Brachyury expression levels predict lineage potential and axis-forming ability of *in vitro* derived neuromesodermal progenitors

running title: Brachyury dictates neuromesodermal lineage specification

Anahí Binagui-Casas*[#1], Anna Granés*[1], Alberto Ceccarelli[2], Filip J. Wymeersch[3], Matthew French[4], Rosa Portero[1], Jen Annoh[1], Yali Huang[5], Eleni Karagianni[6], Frederick C.K. Wong[7], A. Sophie Brumm[1], Daniel Lopez Ramajo[8], Minoru Takasato[3,9], Sally Lowell[1], Osvaldo Chara[2,10], Valerie Wilson[#1]

1. Institute for Regeneration and Repair, Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh EH16 4UU, Scotland.

2. School of Biosciences, University of Nottingham, Sutton Bonington Campus, Nottingham LE12, UK

3. Current: RIKEN Center for Biosystems Dynamics Research, 2-2-3 Minatojima-minamimachi, Kobe 650-0047, Japan.

4. Current: Department of Genetics, University of Cambridge, Cambridge, CB2 3EH UK

5. Current: Department of Cell, Developmental and Integrative Biology, University of Alabama at Birmingham, 1918 University Blvd, Birmingham, AL 35294, USA

6. Current: GSK, Gunnels Wood Road, Stevenage, SG1 2NY K

7. Current: The Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge CB10 1SA, UK

8. Current: Immunology Unit, Department of Pathology and Experimental Therapy, School of Medicine, Universitat de Barcelona, 08007, Barcelona, Spain

9. Laboratory of Molecular Cell Biology and Development. Department of Animal Development and Physiology, Graduate School of Biostudies, Kyoto University, Kyoto 606-8501, Japan.

10. Instituto de Tecnología, Universidad Argentina de la Empresa, Buenos Aires, Argentina.

* joint first authors

co-corresponding authors: a.binagui-casas@ed.ac.uk, v.wilson@ed.ac.uk

Keywords: Neuromesodermal progenitors, Axis elongation, Pluripotent stem cells, Mesoderm, Neurectoderm, In vitro differentiation, Single cell

Abstract

Neuromesodermal progenitors (NMPs) produce the spinal cord and musculoskeleton in the elongating anterior-posterior axis. *In vivo*, NMPs possess dual potency, coinciding with regions coexpressing SOX2 and Brachyury (TBXT). *In vitro*, SOX2/TBXT co-expressing cells can be produced from pluripotent cells and, like their *in vivo* counterparts, can produce neural tube and somitic mesoderm. However, the functional characteristics of *in vitro* SOX2/TBXT co-expressing cells remain unclear, confounding comparisons with *in vivo* data. To address this, we developed a dual *Sox2/Tbxt* reporter mouse ESC line. SOX2/TBXT reporter-positive cells emerge *in vitro* from pluripotent populations with dynamics that mirror their appearance in the embryo. Purified SOX2/TBXT co-expressing populations can differentiate towards neurectoderm or mesoderm, including lateral mesoderm upon BMP stimulation. In gastruloids, quantitative live imaging shows that WNT or NOTCH inhibition rapidly leads to downregulation of TBXT expression and diminished axial extension. We show that clonally plated SOX2/TBXT co-expressing cells are bipotent NMPs that can also self-propagate. By combining clonal analysis with mathematical modelling, we identify two thresholds of SOX2/TBXT expression, switching clonal output from neural- to mesoderm- biased, and from mesoderm-biased to mesoderm-specified. Media and substrate composition alter the lineage outcomes of *in vitro* derived mouse NMPs. Thus, this *Sox2/Tbxt* double reporter cell line provides support for unsuspected heterogeneity in NMPs, together with evidence for a role of these transcription factors in directing cell fate to drive axis elongation.

Introduction

In vertebrates, the head-to-tail (or anteroposterior, AP) body axis forms during early embryonic development. In mice, where this process is well characterised (Wymeersch et al., 2021), a transient population of progenitors, collectively called neuromesodermal competent cells (NMCs), produces both paraxial mesoderm and neurectoderm as the AP axis elongates in post-gastrulation stages. NMC populations across vertebrates harbour cells with functional heterogeneity, including *bona fide* bipotent neuromesodermal progenitors (NMPs) (Binagui-Casas et al., 2021). Importantly, the hallmark of NMPs is their dual fate towards both neuroectodermal and mesodermal lineages, two separate germ layers, at post-gastrulation stages (Guillot et al., 2021; Tzouanacou et al., 2009). Fate mapping has shown that cells with neuromesodermal fate are found at the posterior end of various vertebrate embryo species, and this region coincides with the co-expression of the transcription factors SOX2 and Brachyury (TBXT), reviewed in (Wymeersch et al., 2021).

Our current understanding of the literature suggests that NMPs in amniotes (chick, mouse) and anamniotes (fish, frog) are distinct. In amniotes, individual bi-potent NMPs are dual neural/mesodermal fated, while in anamniotes, neuromesodermal potency is maintained in NMC populations in the tail bud, but individual cells adopt either neural or mesodermal fates (Attardi et al., 2018; Davis & Kirschner, 2000; Guillot et al., 2021; Martin & Kimelman, 2012; Tzouanacou et al., 2009; Wymeersch et al., 2016). Thus, NMCs in the neuromesodermal-fated region in different species can consist of bona fide NMPs, or of a mixture of mono-fated cells of unknown potency.

69 *In vitro*, FGF and WNT treatment of pluripotent stem cells generates cells that resemble *in vivo* NMC
70 populations both phenotypically and functionally: a high percentage of cells co-express SOX2/TBXT
71 transcription factors, and these populations can differentiate into neural and mesodermal lineages
72 (for a list of common protocols see Wymeersch et al., 2021). Indeed, at clonal level a proportion of *in*
73 *vitro*-derived NMP-like cells expressing *Tbxt* (monitored with a fluorescent reporter) have dual
74 potential for neural and mesodermal differentiation (Tsakiridis & Wilson, 2015). However, in these
75 experiments, the knock-in reporter created a heterozygous null allele of TBXT (Fehling et al., 2003), a
76 condition known to affect axial development (Gruneberg, 1958). Furthermore, the expression of SOX2
77 in the starting population was not measured, and therefore the relationship between SOX2 and TBXT
78 coexpression, and bipotency is unknown.

79 Recent single cell RNA-seq studies have found that *in vitro* cultures harbouring NMC populations are
80 transcriptionally heterogeneous, with varying levels of *Sox2* and *Tbxt*, as well as other NMC-specific
81 transcripts such as *Msgn1*, *Nkx1-2,* and *Tbx6* (Dias et al., 2020; Edri, Hayward, Baillie-Johnson, et al.,
82 2019; Edri, Hayward, Jawaid, et al., 2019; Gouti et al., 2017). One interpretation of this data is that *in*
83 *vitro* NMC populations contain contaminating, non-NMP cells. Alternatively, the transcriptional
84 heterogeneity in *Sox2* and *Tbxt* expressing cells themselves could reflect inherent functional
85 heterogeneity at the single cell level given that transcription factors such as SOX2 elicit expression
86 level-dependent function (Corsinotti et al., 2017), and that, in vivo, levels of TBXT and SOX2 protein
87 appear to correlate with variations in NMC fate (French et al., 2025; Romanos et al., 2021; Wymeersch
88 et al., 2016).

89 Addressing whether individual *in vitro* NMP-like cells co-expressing *Sox2* and *Tbxt* are bi-potent, and
90 how heterogeneity in SOX2 and TBXT levels dictates differentiation outcomes requires the use of cells
91 that report on the expression of both markers. Previous reporters have employed knock-in constructs
92 that disrupt the function of these genes (Avilion et al., 2003; Fehling et al., 2003) or are additive
93 bacterial artificial chromosome transgenics where there is a risk that expression does not accurately
94 mirror endogenous expression (Koch et al., 2017).

95 Here we generated a minimally disruptive mouse embryonic stem cell (mESC) line with dual *Sox2/Tbxt*
96 gene expression reporter activity and use it to examine live cell behaviours, population heterogeneity
97 and functional potency of *in vitro* NMP-like cells. Clonal plating demonstrates that individual *Sox2/Tbxt*
98 dual positive cells are *bona fide* bipotent progenitors, which can also self-propagate. Additionally, we
99 show that rapid deceleration of axis elongation in gastruloids on WNT or NOTCH inhibition coincides
100 with *Tbxt* downregulation, while changes in *Sox2* expression follow more slowly. We further uncover

101 previously elusive lateral plate mesoderm potency within descendants of NMPs, an attribute shown
102 only in bulk culture (Row et al., 2018).

103 Interestingly, we find that the *Sox2/Tbxt* co-expressing population is not functionally homogeneous:
104 the levels of *Tbxt*-GFP expression predict clonal lineage bias towards mesoderm in individual
105 progenitors. This intrinsic bias can be tuned by the culture conditions and cell substrates, mirroring
106 the *in vivo* plasticity of NMC populations (Wymeersch et al., 2016).

107

108 **Results**

109 **1. A novel reporter line, STR-KI, accurately reports on SOX2 and TBXT expression and NMP**
110 **differentiation trajectories *in vitro***

111 To characterise the dynamics of NMC emergence from mESCs, we constructed a 2A-peptide-linked
112 dual reporter: SOX2::mCherry; TBXT::GFP, hereafter named SOX2-TBXT Reporter Knock-In (STR-KI).
113 CRISPR knock-in transgenesis was used to create heterozygous knock-in alleles that replace the stop
114 codon of the endogenous gene loci with T2A/P2A 'self-cleaving' (ribosome-skipping (Donnelly et al.,
115 2001)) peptide, retaining the *Tbxt* and *Sox2* open reading frames in their entirety (Dewari et al., 2018).
116 Since even partial loss of TBXT or SOX2 function affects differentiation (Avilion et al., 2003; Corsinotti
117 et al., 2017; Gruneberg, 1958; Masui et al., 2007) we employed this strategy to minimise the chance
118 of adverse phenotypic effects and maximise the faithfulness of the reporters (Fig. 1A). Chimeric
119 embryos containing high contribution from STR-KI mESCs showed a normal phenotype, demonstrating
120 that the gene targeting was indeed minimally disruptive to SOX2 and TBXT protein function.
121 Furthermore, the fluorescent reporters in pre-streak and somite stage embryos recapitulated the *in*
122 *vivo* expression pattern of TBXT and SOX2 in mesodermal and neural lineages, respectively (Fig. 1B,
123 Suppl. Fig. 1C, D) (Wymeersch et al., 2016).In pre-streak chimeras, SOX2::mCherry ($Sox2^{mCh}$ hereafter)
124 was expressed throughout the epiblast, whereas TBXT::GFP ($Tbxt^{GFP}$ hereafter) expression was
125 undetectable (Suppl. Fig. 1C). By early somite stages onwards, $Sox2^{mCh}$ was predominantly restricted
126 to neural tissues, and to presumptive primordial germ cells at the base of the allantois, whereas
127 $Tbxt^{GFP}$-expressing cells were observed in the node, notochord and nascent mesoderm at the posterior
128 end (Fig.1B and Suppl. 1D). Differentiation of STR-KI ESCs into NMP-like cells, following an established
129 protocol (Gouti et al., 2014) produced mCherry and GFP double positive cells ($Sox2^{mCh}$pos/$Tbxt^{GFP}$pos)
130 (Fig. 1C-F). Immunostaining and quantification of transcription factor (TF) intensity in these *in vitro*
131 differentiated cells showed that the reporters faithfully co-localised with endogenous SOX2 and TBXT
132 protein (Fig. 1C, D). For both reporters, we also observed a positive correlation between the levels of
133 reporter and endogenous protein (Fig. 1C). We therefore adopt the working hypothesis that GFP and

134 mCherry levels reflect TBXT and SOX2 protein production respectively, although differential
135 posttranscriptional modifications affecting protein levels cannot be excluded.

136 We next tracked the reporters' dynamics over a time course of NMP differentiation (Fig. 1E). Briefly,
137 this entails differentiation of ESC to an epiblast-like state, followed by treatment with the WNT agonist
138 CHIR99021 (CHIR) and FGF2 from day (d) 2 onwards (Gouti et al., 2014). Cells appeared to move along
139 a differentiation trajectory from Sox2^{mCh}high pluripotent ESC to a Sox2^{mCh}low, putative epiblast state,
140 followed by appearance of Sox2^{mCh}pos/Tbxt^{GFP}pos (double positive cells) within 3 days (d). D3
141 differentiated cells exhibited a spectrum of Tbxt^{GFP} levels before exiting the double positive state
142 towards Sox2^{mCh} negative (neg)/Tbxt^{GFP}high mesoderm from d4. Consistent with previous reports, STR-
143 KI ESC differentiation reproducibly resulted in the production of putative NMCs co-expressing mCherry
144 and GFP at a rate of ~67% of the population by d3 (Fig. 1F).

Figure 1

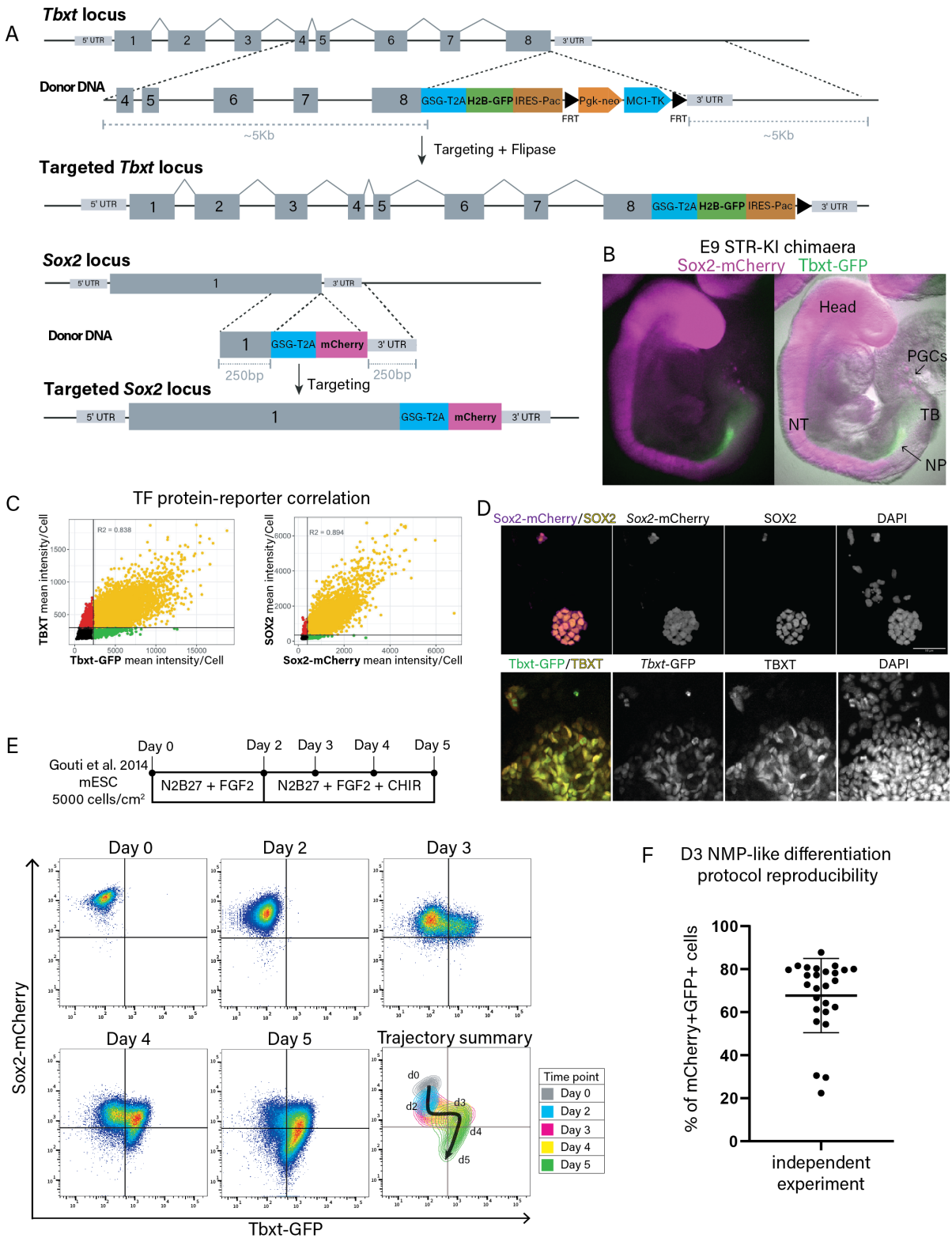


Figure 1. STR-KI accurately reports on SOX2 and TBXT expression and NMP differentiation trajectories *in vitro*. **A.** Schematic illustration of the targeting strategy to knock-in GFP and mCherry in the *Tbxt* and *Sox2* loci, respectively. **B.** Embryonic day 9 (E9) STR-KI chimaera with high STR-KI contribution shows appropriate localisation of fluorescent reporter expression: high mCherry (magenta) in the prospective brain, the neural tube

150 (NT) and in primordial germ cells (PGCs). GFP (green) expression marks the notochordal plate (NP) and tail bud
151 (TB). **C.** Correlation between immunofluorescent staining for SOX2 or TBXT, and their respective reporter protein
152 in individual cells derived from *in vitro* NMP differentiation assays. **D.** Immunofluorescence images of
153 differentiated STR-KI colonies following NMP-like differentiation (Gouti et al., 2014). Scale bar = 50μm. **E.**
154 Representative example from a flow cytometry time-course analysis of STR-KI cells along the differentiation
155 protocol shown in D, from mESCs (day 0) via NMP-like intermediate (at day 3) to day 5 (n=4). A trajectory
156 summary highlights population dynamics. **F.** Quantification of the reproducibility of STR-KI mESC differentiation
157 to day 3 NMP-like cells shows the percentages of $Sox2^{mCh}$/$Tbxt^{GFP}$double positive cells for a cohort of
158 independent experiments (n=23; Mean ± SD: 67.42± 18.28%).

