1

2 **HOXA9 acts as a regulatory switch in acute myeloid**

3 **leukaemia and myeloproliferative neoplasms**

4 **Laure Talarmin¹, Matthew A. Clarke², David Shorthouse³,**

5 **Jasmin Fisher², Benjamin A Hall^{3*}**

6 1. MRC Cancer Unit, University of Cambridge,
7 Hutchison/MRC Research Centre, Box 197, Cambridge
8 Biomedical Campus, Cambridge, CB2 0XZ, United
9 Kingdom

10 2. UCL Cancer Institute, Paul O'Gorman Building, 72
11 Huntley Street,
12 London, WC1E 6BT, United Kingdom

13 3. Department of Medical Physics and Biomedical
14 Engineering, Malet Place Engineering Building,
15 University College London, Gower Street, London
16 WC1E 6BT, United Kingdom

17 *to whom correspondence should be addressed

18 b.hall@ucl.ac.uk

19 The authors declare no competing financial interests.

20

21 **Abstract**

22 Blood malignancies arise from the dysregulation of
23 haematopoiesis. The type of blood cell and the specific order
24 of oncogenic events initiating abnormal growth ultimately
25 determine the cancer subtype and subsequent clinical
26 outcome. HOXA9 plays an important role in acute myeloid
27 leukaemia (AML) prognosis by promoting blood cell expansion
28 and altering differentiation; however, the function of HOXA9 in
29 other blood malignancies is still unclear. Here, we demonstrate
30 the importance of this gene in chronic myeloproliferative
31 neoplasms (MPN) and highlight the biological switch and
32 prognosis marker properties of HOXA9 in AML and MPN. This
33 binary switch function can explain the clinical stratification of
34 these two blood disorders. First, we establish the ability of
35 HOXA9 to stratify AML patients with distinct cellular and
36 clinical outcomes. Then, through the use of a computational
37 network model of MPN, we show that the self-activation of
38 HOXA9 and its relationship to JAK2 and TET2 can explain the
39 branching progression of JAK2/TET2 mutant MPN patients
40 towards divergent clinical characteristics. Finally, we predict a
41 connection between the RUNX1 and MYB genes and a
42 suppressive role for the NOTCH pathway in MPN diseases.

43 **Introduction**

44 Blood cancers are malignancies that can arise from any type of
 45 blood cell and dramatically affect haematopoiesis.
 46 Myeloproliferative Neoplasms (MPNs) are chronic diseases of
 47 the myeloid lineage characterised by an excessive production
 48 of fully functional terminally differentiated blood cells. These
 49 have been classified into 3 types: polycythemia vera (PV),
 50 essential thrombocythemia (ET), and primary myelofibrosis
 51 (PMF) [1]. Despite the relatively good prognosis of these
 52 diseases, MPN patients are at high risk of thrombosis and can
 53 develop a blast phase MPN (MPN-BP) [2]; a subtype of the
 54 blood cancer Acute Myeloid Leukemia (AML) with poor survival
 55 outcomes [3]. The frequency of MPN transformation to blast
 56 phase MPN is highly related to the initial MPN disease type [4,
 57 5,6]. Therefore, a better understanding of the molecular events
 58 driving the different subtypes of MPNs is essential to help
 59 diagnose patients with higher risk of thrombosis and AML
 60 progression.

61

62 AML itself is an aggressive blood and bone marrow
 63 malignancy defined by the uncontrolled growth of myeloid
 64 progenitor cells along with a myeloid-lineage differentiation
 65 arrest [7]. As with MPN, there exist different types of AML with
 66 a broad range of morphologic, cytogenic and immunologic
 67 features, all associated with diverse clinical outcomes [8].
 68 Despite their similarities; prognosis, symptoms, and genetic
 69 alterations differ between AML and MPN. For example, *JAK2*

70 mutation is the main driver event of MPN diseases yet is rarely
 71 found in *de novo* AML [9]. However, myeloid lineage
 72 dysregulation by both MPN and AML, as well as the ability of
 73 MPN to evolve to AML, indicate that both diseases may share
 74 similar biological mechanisms. The identification of these
 75 processes will help identify aberrant genes and pathways in
 76 blood malignancies that could be targets for new drugs.

77

78 Better understanding of the patterns of genetic alterations in
 79 cancer cells can also be used for the classification of blood
 80 diseases and prediction of progression into more severe forms
 81 of the disease [10]. How different combinations and orders of
 82 mutations lead to different subtypes of cancer remains a major
 83 open question [11, 12]. The importance of mutation order has
 84 been demonstrated in MPN by Ortmann et al [13], who show
 85 that two subpopulations of patients with MPN can be
 86 distinguished by the order of mutation acquisition between the
 87 *TET2* and *JAK2* genes and that these subpopulations have
 88 distinct clinical characteristics. Further analyses of these
 89 cohorts show that patients with *JAK2* mutated before *TET2* are
 90 younger at presentation of the disease in clinics, are more
 91 likely to present with PV, have a higher risk of thrombosis and
 92 respond better to *JAK2* inhibitor ruxolitinib. However, the
 93 molecular interplay between both mutations within cancer cells
 94 and how their order rather than their combination triggers
 95 dissimilar clinical characteristics have not been investigated.

96

97 Overexpression of a single homeobox gene, *HOXA9*, has been
 98 reported as sufficient to quickly induce myeloproliferation,
 99 gradually followed by AML progression after a period of time
 100 [14]. Homeobox genes or HOX genes were first identified in
 101 the fruit fly *Drosophila melanogaster* as essential regulators of
 102 early embryogenesis [15] and are thought to have a critical role
 103 in cancer development [16]. In the HOXA family, *HOXA9* is the
 104 most described gene in literature and its expression was
 105 shown to be the single most highly correlating factor, out of
 106 6817 genes tested, for poor prognosis in AML [17]. The
 107 importance of *HOXA9* in AML has been widely explored;
 108 however, this has mainly focused on specific AML subtypes
 109 such as MLL-rearranged leukaemia [18] and NUP98-*HOXA9*
 110 induced leukaemia [19], while its role in other blood
 111 malignancies such as MPN or other AML subtypes is poorly
 112 characterised. Recently, the oncogenic property of *HOXA9* has
 113 been associated with its self-positive feedback loop in myeloid
 114 precursor cells as a result of its ability to bind its own promoter
 115 [20]. We hypothesise in this work that this specific property can
 116 help stratify patients with blood cancers affecting the myeloid
 117 lineage.

