology terms (GO, p.adjust≤0.01, hypergeometric distribution) as nodes. Genes shared between nodes (≥10%) are depicted as edges, based on semantic pairwise-term similarity (Jaccard distance) that is used to cluster nodes and summarize labels. Nodes are colored as pie charts by modules enriched for a given term, and the fraction of genes from each module under that term. **(C)** Module eigengene correlations (kME^2) for each sex-linked gene plotted over its correlation with the $PAX7^{+}SOX10^{+}$ percentage. Colors denote assigned module, sizes indicate pHI rank percentile **(D)** Dotplot of HPO terms linked to Turner syndrome that were significantly-enriched (p.adjust ≤0.1, hypergeometric distribution) in a given module (x-axis, numeric module labels with gene totals in parentheses). Dot color and size denote log-scaled adjusted p-value and ratio of term-associated genes over all genes in a given module. **(E)** Heatplots of TS-associated HPO terms linked to TS and their genes enriched in cluster 6. Colors denote average of DESeq2 quantile-normalized Wald statistic over all 45,X conditions (*top*) or Pearson R correlation with the $PAX7^{+}SOX10^{+}$ rate.

`

409 **SUPPLEMENTARY FIGURE LEGENDS**
410
411 **Figure S1: PAX7 and SOX10 quantification in isogenic monosomy-X and euploid control NCC**
412 **differentiation panels. (A,B)** Representative immunofluorescence (IF) images of hiPSC-derived
413 pairs of monosomy-X alongside euploid control NCCs (from male2 and female donors) show PAX7
414 (green), SOX10 (red), nuclei (Hoechst33342) and merged channels (100 um scale bar). **(C,D)**
415 CellProfiler IF quantification of all images from 4-7 rounds of differentiation across all cell lines.
416 Brackets and p-value (Mann-Whitney U test) indicate groups of cell lines compared within each
417 isogenic panel. Panels separate PAX7 and SOX10 percentages, as well as lines from each donor.
418 Symbols and colors denote karyotype and donor.
419
420 **Figure S2: RNA-seq samples, marker expression and allele-specific analysis for sex-linked genes.**
421 **(A)** Dendrogram of RNA-seq samples segregating by donor and karyotype. **(B)** NC-relevant marker
422 gene sets clustered by median expression across each set. Plot facetted by donor, with karyotype
423 indicated below dendrograms. **(C)** Expression levels (vst) of *HOXA* and *HOXB* cluster genes by
424 condition. **(D)** Vst counts of NC and neuro-ectodermal (NE) markers, distinguished in a human 3D
425 model of folding neural tube from [21], were median-normalized and averaged over euploid and
426 45,X (EU & XO, respectively) NC samples for each (horizontal) donor panel. Significant differences
427 in median expression denoted by Mann-Whitney P-value. **(E)** Heatmap of median-normalized vst
428 differential for PAR, X-pair, and other escapee genes, alongside their reported escapee status
429 [52,55,56]. Left-most barplot annotation panels show each gene's allelic ratio in female euploid XX
430 lines, and the three central barplot panels the log2FoldChange for differentially expressed gene
431 (DESeq2 p.adjust≤0.1) in the three isogenic comparisons (fXO, mXO1, and mXO2, or 0 if not
432 significant). **(F)** Expression levels (vst) of X-linked lncRNA genes *FIRRE, FTX, JPX* and *XIST* by
433 condition.
434
435 **Figure S3: (A)** Gene set enrichment analysis (GSEA) results for the Hallmark collection, ordered
436 the quantile-normalized mean Wald statistic across all monosomy-X conditions (ave.mXO). X axis
437 denotes the log-scaled GSEA adjusted p-value, and colors the normalized enrichment score of
438 up- (red) and down-(blue) regulated genes associated with a given term (bubble size depicts the
439 corresponding number of genes). Results shown for all individual isogenic monosomy-X
440 comparisons (fXO, mXO1/2), averaged (ave.XO/ave.mXO) Wald rankings and representative
441 control comparisons (fmE1/2: female-to-male euploid, and m1m2O: male1/2 X-monosomic
442 samples to each other). **(B)** GSEA results for the KEGG collection, otherwise as in A). **(C)** GSEA for
443 the DESCARTES single-cell atlas, otherwise as in A).
444
445
446
447
448
449
450
451
452

`

Figure S4: WGCNA module preservation, eigengene correlation and term enrichments.
(A) Differential in Z-score preservation statistic after subtracting Z of a sample-size adjusted network (15 samples) across all karyotypes from corresponding Z of a network composed entirely of male-derived monosomy-X samples. Numeric labels and bar widths indicate module labels and size (gene total), respectively. **(B)** Dotplot of Hallmark and KEGG terms that were significantly-enriched (respectively, p.adjust≤0.05, hypergeometric distribution) in a given module (x-axis, numeric module labels with gene totals in parentheses). Dot color and size denote log-scaled adjusted p-value and ratio of term-associated genes over all genes in a given module. **(C)** Pearson correlation matrix of all module eigengenes to each other and $PAX7^{+}SOX10^{+}$ percentage (denoted by arrow). **(D)** Module eigengene correlations (kME, Pearson R) of sex-linked genes shared by X/Y or escaping XCI, alongside their reported escapee status (from [52,55,56]), and X-chromosomal region (PAR, non-PAR). Column annotations plot significantly enriched (log-scaled p-value, Fisher exact test) classes of sex-linked and/or XCI-escaping genes by module. **(E)** Enrichment map of significantly enriched HPO terms linked to TS (p.adjust≤0.1, hypergeometric distribution) as individually labeled nodes with edges drawn for shared genes (≥2%). Nodes are colored as pie charts by modules enriched for a given term, and the fraction of genes from each module under that term.

Supplementary Tables:

Table S1: Gene-set enrichment analysis (GSEA) results for all monosomy X-relevant comparisons ("condition") across GO and selected MSigDB collections (Hallmark, canonical pathways, Human Phenotype Ontology & cell-type signatures). Lists normalized enrichment score (NES) and p.adjust for each comparison included for GSEA (ave.XO, ave.mXO, fXO, mXO1, mXO2, fmE1, fmE2 & m1m2O).

Table S2: Over-representation analysis (ORA) results for all WGCNA modules across selected GO and MSigDB collections (Hallmark, canonical pathways, Human Phenotype Ontology & cell-type signatures). Lists: ID, Description, GeneRatio, BgRatio and p.adjust for each module.