159

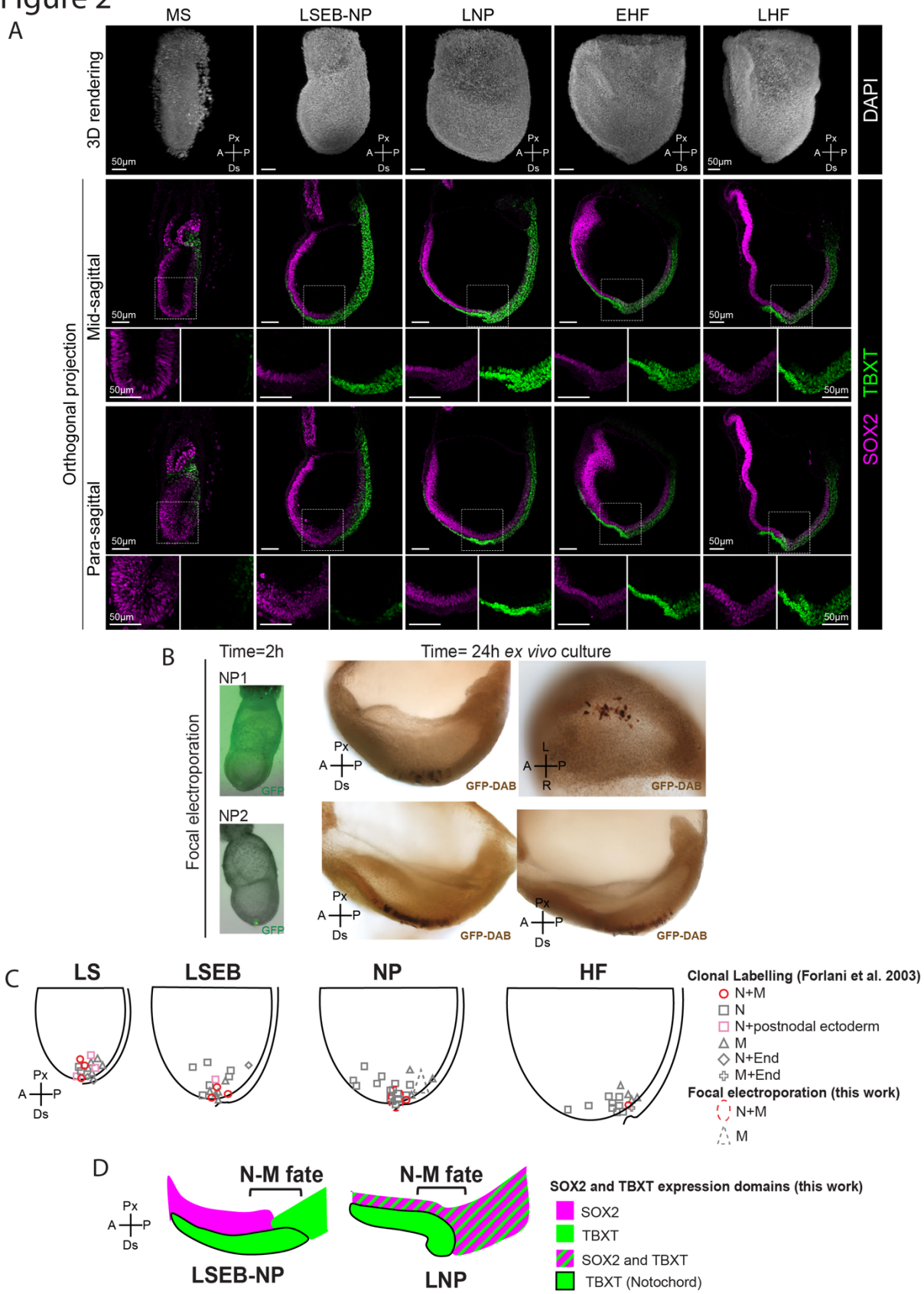
160 **2. Expression dynamics of *Sox2* and *Tbxt in vitro* matches NMP emergence *in vivo***

161 A previous report demonstrated that *Sox2* expression in the caudal lateral epiblast is activated *de novo*
162 in mouse via its N1 enhancer activity (Takemoto et al., 2011), corresponding approximately to the time
163 at which NMCs are a detectable cell population (Wymeersch et al., 2016). Superficially, this would
164 suggest that *Sox2* is secondarily activated in the *Tbxt*-expressing caudal lateral epiblast, contrasting
165 with our observation of the acquisition of *Tbxt* in *Sox2*-expressing cells *in vitro* (Figure 1E). We
166 therefore examined the origin of NMPs in the epiblast.

167

168 For this, we first analysed SOX2 and TBXT immunofluorescence in embryos staged between mid-streak
169 and late headfold (Lawson & Wilson, 2025) (Figure 2A), including the neural plate (also known as pre-
170 headfold) stage at which TBXT+SOX2+ putative NMPs emerge (Wymeersch et al., 2016). Up to early
171 neural plate stage, SOX2 single positive cells were located anterior to the node, while cells posterior
172 to the node expressed only TBXT. Double positive cells emerged at the node and just posteriorly to it
173 during neural plate stage and constituted most of the cells at the posterior node from late neural plate
174 stage onwards. Comparison of this staining pattern with focal electroporations at neural plate stage
175 (n=2 anterior to node, n=3 posterior to node, Figure 2B, C) and previously published *in vivo* single cell
176 clonal lineage tracing around the node (Forlani et al., 2003) (Figure 2C) shows that before neural plate
177 stages, cells with dual neuromesodermal fate are located anterior to the node, while those posterior
178 to it are exclusively mesoderm-fated. Thus, NMPs in the embryo emerge from the SOX2+ cell
179 population anteriorly to the node and subsequently acquire TBXT expression (Figure 2D). This
180 indicates that the acquisition of TBXT expression by SOX2+ cells in NMP differentiation conditions i*n*
181 *vitro* mimics the *in vivo* ontogeny of NMPs.

Figure 2



182

183

184 **Figure 2. NMPs emerge *in vivo* from a pre-existing SOX2+TBXT- population. A**. Immunofluorescent detection
185 of SOX2 and TBXT in the fate-mapped region (boxed) over developmental time. Upper row: DAPI-stained 3D
186 renderings of mouse embryos (dissected on embryonic days E6.5 (midstreak, MS), E7.0 (late streak early bud,
187 LSEB), E7.5 (late neural plate, LNP) and E8.0 (early and late head fold, EHF, LHF); (Lawson & Wilson, 2025)).
188 Lower rows: Orthogonal projections of wholemount immunostained mouse embryos for SOX2 (magenta) and
189 TBXT (green), showing mid- and para-sagittal optical sections of the node area (projections represent 8 µm
190 across the Z axis). All scale bars are 50 µm. **B**. Fate mapping at neural plate (NP) stage via focal electroporation
191 of a ubiquitously-expressed CAG-GFP transgene into cells anterior to the node tracks descendants into the
192 neural plate and anterior caudal lateral epiblast (n=2). Equivalent electroporation in the primitive streak (n=3)
193 produces descendants in the mesoderm (data summarised in C). **C**. Focal electroporations superimposed on
194 data from Forlani et al. 2003, where cells in or near the node at late streak (LS) to head fold (HF) stages were
195 prospectively mapped to segments of the axis. Data has been re-displayed to indicate clones producing
196 descendants in neurectoderm and/or mesoderm. **D**. Summary scheme of data in A-C. Tracing of LSEB-NP and
197 NP stage node region showing expression of SOX2 and TBXT in epiblast and nascent notochord. Expression is
198 non-overlapping at the midline in LSEB-NP transition stage embryo, while in LNP stage embryo onwards,
199 SOX2+TBXT+ cells are evident. The anteroposterior extent of NM fate in the epiblast, extracted from
200 diagrammatic representations in Forlani et al. at each stage is indicated, showing N-M fate is initially localized
201 in the SOX2+ dorsal node region, which later becomes SOX2+TBXT+. Px, proximal; Ds, distal; A, anterior; P,
202 posterior; L, left; R, right.

203

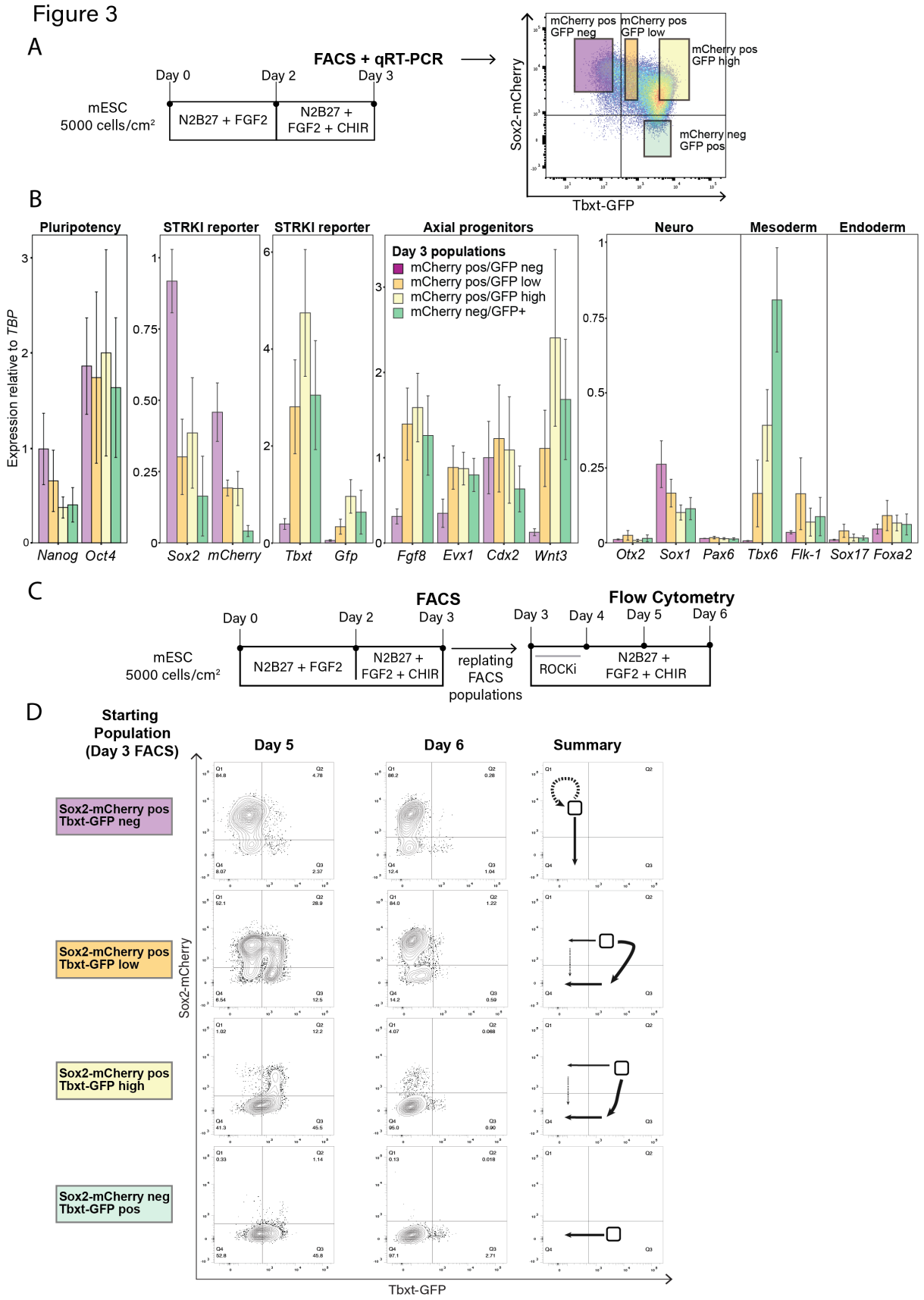
204 **3. *In vitro* derived Sox2^{mCh}/Tbxt^{GFP} co-expressing cells are NMCs**

205 To characterise the cell populations emerging along the ESC-to-NMP-like cell trajectory, we probed
206 the gene expression profiles of various subpopulations of diagnostic marker genes from d2 to d4 via
207 qRT-PCR (Fig. 3A, B; Suppl. 2 A, B). Subpopulations were delineated based on the intensities of
208 mCherry and GFP reporter detection in FACS. Comparison of mRNA expression levels of the
209 endogenous *Sox2* and *Tbxt* with their respective reporters *mCherry* and *GFP* showed that in all sorted
210 populations, mRNA expression reflected endogenous gene expression levels. The levels of *Sox2* and
211 *Tbxt* mRNAs exceeded those of the corresponding reporters, as expected, since the reporter allele,
212 which transcribes both endogenous mRNA and reporter, is heterozygous (Fig 3B). Interestingly, the
213 putative NMC population (Sox2^{mCh}pos/Tbxt^{GFP}pos) displayed relatively constant Sox2^{mCh} fluorescence
214 but a spectrum of Tbxt^{GFP} expression. Since the levels of Tbxt^{GFP} in the double-positive population
215 increased before the appearance of Tbxt^{GFP} single positive mesoderm cells, we hypothesised that low
216 and high Tbxt^{GFP} populations might represent different states in the wider NMC population.
217 Specifically, high Tbxt^{GFP} positive cells might be further along a mesoderm differentiation trajectory

218 compared to Tbxt^{GFP} low cells(French et al., 2025; Romanos et al., 2021; Wymeersch et al.,
219 2016)(French et al., 2025; Romanos et al., 2021; Wymeersch et al., 2016)).
220
221 To test this, we divided the NMC-like population into low-TBXT (Sox2^{mCh}pos/Tbxt^{GFP}low) and high-TBXT
222 (Sox2^{mCh}pos/Tbxt^{GFP}high) expressing subpopulations. We analysed subpopulations based on marker
223 gene expression: pluripotency-associated *Nanog* and *Oct4*; axial progenitor *Fgf8*, *Evx1*, *Cdx2* and
224 *Wnt3a*; paraxial mesoderm *Tbx6;* lateral plate mesoderm (LPM) *Flk-1*; and neural-associated *Otx2,*
225 *Sox1* and *Pax6*. Although both low-TBXT and high-TBXT populations expressed similar levels of NMC
226 markers, high-TBXT cells were enriched for mesodermal markers such as *Wnt3a* and *Tbx6*, while
227 expression of neural genes like *Sox1* were low. The absence of *Sox2*, coupled with high expression of
228 *Tbx6* and *Flk-1* in both Sox2^{mCh}neg/Tbxt^{GFP}pos and Sox2^{mCh}neg/Tbxt^{GFP}neg (double negative)
229 populations, together with the downregulation of NMC/caudal markers like *Fgf8* and *Wnt3a* from d3
230 to d4, suggests that these two populations represent nascent and differentiated mesoderm,
231 respectively (Suppl. Fig. 2A, B). Finally, the expression of endoderm markers *Sox17* and *Foxa2* was
232 universally low, suggesting negligible fractions of contaminant non-NMC cells. Thus, gene expression
233 analysis in populations undergoing differentiation confirms that our double reporter cell line faithfully
234 recapitulates endogenous gene expression and can be used to purify neural, mesodermal and NMC-
235 like cell populations based on Sox2^{mCh} and Tbxt^{GFP} expression levels (Fig. 3B, Suppl. 2A, B).
236 Furthermore, this analysis supports the idea that high levels of Tbxt^{GFP} mark a more mature
237 mesodermal state than low Tbxt^{GFP} levels.
238
239 To test whether the apparently dynamic changes in Sox2^{mCh} and Tbxt^{GFP} cell populations during NMP
240 differentiation (Fig. 1E) represent one or more lineage trajectories, the lineage potentials of sorted
241 subpopulations at d3 of NMP differentiation were assayed by FACS over 3 further days in continuous
242 FGF and CHIR, which are permissive for both neural and mesodermal differentiation (Gouti et al.,
243 2014; Tsakiridis & Wilson, 2015) (Fig 3C, D Suppl. Fig. 2C). Most (>80%) Sox2^{mCh} single positive cells
244 remained single positive. Together with the enrichment of *Sox1* and *Nanog* in the starting population
245 (Fig 3B), this suggests a neural and/or epiblast identity of Sox2^{mCh} single positive cells (Wong et al.,
246 2025). A small proportion of the progeny lost expression of both reporters, likely representing further
247 maturation, since SOX2 is downregulated during neural differentiation . Sox2^{mCh}pos/Tbxt^{GFP}low cells
248 also produced Sox2^{mCh} single positive cells by d6, but also a recapitulation of the pattern shown in Fig.
249 1E: emergence of Sox2^{mCh}pos/Tbxt^{GFP}high cells, followed by a high proportion of
250 Sox2^{mCh}neg/Tbxt^{GFP}pos and later double negative cells. This suggests that Sox2^{mCh}pos/Tbxt^{GFP}low
251 NMCs produce both neural and mesodermal cells. Interestingly, Sox2^{mCh}pos/Tbxt^{GFP}high cells

252 produced few Sox2^{mCh} single positive cells and followed a predominantly mesoderm trajectory,
253 producing Sox2^{mCh}neg/Tbxt^{GFP}pos and then double negative cells. Finally, Sox2^{mCh}neg/Tbxt^{GFP}pos
254 populations only produced double negative cells. Together with its enrichment in expression of
255 mesoderm markers, it suggests this is a mesoderm-committed population. Thus, using our dual
256 reporter, we describe a lineage trajectory in which a largely self-sustaining SOX2-positive population
257 generates a SOX2-positive;TBXT-low dual-potency population, which progresses through a
258 mesoderm-biased SOX2-positive;TBXT-high state towards a nascent mesodermal TBXT-high and
259 finally a double-negative state (Fig 3D, see summary). This indicates that the TBXT-high NMC
260 population is predisposed towards mesoderm differentiation. Thus, within the Sox2^{mCh}pos/Tbxt^{GFP}pos
261 double positive NMC population, there are apparent cells with differing competence to differentiate
262 to neural or mesodermal outcomes.