118

119 Using public datasets from AML patients and MPN studies, we
 120 show that bimodal *HOXA9* expression identifies two distinct
 121 cohorts of patients/mice, reflecting the protein acting as a
 122 binary switch in the cell. We show that this switch-like
 123 behaviour enables a clinical stratification in these blood
 124 diseases leading to separation of individuals into two clinically

125 distinct populations. This leads to distinct prognoses, but also
 126 allows specific disease type classification. First, *HOXA9*
 127 bimodal expression affects the clinical features, such as age
 128 and white blood cell counts, but also patient classification into
 129 specific French-American-British (FAB) or molecular subtypes.
 130 Next, we design a computational network model that offers a
 131 mechanistic explanation of the distinct clinical features of MPN
 132 progression in patients with different orders of *JAK2* and *TET2*
 133 mutations. This computational model predicts that *HOXA9* is
 134 directly downstream of *JAK2* and *TET2* and effectively stores
 135 their mutational history, leading to a phenotypic switch in
 136 double mutant cells dependent upon mutation order and
 137 producing distinct subtypes of the disease. Finally, the network
 138 model also predicts a suppressive role for the NOTCH
 139 pathway in MPN and a new interaction between *RUNX1* and
 140 *MYB*.

141 **Materials and Methods**

142

143 Analysis and visualisation of public cancer datasets

144

145 The AML patient data was generated by TCGA and
 146 downloaded with Firebrowse (RNAseq). The file contains data
 147 from 173 patients. We use the Transcripts Per Million (TPM)
 148 normalised data (TCGA RNAseq) and the R package
 149 multimode [21] to determine the significance of gene
 150 expression bimodality. The *modetest* command is used to
 151 reject unimodality with the default ACR method and a p-value
 152 at 0.05. We use the R package Survival [22] and the AML
 153 clinical data from TCGA to plot the survival curves and
 154 compute the p-values of the log rank test. Overall survival
 155 variables are in months and left as is in the analyses. We use
 156 the Plotly Python Open Source Graphing library to plot Sankey
 157 diagrams (available at <https://plot.ly>).

158

159 Differentially expressed genes in HOXA9 cohorts

160

161 We use a python script to separate patients between the two
 162 *HOXA9* expression peaks found in the AML data from TCGA.
 163 31 patients constitute the low peak and 80 the high peak. We
 164 compute the absolute difference of the mean expression of
 165 each gene between each cohort to find the genes which are
 166 most differentially expressed between the two groups of
 167 patients. We rank the genes from the highest to the lowest
 168 differential expression and take the first 30 genes from this list.

169 This workflow was repeated in the comparison of the two
170 *HOXA9* cohorts. The same arbitrary cut-off, 30 first genes with
171 the highest fold change, mostly identified in the HOX family or
172 genes with no determined role in haematopoiesis.

173

174 Microarray data analysis

175

176 The microarray data set reported in [23] is available in the
177 ArrayExpress repository at European Molecular Biology
178 Laboratory–European Bioinformatics Institute
179 (<http://www.ebi.ac.uk/arrayexpress/>) and is accessible through
180 the ArrayExpress accession number E-MTAB-2986. While 12
181 samples are described in the paper, only 11 could be found in
182 the public data, with one wild type sample missing in the
183 microarray.

184

185 For analysis, from the set of all transcripts in the microarray,
186 the genes with a low detection p-value (below 0.05) were
187 filtered and transformed with a quantile normalisation. The
188 ComplexHeatmap R library was used to plot the heatmaps
189 [24].

190

191 XGBoost

192

193 XGBoost (**eXtreme Gradient Boosting**) was used to rank
194 different gene pathways that have been well described in
195 cancer to identify which pathways and genes amongst these
196 pathways have the highest correlation with a gene of interest

197 and its expression level in the AML patients (TCGA RNASeq)
198 [25]. More details can be found in the Supplementary
199 Information.

200

201 Executable network model of MPN

202

203 Computational models of MPN cancer fate determination were
204 constructed as a qualitative network (QN) in the
205 BioModelAnalyzer [26]. This process is described in more
206 detail in Supplemental methods, but briefly QNs are
207 constructed from reported gene interactions in the wider
208 literature, and refined by testing model behaviour against
209 reported phenotypes.

210

211 Results

212

213 1- *HOXA9* expression separates cohorts of AML 214 patients with distinct clinical features.

215 Ectopic expression of *HOXA9* in AML has been widely
216 demonstrated, but few studies have investigated the biological
217 attributes of this transcription factor contributing to
218 leukemogenesis. Zhong et al [20] have shown that *HOXA9* in
219 cell lines can induce its own expression through a positive
220 feedback loop, which promotes a continuous differentiation
221 block and self-renewal leading to increase of hematopoietic
222 stem cells and development of leukaemia. To validate *HOXA9*
223 self-activation and the oncogenic role in leukaemia in patients,
224 we studied its expression in AML patient data from The Cancer
225 Genome Atlas (TCGA) [27]. We find that *HOXA9* has bimodal
226 expression in these data (Fig. 1A). This bimodality separates
227 patients into two cohorts: 31 patients in the low expression
228 peak, and 80 patients in the high expression peak. A survival
229 analyses of both groups using Kaplan-Meier survival curves
230 and the log-rank test confirms that *HOXA9* can be used as a
231 marker of poor prognosis in AML (Fig. 1B). This patient
232 stratification based upon *HOXA9* expression is consistent with
233 the reported positive feedback loop characteristic of this gene
234 and suggests that once activated or inhibited, the gene would
235 maintain its expression level, leading to divergence in the
236 disease progression.