`

MATERIALS AND METHODS

Human iPSC culture

Euploid and X-monosomic hiPSC were derived from female and male donors, and characterized for pluripotency markers, as well as DNA methylation and X-chromosome inactivation in female euploid hiPSCs via continued *XIST* expression in RNA-seq as previously described [12,57]. Human iPSCs were cultured in feeder-free conditions on GelTrex-coated plates (ThermoFisher Scientific) in mTeSR-1 media (Stem Cell Technologies) in 5% CO_2 at 37°C. iPSCs were passaged with 0.5M EDTA at least weekly in small aggregates.

Neural Crest Cell Differentiation

Confluent wells of iPSCs cultured on Geltrex-coated plates in mTeSR-1 medium were singularized with Accutase (ThermoFisher Scientific) and plated on Geltrex coated plates at 4.5-6 $x10^4$ cells/cm^2 in DMEM/F-12, 1x N2 (recipe modified from Waisman Center Intellectual & Developmental Disabilities Research Center at University of Wisconsin-Madison: Insulin, 1.6 mg/mL; apo-Transferrin, 3.2 mg/mL; Putrescine, 0.5 mg/mL; Progesterone, 0.1 µg/mL; Sodium selenite, 0.16 µg/mL (all from MilliporeSigma)), 1% Glutamax, 1% Non-essential Amino Acids, 1 µM CHIR99021 (StemCell Technologies), 2 µM SB431542 (StemCell Technologies), 1 µM DMH1 (Tocris Bioscience), 20 ng/mL BMP4 (Peprotech), and 10 µM Y-27632 (Tocris Bioscience), for the first two days only. Media was changed daily until day 4 or 5 when the cells were harvested.

Immunocytochemistry

NCCs were fixed with 4% paraformaldehyde for 10 minutes at room temperature, washed with 0.1% Triton-X, permeabilized with 1% Triton-X, blocked with 5% normal goat serum/2% BSA/0.1% Triton-X for 1 hour and incubated with mouse monoclonal PAX7 antibody (abcam, 1:200) and rabbit monoclonal SOX10 antibody (Cell Signaling Technologies, 1:200) overnight at 4°C. The cells were then washed, incubated with AlexaFluor-555 Goat-anti-Mouse and AlexFluor-647 Goat-anti-Rabbit Secondary Antibodies (Thermo Fisher Scientific, 1:500) for 1 hour at room temperature. The cells were then washed, stained with Hoecsht-33342 diluted 1:10,000, and mounted with ProLong Gold Antifade Mountant (ThermoFisher Scientific. The number of PAX7- and SOX10-positive cells was imaged with an EVOS FL Auto system (ThermoFisher Scientific).

Cell Profiler analysis

The proportion of PAX7+ and SOX10+ cells was quantified by automated image analysis of Hoechst-, PAX7-, and SOX10-stained slides using CellProfiler version 4.2.4 [58]. Nuclei were first identified in the Hoechst channel, then the PAX7 and SOX10 channels were used to quantify the number of PAX7+/SOX10+ double-positive, either single-positive type, and double-negative cells per image. Illumination-corrected images were used for all channels, and nuclei outside the 10th to 90th percentile size range for each Hoechst image were excluded. For each biological replicate, the values from 3 fields of view were averaged to determine the percentage of single-positive, and $PAX7^+SOX10^+$ double-positive cells, with an average of ~1000 nuclei counted for each datapoint.

`

RNA-seq analysis

RNA was extracted from NCCs using the PureLink RNA Mini Kit (ThermoFisher Scientific). Libraries were prepared at the UConn Center for Genome Innovation using the Illumina Strand mRNA Kit and 100bp paired-ends reads were sequenced to an average depth of 40 million reads/replicate on the NovaSeq (Illumina).

Read pairs were trimmed using fastp [59], aligned to the human reference genome (hg38) using hisat2 [60] for allele-specific analysis of phased variants using phaser [61], and quantified against GENCODE version (v36) using salmon [62]. For analysis of XCI and escapee, A and B allele counts from phASER were tabulated and calls made by binomial test (lesser allele fraction, LAF > 0.1, $p \leq 0.05$) for all X-linked genes.

For differential expression using DESeq2 [63], salmon count tables were filtered for genes with sufficient expression. Surrogate variables were estimated using the sva package [64], and added to the DESeq2 design testing for condition and correcting for differentiation round. Gene-set enrichment analysis (GSEA) using clusterProfiler [65] was performed on all genes ranked by DESeq2's Wald statistic, as well as the average of the quantile-normalized Wald scores from the three euploid-45,X conditions (fXO, mXO1, mXO2) to ensure equal weighting.

Weighted gene co-expression network analysis (WGCNA) was performed on vst counts, using the WGCNA package [66], as a signed hybrid network using the biweight midcorrelation raised to a soft thresholding power of 12 (scale-free topology fit ≥ 0.85). Modules were correlated to $PAX7^{+}SOX10^{+}$ percentages and to averaged NCC lineage marker sets, which were median vst normalized to ensure equal weights across all sets. Module preservation analysis was performed against published RNA-seq datasets of the developing face [26,27], and heart [25], hiPSC-derived NCC models of Pierre-Robin [29], Waardenburg [28], Familial Dysautonomia [30], Bohring-Opitz [31], Floating-Harbor [32], and Branchio-Oculofacial [33] syndromes, as well as monosomy-X hiPSC-derived trophoblast-like cells [12] and pancreatic tumor stroma [54]. Overrepresentation analysis of gene terms across WGCNA modules was performed with clusterProfiler [65].

ACKNOWLEDGMENTS

We would like to thank the Cotney lab at UConn Health for access to unpublished transcriptome data from the developing face, and Bo Reese at the UConn Center for Genome Innovation for mRNA-seq library preparation and sequencing. This work was supported by NIH grants R35GM124926 and R01HL141324 to S.F.P.

`

585 **REFERENCES**

586 1. Hook, E. B. & Warburton, D. The distribution of chromosomal genotypes associated with
587 Turner's syndrome: livebirth prevalence rates and evidence for diminished fetal mortality
588 and severity in genotypes associated with structural X abnormalities or mosaicism. *Hum.*
589 *Genet.* **64**, 24–7 (1983).

590 2. Hook, E. B. & Warburton, D. Turner syndrome revisited: review of new data supports the
591 hypothesis that all viable 45,X cases are cryptic mosaics with a rescue cell line, implying
592 an origin by mitotic loss. *Hum. Genet.* **133**, 417–24 (2014).

593 3. Bellott, D. W., Hughes, J. F., Skaletsky, H., Brown, L. G., Pyntikova, T., Cho, T.-J., *et al.*
594 Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature*
595 **508**, 494–9 (2014).