Figure 3



263

264

265 **Figure 3. Gene expression and lineage potential analysis of Sox2^{mCh}/Tbxt^{GFP} fluorescent subpopulations during**
266 **differentiation. A.** Differentiation scheme and representative gating strategy for FACS and further qRT-PCR
267 analyses performed at day 3 of the differentiation protocol. **B**. qRT-PCR gene expression analysis of the
268 populations in A. n=5 independent differentiation experiments. **C.** Post-d3 differentiation scheme in which
269 populations in A were isolated by FACS, replated and further differentiated for an additional 3 days. **D.**
270 Representative flow cytometry graphs of one time-course experiment (n=3, see Suppl. Fig. 2C), following the
271 Sox2^{mCh} and Tbxt^{GFP} expression dynamics over three more days. A summary of the lineage outcomes in three
272 experiments over time is provided. The thickness of arrows represents the relative contribution to the indicated
273 trajectory, with dashed arrows indicating a presumed trajectory based on the proportions of cells at the end of
274 the experiment.

275 Previous reports have shown that treatment of pluripotent cells with FGF2 and CHIR leads to the
276 expression of both paraxial and lateral plate mesoderm (LPM) markers (Edri, Hayward, Baillie-Johnson,
277 et al., 2019; Gouti et al., 2017). *In vivo*, while homotopic grafts of NMCs overwhelmingly give rise to
278 paraxial mesoderm (Cambray & Wilson, 2007; Wymeersch et al., 2016), heterotopic grafts to the
279 posterior primitive streak leads to LPM differentiation (Row et al., 2018). This suggests that NMPs may
280 have LPM potency, although it cannot be excluded that this LPM arose from non-NMCs within the
281 graft. To directly test the LPM-potency of NMCs, we examined the potential of FACS-enriched
282 Sox2^{mCh}pos /Tbxt^{GFP}high cells to produce LPM lineages upon BMP treatment, which induces LPM
283 differentiation in bulk NMP cultures (Row et al., 2018). We observed not only FLK-1 protein (a lateral
284 plate/endothelial cell marker) labelling in colonies but also the formation of primitive tubes similar to
285 those seen in classical endothelial assays (Taoudi et al., 2005) (Suppl. Fig. 2D, E). These findings
286 indicate that Sox2^{mCh}pos /Tbxt^{GFP}high NMP-like cells can produce LPM. Thus, the SOX2-pos;TBXT-high
287 mesoderm-fated population shows broad mesodermal potency, supporting the idea that *in vivo*, it is
288 NMCs themselves, and not contaminating populations, that can contribute to both paraxial and lateral
289 mesoderm.

290

291 **4. Live imaging of STR-KI cells uncovers a role for TBXT in WNT and NOTCH-driven axis elongation**

292 In vivo, wildtype TBXT levels are not only critical for NMP differentiation, but also to maintain axial
293 elongation (Wymeersch et al., 2021). However, the dynamic relationship between TBXT, SOX2 and
294 axial elongation remains untested. To determine in real time the relationship between reporter
295 expression, axial elongation and factors known to affect both reporter expression and axial elongation,
296 we next used live-tracking of STR-KI reporter expression. For this, we tracked Sox2^{mCh} and Tbxt^{GFP}
297 expression *in vitro* in a standard gastruloid model of axis elongation, (Baillie-Johnson et al., 2015; van
298 den Brink et al., 2014), in which aggregates of mESCs grown in N2B27 remain SOX2 pos/TBXT negative

299 until 48 hours post aggregation (hpa), then exit pluripotency. At 48 hours, a 24-hour pulse of CHIR is
300 used to activate WNT globally and ensure robust symmetry breaking. WNT activation transiently
301 induces TBXT expression in all cells (van den Brink et al., 2014), before it resolves into polarized
302 expression at the posterior pole by 72 hpa and is then followed by elongation. We first compared the
303 dynamics of mesoderm differentiation in STR-KI-gastruloids with previous reports (Suppl. Fig. 2F).
304 After treatment with CHIR, most cells became Tbxt^{GFP}pos by 72 hpa, and down-regulated Sox2^{mCh}. In
305 the absence of CHIR, STR-KI cells remained Sox2^{mCh}high, but a small proportion of cells began to
306 express Tbxt^{GFP}. These trends were maintained at the population level by 96 hpa (Suppl. Fig. 2F). We
307 also observed downregulation of Sox2^{mCh}, symmetry breaking and polarisation of Tbxt^{GFP} to the
308 posterior growing axis only upon the CHIR pulse (Suppl. Fig. 2G). We conclude that our STR-KI reporter
309 can generate gastruloids similar to other lines and thus it is ideally suited for studying TF dynamics
310 within them.

311 We next assessed how the manipulation of WNT and NOTCH signalling pathways, which regulate
312 mesoderm differentiation from NMPs (Cooper et al., 2024; Wymeersch et al., 2016), impact gastruloid
313 axial elongation. We treated elongating gastruloids (96 hpa) with either WNT or NOTCH inhibitors
314 (IWP-2 and LY411575 (LY), respectively) and live imaged them every hour for 24 hours to monitor
315 major axis length and reporter expression (Fig. 4A, B). Both WNT and NOTCH inhibition decelerated
316 major axis growth (elongation) starting from as early as 2-3h post treatment (Fig. 4C).

317 To determine how reporter expression responded to WNT or NOTCH inhibition during this period,
318 whole gastruloid tissues were segmented and traces extracted along the extending long axis marked
319 at the presumptive posterior end by Tbxt^{GFP} expression. Treatment with IWP2 or LY rapidly decreased
320 Tbxt^{GFP} signal in the NMP region (within ~2-3h). In contrast, Sox2^{mCh} signal in this region increased after
321 ~10-15 hours following an initial minor decrease compared to the DMSO control (Fig. 4D). This
322 coincides in time with growth deceleration (Fig. 4D, E). Thus, the rapid Tbxt^{GFP} signal downregulation
323 coupled with slowdown of axis elongation in response to WNT or NOTCH inhibition suggests AP axis
324 elongation is driven predominantly by *Tbxt*-expressing cells.

Figure 4

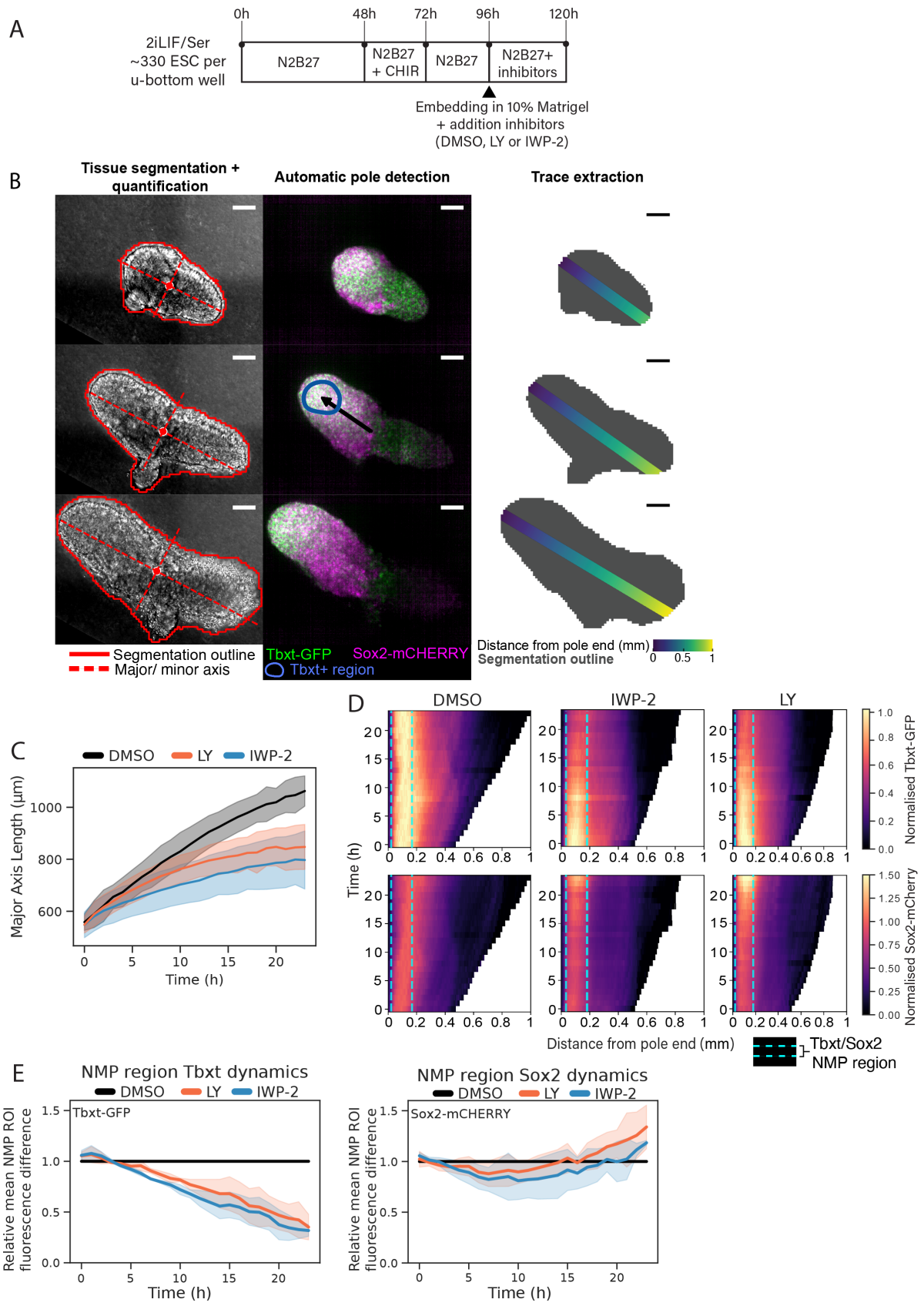


Figure 4. STR-KI cell analysis indicates that Tbxt-expressing cells are responsible for WNT and NOTCH driven STR-KI gastruloid elongation. A. Experimental setup for the generation of STR-KI gastruloids in 10% matrigel at

328 96 hours with DMSO, Wnt signalling inhibitor IWP2 (2.5 µM) or Notch signaling inhibitor LY411575 (150nM).
329 Gastruloids were imaged on a confocal microscope every hour for 24 hours after embedding. **B.** Quantification
330 pipeline of 3D gastruloid confocal image stacks. All scale bars indicate 100 µm. **C.** Axial length after treatment
331 of gastruloids with IWP-2 or LY411575. **D.** Heatmaps showing the average Tbxt-GFP or Sox2-mCherry signal per
332 binned distance from the pole along time post Matrigel embedding. **E.** Average difference in fluorescence signal
333 of the Sox2/Tbxt region between control and treated conditions. N=3 independent experiments. Total
334 gastruloids analysed: DMSO=35, LY= 38, IWP-2=41. Shaded areas indicate the mean +/- 0.95 confidence
335 intervals, calculated by $\bar{x}$ +/- z*(σ/sqrt(n)).

336

337 **5. Clonal tracing of in vitro derived NMCs reveals lineage bias of NMPs dependent on matrix and**
338 **TBXT levels**

339

340 Our findings at the bulk level suggest that differences in *Tbxt* levels influence lineage outcomes in
341 Sox2/Tbxt co-expressing (NMC) subpopulations (Fig. 3D). However these experiments did not
342 distinguish between neural or mesoderm lineage biases across the whole *in vitro* NMC-like population
343 or a subset of already lineage-specified cells expressing the very highest or lowest reporter levels. To
344 distinguish these possibilities, we sorted Sox2^{mCh}pos;Tbxt^{GFP}pos cells by FACS into low and high gates
345 and deposited single cells into 96-well plates (Fig. 5A). We measured their differentiation outcome in
346 FGF/CHIR, which supports both neural and mesodermal differentiation (Fig. 3D). A variety of
347 substrates has been used in different NMC-generating protocols, and the influence of these is
348 currently unclear. We therefore compared clonal outcomes on four commonly used matrices:
349 fibronectin, gelatin, Geltrex and Matrigel. We then correlated this information with the lineage output
350 using immunofluorescence and high throughput microscopy to quantify the SOX2/TBXT cellular
351 phenotype in descendant cells in each well after five days (Fig. 5A, B).

352

353 We first considered the characteristics of all clones, sorted into 'low' and 'high' GFP gates as in Fig. 3A.
354 In all clonal conditions tested, Sox2^{mCh}pos;Tbxt^{GFP}pos cells produced single positive neural or
355 mesodermal descendants, as well as double positive cells, confirming that these cells are NMPs.
356 Sox2^{mCh}pos;Tbxt^{GFP}low and Sox2^{mCh}pos;Tbxt^{GFP}high input cells showed distinct lineage output profiles.
357 Sox2^{mCh}pos;Tbxt^{GFP}high cells gave rise to a greater proportion of Tbxt^{GFP} only colonies, confirming that
358 high TBXT expression in individual Sox2^{mCh}pos;Tbxt^{GFP}high cells strongly predisposes them to a
359 mesodermal outcome (Fig. 5C). From the population analysis (Fig. 3D), we can infer that double-
360 negative descendants of Sox2^{mCh}pos;Tbxt^{GFP}pos cells are predominantly differentiated descendants of
361 Tbxt^{GFP}pos mesoderm. In these clonal assays, double negative cells were mainly produced by

362 Sox2^{mCh}pos;Tbxt^{GFP}high cells, consistent with this interpretation. Interestingly, Sox2^{mCh}pos;Tbxt^{GFP}high
363 cells showed higher plating efficiency than Sox2^{mCh}pos;Tbxt^{GFP}low cells, suggesting that increased TBXT
364 promotes higher survival and/or adhesion to the substrate (Fig. 5C). Thus, at clonal level, high levels
365 of GFP predict a high frequency of mesoderm-only outcomes, and an overall increased survival on
366 multiple substrates.
367
368 Substrate composition also seemed to differentially affect cell phenotype. Strikingly, Matrigel and
369 Geltrex reduced the overall mesoderm predisposition of Sox2^{mCh}pos;Tbxt^{GFP}high cells. This suggests
370 that despite their propensity to differentiate towards mesoderm, Sox2^{mCh}pos;Tbxt^{GFP}high cells retain
371 a substrate-dependent potential for neuromesodermal differentiation. Intriguingly, although the
372 plating efficiency of both Sox2^{mCh}pos;Tbxt^{GFP}low and high cells was particularly poor on gelatin, this
373 substrate appeared to enrich for growth of double positive cells from Sox2^{mCh}pos;Tbxt^{GFP}low cells (Fig.
374 5C, Suppl. Fig. 3C-E). Moreover, the substrate impacted the morphology of the colonies in culture: on
375 gelatin, cells grew as compact colonies and produced an overall higher number of cells per colony,
376 compared to Matrigel culture, where cells formed flat, dispersed colonies (Suppl. Fig. 3D). These
377 observations suggest that once established from Sox2^{mCh}pos;Tbxt^{GFP}low cells, double positive cells are
378 favoured by self-adherent growth. Lastly, we found that Sox2^{mCh}pos;Tbxt^{GFP}high colony survival on
379 Geltrex and Matrigel was comparable, but Matrigel-coating increased the clone survival of
380 Sox2^{mCh}pos;Tbxt^{GFP}low cells. Thus Sox2^{mCh}pos;Tbxt^{GFP}low and high cells exhibit differential behaviour
381 that is modified by the substrate, which can alter both the cell phenotype and the extent of double-
382 positive cell propagation.
383
384 We next compared individual colony composition with input reporter level (Fig. 5A, D). Cells in the
385 Sox2^{mCh}pos;Tbxt^{GFP}high gate over two experiments predominantly produced only Tbxt^{GFP} positive
386 mesoderm and/or double negative presumptive mesoderm (Fig. 5D, high GFP 'plateau' colonies).
387 These two experimental repeats derived from populations where differentiation was relatively
388 advanced and most cells lay in a Tbxt^{GFP} high;Sox2^{mCh} low state just prior to exit from the double
389 positive population (Suppl. Fig. 3A, B – Exp 1-2). This suggests double positive cells reach a 'mesoderm
390 specification' threshold beyond which mesoderm formation is virtually the only outcome and implying
391 that some Tbxt^{GFP};Sox2^{mCh}high cells are not NMPs, recalling our previous observation that, in vivo, a
392 minority of TBXT/SOX2 immunofluorescent cells are present at the midline where we detect only
393 mesoderm potential (Cambray & Wilson, 2007; French et al., 2024).
394

395 To investigate the relationship between input and output levels of Tbxt^{GFP} we then indexed a
396 proportion of the input population, expanding the 'Tbxt-high' gate to include medium-level GFP-
397 expressing cells, to assess whether the clone composition varies as a continuum or shows threshold
398 responses to either reporter.
399
400 We used the indexed cells to mathematically determine whether the clonal lineage of individual NMP
401 progeny is controlled by the initial levels of *Sox2* or *Tbxt* expression and the cell substrate. Sox2^{mCh} and
402 Tbxt^{GFP} expression levels recorded at the time of plating (time = 0), were designated *S* and *T*,
403 respectively. After five days, each seeded cell gave rise to varying numbers of cells (*n*) with four
404 phenotypes based on their expression of TBXT and SOX2: n_{S-T+}, n_{S+T-}, n_{S+T+} and n_{S-T-}.
405 To determine whether the clonal lineage at day 5 post index sort was predominantly mesodermal or
406 neural, we evaluated the difference between numbers of Sox2^{mCh} and Tbxt^{GFP} single positive cells per
407 colony (n_{S-T+}- n_{S+T-}). The histogram of n_{S-T+}- n_{S+T-} determined at time = 5 days, resembled a Gaussian
408 distribution centred at n_{S-T+}- n_{S+T-}=zero (Suppl. Fig. 3F).