237

238 To investigate if the switch role property of *HOXA9* impacts
239 AML subtypes, we looked at the distribution of French-
240 American-British (FAB, named M0-M7), and molecular
241 classifications among the two *HOXA9* cohorts. We show that
242 different *HOXA9* expression cohorts exclude specific FAB
243 subtypes (Fig. 1C). This suggests that *HOXA9* expression is
244 strongly coupled to FAB subtypes.

245

246 In light of these findings, we looked to characterise the
247 common features of *HOXA9* expression cohorts. Cytogenic
248 aberrations and gene rearrangements are frequent in AML and
249 are known to alter the disease morphology as well as the
250 clinical features and prognosis [27]. We find that *HOXA9*
251 stratifies patients with different molecular classification (Fig.
252 1D). MLL-induced leukaemia has been linked to high *HOXA9*
253 [18], while M3 AML subtype is characterised by PML-RAR α
254 translocation and low *HOXA9* in the literature [28]. Low
255 *HOXA9* expression in AML with RUNX1-RUNXT1 and CBF β -
256 MYH11 abnormalities, which constitute the core binding factor
257 (CBF) AML, was also established but unexplained in literature
258 [29]. With this work, we confirm these findings and further
259 establish the correlation between high *HOXA9* expression and
260 the M0 and M5 subtypes as well as complex cytogenetics.
261 Finally, we expand the *HOXA9* stratification effect by looking at
262 other available clinical features in the data and find that
263 *HOXA9* expression correlates with age, white blood cell count
264 (WBC) and blast percentage in the bone marrow (Fig. 1E).

265

266 PML-RAR α , RUNX1-RUNXT1 (AML1-ETO) or CBFB-MYH11
 267 chromosomal abnormalities confer good prognosis in AML
 268 patients [30, 31]. All these aberrations are linked to low
 269 *HOXA9* expression which also exhibits good survival prognosis
 270 among patients compared to high expression. To confirm that
 271 high *HOXA9* is a poor prognosis marker independently of its
 272 associated molecular aberrations or FAB subtypes, we look at
 273 survival outcomes within FAB classes. M0, M3 and M5 being
 274 all specific to one cohort, we examine the survival of patients
 275 within the M2 and M4 subtypes for high and low *HOXA9*
 276 expression. Survival curves and log rank tests within both
 277 subtypes confirm the poor prognosis marker function of high
 278 *HOXA9* (Fig. S1 and S2). These results as well as the
 279 diverging clinical features among patient cohorts are consistent
 280 with *HOXA9* becoming trapped in high- or low- expression
 281 states through self-activation in AML diseases.

282

283 **2- The *JAK2/TET2/HOXA9* motif explains divergent** 284 **disease clinical outcomes in MPN.**

285 *JAK2* is the most commonly mutated gene in many MPN
 286 patients, but different subtypes of the disease with distinct
 287 clinical traits are observed [32]. In contrast, *TET2* was only
 288 recently identified in blood studies. First discovered in MPN in
 289 2008 by Delhommeau et al [33], *TET2* mutation resulting in its
 290 function loss has been associated with diverse hematologic
 291 malignancies [34]. We have shown that *HOXA9* can enable
 292 clinical stratification in AML through its positive feedback loop.

293 Ortmann [13] describes a bifurcation among MPN patients that
294 acquire *JAK2* and *TET2* mutations in different orders. This
295 raises the question whether *HOXA9* expression could also
296 explain this divergent clinical symptoms.

297

298 To address this question, we constructed a computational
299 network model in a multistep process. In order to reproduce
300 the branching in MPN patients, the underlying network of gene
301 interactions must include genes that are sensitive to the
302 mutation order [35]. This requires that parts of the network act
303 as a switch, capable of storing "memory" of previous events. This
304 "memory" property can be encoded by a positive feedback
305 loop acting on a gene that is downstream of both mutated
306 genes [36]. This hypothetical gene must additionally respond
307 differently to each of the mutations. That is to say, one
308 mutation must activate the gene whilst the other inhibits it, so
309 that the gene can maintain its change in activity after the
310 occurrence of the second mutation. The loop is necessary to
311 induce this inheritable change in the presence of constitutive
312 reset processes such as protein and RNA degradation.

313

314 We developed a computational model of this simple gene motif
315 with *JAK2* and *TET2* genes and a hypothetical gene target with
316 a positive feedback loop (Fig. 2A). We identified *HOXA9* as a
317 gene with connections to *TET2* and *JAK2* that was consistent
318 with our computational motif. *TET2* and *JAK2* have been
319 indirectly and directly linked to *HOXA9* activity (Fig. 2B).
320 *STAT5* is a well-known downstream target of *JAK2* [37], and it

321 is also established that *STAT5* and *HOXA9* act as binding
322 partners in hematopoietic cells [38]. Furthermore, it was
323 recently shown that tyrosine phosphorylation of *HOXA9* is
324 *JAK2*-dependent [39] and seems to increase the effect of
325 *HOXA9* on its downstream targets [39]. Regarding the
326 interaction of *TET2* with *HOXA9*, Bocker et al found
327 significantly reduced expression of *HOXA* genes when *TET2*
328 expression is lost [40]. In particular *HOXA9* expression in
329 kidney is significantly decreased by *TET2* loss. *HOXA9* is
330 therefore activated by both *JAK2* and *TET2* and possesses a
331 self-positive feedback loop property [20]. Therefore, the
332 *JAK2/TET2/HOXA9* motif shares all the required properties for
333 observing a clinical divergence in blood diseases.