596 4. Posynick, B. J. & Brown, C. J. Escape From X-Chromosome Inactivation: An Evolutionary
597 Perspective. *Front. Cell Dev. Biol.* **7**, 241 (2019).

598 5. Gravholt, C. H., Viuff, M. H., Brun, S., Stochholm, K. & Andersen, N. H. Turner syndrome:
599 mechanisms and management. *Nat. Rev. Endocrinol.* **15**, 601–614 (2019).

600 6. Tang, W. & Bronner, M. E. Neural crest lineage analysis: From past to future trajectory.
601 *Dev.* **147**, (2021).

602 7. Gravholt, C. H., Viuff, M., Just, J., Sandahl, K., Brun, S., van der Velden, J., *et al.* The
603 Changing Face of Turner Syndrome. *Endocr. Rev.* **44**, 33–69 (2023).

604 8. Perry, J., Palmer, S., Gabriel, A. & Ashworth, A. A short pseudoautosomal region in
605 laboratory mice. *Genome Res.* **11**, 1826–32 (2001).

606 9. Raudsepp, T. & Chowdhary, B. P. The Eutherian Pseudoautosomal Region. *Cytogenet.*
607 *Genome Res.* (2016) doi:10.1159/000443157.

608 10. Lynn, P. M. Y. & Davies, W. The 39,XO mouse as a model for the neurobiology of Turner
609 syndrome and sex-biased neuropsychiatric disorders. *Behavioural Brain Research* vol.
610 179 173–182 (2007).

611 11. Probst, F. J., Cooper, M. L., Cheung, S. W. & Justice, M. J. Genotype, phenotype, and
612 karyotype correlation in the XO mouse model of turner syndrome. *J. Hered.* **99**, 512–517
613 (2008).

614 12. Ahern, D. T., Bansal, P., Armillei, M. K., Faustino, I. V., Kondaveeti, Y., Glatt-Deeley, H. R.,
615 *et al.* Monosomy X in isogenic human iPSC-derived trophoblast model impacts expression
616 modules preserved in human placenta. *Proc. Natl. Acad. Sci.* **119**, e2211073119 (2022).

617 13. Rothstein, M., Bhattacharya, D. & Simoes-Costa, M. The molecular basis of neural crest
618 axial identity. *Dev. Biol.* **444**, S170–S180 (2018).

619 14. Etchevers, H. C., Dupin, E. & Le Douarin, N. M. The diverse neural crest: From embryology
620 to human pathology. *Dev.* **146**, (2019).

621 15. Lee, G., Kim, H., Elkabetz, Y., Al Shamy, G., Panagiotakos, G., Barberi, T., *et al.* Isolation
622 and directed differentiation of neural crest stem cells derived from human embryonic

stem cells. *Nat. Biotechnol.* **25**, 1468–1475 (2007).

16. Leung, A. W., Leung, A. W., Murdoch, B., Salem, A. F., Prasad, M. S., Gomez, G. A., *et al.* WNT/β-catenin signaling mediates human neural crest induction via a pre-neural border intermediate. *Dev.* **143**, 398–410 (2016).

17. Hackland, J. O. S., Frith, T. J. R., Thompson, O., Marin Navarro, A., Garcia-Castro, M. I., Unger, C., *et al.* Top-Down Inhibition of BMP Signaling Enables Robust Induction of hPSCs Into Neural Crest in Fully Defined, Xeno-free Conditions. *Stem Cell Reports* **9**, 1043–1052 (2017).

18. Frith, T. J., Granata, I., Wind, M., Stout, E., Thompson, O., Neumann, K., *et al.* Human axial progenitors generate trunk neural crest cells in vitro. *Elife* **7**, (2018).

19. Simões-Costa, M. & Bronner, M. E. Establishing neural crest identity: a gene regulatory recipe. *Development* **142**, 242–257 (2015).

20. Soldatov, R., Kaucka, M., Kastriti, M. E., Petersen, J., Chontorotzea, T., Englmaier, L., *et al.* Spatiotemporal structure of cell fate decisions in murine neural crest. *Science (80-.).* **364**, (2019).

21. Karzbrun, E., Khankhel, A. H., Megale, H. C., Glasauer, S. M. K., Wyle, Y., Britton, G., *et al.* Human neural tube morphogenesis in vitro by geometric constraints. *Nat. 2021 5997884* **599**, 268–272 (2021).

22. Vega, S., Morales, A. V., Ocaña, O. H., Valdés, F., Fabregat, I. & Nieto, M. A. Snail blocks the cell cycle and confers resistance to cell death. *Genes Dev.* **18**, 1131–1143 (2004).

23. Thiery, J. P., Acloque, H., Huang, R. Y. J. & Nieto, M. A. Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* **139**, 871–890 (2009).

24. Lebenthal, Y., Levy, S., Sofrin-Drucker, E., Nagelberg, N., Weintrob, N., Shalitin, S., *et al.* The Natural History of Metabolic Comorbidities in Turner Syndrome from Childhood to Early Adulthood: Comparison between 45,X Monosomy and Other Karyotypes. *Front. Endocrinol. (Lausanne).* **9**, 27 (2018).

25. Vanoudenhove, J., Yankee, T. N., Wilderman, A. & Cotney, J. Epigenomic and Transcriptomic Dynamics During Human Heart Organogenesis. *Circ. Res.* **127**, E184–E209 (2020).

26. Yankee, T. N., Wilderman, A., Winchester, E. W., VanOudenhove, J. & Cotney, J. Integrative analysis of transcriptomics in human craniofacial development reveals novel candidate disease genes. *bioRxiv* 2022.02.28.482338 (2022) doi:10.1101/2022.02.28.482338.

27. Brinkley, J. F., Fisher, S., Harris, M. P., Holmes, G., Hooper, J. E., Jabs, E. W., *et al.* The facebase consortium: A comprehensive resource for craniofacial researchers. *Dev.* **143**, 2677–2688 (2016).

28. Wen, J., Song, J., Bai, Y., Liu, Y., Cai, X., Mei, L., *et al.* A Model of Waardenburg Syndrome Using Patient-Derived iPSCs With a SOX10 Mutation Displays Compromised Maturation

661 and Function of the Neural Crest That Involves Inner Ear Development. *Front. cell Dev.*
662 *Biol.* **9**, (2021).

663 29. Long, H. K., Osterwalder, M., Welsh, I. C., Hansen, K., Davies, J. O. J., Liu, Y. E., *et al.* Loss
664 of Extreme Long-Range Enhancers in Human Neural Crest Drives a Craniofacial Disorder.
665 *Cell Stem Cell* **27**, 765-783.e14 (2020).