409 Of note, many NMPs had small n_{S-T+}- n_{S+T-} values (close to zero) and it would be rather arbitrary to
410 identify these NMPs as mesodermal or neural biased. Instead, the tails of the distribution should
411 correspond with high confidence to mesodermal ($n_{S-T+} >> n_{S+T-}$) or neural ($n_{S-T+} << n_{S+T-}$) biased clones.
412 Therefore, we used the standard error of the mean (SEM) of the distribution as a cut-off to assign NMP
413 lineages, even though this excludes colonies from our analyses. Therefore, we classified the lineage of
414 the progeny as mesodermal if $n_{S-T+} > n_{S+T+}$ by 27 cells (where 27 corresponds to the SEM). Conversely,
415 we consider the clone to be neural biased when $n_{S+T-} > n_{S+T+}$ by 27 cells. To investigate whether the
416 lineage of the NMP progeny at time = 5 days depends on the Tbxt^{GFP} or Sox2^{mCh} level of the initial index
417 sorted NMP cell at time = 0, we plotted the lineage classification of each seeded NMP as a function of
418 both Sox2^{mCh} and Tbxt^{GFP} (Supp. Fig. 3G and Fig. 5E, respectively). To do this, we discretized the scales
419 of Sox2^{mCh} and Tbxt^{GFP} by selecting bins of equal numbers of colonies.

420 The mesoderm or neural bias of the NMP progeny did not appear to be affected by the level of Sox2^{mCh}
421 in the Sox2^{mCh}pos;Tbxt^{GFP}pos starting population (Suppl. Fig. 3G). In contrast, the starting level of
422 Tbxt^{GFP} strongly influenced neural/mesodermal bias of the clones (Fig. 5D, E). At low values of Tbxt^{GFP},
423 most NMPs gave rise to neural-biased clones (Fig. 5E). As Tbxt^{GFP} approached a value of 1745 arbitrary
424 units in the index sort, the proportion of mesoderm-biased versus neural biased clones rapidly
425 increased (Fig. 5D, E). Although the ratio of Sox2 to Tbxt has been cited as an important factor
426 determining NMP fate (Koch et al., 2017; Morabito et al., 2025), this ratio did not predict the clonal
427 outcome better than the levels of Tbxt^{GFP} alone (Suppl. Fig. 4 C-F). The substrate, however, did appear
428 to modulate differentiation: neural-biased colonies were more abundant on fibronectin and Matrigel,

429 whereas we did not observe a particular bias towards either lineage when cells were plated on gelatin
430 or Geltrex of lineage-biased Sox2^{mCh}pos;Tbxt^{GFP}pos cells (Suppl. Fig. 3H). These results suggest the
431 existence of a substrate-sensitive Tbxt^{GFP} expression threshold value ('N-M bias swap threshold')
432 above which NMP cells are more likely to produce mesodermal cells, in addition to the mesodermal
433 specification threshold described above.

434

435 To gain mechanistic insight into fate decisions in *in vitro* derived NMPs, we developed a minimal
436 mathematical model of the proliferation and differentiation dynamics of individual NMPs (see
437 Materials and Methods for details). We assumed that each NMP cell can divide symmetrically to: i)
438 self-renew, giving rise to 2 NMPs (SOX2pos;TBXTpos); ii) differentiate into 2 mesoderm cells
439 (SOX2neg;TBXTpos; iii) differentiate into neurectoderm (SOX2pos; TBXTneg); and iv) differentiate into
440 double negative cells (SOX2neg;TBXTneg). Thus, our model has 4 parameters, which are the rates by
441 which NMP cells give rise to each outcome. To check whether the model is sufficient to reproduce
442 the dependence of the neural/mesodermal decision on the initial Tbxt^{GFP} levels in NMPs, we fitted our
443 model to the experimental data.

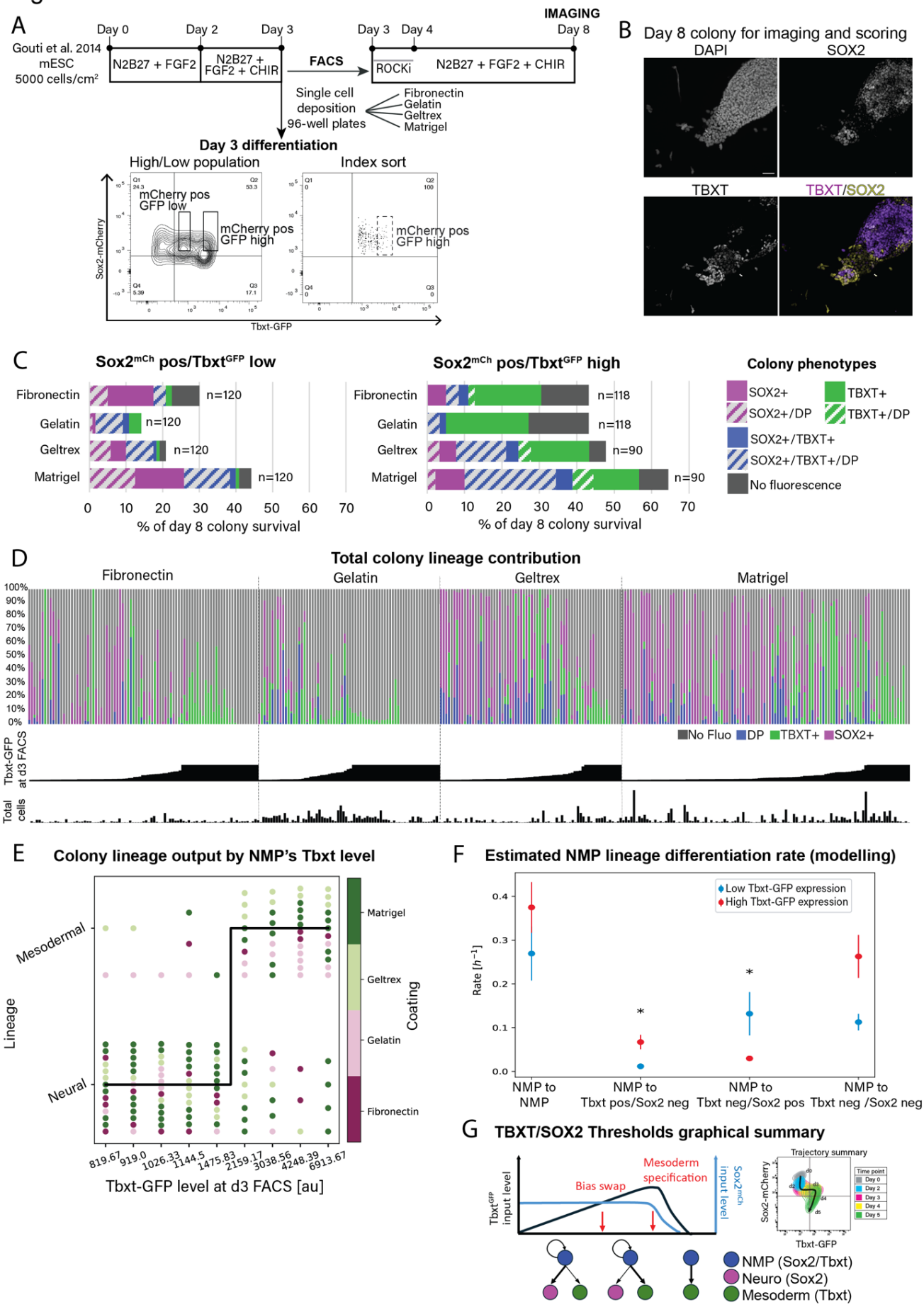
444 When we grouped NMPs according to cell substrate, we observed that gelatin was associated with
445 the highest rates of self-propagation of NMPs and their differentiation towards mesoderm
446 (SOX2neg;TBXTpos), as well as descendants of nascent mesoderm, which are double negative (Suppl.
447 Fig. 4A). Matrigel, on the other hand, produced the highest rate of differentiation towards
448 neuroectoderm (SOX2pos; TBXTneg) (Suppl. Fig. 4A). Classifying NMPs according to whether they had
449 Tbxt^{GFP} levels below or above the Tbxt^{GFP} threshold reported above (regardless of substrate), showed
450 that Tbxt^{GFP} values above the threshold lead to a higher differentiation rate towards mesoderm,
451 whereas Tbxt^{GFP} values below the threshold lead to neuroectoderm bias (Fig. 5F and Suppl. Fig. 4E),
452 consistent with our results shown in Fig. 5E. Interestingly, for both populations, cells produced more
453 NMPs at higher rates than either neural or mesoderm cells. This suggests that the threshold in
454 differentiation bias does not affect the ability of NMPs to self-propagate and distinguishes this lower
455 'N-M bias swap' threshold from the high-Tbxt^{GFP} 'specification threshold', where cells predominantly
456 exit the NMP state towards mesoderm differentiation. Despite the ability of fibronectin and Matrigel
457 to bias cells towards neural differentiation, the mid-level threshold still holds when the experimental
458 data are disaggregated according to the coating used (Suppl. Fig. 4B, F).

459

460 In summary, clonal analysis and mathematical modelling indicates that individual
461 Sox2^{mCh}pos;Tbxt^{GFP}pos double positive cells *in vitro* are NMPs, and that threshold levels of *Tbxt*

462 determine whether they are neural-biased or mesoderm-biased (N-M bias swap threshold). Levels of
463 Tbxt^{GFP} in combination with Sox2^{mCh} determine mesoderm specification (mesoderm specification
464 threshold; summarised in Fig. 5G). The cell substrate can modulate the proportions of neural versus
465 mesoderm cells formed, while gelatin, a relatively challenging substrate for NMP adherence, once
466 cells are attached, favours self-adherence and more efficient NMP self-propagation.

Figure 5


A
Gouti et al. 2014
mESC
5000 cells/cm²
Day 0
Day 2
Day 3
N2B27 + FGF2
N2B27 + FGF2 + CHIR
FACS
Day 3
Day 4
IMAGING
Day 8
ROCKi
N2B27 + FGF2 + CHIR
Single cell deposition 96-well plates
Fibronectin
Gelatin
Geltrex
Matrigel
Day 3 differentiation
High/Low population
Index sort
mCherry pos GFP low
mCherry pos GFP high
Sox2-mCherry
Tbxt-GFP
B
Day 8 colony for imaging and scoring
DAPI
SOX2
TBXT
TBXT/SOX2
C
Sox2mCh pos/Tbxtgfp low
Sox2mCh pos/Tbxtgfp high
Fibronectin
Gelatin
Geltrex
Matrigel
n=120
n=118
n=90
% of day 8 colony survival
Colony phenotypes
SOX2+
SOX2+/DP
TBXT+
TBXT+/DP
SOX2+/TBXT+
SOX2+/TBXT+/DP
No fluorescence
D
Total colony lineage contribution
Fibronectin
Gelatin
Geltrex
Matrigel
No Fluo
DP
TBXT+
SOX2+
Tbxt-GFP at d3 FACS
Total cells
E
Colony lineage output by NMP's Tbxt level
Mesodermal
Neural
Lineage
Matrigel
Geltrex
Gelatin
Fibronectin
Coating
Tbxt-GFP level at d3 FACS [au]
F
Estimated NMP lineage differentiation rate (modelling)
Low Tbxt-GFP expression
High Tbxt-GFP expression
Rate [h⁻¹]
NMP to NMP
NMP to Tbxt pos/Sox2 neg
NMP to Tbxt neg/Sox2 pos
NMP to Tbxt neg /Sox2 neg
G
TBXT/SOX2 Thresholds graphical summary
Tbxtgfp input level
Bias swap
Mesoderm specification
Sox2mCh input level
Trajectory summary
Sox2-mCherry
Tbxt-GFP
NMP (Sox2/Tbxt)
Neuro (Sox2)
Mesoderm (Tbxt)

469 **Figure 5. *In vitro* derived *Sox2*/*Tbxt* co-expressing cells are bipotent at clonal level and their lineage output is**
470 **determined by *Tbxt* expression. A.** Differentiation protocol scheme and representative gating strategy for FACS
471 and index sorting according to Sox2^{mCh} and Tbxt^{GFP} expression. Cells were individually seeded into different
472 substrates and cultured a further five days. Colony lineage output was assessed by immunofluorescence of
473 endogenous SOX2 and TBXT proteins, after which the imaging data was segmented for individual cells in colonies
474 and scored for their lineage output. **B.** Example of a colony on gelatin at the end of the experiment
475 immunostained for SOX2 and TBXT. **C.** Total lineage output of colonies derived from single sorted
476 Sox2^{mCh}pos;Tbxt^{GFP}high or Sox2^{mCh}pos;Tbxt^{GFP}low founder cells. Colonies were scored as containing SOX2pos,
477 SOX2pos;TBXTpos (double positive (DP)), TBXTpos, and non-fluorescent (*i.e.* no signal detected) with n the
478 number of colonies scored for each substrate. **D.** Overview of all individual colonies scored on their percent
479 lineage output, classified by substrate and ordered by Tbxt^{GFP} level at the time of FACS (day 3). The total number
480 of cells per colony ranged from 2- 5335 cells. Note that the plateau in GFP level is arbitrary and corresponds to
481 progenitors that were gated for Sox2^{mCh}pos;Tbxt^{GFP}high but not index sorted and demarcates a mesoderm
482 commitment threshold. **E.** NMP lineage output as a function of initial Tbxt^{GFP} levels. Each dot represents one
483 colony, coloured by substrate used: matrigel (dark green), geltrex (light green), gelatin (pink) and fibronectin
484 (fuchsia). The Tbxt^{GFP} level on the X-axis is the middle value in the bin. **F.** Mathematical model to estimate fate
485 in index sorted NMPs. The rate indicates whether single founder NMPs were above (red) or below (blue) the
486 initial Tbxt level threshold of 1745 arbirtary units (au) as function of their colony output at the end of the
487 experiment. The dots and the error bars correspond to the mean and the standard error of the mean,
488 respectively. Asterisks indicate significant differences obtained after applying a Mann-Whitney statistical test (p
489 < 0.05). **G.** 'two-threshold model' summarising our findings.

490

491 **Discussion**

492 We generated a minimally disruptive embryonic stem cell line, STR-KI, that faithfully reports on *Tbxt*
493 and *Sox2* expression. This cell line mirrors endogenous protein expression and minimises adverse
494 phenotypic effects, lending confidence to the studies of lineage dynamics performed here. Reduction
495 in SOX2 function negatively affects ESC maintenance and differentiation (Corsinotti et al., 2017), while
496 reduction of *Tbxt* levels disrupts tail length (Gruneberg, 1958) and cell movement of *Tbxt* +/null cells
497 in chimeras is defective (Wilson et al., 1995). Therefore, it is likely that normal protein function is
498 preserved in this cell line since no abnormal phenotype or distribution of cells was detected *in vivo* or
499 *in vitro*.

500 We further show that NMCs generated *in vitro* from mESCs (1) are bi-potent and (2) can self-renew in
501 clonal analyses. These are functional attributes used to define *in vivo* NMPs (Tzouanacou et al., 2009).
502 Our study of the differentiating population over time indicates that ESCs during NMP-like
503 differentiation first reduce Sox2^{mCh} expression when exposed to FGF2, as expected in the epiblast *in*

504 *vivo* (Corsinotti et al., 2017). Then, after CHIR addition, the double positive population
505 (Sox2^{mCh}pos;Tbxt^{GFP}pos) emerges from these Sox2^{mCh} low cells, mimicking in vivo NMC emergence.
506 Over the next 48 hours, Tbxt^{GFP} expression increases, until Sox2^{mCh} expression disappears from Tbxt^{GFP}
507 high cells. Tbxt^{GFP} is later extinguished, resulting in a double-negative population. This cellular
508 trajectory fits well with the known fate of epiblast cells that will eventually become NMPs *in vivo*:
509 distal epiblast cells express relatively low *Sox2* levels and exit pluripotency to enter a bipotent NMP
510 state in the caudal lateral epiblast or node-streak border. If NMPs encounter the primitive streak at
511 the midline they will upregulate *Tbxt*, undergo epithelial to mesenchymal transition and exit the streak
512 and its WNT-high environment as mesoderm, thereafter downregulating *Tbxt* (French et al., 2024;
513 Kispert et al., 1995; Wymeersch et al., 2016).