334

335 Based on this *JAK2/TET2/HOXA9* motif, we refine our
336 computational model to reproduce the observed biological
337 differences between patients with different combinations of
338 *JAK2* and *TET2* mutations. To do so, we extend our
339 computational network model featuring three main genes:
340 *JAK2*, *TET2*, and *HOXA9* (Fig. 2C). Our model includes six
341 phenotypes relevant to cancer development: stem cell self-
342 renewal, common myeloid progenitor (CMP) expansion,
343 granulocyte-monocyte progenitor (GMP) expansion, GMP
344 differentiation, erythroid differentiation and megakaryocyte-
345 erythroid progenitor (MEP) expansion. To capture the wider
346 biology of MPN progression, we further included important
347 hematopoietic markers in our computational model. A detailed
348 literature review and full description of how we built the

349 network model are available in Table S1, Table S2 and the
350 Supplemental Methods.

351

352 Finally, four fundamental cancer genotypes are defined: the
353 wild type (no mutation), the *TET2* single mutant, the *JAK2*
354 single mutant and the double mutant (in either order) (Table 1).

355 The wild type model illustrates haematopoiesis in its healthy
356 state. The single mutants are defined using the literature (see
357 Supplementary Information). The final genotype is the double
358 mutant which can lead to one of two cancer endpoints (fixpoint
359 attractors that represent one of the two clinical outcomes).
360 Each fixpoint represents either *TET2* first or *JAK2* first double
361 mutants and are defined from results presented by Ortmann et
362 al [13]. Our computational model as shown in Figure 2C
363 reproduces the expected behaviours described in Table 1 and
364 therefore the clinical stratification observed in Ortmann et al
365 [13]. The model suggests that the elevated differentiation
366 observed in the *JAK2* first double mutants [13] is induced by
367 the increased expression of *RUNX1*, *KLF1* and *GATA1* as well
368 as the downregulation of *MYC* not found in *TET2* first double
369 mutant. This gene expression difference between double
370 mutants can partly explain the divergent clinical behaviours
371 between the two groups of patients, including the increased
372 risk of thrombosis and the faster diagnosis as a result of the
373 abnormally high number of differentiated cells in these
374 patients.

375 3- **MPN network predicts new gene dynamics and** 376 **interactions.**

377 The computational model identifies new gene dynamics as part
378 of the MPN disease progression. In the complete network
379 model, *HOXA9* requires both *JAK2* and *TET2* expression to
380 remain active (Table S2). Upregulation of either *JAK2* or
381 *HOXA9* results in the hyperactivation of *HOXA9* while *TET2*
382 loss causes inactivation. Wild type activity is maintained by the
383 balance of these two genes. *JAK2* activation mutation and
384 *TET2* loss both drive the system into a committed state. *JAK2*
385 activation raises *HOXA9* activity to a level at which it can
386 maintain its activity through control of its own expression.
387 Subsequent loss of *TET2* does not impact its activity as this
388 hyperactivation makes it independent of *TET2*. Conversely,
389 *TET2* loss causes a loss of *HOXA9* expression in the cell,
390 rendering it insensitive to subsequent *JAK2* activation. This
391 occurs as *HOXA9* expression drops, preventing a subsequent
392 response to *JAK2* activation due to low concentration of the
393 protein in the cell. Therefore, a possible explanation of the
394 order dependence could be a combination of mutual
395 dependency between *JAK2* and *TET2* to activate *HOXA9*,
396 combined with the positive feedback self-loop of *HOXA9* itself.

397

398 Whilst building single mutant phenotypes, we noticed a
399 relationship between *JAK2* and GMP expansion is required to
400 match the increased number of myeloid progenitors observed
401 in organisms with a *JAK2* mutation. To explore possible

402 pathways downstream of *JAK2* that could explain this link to
 403 myeloid diseases, we applied a machine learning approach
 404 (XGBoost) to AML TCGA data as a relevant and closely
 405 related blood cancer. We found that *JAK2* is highly correlated
 406 with the NOTCH pathway (Fig. 2D), which has been found to
 407 act as a tumour suppressor in leukemia due to the great
 408 expansion of GMP cells after loss of NOTCH signalling [41].
 409 From the SHAP scores of NOTCH genes plotted in Figure 2E,
 410 we identified *ITCH* to be among the genes with a high score in
 411 the NOTCH pathway for *JAK2*. *ITCH* controls the degradation
 412 of NOTCH [42] and is found to be induced by JNK1 [43] from
 413 the MAPK pathway which is a well-known downstream
 414 pathway of *JAK2/STAT5* [44]. We therefore suggest that the
 415 *JAK2* path to GMP expansion could be MAPK and NOTCH
 416 dependent.

417

418 Another new interaction predicted by our network model is the
 419 inhibition of *MYB* by *RUNX1*. The Common Myeloid
 420 Progenitors (CMP) are found to be differentially expanded
 421 between *JAK2* and *TET2* first patients in Ortmann et al [13]. In
 422 our initial model, we integrated an inhibition interaction
 423 between *SPI1* and *MYB*, our CMP expansion marker, a
 424 connection which has been found experimentally [45]. This
 425 inhibition and the stable *SPI1* expression in the double mutant
 426 states prevented the known bifurcation in CMP expansion in
 427 double mutants. Further investigations lead us to suggest that
 428 the bifurcation could be obtained by replacing *SPI1* by *RUNX1*
 429 for the *MYB* inhibition which is supported by different studies.

430 *RUNX1* activates *SPI1* and *GATA1*, and both are found to be
431 inhibitors of *MYB* [45, 46]. Additionally, conditional knockout of
432 *RUNX1* in mice results in enhanced CMP frequencies [47][48].
433 All together, these findings suggest that *RUNX1* can be linked
434 to CMP expansion via *MYB* inhibition.

435

436 **4- Analyses of public MPN datasets validate NOTCH** 437 **role in MPN as well as HOXA9 bimodality and** 438 **prognosis role in blood diseases.**

439 To validate the predictions arising from our MPN computational
440 model, we compared our findings to public MPN data not used
441 in the model construction. Chen et al compare MPN with
442 different *JAK2* and *TET2* mutational profiles using
443 transcriptomic mouse data [23]. We compare the gene
444 expression of pathways/gene subsets to those we have
445 included in our model to determine if our model fits their data.
446 We first find that the NOTCH pathway behaves as predicted
447 (Fig. 2F). We also find that the trend in the expression of
448 *RUNX1*, *MYC* and *MYB* support our model (Fig. S3, S4 and
449 S5). Confusingly *HOXA9* expression showed a “switching”
450 behaviour in this mouse model but associated a low
451 expression of *HOXA9* with *JAK2* mutations and high
452 expression with *TET2* mutations.