666 30. Zeltner, N., Fattahi, F., Dubois, N. C., Saurat, N., Lafaille, F., Shang, L., *et al.* Capturing the
667 biology of disease severity in a PSC-based model of familial dysautonomia. *Nat. Med.* **22**,
668 1421–1427 (2016).

669 31. Matheus, F., Rusha, E., Rehimi, R., Molitor, L., Pertek, A., Modic, M., *et al.* Pathological
670 ASXL1 Mutations and Protein Variants Impair Neural Crest Development. *Stem cell*
671 *reports* **12**, 861–868 (2019).

672 32. Greenberg, R. S., Long, H. K., Swigut, T. & Wysocka, J. Single Amino Acid Change
673 Underlies Distinct Roles of H2A.Z Subtypes in Human Syndrome. *Cell* **178**, 1421-1436.e24
674 (2019).

675 33. Laugsch, M., Bartusel, M., Rehimi, R., Alirzayeva, H., Karaolidou, A., Crispatzu, G., *et al.*
676 Modeling the Pathological Long-Range Regulatory Effects of Human Structural Variation
677 with Patient-Specific hiPSCs. *Cell Stem Cell* **24**, 736-752.e12 (2019).

678 34. Roman, A. K. S., Godfrey, A. K., Skaletsky, H., Bellott, D. W., Groff, A. F., Harris, H. L., *et al.*
679 The human inactive X chromosome modulates expression of the active X chromosome.
680 *Cell Genomics* **3**, 100259 (2023).

681 35. Huang, N., Lee, I., Marcotte, E. M. & Hurles, M. E. Characterising and Predicting
682 Haploinsufficiency in the Human Genome. *PLOS Genet.* **6**, e1001154 (2010).

683 36. Petrovski, S., Wang, Q., Heinzen, E. L., Allen, A. S. & Goldstein, D. B. Genic Intolerance to
684 Functional Variation and the Interpretation of Personal Genomes. *PLOS Genet.* **9**,
685 e1003709 (2013).

686 37. Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., *et al.* The
687 mutational constraint spectrum quantified from variation in 141,456 humans. *Nat. 2020*
688 *5817809* **581**, 434–443 (2020).

689 38. Rao, E., Weiss, B., Fukami, M., Rump, A., Niesler, B., Mertz, A., *et al.* Pseudoautosomal
690 deletions encompassing a novel homeobox gene cause growth failure in idiopathic short
691 stature and Turner syndrome. *Nat. Genet.* **16**, 54–63 (1997).

692 39. Cobb, J., Dierich, A., Huss-Garcia, Y. & Duboule, D. A mouse model for human short-
693 stature syndromes identifies Shox2 as an upstream regulator of Runx2 during long-bone
694 development. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 4511–5 (2006).

695 40. Prakash, V., Carson, B. B., Feenstra, J. M., Dass, R. A., Sekyrova, P., Hoshino, A., *et al.*
696 Ribosome biogenesis during cell cycle arrest fuels EMT in development and disease. *Nat.*
697 *Commun. 2019 101* **10**, 1–16 (2019).

698 41. Bhattacharya, D., Azambuja, A. P. & Simoes-Costa, M. Metabolic Reprogramming

Promotes Neural Crest Migration via Yap/Tead Signaling. *Dev. Cell* **53**, 199-211.e6 (2020).

42. Costa, R., Muccioli, S., Brillo, V., Bachmann, M., Szabò, I. & Leanza, L. Mitochondrial dysfunction interferes with neural crest specification through the FoxD3 transcription factor. *Pharmacol. Res.* **164**, (2021).

43. Castro, V. L., Reyes-Nava, N. G., Sanchez, B. B., Gonzalez, C. G., Paz, D. & Quintana, A. M. Activation of WNT signaling restores the facial deficits in a zebrafish with defects in cholesterol metabolism. *genesis* **58**, e23397 (2020).

44. Bhattacharya, D., Khan, B. & Simoes-Costa, M. Neural crest metabolism: At the crossroads of development and disease. *Dev. Biol.* **475**, 245–255 (2021).

45. Omoe, K. & Endo, A. Relationship between the monosomy X phenotype and Y-linked ribosomal protein S4 (Rps4) in several species of mammals: a molecular evolutionary analysis of Rps4 homologs. *Genomics* **31**, 44–50 (1996).

46. Geerkens, C., Just, W., Held, K. R. & Vogel, W. Ullrich-Turner syndrome is not caused by haploinsufficiency of RPS4X. *Hum. Genet. 1996 971* **97**, 39–44 (1996).

47. Balaton, B. P., Fornes, O., Wasserman, W. W. & Brown, C. J. Cross-species examination of X-chromosome inactivation highlights domains of escape from silencing. *Epigenetics and Chromatin* **14**, 12 (2021).

48. Bugno-Poniewierska, M. & Raudsepp, T. Horse Clinical Cytogenetics: Recurrent Themes and Novel Findings. *Anim. an Open Access J. from MDPI* **11**, 1–26 (2021).

49. Szczerbal, I. & Switonski, M. Genetic disorders of sex development in cats: An update. *Anim. Reprod. Sci.* **216**, 106353 (2020).

50. Li, G., Davis, B. W., Raudsepp, T., Pearks Wilkerson, A. J., Mason, V. C., Ferguson-Smith, M., *et al.* Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. *Genome Res.* **23**, 1486–95 (2013).

51. Prakash, S. K., Bondy, C. A., Maslen, C. L., Silberbach, M., Lin, A. E., Perrone, L., *et al.* Autosomal and X chromosome structural variants are associated with congenital heart defects in Turner syndrome: The NHLBI GenTAC registry. *Am. J. Med. Genet. Part A* **170**, 3157–3164 (2016).

52. Tukiainen, T., Villani, A.-C., Yen, A., Rivas, M. A., Marshall, J. L., Satija, R., *et al.* Landscape of X chromosome inactivation across human tissues. *Nature* **550**, 244–248 (2017).

53. Hsiao, C. J., Tung, P. Y., Blischak, J. D., Burnett, J. E., Barr, K. A., Dey, K. K., *et al.* Characterizing and inferring quantitative cell cycle phase in single-cell RNA-seq data analysis. *Genome Res.* **30**, 611–621 (2020).

54. Birnbaum, D. J., Begg, S. K. S., Finetti, P., Vanderburg, C., Kulkarni, A. S., Neyaz, A., *et al.* Transcriptomic Analysis of Laser Capture Microdissected Tumors Reveals Cancer- and Stromal-Specific Molecular Subtypes of Pancreatic Ductal Adenocarcinoma. *Clin. Cancer Res.* **27**, 2314–2325 (2021).