514 *In vivo*, axial elongation is sensitive to perturbation of both WNT and NOTCH signalling, at least partly
515 because NMPs require WNT for maintenance (Wymeersch et al., 2016) and both WNT and NOTCH for
516 continued mesoderm production (Cooper et al., 2024; Martins-Costa et al., 2024; Wymeersch et al.,
517 2016). TBXT is directly activated by WNT signalling (Yamaguchi et al., 1999), and optimal levels of TBXT
518 require NOTCH signalling (Cooper et al., 2024). Our observation that inhibiting WNT or NOTCH
519 signalling result in a decrease in TBXT is consistent with this published data. We show, however, that
520 gastruloids undergo a particularly rapid slowing of axis elongation, concomitant with the suppression
521 of Tbxt^{GFP} by either inhibitor (~2-3h), while Sox2^{mCh} levels change only minimally for 10-15h. This
522 suggests that adequate levels of both WNT and NOTCH are directly required to maintain TBXT-driven
523 axial elongation.

524 We have shown that single *Tbxt/Sox2* co-expressing cells can differentiate towards SOX2+
525 neurectoderm and TBXT+ (and subsequently double-negative) mesoderm. This formally demonstrates
526 that double positive cells generated using standard protocols from mESCs are indeed bipotent NMPs.
527 Our mathematical modelling of index sorted founder cells and the fate in their respective colonies
528 shows that higher Tbxt^{GFP} levels predispose cells towards mesoderm differentiation. Interestingly, this
529 seems to occur at two distinct threshold levels of Tbxt^{GFP}. Firstly, double positive cells with high Tbxt^{GFP}
530 and diminishing Sox2^{mCh} fluorescence appear specified for mesoderm differentiation, reminiscent of
531 a fraction of cells at the primitive streak which express high levels of TBXT and still express SOX2 but
532 are not bipotent on transplantation to NM-fated regions (Cambray & Wilson, 2007; French et al.,
533 2024). These cells are enriched in vivo for TBX6, a factor required to commit cells to a mesodermal
534 fate (Takemoto et al., 2011). In support of this idea, TBX6 expression is elevated in TBXT^{GFP} high cells.

535

536 A second mid-level threshold of Tbxt^{GFP} appears to switch cells from a neural-biased to a mesoderm-
537 biased state. This is somewhat unexpected because *in vivo* NMCs, even from predominantly
538 mesoderm-fated regions (which express higher TBXT than those in neural-fated regions), can still
539 produce neurectoderm on transplantation to more neural-fated regions (Wymeersch et al., 2016).
540 However, our observation that cell substrate can modify the rate of differentiation suggests that the
541 neural/mesodermal outcome of a clone/colony can also be influenced by environmental cues.
542

543 Unexpectedly, we did not find evidence that the ratio of Tbxt^{GFP} to Sox2^{mCh} affected the bias-swap
544 threshold. This contrasts with previous studies suggesting that the Tbxt/Sox2 ratio or the ratio of WNT
545 to Sox2 determines the outcome of differentiation (Koch et al., 2017; Morabito et al., 2025; Romanos
546 et al., 2021). However, the levels of *Sox2* in the NMP compartment in those experiments are relatively
547 constant. Indeed, in the mouse caudal lateral epiblast, the gradient of SOX2 expression is much
548 shallower compared to that of TBXT (French et al., 2025). If the TBXT:SOX2 ratio determined NMP
549 outcome, loss of TBXT function would be expected to bias cells towards neural differentiation.
550 However, *in vivo*, chimaera analysis of TBXT$^{-/-}$ cells does not seem to show this bias (Guibentif et al.,
551 2021). Therefore, it is conceivable that against a baseline of relatively uniform SOX2, the TBXT:SOX2
552 ratio might appear to predict neural versus mesodermal outcome, when TBXT levels might be the
553 main determinant of differentiation bias. Nevertheless, SOX2 may serve an important function in the
554 NMC population. In zebrafish, ectopic expression of SOX2 in NMCs prevents their exit to the
555 mesoderm (Kinney et al., 2020). Together with the observation that TBX6 activation upon entry to the
556 primitive streak suppresses SOX2 to prevent ectopic neural differentiation in the paraxial mesoderm
557 (Takemoto et al., 2011), this suggests SOX2 threshold levels operate at the point of exit of cells from
558 the NMC population. Interestingly, ordering *in vivo* single NMP transcriptomes along a differentiation
559 trajectory shows that peak *Tbxt* mRNA levels occur in double-positive cells, and *Tbx6* mRNA, signifying
560 mesodermal commitment, begins to accumulate only in cells with lower levels of both mRNAs (Gouti
561 et al., 2017). Our observation that during in vitro NMP differentiation, TBXT^{GFP} increases while Sox2^{mCh}
562 levels remain constant, then the levels of both reporters drop before cells exit the double positive
563 population is consistent with such a gatekeeper function of SOX2, but not with a continuously variable
564 mutual inhibition of SOX2 and TBXT.

565

566 The primacy of TBXT dosage in influencing the outcome of differentiation recalls previous experiments
567 *in vivo* where levels of TBXT dictate cell behaviours: low-TBXT expressing cells stay in the primitive
568 streak/tail bud,and exit as mesoderm if TBXT levels are increased (Wilson & Beddington, 1997; Wilson

569 et al., 1995). Our study indicates that this works at the individual cell level, whereby cells sense cues
570 in their environment (e.g. WNT, NOTCH), respond by tuning TBXT up or down, which in turn influences
571 the differentiation outcome at the single cell level. Furthermore, quantitative reduction in wildtype
572 TBXT gene dosage progressively shortens the axis (Gruneberg, 1958; Stott et al., 1993; Xia et al., 2024).
573 This suggests that continued axial elongation (through the persistence of NMPs) is sensitive to TBXT
574 levels through modulating these probabilistic states. Since gastruloid manipulation during elongation
575 shows this TBXT sensitivity, such stem-cell based models will prove useful in the future to investigate
576 the mechanistic bases of these molecular thresholds during developmental processes, such as axis
577 elongation.

578

579 **Materials and methods**

580 Reporter mESC generation

581 To generate the double reporter cell line, we performed CRISPR-based transgenesis in two targeting
582 rounds (one for each gene locus, starting by *Tbxt*) in E14tg2a ESCs following an established CRISPR-
583 Cas9 protocol (Dewari et al., 2018). crRNA and sgRNA guides are listed in Table S1. Briefly, we replaced
584 the stop codon at the C-terminus of the targeted genes by homologous recombination with a reporter
585 construct containing T2A-mCherry (in the case of *Sox2*) or P2A-H2B-GFP (in the case of *Tbxt*) (Fig.1A).
586 The *Tbxt* targeting vector contained a puromycin resistance cassette, which was used for clone
587 selection. *Sox2*+ cells were selected by FACS, sorting individual mCherry+ cells into 96-well plates.
588 Single clones were picked, expanded, and verified for correct transgene integration by PCR and
589 nanopore sequencing (Suppl. Fig. 1A, B).

590

591 PCR and nanopore sequencing

592 Correct integration of the fluorescent protein sequences into the endogenous loci of *Tbxt* and *Sox2*
593 was confirmed by PCR amplification of the targeted site. The primer sequences used are listed in Table
594 S1. To validate the sequence, PCR products were subcloned into a plasmid using the Zero Blunt™
595 TOPO™ PCR Cloning Kit and sequenced via Nanopore sequencing.

596

597 Mouse husbandry

598 C57BL/6 female mice (Charles River) were used for chimaera generation. MF1 mice were used for
599 electroporation experiments. ICR wildtype mice (JAX stock #009122, The Jackson Laboratory) were

600 used for immunostaining. Mice were maintained on a 12 hr-light/12 hr-dark cycle. To collect embryos
601 at specific developmental stages, timed matings were set up overnight. Noon on the day of finding a
602 vaginal plug was designated as embryonic day (E) 0.5. Embryos dissected for immunostaining were
603 approved by the Ethical Committee of the RIKEN Center for Biosystems Dynamics Research (A2016-
604 03-13).

605

606 Chimaera generation

607 C57BL/6 female mice were superovulated (100 IU/mL PMSG and 100 IU/mL hCG intraperitoneal
608 injections 48 h apart) and crossed with wild type studs. Pregnant mice were culled at E2.5 by cervical
609 dislocation, ovaries with oviducts were dissected and collected in pre-warmed M2 medium. Oviducts
610 were flushed using PBS and a 20-gauge needle attached to a 1 mL syringe and filled with PB1
611 (Whittingham, 1974). E2.5 embryos were collected and washed in PB1, their zona pellucida removed
612 using acidic Tyrode's solution and transferred to a plate with incisions where of 8-15 cells were added
613 to each embryo. Embryos were then incubated at 37°C in 5% CO2 for 24h prior to transfer to
614 pseudopregnant females. Blastocysts were selected and collected to be transferred into the uterus of
615 a pseudopregnant CD-1 female. Females were culled according to Schedule 1 (Animals (Scientific
616 Procedures) Act 1986), and embryos were dissected at E6.5-9.0 in M2 medium and observed for
617 chimeric ESC contribution under an Olympus IX51 microscope. The sex of embryos used in this study
618 was not determined. All reagents are listed in Table S1.

619

620 Focal electroporation and fate mapping

621 A more detailed description can be found in (Huang et al., 2015). E7.5 neural plate (NP) stage embryos
622 were dissected in M2 medium with their extraembryonic cavities intact, but Reichert's membrane
623 removed. A PBS-filled 30 mm petri dish was prepared under a stereomicroscope (Zeiss Stemi 2000-C)
624 with a PBS-filled, capillary-sheathed, platinum point electrode (anode; final diameter of capillary
625 sheath 20-30 µm) and an 0.2mm L-shaped platinum wire (cathode) connected to an ECM 830 square
626 wave pulse generator (BTX). Using a pneumatic pico pump (World Precision Instruments, PV830) and
627 micropulled injection needles, a small volume (<5µL) of DNA solution (pCAG::GFP plasmid at 1-1.5
628 µg/mL, 0.01% Fast Green food colouring dye in PBS) was injected in the amniotic cavity of an NP-stage
629 embryo. Embryos were briefly transferred from M2 medium to the PBS-filled electroporation dish,
630 and electroporated using 200 Volts in 6 pulses, each of 50ms duration with a 1s interval between each
631 pulse. Embryos were then immediately transferred to pre-equilibrated culture medium and cultured

632 in 4-well plates in an incubator supplied with 5% CO2 in air at 37°C for 24h (Cambray & Wilson, 2007).
633 GFP contribution was assessed at 2h post-electroporation using a fluorescence compound dissecting
634 microscope (Nikon AZ100). For assessment of contribution, samples were fixed, immunostained as
635 wholemount as described below using an anti-GFP primary antibody (Abcam ab13970, 1:800) and
636 Liquid DAB+ Substrate Chromogen System (Dako K3467). Imaging and scoring of contribution were
637 done under a Nikon AZ100 microscope.

638

639 Immunofluorescence on wholemount embryos:
640 E6.5-E8.0 embryos were dissected in home-made M2 medium (Nowotschin et al., 2010), their yolk sac
641 and amnion membranes removed, after which they were fixed for 25min at 4°C in 4% PFA solution
642 (Nacalai Tesque #09154-85). After three short washes in PBS in 0.1% Triton X-100 (now referred to as
643 PBST), samples were permeabilized in 0.5% Triton X-100 in PBS for 15 min, followed by 20min in 0.5
644 M glycine in PBST, two 10min washes in PBST and blocked overnight at 4°C in 10% donkey serum
645 (Merck) in PBS/0.3% TritonX100. Primary antibodies were diluted in blocking buffer for 48h on a
646 rocking platform at 4°C. Antibodies used (supplier, final concentration): anti-Sox2 (Abcam, ab92494,
647 1:200); anti-TBXT (R&D, AF2085, 1:200). After two days, four 25min washes were performed with PBST
648 on a rocking platform at room temperature. Alexa Fluor®-conjugated secondary antibodies were
649 diluted in blocking buffer at 2 µg/ml final concentration (Thermo Fisher Scientific) and incubated for
650 48h at 4°C, then washed in PBST (4x25min) with the last wash containing DAPI (Invitrogen, D3571,
651 final concentration at 2.5 µg/mL). To image, samples were cleared in BA:BB (2:1 benzyl alcohol:benzyl
652 benzoate; Sigma) after dehydration through an increasing methanol (Nacalai Tesque) /PBS steps (25%,
653 50%, 75%, 2x 100%, 5min each).

654 Wholemount embryo samples were imaged under a LSM800 confocal system using GaAsP detectors
655 (Zeiss). Three replicate embryos were imaged per developmental stage. Whole-mount
656 immunostaining data was processed using Zeiss software (Zeiss). For DAPI channel processing, a
657 median filter and background subtraction were applied. Orthogonal projections were generated as X-
658 Y weighted averages of 8 consecutive slices, representing a total of 8 µm across the Z-axis. IMARIS
659 software (Oxford Instruments) was used to generate 3D renderings.

660

661 Cell culture, differentiation and flow cytometry/FACS

662 STR-KI mESCs were maintained on 0.1% gelatin-coated plates in Glasgow Minimum Essential Medium
663 (GMEM, Sigma-Aldrich #G5154) supplemented with 10% foetal calf serum (Gibco #10270-106), 100
664 U/mL LIF (made in-house), 100 µM 2-mercaptoethanol (Gibco #31350-010), 1X non-essential amino

665 acids (Gibco #11140-035), 2 mM L-Glutamine (Invitrogen #25030-024) and 1 mM Sodium Pyruvate
666 (Invitrogen #11360-039). mESCs were passaged every other day using 1X trypsin (Sigma #59429C)
667 after washing with PBS.

668 For NMP differentiation, following Gouti et al. 2014, STR-KI mESCs were plated at a density of 5,000
669 cells/cm^2 on 0.1% gelatin-coated plates in N2B28 medium. N2B27 medium consisted of a 1:1 mixture
670 of Neurobasal (Thermofisher #21103049) and Advanced DMEM/F12 (Thermofisher #12634028)
671 supplemented with 2 mM L-Glutamine, 50 µM 2-mercaptoethanol, 0.5X B27 (Thermofisher
672 #17504001) and 0.5X N2 (Thermofisher #17502048). Cells were supplemented with 10 ng/mL FGF2
673 (R&D Systems #3718-FB) for 48 hours, and for 24 additional hours with both 10 ng/mL FGF2 and 5µM
674 CHIRON99021 (Axon #1386).

675 For flow cytometry analyses, after TrypLE Express Enzyme 1X treatment and centrifugation, cells were
676 resuspended in PBS containing 2% FCS and 0.1 µg/mL Draq7. For FACS resuspension was done in
677 N2B27 medium containing 0.1 µg/mL Draq7.

678 Cells were analysed using a 4 laser LSR Fortessa (BD) flow cytometer, employing B 530/30-A, Y/G
679 610/20-A, and R780/60-A laser/filter combinations. For cell sorting, a BD FACS Aria II Cell Sorter was
680 used with B 525/50-A, Y/G 610/20-A, and R780/60-A laser/filter combinations.

681 For single cell plating experiments, differentiated cells were detached using TrypLE Express Enzyme
682 1X (Gibco 12604013) and sorted into CellCarrier 96-well plates coated with one of the following
683 substrates: 0.1% gelatin, fibronectin (16.6 µL/mL in PBS; Sigma, #F1141), Matrigel (200 µg/mL in PBS;
684 Corning, #354277), or Geltrex (120 µg/mL in Advanced DMEM/F12; Thermofisher, #A1413201). Plates
685 were coated and incubated for 1 hour at 37ºC prior to cell seeding. Cells were cultured in N2B27
686 supplemented with 100 U/mL penicillin/streptomycin (Invitrogen, #15140-122), 10 µM ROCK inhibitor
687 Y-27632 (Tocris, #1254/10) and either 10 ng/mL FGF2 or in combination with 5µM CHIRON99021.
688 After 24 hours, the medium was replaced without ROCK inhibitor. The ROCK inhibitor Y-27632 was
689 used to enhance single-cell survival (Watanabe et al., 2007) and was used in FACS experiments to
690 counteract dissociation and cell-stress-induced apoptosis.

691 All cells were maintained under standard culture condition (37°C and 5% CO2).

692

693 <u>Quantitative reverse transcription polymerase chain reaction (qRT-PCR)</u>

694 A total of 200 cells were sorted into PCR tubes containing 10 µL of 2X Reaction Mix from the CellsDirect
695 One-Step qRT-PCR Kit (Invitrogen #11753100) with 0.4 U/µl of RNAse inhibitor (Invitrogen #AM2694)

696 and flash-frozen on dry ice. Reverse transcription reactions were performed using the SuperScript III
697 RT/Platinum Taq mix provided in the same kit. qPCR was performed using LightCycler 480 SYBR Green
698 I Master (Roche, #4887352001) in 384-well LightCycler 480 Multiwell Plates (Roche #04729749001)
699 and using the Roche LightCycler 480 Real-Time PCR System. The *Tbp* gene was used as an endogenous
700 control, and relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method.

701

702 Immunocytochemistry and confocal microscopy

703 Cells were fixed with 4% paraformaldehyde for 10 minutes, then washed twice with PBS over a total
704 of 10 minutes and permeabilised with 0.5% Triton X-100 in PBS for 10 minutes. Following two
705 additional PBS washes, cells were blocked for 30 minutes in a solution containing 5% donkey serum
706 and 0.1% Triton X-100 in PBS. Primary antibodies, diluted in the same blocking solution, were added
707 overnight at 5ºC. After washing twice with PBS, secondary antibodies, diluted in blocking solution,
708 were added for 2 hours at room temperature. After two further PBS washes, DAPI (Invitrogen, D3571,
709 final concentration at 2.5 µg/mL) was added for 5 minutes at room temperature, followed by two PBS
710 washes. Imaging of multi-well plates was performed using the High-Content Imaging Opera Phenix
711 Plus from the IRR Imaging Core Facility. Single cell colony lineage scoring was performed using Signals
712 Image Artist software (Revvity), using the 3D Analysis stack processing pipeline. Image stacks were
713 processed to identify and segment colonies. For cell counting, the Find Nuclei – Method C function
714 was utilized. Following segmentation, both cell number and fluorescence intensity data were
715 extracted from the analysed image sets.