453

454 Jeong et al [49] previously demonstrated the direct
455 phosphorylation of *TET2* by *JAK2* in a combination of *in vitro*
456 human/murine hematopoietic cell lines with erythroid

457 characteristics. Once phosphorylated *TET2* activates *KLF1*, an
 458 important positive regulator of erythropoiesis [50]. In this
 459 context, loss of *TET2* implies reduced erythroid differentiation
 460 which is in agreement with our model. In the same study, the
 461 authors show in a murine cell line that *JAK2* mutation leads to
 462 *HOXA9* upregulation. These findings are consistent with our
 463 *JAK2/TET2/HOXA9* motif but disagree with Chen's microarray
 464 experiments where *HOXA9* expression is lowered in *JAK2*
 465 single and double mutants (Fig. 2F). Given the downstream
 466 genes follow the expected expression, this raises the question
 467 of whether the interactions in the original motif should be
 468 replaced by a pair of inhibitions expected by Chen data rather
 469 than activations. Whilst there exist possible routes to connect
 470 *TET2* and *HOXA9* through an inhibition, and to change the
 471 sign of the downstream interactions for *HOXA9*, we are
 472 however unable to find evidence of inhibition of *HOXA9* by
 473 *JAK2*. We further note that as the Jeong data is human
 474 derived, it may be a more representative experimental model
 475 system. Both datasets however support the role of *HOXA9* as
 476 a binary switch in MPN.

477 Discussion

478 Out of 6817 genes tested *HOXA9* is the single most highly
 479 correlating factor for poor prognosis due to treatment failure in
 480 AML [17]. This finding has made *HOXA9* the most studied
 481 gene in the *HOXA* family. Here we have demonstrated that it
 482 acts in AML as a discrete switch rather than a spectrum, which
 483 impact AML clinical characteristics such as classification and
 484 survival. Our study shows that despite the existing correlation
 485 between *HOXA9* expression and some AML subtypes, its poor
 486 prognosis characteristic is independent of patient classification.
 487 We further extend the prognosis marker role of the *HOXA9*
 488 gene to another blood disorder, MPN. We show that in MPN
 489 diseases with *JAK2/TET2* mutations, *HOXA9* high expression
 490 is found in the *JAK2* first patients while *TET2* first patients
 491 display lower *HOXA9* expression. As *JAK2* first patients have a
 492 higher risk of developing thrombosis compared to *TET2* first
 493 patients, and as thrombotic events are the main causes of
 494 death in MPNs patients [51], this further suggests a deleterious
 495 influence of *HOXA9* high expression on patient clinical
 496 outcomes in another myeloid disease and emphasise the role
 497 of *HOXA9* as a poor prognosis marker in blood malignancies.

498

499 In addition to providing new insights into the regulatory control
 500 of cancer cell fate through *HOXA9*, our computational network
 501 model recapitulates the disease symptoms using well-known
 502 hematopoietic transcription factors. Further investigations of
 503 these genes could benefit clinicians by designing new drugs or

504 applying already existing treatments to reduce symptoms and
505 the risk of developing blast phase MPN. In addition to the
506 specific claims of the model, several other clinical implications
507 arise. One key feature of *TET2*-first MPN patients is their
508 reduced sensitivity to Ruxolitinib, a JAK2 inhibitor drug [13].
509 Interestingly our computational model suggests that after *TET2*
510 loss, most common *JAK2* targets are unchanged by *JAK2*
511 activation mutation due to the “switching” property exerted by
512 *HOXA9* self-loop. It follows that *JAK2* inhibition is therefore
513 inefficient for those genes. Finally, whilst *JAK2* is the main
514 driver mutation found in all MPN patients, different diseases
515 with distinct clinical traits can be observed [32]. Until now, the
516 source of this clinical diversity following *JAK2* mutation was
517 unclear. Here, we demonstrate that patients who first had a
518 *TET2* mutation have a reduced number of erythroid cells as a
519 result of *TET2* indirect downregulation of *GATA1* and *KLF1*
520 which explain the reduced number of PV diseases in *TET2* first
521 patients despite the presence of *JAK2* mutation [13]. While
522 *JAK2* dysregulation may be the principal driver of MPNs, other
523 mutations shape the disease clinical type by altering the
524 normal development of distinct hematopoietic subpopulations.
525 Finally, we predict the involvement of the NOTCH pathway in
526 MPN diseases. NOTCH shows both oncogenic and tumour
527 suppressor roles in different tissues and in the hematopoietic
528 system: NOTCH favours cancer growth in T acute
529 lymphoblastic leukemia (T-ALL) through its *MYC* activation but
530 is also found to augment the host immune response against
531 cancer by activation of M1 macrophages [52]. The role of

532 NOTCH in hematopoietic stem and progenitor cells is still an
533 on-going debate, however, it seems that a certain level of
534 NOTCH signalling is required to protect individuals from
535 hematological malignancies [53]. We suggest that *JAK2*
536 increases GMP expansion through its inhibitory effect on
537 NOTCH via the MAPK pathway and ITCH and so predict a
538 tumour suppressor role for NOTCH in the GMP cell population.
539 Our network model suggests a novel mechanism for
540 understanding how cancer fate can be determined through
541 regulatory switches and highlights several new areas for
542 further studies.

543

544 **Acknowledgements**

545 We thank members of the Hall group at the MRC-Cancer Unit
546 and the Fisher group at the University College London Cancer
547 Institute, and Aleksandra Watson at the University of
548 Cambridge for valuable discussions. BAH acknowledges
549 support from the Royal Society (grant no. UF130039). DS and
550 BAH were further supported by the Medical Research Council
551 (grant no. MR/S000216/1) LT acknowledges support from
552 Microsoft Research. JF was supported by the National Institute
553 for Health Research University College London Hospitals
554 Biomedical Research Centre and Cancer Research UK.