55. Wainer Katsir, K. & Linial, M. Human genes escaping X-inactivation revealed by single cell

expression data. *BMC Genomics* **20**, 1–17 (2019).

56. Sauteraud, R., Stahl, J. M., James, J., Englebright, M., Chen, F., Zhan, X., *et al.* Inferring genes that escape X-Chromosome inactivation reveals important contribution of variable escape genes to sex-biased diseases. *Genome Res.* **31**, 1629–1637 (2021).

57. Bansal, P., Ahern, D. T., Kondaveeti, Y., Qiu, C. W. & Pinter, S. F. Contiguous erosion of the inactive X in human pluripotency concludes with global DNA hypomethylation. *Cell Rep.* **35**, 109215 (2021).

58. Mcquin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, A., Karhohs, K. W., *et al.* Cellprofiler 3.0. *PLoS Biol.* 1–17 (2018).

59. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).

60. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).

61. Castel, S. E., Mohammadi, P., Chung, W. K., Shen, Y. & Lappalainen, T. Rare variant phasing and haplotypic expression from RNA sequencing with phASER. *Nat. Commun.* **7**, 1–6 (2016).

62. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

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

64. Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E. & Storey, J. D. The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012).

65. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology* vol. 16 284–287 (2012).

66. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 1–13 (2008).

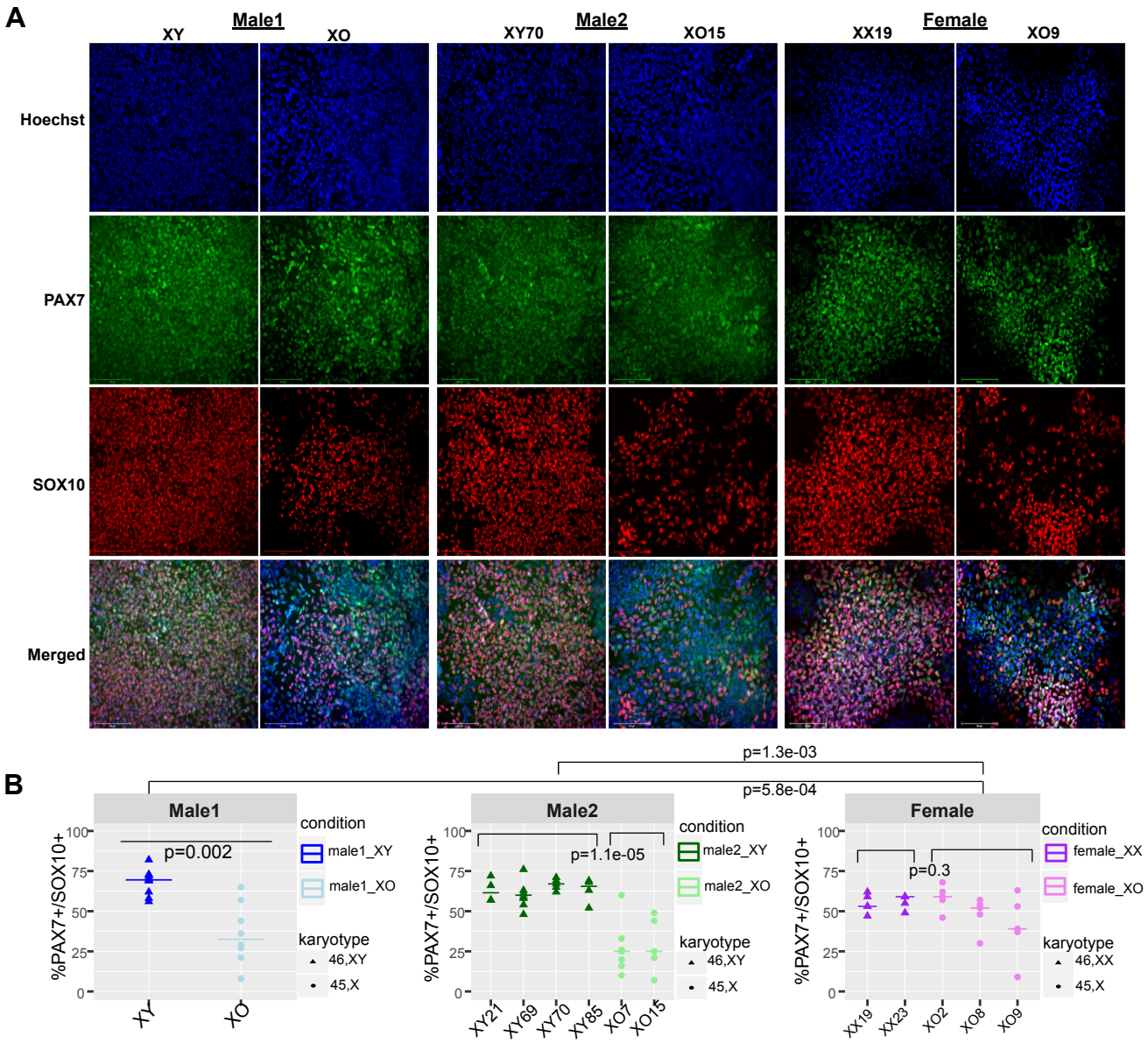
A
Male1
XY
XO
Male2
XY70
XO15
Female
XX19
XO9
Hoechst
PAX7
SOX10
Merged
B
p=1.3e-03
p=5.8e-04
Male1
p=0.002
%PAX7+/SOX10+
condition
male1_XY
male1_XO
karyotype
46,XY
45,X
XY
XO
Male2
p=1.1e-05
male2_XY
male2_XO
XY21
XY69
XY70
XY85
XO7
XO15
Female
p=0.3
female_XX
female_XO
46,XX
XX19
XX23
XO2
XO8
XO9