716

717 Gastruloid culture, live imaging and quantitative analysis

718 Gastruloid culture was carried out as previously described (van den Brink et al., 2020 Vianello et al.,
719 2020 [dx.doi.org/10.17504/protocols.io.9j5h4q6]). In summary, 2i/LIF/FCS mESCs were washed in PBS
720 before adding 0.05% Trypsin EDTA solution. After cell detachment, 5-10 volumes of fresh 2i/LIF/FCS
721 were added to quench the Trypsin EDTA solution. Cells were transferred to a universal tube and
722 pelleted by centrifugation at 300g for 3 minutes. The media was aspirated, and the pellet was
723 resuspended in cold PBS to make a single cell suspension. A PBS wash was repeated once more, then
724 the cell pellet was resuspended in prewarmed N2B27 to a single cell solution. Cells were counted and
725 diluted to a solution of 8,250 cells/mL in N2B27 medium, after which 40µL of this medium (containing
726 ~330 cells) was plated into untreated u-bottom 96 wells and incubated for 48 hours at 37°C and 5%
727 CO2. At 48 hours, an additional 150µL of N2B27 and 3 µM CHIR99021 were added to the well. The last

728 half of the additional media was expelled forcefully to dislodge the aggregate but without spilling the
729 medium. The aggregates were cultured for a further 24 hours to d3, after which 150µL of media was
730 removed and replaced with 150µL fresh N2B27. At d4, gastruloids were washed in N2B27 and then
731 transferred to flat-bottom and black-walled 96-well plates (PerkinElmer) containing ice-cold phenol-
732 red free N2B27 (Thermofisher #21041025, #12348017) + 10% matrigel supplemented with LY 411575
733 (150nM), IWP-2 (2.5µM), or an equal amount of DMSO. Gastruloids were left for 5 minutes to drop to
734 the bottom of the well and then placed into an incubation chamber at 37°C and 5% CO2 to set the
735 matrigel. After this period, gastruloids were imaged on a confocal microscope (Opera Phenix Plus)
736 every hour for 24 hours. Transmitted light and fluorescence signal was collected using epifluorescence
737 and a 20x objective. Excitations of 488nm and 561nm, emissions of 522nm and 599nm, and exposure
738 times of 0.1s and 0.2s were used for GFP and mCherry respectively. XYZ voxel dimensions were
739 0.89x0.89x5 microns.

740 Image analysis was performed in python using custom scripts and open-source libraries that can be
741 found at 10.5281/zenodo.17192095. Whole tissue segmentation masks were generated using with
742 the Segment Anything Model (SAM) (arXiv:2304.02643). Extending poles of gastruloids were inferred
743 by median projecting $Tbxt^{GFP}$ signal z-stacks to 2D, smoothing with a Gaussian filter, and isolating the
744 largest domain of pixels over the 99th percentile. The vector between the centroid of this domain and
745 the centroid of the whole gastruloid mask determines the orientation to the *Sox2/Tbxt* co-expressing
746 region. The orientation calculated at 10 hours post embedding was applied to all other timepoints.
747 Traces were collected in a 90µm wide linear line orthogonal to the vector orientation. Extending pole
748 "ends" containing Sox2/Tbxt coexpressing domains were isolated as a bin between 35µm and 178µm
749 from the end of the pole. The limits of this region were manually determined by the overlap of
750 $Tbxt^{GFP}$/$Sox2^{mCh}$ signal gradients and previous observations of the location of NMP regions in vivo (i.e
751 the tail tip) (French et al., 2025). Signal of $Tbxt^{GFP}$ and $Sox2^{mCh}$ was measured in 2D by the average
752 intensity from median z-projections of both channels in 0.89µm wide bins along the gastruloid trace.
753 Channel signal was normalised per gastruloid to the 5th and 99th percentile of trace signal values at 0h,
754 then per timepoint and condition normalised to the 99th percentile of DMSO control. For the NMP
755 region Tbxt/Sox2 dynamics, the difference between signal from treated and control conditions was
756 calculated per replicate and timepoint.

757 For flow cytometry analyses, an average of 60 gastruloids were pooled per experiment and bulk
758 analysed at 72hpa and 30 gastruloids at 96hpa. Analyses were done using the NovoCyte Penteon
759 analyser (Agilent). Flow data was analysed in R using the CytoExploreR plugin (Dillon Hammill (2021),
760 CytoExploreR: Interactive Analysis of Cytometry Data. R package version
761 1.1.0. https://github.com/DillonHammill/CytoExploreR).

762 For imaging, gastruloids were fixed with 4% PFA for 40 min at room temperature and stained with
763 DAPI overnight. Gastruloids were mounted in low melting point agarose in a 96-well ibidi plate and
764 clarified with RapiClear 1.49 overnight. An average of 3-5 gastruloids per condition were imaged using
765 an OperaPhenix confocal microscope, a 300 µm section was imaged in 2 µm steps and representative
766 examples were selected. Image analysis was performed in Fiji. E14tg2a gastruloids were used as a
767 negative control to set fluorescence thresholds. A MAX intensity projection of the DAPI signal was
768 used to create outlines of the gastruloids using a custom Fiji macro.

769

770 Mathematical modelling

771 In this study, we propose a minimal mathematical model of the dynamics of NMP cells and their
772 progeny to investigate the cell fate decision process. We assume that each NMP cell divides
773 symmetrically, resulting in i) two NMPs, or ii) two cells that are simultaneously Tbxt+ and Sox2-, or iii)
774 two cells that are simultaneously Tbxt- and Sox2+, or iv) two cells that are simultaneously Tbxt- and
775 Sox2-. We also assume that, with the exception of NMPs, the other cells do not divide within the time
776 window of our study. The model is encoded in the following linear system of ordinary differential
777 equations describing the time course of NMPs and their progeny:

778 $$\frac{dNMP}{dt} = \alpha * NMP - \beta * NMP - \gamma * NMP - \delta * NMP \quad \text{(Eq. 1)}$$

779 $$\frac{dn_{S+T-}}{dt} = 2 * \beta * NMP \quad \text{(Eq. 2)}$$

780 $$\frac{dn_{S-T+}}{dt} = 2 * \gamma * NMP \quad \text{(Eq. 3)}$$

781 $$\frac{dn_{S-T-}}{dt} = 2 * \delta * NMP \quad \text{(Eq. 4)}$$

782 Where *t* is time, *NMP*, n_{S+T-}, n_{S-T+} and n_{S-T-} are the number of NMPs, the number of cells that are
783 simultaneously Sox2+ and T-, the number of cells that are simultaneously Sox2- and T+ and the number
784 of cells that are simultaneously Sox2- and T-, respectively. The model has 4 parameters: α, β, γ and δ,
785 which correspond to the rates at which NMP cells give rise to more NMP cells, cells positive for Sox2
786 but negative for T (*i.e.*, neural lineage), cells negative for Sox2 but positive for T (*i.e.*, mesodermal
787 lineage) and cells negative for both Sox2 and T (double negative).

788 To test whether the model was sufficient to reproduce the dependence of the neural/mesodermal
789 decision on the *T* level of the NMP, we fitted the model to the experimental data.

790 This system has an analytical solution given by

$$NMP(t) = NMP_0 * e^{(\alpha-\beta-\gamma-\delta)t} \quad \text{(Eq. 5)}$$

$$n_{S+T-}(t) = \frac{2*\beta*NMP_0*e^{(\alpha-\beta-\gamma-\delta)t}}{(\alpha-\beta-\gamma-\delta)} + n_{S+T-0} \quad \text{(Eq. 6)}$$

$$n_{S-T+}(t) = \frac{2*\gamma*NMP_0*e^{(\alpha-\beta-\gamma-\delta)t}}{(\alpha-\beta-\gamma-\delta)} + n_{S-T+0} \quad \text{(Eq. 7)}$$

$$n_{S-T-}(t) = \frac{2*\delta*NMP_0*e^{(\alpha-\beta-\gamma-\delta)t}}{(\alpha-\beta-\gamma-\delta)} + n_{S-T-0} \quad \text{(Eq. 8)}$$

Where *NMP₀* is the initial number of NMP cells, n_{S+T-0} is the initial number of Sox2 positive but T negative cells, n_{S-T+0} is the initial number of Sox2 negative but T positive cells and n_{S-T-0} is the initial number of double negative cells.

Finally, we assume the following initial condition, reflecting the experiments described in section 3.

$$NMP_0 = 1 \quad \text{(Eq. 9)}$$

$$n_{S+T-0} = 0 \quad \text{(Eq. 10)}$$

$$n_{S-T+0} = 0 \quad \text{(Eq. 11)}$$

$$n_{S-T-0} = 0 \quad \text{(Eq. 12)}$$

To fit the model to the experimental data, we minimized the distance function δ between the experimental number of cells and the model predicted number of all cell types for each experiment. The distance d was defined as:

$$\delta = \sqrt{(NMP_e - NMP_t)^2 + \left(n_{S+T-e} - n_{S+T-t}\right)^2 + \left(n_{S-T+e} - n_{S-T+t}\right)^2 + \left(n_{S-T-e} - n_{S-T-t}\right)^2} \quad \text{(Eq. 13)}$$

Where NMP_e and NMP_t are the experimental and model predicted number of NMP cells, n_{S+T-e} and n_{S+T-t} are the experimental and model predicted number of cells positive for SOX2 but negative for TBXT, n_{S-T+e} and n_{S-T+t} are the experimental and model predicted number of cells positive for T but negative for SOX2, and finally n_{S-T-e} and n_{S-T-t} are the experimental and model predicted number of cells negative for both SOX2 and TBXT. From all best-fitting parameter values that minimized d for each experiment, we calculated the mean and standard error of the mean.

A Jupyter Notebook (http://jupyter.org/) containing the source code used to generate the mathematical model plots in Fig. 5 and Suppl. Fig. 3, 4 can be found at 10.5281/zenodo.17154933 (Ceccarelli & Chara, 2025).

817

818 Table S1

Reagent	Source	Identifier or catalog number
DNeasy Blood & Tissue Kit	QIAGEN	69504
Q5® Hot Start High-Fidelity 2X Master Mix	BioLabs	M0494S
Agarose	Invitrogen	16500-500
SYBR™ Safe DNA Gel Stain	Invitrogen	S33102
Zymoclean Gel DNA Recovery Kit Protocol	Zymo Research	D4001
Qubit 1X dsDNA HS Assay Kit	Invitrogen	Q33230
Glasgow Minimum Essential Medium	Sigma-Aldrich	G5154
Foetal calf serum	Gibco	10270-106
LIF	Made in-house	N/A
2-mercaptoethanol	Gibco	31350-010
Non-essential amino acids	Gibco	11140-035
L-Glutamine	Invitrogen	25030-024
Sodium Pyruvate	Invitrogen	11360-039
Trypsin	Sigma	59429C
Neurobasal	Thermofisher	21103049
M2 medium	Sigma	M7167
Tyrode's solution, acidic	Sigma	1788
Advanced Dulbecco's Modified Eagle Medium F12	Thermofisher	12634028
B27	Thermofisher	17504001
N2	Thermofisher	17502048
FGF2	R&D Systems	3718-FB
DMEM/F-12, no phenol red	Thermofisher	21041025
Neurobasal™ Medium, minus phenol red	Thermofisher	12348017
CHIRON99021	Axon	1386
TrypLE Express Enzyme 1X	Gibco	12604013
Fibronectin	Sigma	F1141
Matrigel	Corning	354277
Geltrex	Thermofisher	A1413201
Penicillin/streptomycin	Invitrogen	15140-122
ROCK inhibitor Y-27632	Tocris	1254/10
2X Reaction Mix (CellsDirect One-Step qRT-PCR Kit)	Invitrogen	11753100
RNAse inhibitor	Invitrogen	AM2694
Superscript III RT/Platinum Taq mix (CellsDirect One-Step qRT-PCR Kit)	Invitrogen	11753100
LightCycler 480 SYBR Green I Master	Roche	4887352001
LightCycler 480 Multiwell Plate 384	Roche	04729749001
Cas9 protein	Gift from Pollard Lab	Dewari et al., 2018

crRNA guide Sox2	Twist Bioscience	5'-TGCCCCTGTCGCACATGTGA-3'
sgRNA guide pairs Tbxt	IDT	5'-GGCTTTGGGCCGTGTCATAC-3' 5'-AAAGCCTCCTCATAGCCTCG-3'
Zero Blunt™ TOPO™ PCR Cloning Kit	Invitrogen	450245
Sox2-mcherry PCR primers	IDT	5'-AGCTCGCAGACCTACATGAA-3' 5'- GACTTTTGCGAACTCCCTGC-3'
Tbxt-GFP PCR primers	IDT	5'-CGGCACACTAAACAGTCACTGTGC-3' 5'-TCAGTTACAATCTTCTGGCTATGCT-3'

Author contributions

Conceptualization: A.B-C., V.W;

Methodology: A.B.-C., A.G, A.C, O.C, V.W;

Formal analysis: A.G, A.C, O.C, M.F

Investigation: A.B.-C., A.G, A.C, E.K, F.J.W, R.P, M.F, J.A, Y.H, A.S.B, D.L.R,

Resources: F.C.K.W, F.J.W, M.T, S.L

Data curation: A.G, A.B-C, V.W. O.C, A.C;

Writing - original draft: A.B-C, V.W;

Writing - editing: A.B-C, V.W, F.J.W, S.L, M.Y, A.G, F.C.K.W, A.C, O.C;

Visualization: A.B.-C., A.G, A.C, F.W, M.F, J.A, V.W

Supervision: A.B-C, O.C, V.W;

Project administration: A.B-C, V.W;

Funding acquisition: A.B-C, V.W

Acknowledgements

The authors thank the IRR core facilities at the University of Edinburgh for their important contributions to this work, particularly the IRR flow cytometry, tissue culture and high content imaging facilities. We thank Augusto Borges for technical advice on modelling aspects of the study. We would also like to thank André Dias for critically reading an early version of the manuscript.

Competing interests

No competing interests declared.

Funding

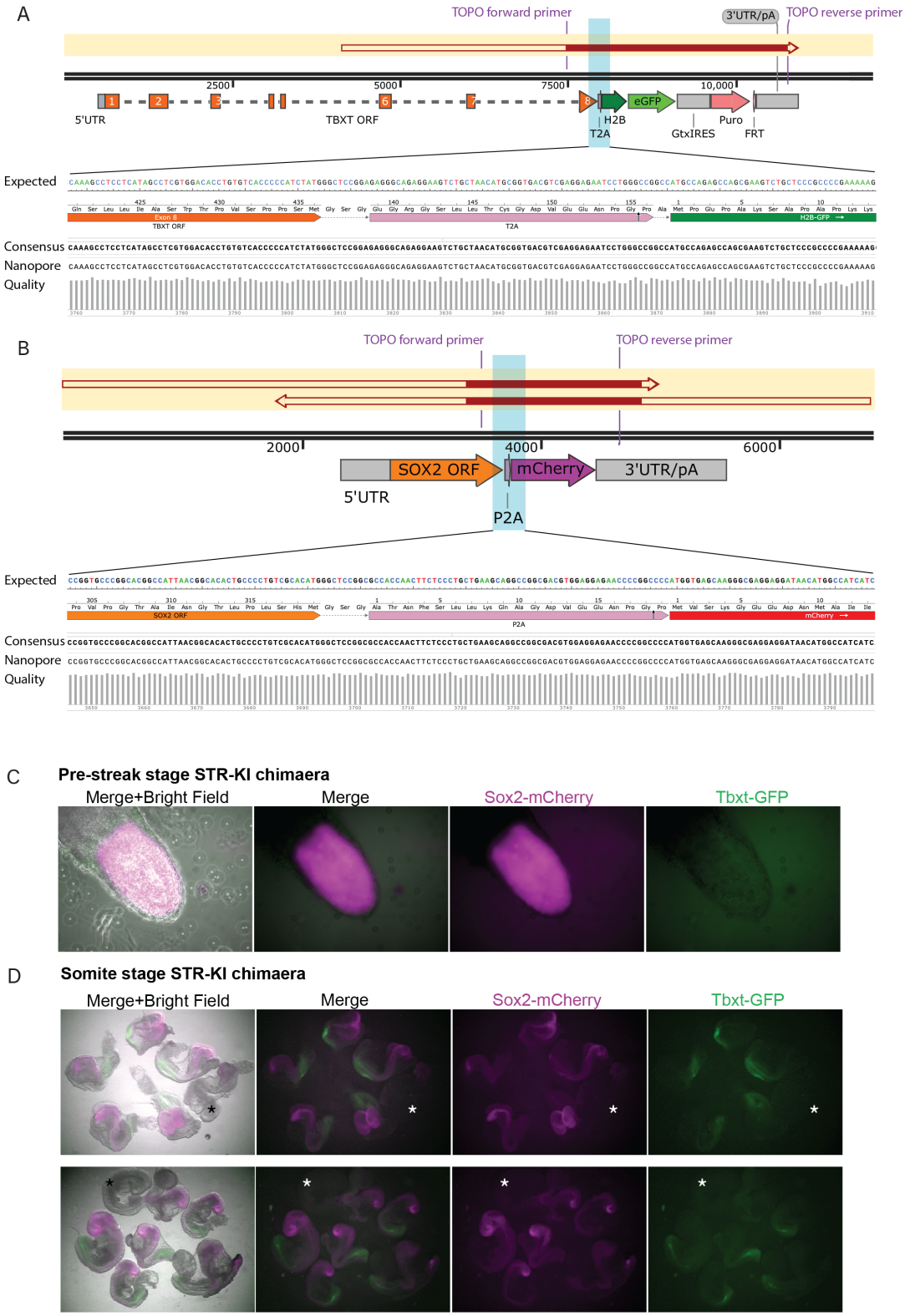
This work was supported by the Medical Research Council (MR/S008799/1 to A.B-C, A.G., V.W, MR/K011200 to E.K, F.C.K.W., Y.H), Carnegie Trust for the Universities of Scotland (RIG012666 to A.B-C), the Agencia Nacional de Promoción Científica y Tecnológica of Argentina (PICT-2019-03828 to A.S.C. and O.C), Biotechnology and Biological Sciences Research Council (BB/X014908/1 to O.C.), Wellcome Trust Senior Fellowship (220298 to M.F., J.A., S.B. and S.L), Japan Society for the Promotion of Science KAKENHI grant (JP19K16157 to F.J.W), Erasmus Exchange fund (to A.G. and D.L.R), Gurdon summer studentship (to A.G).