555 **Competing Interests**

556 The authors declare no competing financial interests.

557 **References**

558

559 [1] J. L. Spivak, “Myeloproliferative neoplasms,” *New*
560 *England Journal of Medicine*, vol. 376, no. 22, pp. 2168–2181,
561 2017.

562 [2] A. Tefferi, M. Mudireddy, F. Mannelli, K. H. Begna,
563 M. M. Patnaik, C. A. Hanson, R. P. Ketterling, N. Gangat,
564 M. Yogarajah, V. De Stefano *et al.*, “Blast phase
565 myeloproliferative neoplasm: Mayo-agimm study of 410
566 patients from two separate cohorts,” *Leukemia*, vol. 32, no. 5,
567 pp. 1200–1210, 2018.

568 [3] M. Yogarajah and A. Tefferi, “Leukemic transformation
569 in myeloproliferative neoplasms: a literature review on risk,
570 characteristics, and outcome,” in *Mayo Clinic Proceedings*,
571 vol. 92, no. 7. Elsevier, 2017, pp. 1118–1128.

572 [4] A. Tefferi, P. Guglielmelli, D. R. Larson, C. Finke, E. A.
573 Wassie, L. Pieri, N. Gangat, R. Fjerza, A. A. Belachew, T. L.
574 Lasho *et al.*, “Long-term survival and blast transformation in
575 molecularly annotated essential thrombocythemia,
576 polycythemia vera, and myelofibrosis,” *Blood, The Journal of*
577 *the American Society of Hematology*, vol. 124, no. 16, pp.
578 2507–2513, 2014.

579 [5] A. Tefferi, E. Rumi, G. Finazzi, H. Gisslinger,
580 A. Vannucchi, F. Rodeghiero, M. Randi, R. Vaidya,
581 M. Cazzola, A. Rambaldi *et al.*, “Survival and prognosis among
582 1545 patients with contemporary polycythemia vera: an

international study," *Leukemia*, vol. 27, no. 9, pp. 1874–1881,
2013.

[6] T. Barbui, J. Thiele, F. Passamonti, E. Rumi, E. Boveri,
M. Ruggeri, F. Rodeghiero, E. S. d'Amore, M. L. Randi,
I. Bertozzi *et al.*, "Survival and disease progression in essential
thrombocythemia are significantly influenced by accurate
morphologic diagnosis: an international study," *Journal of*
clinical oncology, vol. 29, no. 23, pp. 3179–3184, 2011.

[7] C. S. Grove and G. S. Vassiliou, "Acute myeloid
leukaemia: a paradigm for the clonal evolution of cancer?"
Disease models & mechanisms, vol. 7, no. 8, pp. 941–951,
2014.

[8] U. Bacher, W. Kern, S. Schnittger, W. Hiddemann,
C. Schoch, and T. Haferlach, "Further correlations of
morphology according to fab and who classification to
cytogenetics in de novo acute myeloid leukemia: a study on
2,235 patients," *Annals of hematology*, vol. 84, no. 12, pp.
785–791, 2005.

[9] J. Aynardi, R. Manur, P. R. Hess, S. Chekol, J. J.
Morrissette, D. Babushok, E. Hexner, H. J. Rogers, E. D. Hsi,
E. Margolskee *et al.*, "Jak2 v617f-positive acute myeloid
leukaemia (aml): a comparison between de novo aml and
secondary aml transformed from an underlying
myeloproliferative neoplasm. a study from the bone marrow
pathology group," *British journal of haematology*, vol. 182,
no. 1, pp. 78–85, 2018.

- 609 [10] R. C. Lindsley, “Uncoding the genetic heterogeneity of
610 myelodysplastic syndrome,” *Hematology 2014, the American*
611 *Society of Hematology Education Program Book*, vol. 2017,
612 no. 1, pp. 447–452, 2017.
- 613 [11] D. G. Kent and A. R. Green, “Order matters: the order
614 of somatic mutations influences cancer evolution,” *Cold Spring*
615 *Harbor perspectives in medicine*, vol. 7, no. 4, p. a027060,
616 2017.
- 617 [12] M. A. Clarke, S. Woodhouse, N. Piterman, B. A. Hall,
618 and J. Fisher, “Using state space exploration to determine how
619 gene regulatory networks constrain mutation order in cancer
620 evolution,” in *Automated reasoning for systems biology and*
621 *medicine*. Springer, 2019, pp. 133–153.
- 622 [13] C. A. Ortmann, D. G. Kent, J. Nangalia, Y. Silber, D. C.
623 Wedge, J. Grinfeld, E. J. Baxter, C. E. Massie,
624 E. Papaemmanuil, S. Menon *et al.*, “Effect of mutation order on
625 myeloproliferative neoplasms,” *New England Journal of*
626 *Medicine*, vol. 372, no. 7, pp. 601–612, 2015.
- 627 [14] C. Bach, S. Buhl, D. Mueller, M.-P. Garcá-Cuéllar,
628 E. Maethner, and R. K. Slany, “Leukemogenic transformation
629 by hoxa cluster genes,” *Blood, The Journal of the American*
630 *Society of Hematology*, vol. 115, no. 14, pp. 2910–2918, 2010.
- 631 [15] E. B. Lewis, “A gene complex controlling segmentation
632 in drosophila,” in *Genes, development and cancer*. Springer,
633 1978, pp. 205–217.