Figure 1

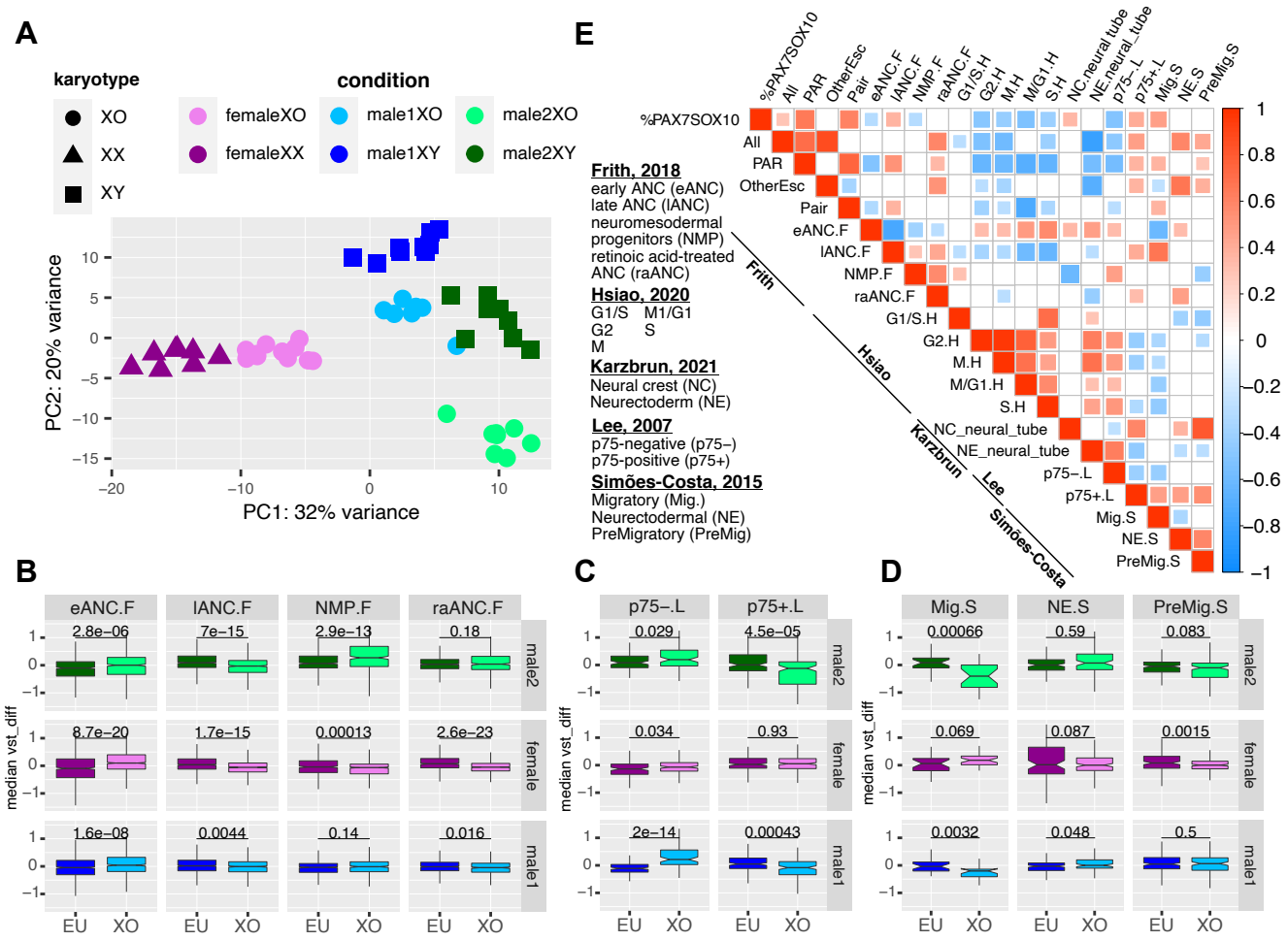
A
karyotype
XO
XX
XY
condition
femaleXO
femaleXX
male1XO
male1XY
male2XO
male2XY
PC2: 20% variance
PC1: 32% variance
E
%PAX7SOX10
All
PAR
OtherEsc
Pair
eANC.F
lANC.F
NMP.F
raANC.F
G1/S.H
G2.H
M.H
M/G1.H
S.H
NC_neural_tube
NE_neural_tube
p75−.L
p75+.L
Mig.S
NE.S
PreMig.S
Frith, 2018
early ANC (eANC)
late ANC (lANC)
neuromesodermal progenitors (NMP)
retinoic acid-treated ANC (raANC)
Hsiao, 2020
G1/S M1/G1
G2 S
M
Karzbrun, 2021
Neural crest (NC)
Neurectoderm (NE)
Lee, 2007
p75-negative (p75−)
p75-positive (p75+)
Simões-Costa, 2015
Migratory (Mig.)
Neurectodermal (NE)
PreMigratory (PreMig)
B
C
D
median vst_diff
male2
female
male1
EU
XO