Data availability

Data availability: All relevant data can be found within the article and its supplementary information.

Supplementary Figures

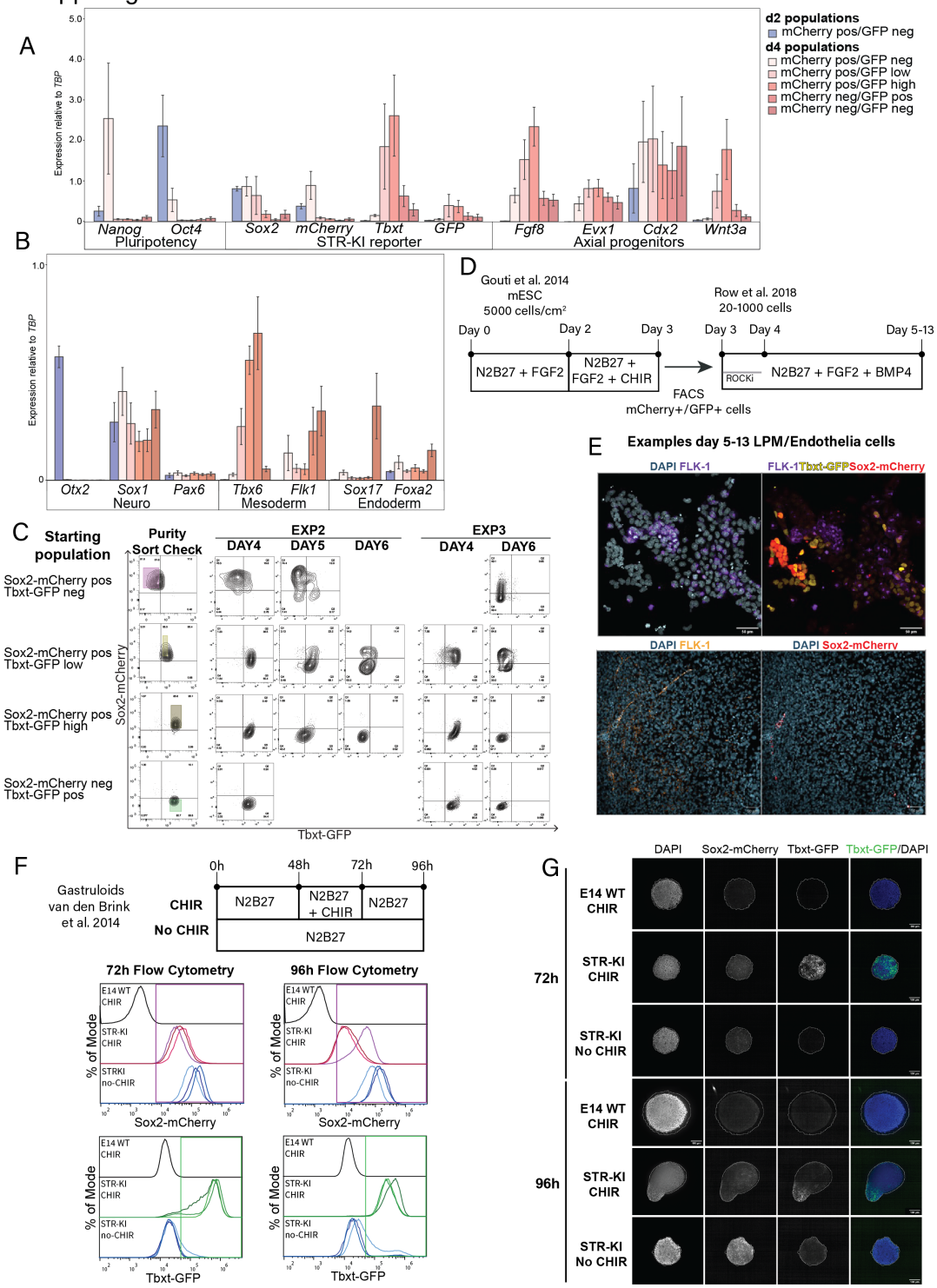
Suppl. Figure 1



Suppl. Figure 1. STR-KI cell line contributes highly to chimaeras. A-B. Schematics of targeted region of the Tbxt **(A)** and Sox2 **(B)** loci in the STR-KI cell line showing correct integration and sequence alignment to the consensus after nanopore sequencing. **C.** mouse pre-streak and **D.** E8.5 (2-12) somite stage chimeras appear

858 morphologically normal, show high contribution of STR-KI cells, and correct location of the reporter fluorescent
859 proteins. Asterisk in D shows a non-contributing littermate as a negative control.

Suppl. Figure 2



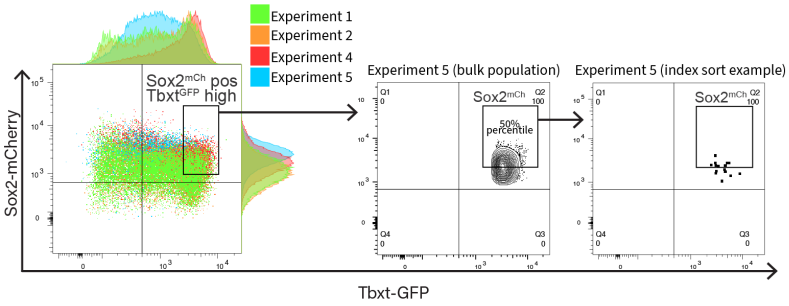
860

861

862 **Suppl. Figure 2. *In vitro* derived *Sox2*-mCherry and *Tbxt*-GFP co-expressing cells are NMPs, and their lineage**
863 **trajectory can be captured for potency studies. A-B.** qRT-PCR analysis of populations on day 2 and day 4 using
864 the same gating strategy and differentiation protocol in Fig. 2A. Gene markers of pluripotency, axial progenitors
865 (inset A), and neural, mesodermal, endodermal differentiation markers (inset B) with n=3 independent
866 experiments for each population. **C.** Repeats of the time-course experiment described in Fig. 2C-D. A factor of
867 variability between these three experiments can be noted and is inherent to any differentiation protocol.
868 However, cell populations of all three experiments followed the same trajectory trends as summarised in Fig.
869 2D. A representative purity check of the initial FACS populations is shown. **D.** Scheme of the experimental set up
870 for FACS *Sox2*^{mCherry}*Tbxt*^{GFP} double positive progenitors and testing of lateral plate mesoderm (LPM) potency. **E.**
871 Colonies stained for the LPM marker Flk-1. In some Flk-1+ colonies changes in morphology where minimal (top)
872 whereas in others elongated FLK-1+ cells were observed, indicating potentially more progressed endothelial
873 differentiation. **F.** Experimental setup for the generation of gastruloids to test symmetry breaking and *Tbxt*^{GFP}
874 activation upon CHIR activation. Flow cytometry on an average of 60 gastruloids were pooled and bulk analysed
875 for 72 hpa and 30 gastruloids for 96 hpa (n=1). **G.** Immunostaining of wildtype (WT) or STR-KI derived gastruloids
876 for *Sox2*^{mCherry} and *Tbxt*^{GFP} at 72 and 96 hpa. Note the absence of Tbxt^{GFP} in both WT and no CHIR treated
877 gastruloids. Axis elongation is visible between 96-120 hpa, depending on the cell line use, with our STR-KI
878 reporter showing elongation from 96 hpa upon CHIR treatment representative image of 3-5 gastruloids analysed
879 per sample.

Suppl. Figure 3

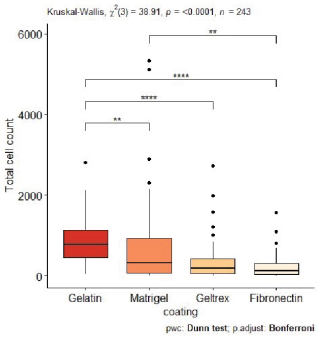
A **Sox2^{mCh}pos/Tbxt^{GFP} high and Sox2^{mch} trends gating strategy for data in Fig. 5**



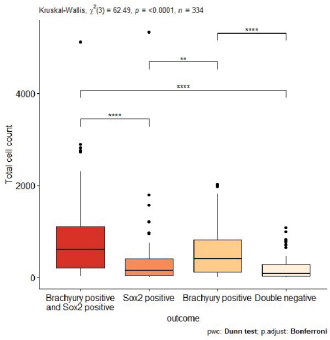
B **% of cells in the 50% percentile Sox2^{mCh} gate**

Experiment	% in Sox2^{mCh} gate
Exp 1	20.9%
Exp 2	16.7%
Exp 4-fibronectin	21.7%
Exp 4-gelatin	36.0%
Exp 4-geltrex	24.0%
Exp 4-matrigel	33.3%
Exp 5-fibronectin	40.0%
Exp 5-gelatin	53.3%
Exp 5-geltrex	60.0%
Exp 5-matrigel	53.3%

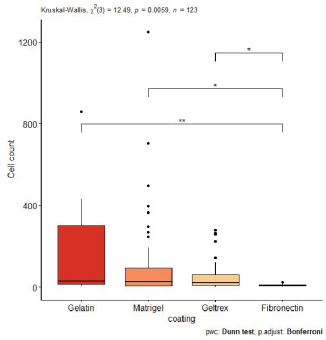
C **Colony cell count by coating**



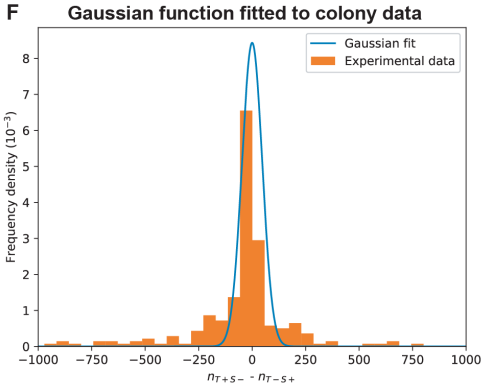
D **Colony cell count by outcome**



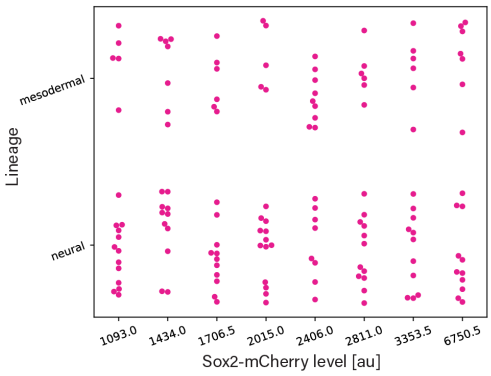
E **Sox2/Tbxt (NMP) cell count by coating**



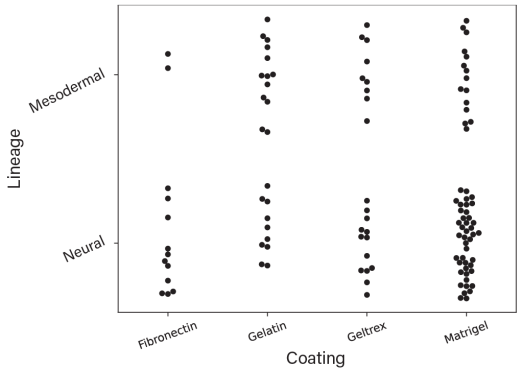
F **Gaussian function fitted to colony data**



G **Colony lineage output by NMP's Sox2 level**



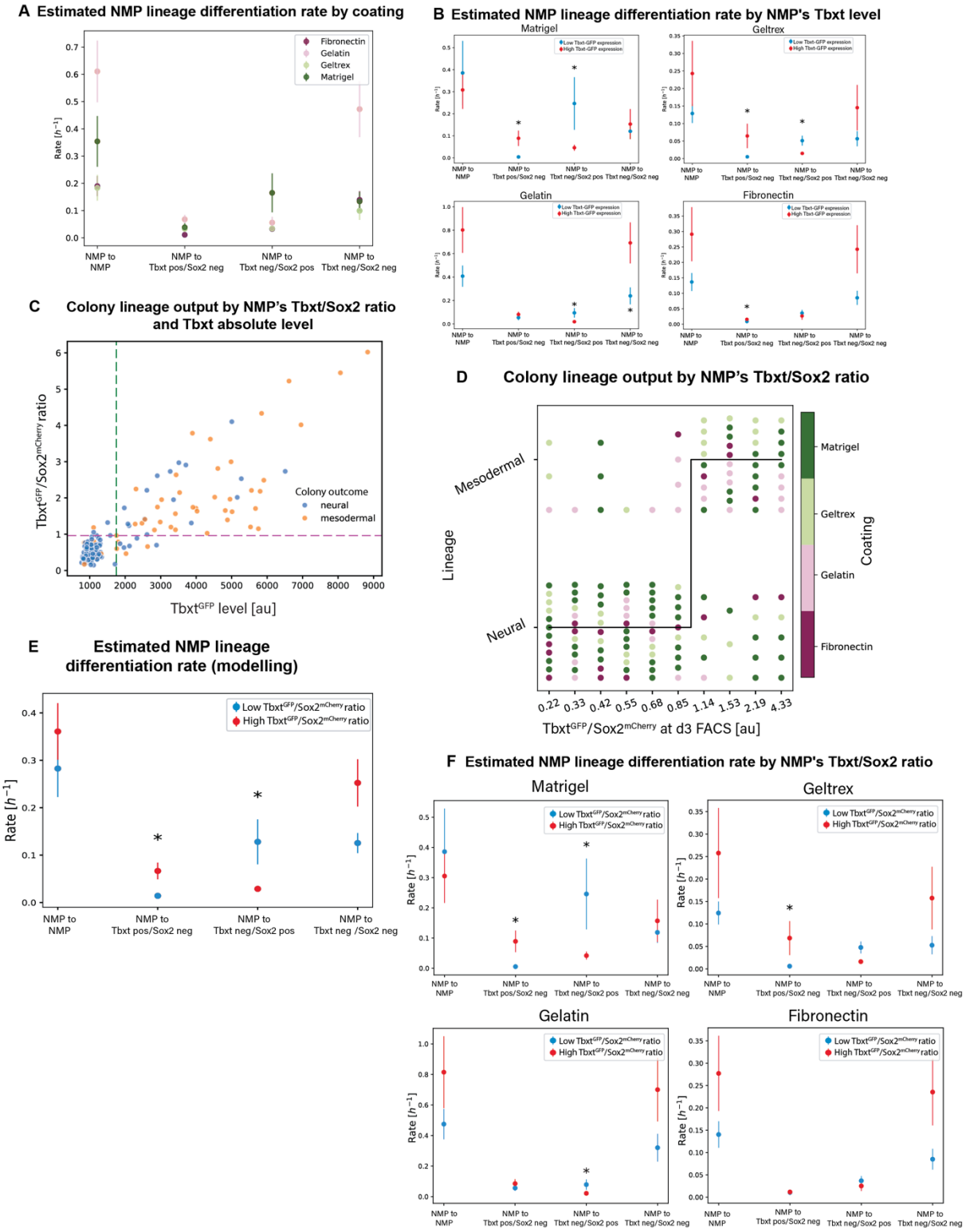
H **Colony lineage output by coating**



882 **Suppl. Figure 3. Quantitative analysis and mathematical modelling of clonal data.**

883 **A.** Overlap of bulk-flow data from NMP differentiations where single cells that fell in the
884 Sox2mChpos/TbxtGFPhigh sorting gate were reanalysed by level of Sox2. Exp 1,2 were unindexed, and Exp 4,5
885 were indexed. The % of cells above the 50% percentile of mCherry were calculated based on the least advanced
886 differentation repeat (exp5). **B.** Summary table listing the % of cells within the Sox2mCh/TbxtGFP high
887 population falling in the top 50% percentile of Sox2mCherry across all experiments. **C.** Total number of cells per
888 colony measured after clonal expansion of single cells grown on different extracellular matrices or classified by
889 colony phenotype (**D**). **E.** Total number of SOX2+TBXT+ cells per colony cultured on different coating matrices.
890 **F.** Histogram for n_{S-T+} - n_{S+T-} (orange bars). Blue line: Gaussian function fitted to the data. **G.** Clonally plated Sox2
891 ^{mCherry} Tbxt^{GFP} (NMP) lineage output as a function of Sox2^{mCherry} levels in individually plated NMPs. Each dot
892 represents one colony. The Sox2^{mCherry} level on the X-axis is the middle value in the bin. **H.** Clonally plated Sox2
893 ^{mCherry} Tbxt^{GFP} (NMP) lineage output as a function of the extracellular coating used. For C-E: Dots are the mean
894 with error bars, the standard error of the mean. Asterisks indicate statistically significant differences ($p < 0.05$)
895 obtained after applying a Mann-Whitney statistical test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** < 0.0001).