- 634 [16] S. Bhatlekar, J. Z. Fields, and B. M. Boman, "Hox
635 genes and their role in the development of human cancers,"
636 *Journal of molecular medicine*, vol. 92, no. 8, pp. 811–823,
637 2014.
- 638 [17] T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard,
639 M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R.
640 Downing, M. A. Caligiuri *et al.*, "Molecular classification of
641 cancer: class discovery and class prediction by gene
642 expression monitoring," *science*, vol. 286, no. 5439, pp. 531–
643 537, 1999.
- 644 [18] J. Faber, A. V. Krivtsov, M. C. Stubbs, R. Wright, T. N.
645 Davis, M. van den Heuvel-Eibrink, C. M. Zwaan, A. L. Kung,
646 and S. A. Armstrong, "Hoxa9 is required for survival in human
647 mll-rearranged acute leukemias," *Blood, The Journal of the*
648 *American Society of Hematology*, vol. 113, no. 11, pp. 2375–
649 2385, 2009.
- 650 [19] Y. Shima, M. Yumoto, T. Katsumoto, and
651 I. Kitabayashi, "Mll is essential for nup98-hoxa9-induced
652 leukemia," *Leukemia*, vol. 31, no. 10, pp. 2200–2210, 2017.
- 653 [20] X. Zhong, A. Prinz, J. Steger, M.-P. Garcia-Cuellar,
654 M. Radsak, A. Bentaher, and R. K. Slany, "Hoxa9 transforms
655 murine myeloid cells by a feedback loop driving expression of
656 key oncogenes and cell cycle control genes," *Blood advances*,
657 vol. 2, no. 22, pp. 3137–3148, 2018.

- 658 [21] J. Ameijeiras-Alonso, R. M. Crujeiras, and
659 A. Rodriguez-Casal, “multimode: An r package for mode
660 assessment,” *arXiv preprint arXiv:1803.00472*, 2018.
- 661 [22] T. Therneau and T. Lumley, “R survival package,”
662 2013.
- 663 [23] E. Chen, R. K. Schneider, L. J. Breyfogle, E. A. Rosen,
664 L. Poveromo, S. Elf, A. Ko, K. Brumme, R. Levine, B. L. Ebert
665 *et al.*, “Distinct effects of concomitant jak2v617f expression
666 and tet2 loss in mice promote disease progression in
667 myeloproliferative neoplasms,” *Blood, The Journal of the*
668 *American Society of Hematology*, vol. 125, no. 2, pp. 327–335,
669 2015.
- 670 [24] Z. Gu, R. Eils, and M. Schlesner, “Complex heatmaps
671 reveal patterns and correlations in multidimensional genomic
672 data,” *Bioinformatics*, vol. 32, no. 18, pp. 2847–2849, 2016.
- 673 [25] T. Chen and C. Guestrin, “Xgboost: A scalable tree
674 boosting system,” in *Proceedings of the 22nd acm sigkdd*
675 *international conference on knowledge discovery and data*
676 *mining*, 2016, pp. 785–794.
- 677 [26] D. Benque, S. Bourton, C. Cockerton, B. Cook,
678 J. Fisher, S. Ishtiaq, N. Piterman, A. Taylor, and M. Y. Vardi,
679 “Bma: Visual tool for modeling and analyzing biological
680 networks,” in *International Conference on Computer Aided*
681 *Verification*. Springer, 2012, pp. 686–692.
- 682 [27] C. G. A. R. Network, “Genomic and epigenomic
683 landscapes of adult de novo acute myeloid leukemia,” *New*

684 *England Journal of Medicine*, vol. 368, no. 22, pp. 2059–2074,
685 2013.

686 [28] K. Rejlova, A. Musilova, K. S. Kramarzova, M. Zaliova,
687 K. Fiser, M. Alberich-Jorda, J. Trka, and J. Starkova, “Low hox
688 gene expression in pml-rara-positive leukemia results from
689 suppressed histone demethylation,” *Epigenetics*, vol. 13, no. 1,
690 pp. 73–84, 2018.

691 [29] A. Lasa, M. Carnicer, A. Aventin, C. Estivill, S. Brunet,
692 J. Sierra, and J. Nomdedeu, “Meis 1 expression is
693 downregulated through promoter hypermethylation in aml1-eto
694 acute myeloid leukemias,” *Leukemia*, vol. 18, no. 7, pp. 1231–
695 1237, 2004.

696 [30] J. C. Byrd, K. Mrózek, R. K. Dodge, A. J. Carroll, C. G.
697 Edwards, D. C. Arthur, M. J. Pettenati, S. R. Patil, K. W. Rao,
698 M. S. Watson *et al.*, “Pretreatment cytogenetic abnormalities
699 are predictive of induction success, cumulative incidence of
700 relapse, and overall survival in adult patients with de novo
701 acute myeloid leukemia: results from cancer and leukemia
702 group b (calgb 8461) presented in part at the 43rd annual
703 meeting of the american society of hematology, orlando, fl,
704 december 10, 2001, and published in abstract form. 59,”
705 *Blood, The Journal of the American Society of Hematology*,
706 vol. 100, no. 13, pp. 4325–4336, 2002.

707 [31] Z.-Y. Wang and Z. Chen, “Acute promyelocytic
708 leukemia: from highly fatal to highly curable,” *Blood, The*

- 709 *Journal of the American Society of Hematology*, vol. 111,
710 no. 5, pp. 2505–2515, 2008.
- 711 [32] R. L. Levine, M. Wadleigh, J. Cools, B. L. Ebert,
712 G. Wernig, B. J. Huntly, T. J. Boggon, I. Wlodarska, J. J. Clark,
713 S. Moore *et al.*, “Activating mutation in the tyrosine kinase jak2
714 in polycythemia vera, essential thrombocythemia, and myeloid
715 metaplasia with myelofibrosis,” *Cancer cell*, vol. 7, no. 4, pp.
716 387–397, 2005.
- 717 [33] F. Delhommeau, S. Dupont, V. D. Valle, C. James,
718 S. Trannoy, A. Masse, O. Kosmider, J.-P. Le Couedic,
719 F. Robert, A. Alberdi *et al.*, “Mutation in tet2 in myeloid
720 cancers,” *New England Journal of Medicine*, vol. 360, no. 22,
721 pp. 2289–2301, 2009.
- 722 [34] S. Chiba, “Dysregulation of tet2 in hematologic
723 malignancies,” *International journal of hematology*, vol. 105,
724 no. 1, pp. 17–22, 2017.
- 725 [35] Y. Z. Paterson, D. Shorthouse, M. W. Pleijzier,
726 N. Piterman, C. Bendtsen, B. A. Hall, and J. Fisher, “A toolbox
727 for discrete modelling of cell signalling dynamics,” *Integrative
728 Biology*, vol. 10, no. 6, pp. 370–382, 2018.
- 729 [36] W. Xiong and J. E. Ferrell, “A positive-feedback-based
730 bistable memory module that governs a cell fate decision,”
731 *Nature*, vol. 426, no. 6965, pp. 460–465, 2003.
- 732 [37] S. Zhao, K. Zoller, M. Masuko, P. Rojnuckarin, X. O.
733 Yang, E. Parganas, K. Kaushansky, J. N. Ihle,
734 T. Papayannopoulou, D. M. Willerford *et al.*, “Jak2,