Figure 2

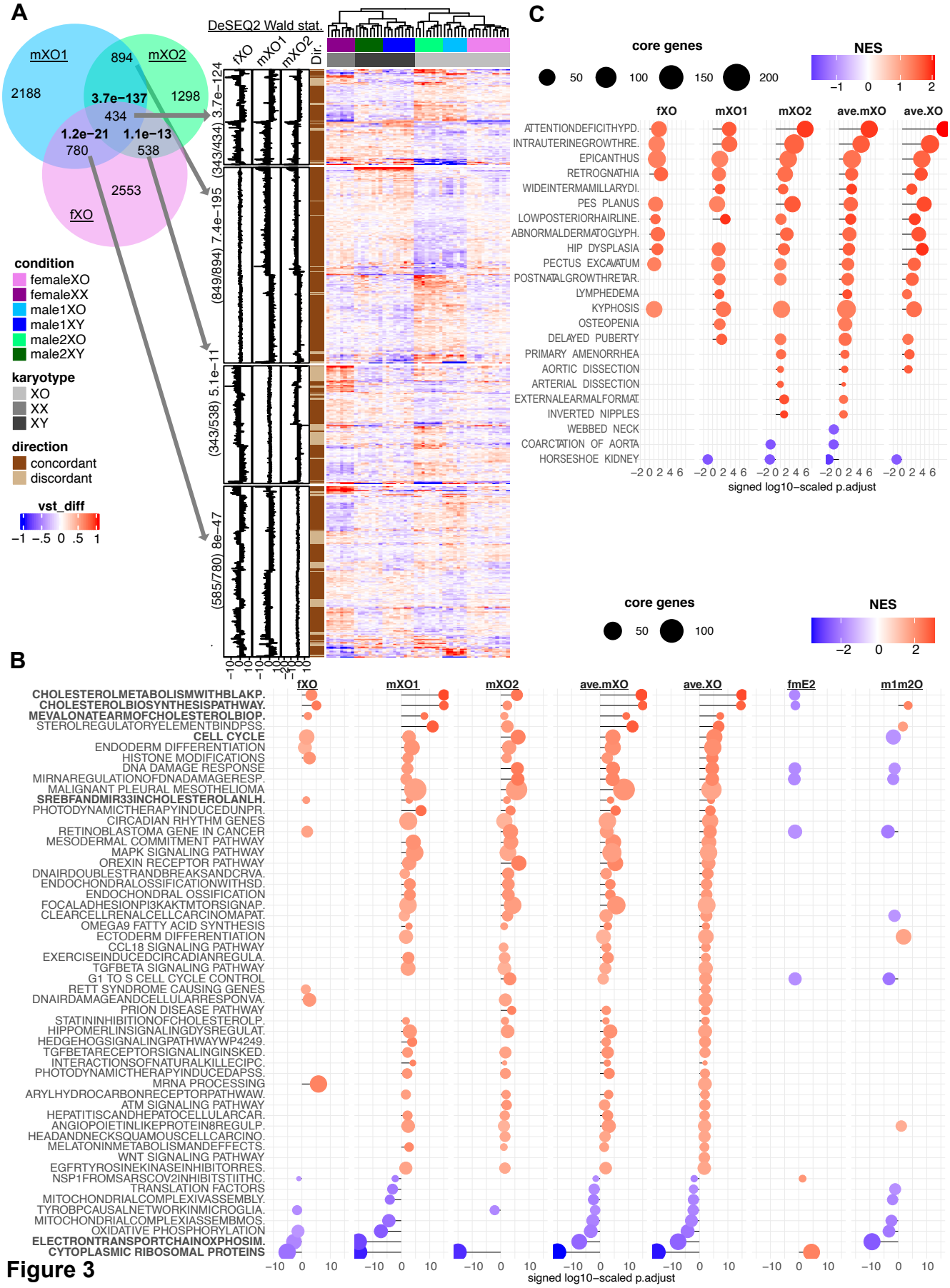
A
mXO1
mXO2
fXO
2188
894
1298
3.7e−137
434
1.2e−21
780
1.1e−13
538
2553
condition
femaleXO
femaleXX
male1XO
male1XY
male2XO
male2XY
karyotype
XO
XX
XY
direction
concordant
discordant
vst_diff
−1 −.5 0 .5 1
DeSEQ2 Wald stat.
fXO
mXO1
mXO2
Dir.
(343/434) 3.7e−124
(849/894) 7.4e−195
(343/538) 5.1e−11
(585/780) 8e−47
C
core genes
50 100 150 200
NES
−1 0 1 2
fXO
mXO1
mXO2
ave.mXO
ave.XO
ATTENTIONDEFICITHYPD.
INTRAUTERINEGROWTHRE.
EPICANTHUS
RETROGNATHIA
WIDEINTERMAMILLARYDI.
PES PLANUS
LOWPOSTERIORHAIRLINE.
ABNORMALDERMATOGLYPH.
HIP DYSPLASIA
PECTUS EXCAVATUM
POSTNATALGROWTHRETAR.
LYMPHEDEMA
KYPHOSIS
OSTEOPENIA
DELAYED PUBERTY
PRIMARY AMENORRHEA
AORTIC DISSECTION
ARTERIAL DISSECTION
EXTERNALEARMALFORMAT.
INVERTED NIPPLES
WEBBED NECK
COARCTATION OF AORTA
HORSESHOE KIDNEY
−2 0 2 4 6
signed log10−scaled p.adjust
B
core genes
50 100
NES
−2 0 2
fXO
mXO1
mXO2
ave.mXO
ave.XO
fmE2
m1m2O
CHOLESTEROLMETABOLISMWITHBLAKP.
CHOLESTEROLBIOSYNTHESISPATHWAY.
MEVALONATEARMOFCHOLESTEROLBIOP.
STEROLREGULATORYELEMENTBINDPSS.
CELL CYCLE
ENDODERM DIFFERENTIATION
HISTONE MODIFICATIONS
DNA DAMAGE RESPONSE
MIRNAREGULATIONOFDNADAMAGERESP.
MALIGNANT PLEURAL MESOTHELIOMA
SREBFANDMIR33INCHOLESTEROLANLH.
PHOTODYNAMICTHERAPYINDUCEDUNPR.
CIRCADIAN RHYTHM GENES
RETINOBLASTOMA GENE IN CANCER
MESODERMAL COMMITMENT PATHWAY
MAPK SIGNALING PATHWAY
OREXIN RECEPTOR PATHWAY
DNAIRDOUBLESTRANDBREAKSANDCRVA.
ENDOCHONDRALOSSIFICATIONWITHSD.
ENDOCHONDRAL OSSIFICATION
FOCALADHESIONPI3KAKTMTORSIGNAP.
CLEARCELLRENALCELLCARCINOMAPAT.
OMEGA9 FATTY ACID SYNTHESIS
ECTODERM DIFFERENTIATION
CCL18 SIGNALING PATHWAY
EXERCISEINDUCEDCIRCADIANREGULA.
TGFBETA SIGNALING PATHWAY
G1 TO S CELL CYCLE CONTROL
RETT SYNDROME CAUSING GENES
DNAIRDAMAGEANDCELLULARRESPONVA.
PRION DISEASE PATHWAY
STATININHIBITIONOFCHOLESTEROLP.
HIPPOMERLINSIGNALINGDYSREGULAT.
HEDGEHOGSIGNALINGPATHWAYWP4249.
TGFBETARECEPTORSIGNALINGINSKED.
INTERACTIONSOFNATURALKILLECIPC.
PHOTODYNAMICTHERAPYINDUCEDAPSS.
MRNA PROCESSING
ARYLHYDROCARBONRECEPTORPATHWAW.
ATM SIGNALING PATHWAY
HEPATITISCANDHEPATOCELLULARCAR.
ANGIOPOIETINLIKEPROTEIN8REGULP.
HEADANDNECKSQUAMOUSCELLCARCINO.
MELATONINMETABOLISMANDEFFECTS.
WNT SIGNALING PATHWAY
EGFRTYROSINEKINASEINHIBITORRES.
NSP1FROMSARSCOV2INHIBITSTIITHC.
TRANSLATION FACTORS
MITOCHONDRIALCOMPLEXIVASSEMBLY.
TYROBPCAUSALNETWORKINMICROGLIA.
MITOCHONDRIALCOMPLEXIASSEMBMOS.
OXIDATIVE PHOSPHORYLATION
ELECTRONTRANSPORTCHAINOXPHOSIM.
CYTOPLASMIC RIBOSOMAL PROTEINS
−10 0 10
signed log10−scaled p.adjust