Suppl. Figure 4



Suppl. Figure 4. Quantitative analysis and mathematical modelling of clonal data using the Tbxt absolute levels Tbxt/Sox2 levels ratio. A. Lineage differentiation rates estimated with our

901 mathematical model (Tbxt absolute levels), coloured by coating: matrigel (dark green), geltrex (light
902 green), gelatin (pink) and fibronectin (fuchsia). Dots are the mean with error bars, the standard error
903 of the mean**. B.** Differentation lineage rates estimated with our mathematical model (Tbxt absolute
904 levels) sorted by coating used. The initial NMP cells were above (in red) or below (in blue) a TbxtGFP
905 level threshold of 1745 arbitrary units (au) **C.** Colony lineage output when plotting Tbxt/Sox2 ratio
906 over Tbxt absolute numbers. GFP/mCherry threshold value = 0.97 **D.** NMP lineage output as a function
907 of initial $Tbxt^{GFP}$ /$Sox2^{mCh}$ ratio. Each dot represents one colony/experiment, with substrates used:
908 matrigel (dark green), geltrex (light green), gelatin (pink) and fibronectin (fuchsia). The ratio reported
909 is the middle value in the bin. **E.** Mathematical model to estimate fate in index sorted NMPs. The rate
910 indicates whether single founder NMPs were above (red) or below (blue) the initial $Tbxt^{GFP}$ /$Sox2^{mCh}$
911 threshold ratio (0.97), as a function of their colony output at the end of the experiment. The dots and
912 the error bars correspond to the mean and the standard error of the mean, respectively. Asterisks
913 indicate significant differences obtained after applying a Mann-Whitney statistical test ($p < 0.05$). **F.**
914 Differentiation rates of individual lineages estimated with our mathematical model sorted by coating
915 used. The initial NMP cells were above (in red) or below (in blue) the $Tbxt^{GFP}$ /$Sox2^{mCh}$ threshold ratio.
916 Dots are the mean with error bars, the standard error of the mean. Asterisks indicate statistically
917 significant differences ($p < 0.05$) obtained after applying a Mann-Whitney statistical test.

918 **References**

919 Attardi, A., Fulton, T., Florescu, M., Shah, G., Muresan, L., Lenz, M. O., Lancaster, C., Huisken, J., van
920 Oudenaarden, A., & Steventon, B. (2018). Neuromesodermal progenitors are a conserved
921 source of spinal cord with divergent growth dynamics. *Development*, *145*(21).
922 https://doi.org/10.1242/dev.166728
923 Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N., & Lovell-Badge, R. (2003). Multipotent
924 cell lineages in early mouse development depend on SOX2 function. *Genes Dev*, *17*(1), 126-
925 140. https://doi.org/10.1101/gad.224503
926 Binagui-Casas, A., Dias, A., Guillot, C., Metzis, V., & Saunders, D. (2021). Building consensus in
927 neuromesodermal research: Current advances and future biomedical perspectives. *Current*
928 *Opinion in Cell Biology*, *73*, 133-140. https://doi.org/10.1016/j.ceb.2021.08.003
929 Cambray, N., & Wilson, V. (2007). Two distinct sources for a population of maturing axial
930 progenitors. *Development*, *134*(15), 2829-2840. https://doi.org/10.1242/dev.02877
931 Cooper, F., Souilhol, C., Haston, S., Gray, S., Boswell, K., Gogolou, A., Frith, T. J. R., Stavish, D., James,
932 B. M., Bose, D., Kim Dale, J., & Tsakiridis, A. (2024). Notch signalling influences cell fate
933 decisions and HOX gene induction in axial progenitors. *Development*, *151*(3).
934 https://doi.org/10.1242/dev.202098
935 Corsinotti, A., Wong, F. C., Tatar, T., Szczerbinska, I., Halbritter, F., Colby, D., Gogolok, S., Pantier, R.,
936 Liggat, K., Mirfazeli, E. S., Hall-Ponsele, E., Mullin, N. P., Wilson, V., & Chambers, I. (2017).
937 Distinct SoxB1 networks are required for naive and primed pluripotency. *Elife*, *6*.
938 https://doi.org/10.7554/eLife.27746
939 Davis, R. L., & Kirschner, M. W. (2000). The fate of cells in the tailbud of Xenopus laevis.
940 *Development*, *127*(2), 255-267. https://doi.org/10.1242/dev.127.2.255

941 Dewari, P. S., Southgate, B., McCarten, K., Monogarov, G., O'Duibhir, E., Quinn, N., Tyrer, A., Leitner,
942 M. C., Plumb, C., Kalantzaki, M., Blin, C., Finch, R., Bressan, R. B., Morrison, G., Jacobi, A. M.,
943 Behlke, M. A., von Kriegsheim, A., Tomlinson, S., Krijgsveld, J., & Pollard, S. M. (2018). An
944 efficient and scalable pipeline for epitope tagging in mammalian stem cells using Cas9
945 ribonucleoprotein. *Elife*, *7*. https://doi.org/10.7554/eLife.35069
946 Dias, A., Lozovska, A., Wymeersch, F. J., Novona, A., Binagui-Casas, A., Sobral, D., Martins, G. G.,
947 Wilson, V., & Mallo, M. (2020). A Tgfbr1/Snai1-dependent developmental module at the
948 core of vertebrate axial elongation. *Elife*, *9*. https://doi.org/ARTN e56615

949 10.7554/eLife.56615
950 Donnelly, M. L. L., Luke, G., Mehrotra, A., Li, X., Hughes, L. E., Gani, D., & Ryan, M. D. (2001). Analysis
951 of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic
952 reaction, but a novel translational effect: a putative ribosomal 'skip'. *J Gen Virol*, *82*(Pt 5),
953 1013-1025. https://doi.org/10.1099/0022-1317-82-5-1013
954 Edri, S., Hayward, P., Baillie-Johnson, P., Steventon, B. J., & Martinez Arias, A. (2019). An epiblast
955 stem cell-derived multipotent progenitor population for axial extension. *Development*,
956 *146*(10). https://doi.org/10.1242/dev.168187
957 Edri, S., Hayward, P., Jawaid, W., & Arias, A. M. (2019). Neuro-mesodermal progenitors (NMPs): a
958 comparative study between pluripotent stem cells and embryo-derived populations.
959 *Development*, *146*(12). https://doi.org/ARTN dev180190

960 10.1242/dev.180190
961 Fehling, H. J., Lacaud, G., Kubo, A., Kennedy, M., Robertson, S., Keller, G., & Kouskoff, V. (2003).
962 Tracking mesoderm induction and its specification to the hemangioblast during embryonic
963 stem cell differentiation. *Development*, *130*(17), 4217-4227.
964 https://doi.org/10.1242/dev.00589
965 Forlani, S., Lawson, K. A., & Deschamps, J. (2003). Acquisition of Hox codes during gastrulation and
966 axial elongation in the mouse embryo. *Development*, *130*(16), 3807-3819.
967 https://doi.org/10.1242/dev.00573
968 French, M., Migueles, R. P., Neaverson, A., Chakraborty, A., Pettini, T., Steventon, B., Clark, E., Dale,
969 J. K., Blin, G., Wilson, V., & Lowell, S. (2025). A toolkit for mapping cell identities in relation to
970 neighbors reveals conserved patterning of neuromesodermal progenitor populations. *PLoS*
971 *Biol*, *23*(7), e3003244. https://doi.org/10.1371/journal.pbio.3003244
972 French, M., Portero, R., Dale, J. K., Blin, G., Wilson, V., & Lowell, S. (2024). A toolkit for mapping cell
973 identities in relation to neighbours reveals Notch-dependent heterogeneity within
974 neuromesodermal progenitor populations. *bioRxiv*, 2024.2009.2003.610492.
975 https://doi.org/10.1101/2024.09.03.610492
976 Gouti, M., Delile, J., Stamataki, D., Wymeersch, F. J., Huang, Y. L., Kleinjung, J., Wilson, V., & Briscoe,
977 J. (2017). A Gene Regulatory Network Balances Neural and Mesoderm Specification during
978 Vertebrate Trunk Development. *Developmental Cell*, *41*(3), 243-+.
979 https://doi.org/10.1016/j.devcel.2017.04.002
980 Gouti, M., Tsakiridis, A., Wymeersch, F. J., Huang, Y. L., Kleinjung, J., Wilson, V., & Briscoe, J. (2014).
981 In Vitro Generation of Neuromesodermal Progenitors Reveals Distinct Roles for Wnt
982 Signalling in the Specification of Spinal Cord and Paraxial Mesoderm Identity. *Plos Biology*,
983 *12*(8). https://doi.org/ARTN e1001937

984 10.1371/journal.pbio.1001937
985 Gruneberg, H. (1958). Genetical studies on the skeleton of the mouse. XXIII. The development of
986 brachyury and anury. *J Embryol Exp Morphol*, *6*(3), 424-443.
987 https://www.ncbi.nlm.nih.gov/pubmed/13575656
988 Guibentif, C., Griffiths, J. A., Imaz-Rosshandler, I., Ghazanfar, S., Nichols, J., Wilson, V., Gottgens, B.,
989 & Marioni, J. C. (2021). Diverse Routes toward Early Somites in the Mouse Embryo.
990 *Developmental Cell*, *56*(1), 141-+. https://doi.org/10.1016/j.devcel.2020.11.013

991 Guillot, C., Djeffal, Y., Michaut, A., Rabe, B., & Pourquie, O. (2021). Dynamics of primitive streak
992 regression controls the fate of neuromesodermal progenitors in the chicken embryo. *Elife*,
993 *10*. https://doi.org/10.7554/eLife.64819
994 Huang, Y., Wilkie, R., & Wilson, V. (2015). Methods for Precisely Localized Transfer of Cells or DNA
995 into Early Postimplantation Mouse Embryos. *J Vis Exp*(106), e53295.
996 https://doi.org/10.3791/53295
997 Kinney, B. A., Al Anber, A., Row, R. H., Tseng, Y. J., Weidmann, M. D., Knaut, H., & Martin, B. L.
998 (2020). Sox2 and Canonical Wnt Signaling Interact to Activate a Developmental Checkpoint
999 Coordinating Morphogenesis with Mesoderm Fate Acquisition. *Cell Rep*, *33*(4), 108311.
1000 https://doi.org/10.1016/j.celrep.2020.108311
1001 Kispert, A., Ortner, H., Cooke, J., & Herrmann, B. G. (1995). The chick Brachyury gene: developmental
1002 expression pattern and response to axial induction by localized activin. *Dev Biol*, *168*(2), 406-
1003 415. https://doi.org/10.1006/dbio.1995.1090
1004 Koch, F., Scholze, M., Wittler, L., Schifferl, D., Sudheer, S., Grote, P., Timmermann, B., Macura, K., &
1005 Herrmann, B. G. (2017). Antagonistic Activities of Sox2 and Brachyury Control the Fate
1006 Choice of Neuro-Mesodermal Progenitors. *Dev Cell*, *42*(5), 514-526 e517.
1007 https://doi.org/10.1016/j.devcel.2017.07.021
1008 Lawson, K. A., & Wilson, V. (2025). Chapter 3 - A revised staging of mouse development before
1009 organogenesis. In G. Morriss-Kay & S. Srinivas (Eds.), *Kaufman's Atlas of Mouse Development*
1010 *Supplement (Second Edition)* (pp. 51-65). Academic Press.
1011 https://doi.org/https://doi.org/10.1016/B978-0-443-23739-3.00009-2
1012 Martin, B. L., & Kimelman, D. (2012). Canonical Wnt signaling dynamically controls multiple stem cell
1013 fate decisions during vertebrate body formation. *Dev Cell*, *22*(1), 223-232.
1014 https://doi.org/10.1016/j.devcel.2011.11.001
1015 Martins-Costa, C., Wilson, V., & Binagui-Casas, A. (2024). Neuromesodermal specification during
1016 head-to-tail body axis formation. *Curr Top Dev Biol*, *159*, 232-271.
1017 https://doi.org/10.1016/bs.ctdb.2024.02.012
1018 Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A.,
1019 Matoba, R., Sharov, A. A., Ko, M. S., & Niwa, H. (2007). Pluripotency governed by Sox2 via
1020 regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol*, *9*(6), 625-635.
1021 https://doi.org/10.1038/ncb1589
1022 Morabito, R. D., Tatarakis, D., Swick, R., Stettnisch, S., Schilling, T. F., Horsfield, J. A., & Martin, B. L.
1023 (2025). The ratio of Wnt signaling activity to Sox2 transcription factor levels predicts
1024 neuromesodermal fate potential. *bioRxiv*. https://doi.org/10.1101/2025.01.16.633481
1025 Nowotschin, S., Ferrer-Vaquer, A., & Hadjantonakis, A. K. (2010). Imaging mouse development with
1026 confocal time-lapse microscopy. *Methods Enzymol*, *476*, 351-377.
1027 https://doi.org/10.1016/S0076-6879(10)76020-1
1028 Romanos, M., Allio, G., Roussigne, M., Combres, L., Escalas, N., Soula, C., Medevielle, F., Steventon,
1029 B., Trescases, A., & Benazeraf, B. (2021). Cell-to-cell heterogeneity in Sox2 and Bra
1030 expression guides progenitor motility and destiny. *Elife*, *10*.
1031 https://doi.org/10.7554/eLife.66588
1032 Row, R. H., Pegg, A., Kinney, B. A., Farr, G. H., Maves, L., Lowell, S., Wilson, V., & Martin, B. L. (2018).
1033 BMP and FGF signaling interact to pattern mesoderm by controlling basic helix-loop-helix
1034 transcription factor activity. *Elife*, *7*. https://doi.org/ARTN e31018

1035 10.7554/eLife.31018
1036 Stott, D., Kispert, A., & Herrmann, B. G. (1993). Rescue of the tail defect of Brachyury mice. *Genes*
1037 *Dev*, *7*(2), 197-203. https://doi.org/10.1101/gad.7.2.197
1038 Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D. M., Lovell-Badge, R., Papaioannou, V. E., & Kondoh,
1039 H. (2011). Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial
1040 stem cells. *Nature*, *470*(7334), 394-398. https://doi.org/10.1038/nature09729

Tsakiridis, A., & Wilson, V. (2015). Assessing the bipotency of in vitro-derived neuromesodermal progenitors. *F1000Res*, *4*, 100. https://doi.org/10.12688/f1000research.6345.2

Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V., & Nicolas, J. F. (2009). Redefining the Progression of Lineage Segregations during Mammalian Embryogenesis by Clonal Analysis. *Developmental Cell*, *17*(3), 365-376. https://doi.org/10.1016/j.devcel.2009.08.002

Whittingham, D. G. (1974). Embryo banks in the future of developmental genetics. *Genetics*, *78*(1), 395-402. https://doi.org/10.1093/genetics/78.1.395

Wilson, V., & Beddington, R. (1997). Expression of T protein in the primitive streak is necessary and sufficient for posterior mesoderm movement and somite differentiation. *Developmental Biology*, *192*(1), 45-58. https://doi.org/DOI 10.1006/dbio.1997.8701

Wilson, V., Manson, L., Skarnes, W. C., & Beddington, R. S. P. (1995). The T-Gene Is Necessary for Normal Mesodermal Morphogenetic Cell Movements during Gastrulation. *Development*, *121*(3), 877-886. <Go to ISI>://WOS:A1995QN06700023

Wong, F. C. K., Zhang, M., Thomson, E., Schumacher, L. J., Tsakiridis, A., Ashmore, J., Li, T., Blin, G., Karagianni, E., Mullin, N. P., Chambers, I., & Wilson, V. (2025). NANOG is repurposed after implantation to repress Sox2 and begin pluripotency extinction. *EMBO J*. https://doi.org/10.1038/s44318-025-00527-9

Wymeersch, F. J., Huang, Y. L., Blin, G., Cambray, N., Wilkie, R., Wong, F. C. K., & Wilson, V. (2016). Position-dependent plasticity of distinct progenitor types in the primitive streak. *Elife*, *5*. https://doi.org/ARTN e10042

10.7554/eLife.10042

Wymeersch, F. J., Wilson, V., & Tsakiridis, A. (2021). Understanding axial progenitor biology in vivo and in vitro. *Development*, *148*(4). https://doi.org/ARTN dev180612

10.1242/dev.180612

Xia, B., Zhang, W., Zhao, G., Zhang, X., Bai, J., Brosh, R., Wudzinska, A., Huang, E., Ashe, H., Ellis, G., Pour, M., Zhao, Y., Coelho, C., Zhu, Y., Miller, A., Dasen, J. S., Maurano, M. T., Kim, S. Y., Boeke, J. D., & Yanai, I. (2024). On the genetic basis of tail-loss evolution in humans and apes. *Nature*, *626*(8001), 1042-1048. https://doi.org/10.1038/s41586-024-07095-8

Yamaguchi, T. P., Takada, S., Yoshikawa, Y., Wu, N., & McMahon, A. P. (1999). T (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev*, *13*(24), 3185-3190. https://doi.org/10.1101/gad.13.24.3185