Figure 3

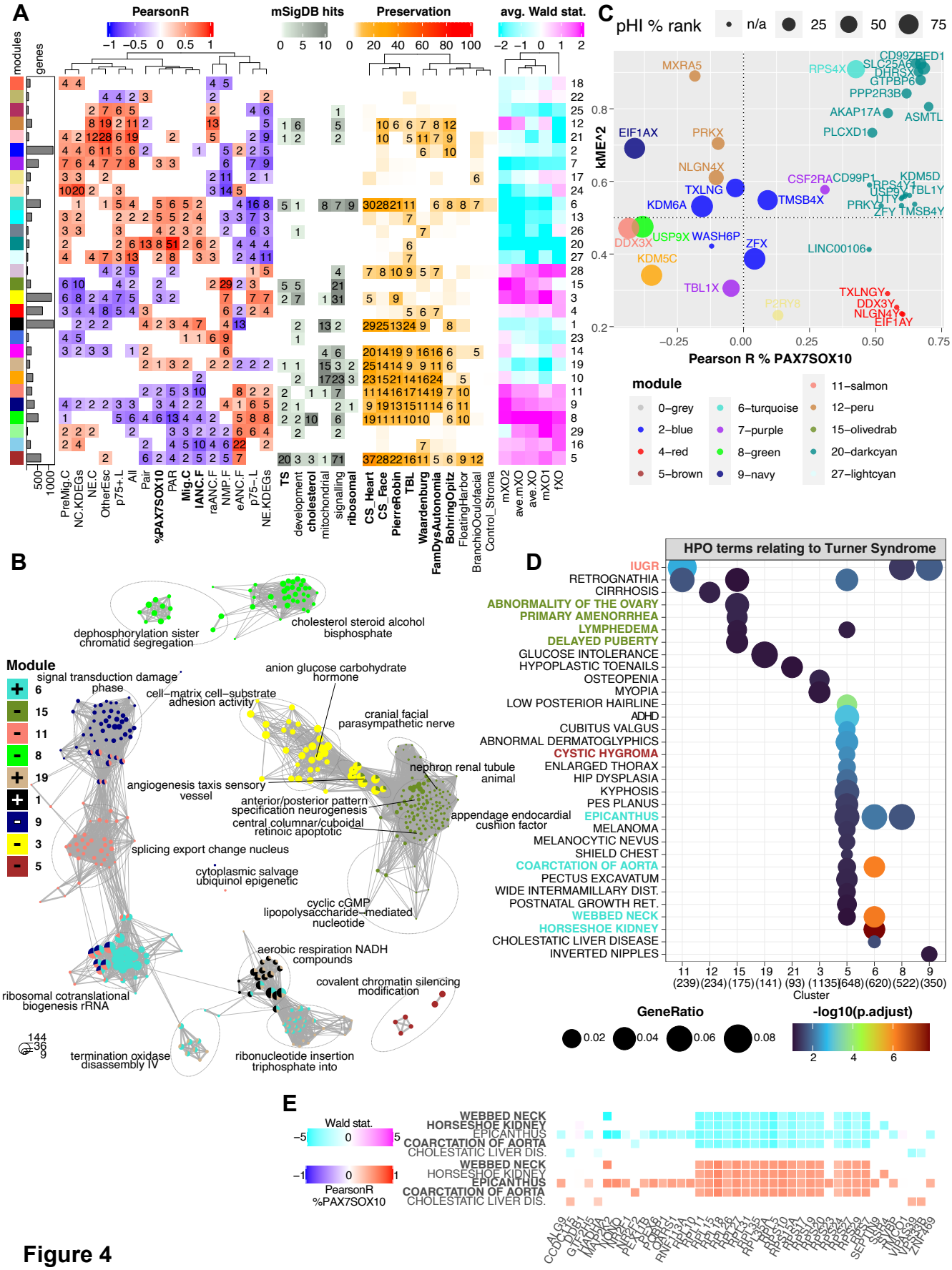
A
PearsonR
mSigDB hits
Preservation
avg. Wald stat.
modules
genes
PreMig.C
NC.KDEGs
NE.C
OtherEsc
p75+.L
All
Pair
%PAX7SOX10
PAR
Mig.C
IANC.F
raANC.F
NMP.F
eANC.F
p75−.L
NE.KDEGs
TS
development
cholesterol
mitochondrial
signalling
ribosomal
CS_Heart
CS_Face
PierreRobin
TBL
Waardenburg
FamDysAutonomia
BohringOpitz
FloatingHarbor
BranchioOculofacial
Control_Stroma
mXO2
ave.mXO
ave.XO
mXO1
fXO
C
pHI % rank
n/a
25
50
75
kME^2
Pearson R % PAX7SOX10
module
0–grey
2–blue
4–red
5–brown
6–turquoise
7–purple
8–green
9–navy
11–salmon
12–peru
15–olivedrab
20–darkcyan
27–lightcyan
B
dephosphorylation sister chromatid segregation
cholesterol steroid alcohol bisphosphate
Module
signal transduction damage phase
cell–matrix cell–substrate adhesion activity
anion glucose carbohydrate hormone
cranial facial parasympathetic nerve
nephron renal tubule animal
angiogenesis taxis sensory vessel
anterior/posterior pattern specification neurogenesis
central columnar/cuboidal retinoic apoptotic
appendage endocardial cushion factor
splicing export change nucleus
cytoplasmic salvage ubiquinol epigenetic
cyclic cGMP lipopolysaccharide–mediated nucleotide
aerobic respiration NADH compounds
covalent chromatin silencing modification
ribosomal cotranslational biogenesis rRNA
termination oxidase disassembly IV
ribonucleotide insertion triphosphate into
D
HPO terms relating to Turner Syndrome
IUGR
RETROGNATHIA
CIRRHOSIS
ABNORMALITY OF THE OVARY
PRIMARY AMENORRHEA
LYMPHEDEMA
DELAYED PUBERTY
GLUCOSE INTOLERANCE
HYPOPLASTIC TOENAILS
OSTEOPENIA
MYOPIA
LOW POSTERIOR HAIRLINE
ADHD
CUBITUS VALGUS
ABNORMAL DERMATOGLYPHICS
CYSTIC HYGROMA
ENLARGED THORAX
HIP DYSPLASIA
KYPHOSIS
PES PLANUS
EPICANTHUS
MELANOMA
MELANOCYTIC NEVUS
SHIELD CHEST
COARCTATION OF AORTA
PECTUS EXCAVATUM
WIDE INTERMAMILLARY DIST.
POSTNATAL GROWTH RET.
WEBBED NECK
HORSESHOE KIDNEY
CHOLESTATIC LIVER DISEASE
INVERTED NIPPLES
Cluster
GeneRatio
-log10(p.adjust)
E
Wald stat.
PearsonR %PAX7SOX10
WEBBED NECK
HORSESHOE KIDNEY
EPICANTHUS
COARCTATION OF AORTA
CHOLESTATIC LIVER DIS.

Figure 4