- 735 complemented by a second signal from c-kit or flt-3, triggers
736 extensive self-renewal of primary multipotential hemopoietic
737 cells," *The EMBO journal*, vol. 21, no. 9, pp. 2159–2167, 2002.
- 738 [38] C. E. de Bock, S. Demeyer, S. Degryse, D. Verbeke,
739 B. Sweron, O. Gielen, R. Vandepoel, C. Vicente, M. V. Bempt,
740 A. Dagklis *et al.*, "Hoxa9 cooperates with activated jak/stat
741 signaling to drive leukemia development," *Cancer discovery*,
742 vol. 8, no. 5, pp. 616–631, 2018.
- 743 [39] L. Bei, C. Shah, H. Wang, W. Huang, L. C. Platanias,
744 and E. A. Eklund, "Regulation of cdx4 gene transcription by
745 hoxa9, hoxa10, the mll-ell oncogene and shp2 during
746 leukemogenesis," *Oncogenesis*, vol. 3, no. 12, pp. e135–e135,
747 2014.
- 748 [40] M. T. Bocker, F. Tuorto, G. Raddatz, T. Musch, F.-C.
749 Yang, M. Xu, F. Lyko, and A. Breiling, "Hydroxylation of 5-
750 methylcytosine by tet2 maintains the active state of the
751 mammalian hoxa cluster," *Nature communications*, vol. 3,
752 no. 1, pp. 1–12, 2012.
- 753 [41] A. Klinakis, C. Lobry, O. Abdel-Wahab, P. Oh,
754 H. Haeno, S. Buonamici, I. van De Walle, S. Cathelin,
755 T. Trimarchi, E. Araldi *et al.*, "A novel tumour-suppressor
756 function for the notch pathway in myeloid leukaemia," *Nature*,
757 vol. 473, no. 7346, pp. 230–233, 2011.
- 758 [42] P. Chastagner, A. Israel, and C. Brou, "Aip4/itch
759 regulates notch receptor degradation in the absence of ligand,"
760 *PloS one*, vol. 3, no. 7, 2008.

- 761 [43] E. Gallagher, M. Gao, Y.-C. Liu, and M. Karin,
762 “Activation of the e3 ubiquitin ligase itch through a
763 phosphorylation-induced conformational change,” *Proceedings*
764 *of the National Academy of Sciences*, vol. 103, no. 6, pp.
765 1717–1722, 2006.
- 766 [44] R. M. de Freitas and C. M. da Costa Maranduba,
767 “Myeloproliferative neoplasms and the jak/stat signaling
768 pathway: an overview,” *Revista brasileira de hematologia e*
769 *hemoterapia*, vol. 37, no. 5, pp. 348–353, 2015.
- 770 [45] T. Bellon, D. Perrotti, and B. Calabretta, “Granulocytic
771 differentiation of normal hematopoietic precursor cells induced
772 by transcription factor pu. 1 correlates with negative regulation
773 of the c-myb promoter,” *Blood, The Journal of the American*
774 *Society of Hematology*, vol. 90, no. 5, pp. 1828–1839, 1997.
- 775 [46] W. Zhao, C. Kitidis, M. D. Fleming, H. F. Lodish, and
776 S. Ghaffari, “Erythropoietin stimulates phosphorylation and
777 activation of gata-1 via the pi3-kinase/akt signaling pathway,”
778 *Blood*, vol. 107, no. 3, pp. 907–915, 2006.
- 779 [47] M. Ichikawa, T. Asai, T. Saito, G. Yamamoto, S. Seo,
780 I. Yamazaki, T. Yamagata, K. Mitani, S. Chiba, H. Hirai *et al.*,
781 “Aml-1 is required for megakaryocytic maturation and
782 lymphocytic differentiation, but not for maintenance of
783 hematopoietic stem cells in adult hematopoiesis,” *Nature*
784 *medicine*, vol. 10, no. 3, pp. 299–304, 2004.
- 785 [48] M. Ichikawa, S. Goyama, T. Asai, M. Kawazu,
786 M. Nakagawa, M. Takeshita, S. Chiba, S. Ogawa, and

787 M. Kurokawa, “Aml1/runx1 negatively regulates quiescent
788 hematopoietic stem cells in adult hematopoiesis,” *The Journal*
789 *of Immunology*, vol. 180, no. 7, pp. 4402–4408, 2008.

790 [49] J. J. Jeong, X. Gu, J. Nie, S. Sundaravel, H. Liu, W.-L.
791 Kuo, T. D. Bhagat, K. Pradhan, J. Cao, S. Nischal *et al.*,
792 “Cytokine-regulated phosphorylation and activation of tet2 by
793 jak2 in hematopoiesis,” *Cancer discovery*, vol. 9, no. 6, pp.
794 778–795, 2019.

795 [50] F. Lohmann and J. J. Bieker, “Activation of ekf
796 expression during hematopoiesis by gata2 and smad5 prior to
797 erythroid commitment,” *Development*, vol. 135, no. 12, pp.
798 2071–2082, 2008.

799 [51] F. Cervantes, F. Passamonti, and G. Barosi, “Life
800 expectancy and prognostic factors in the classic bcr/abl-
801 negative myeloproliferative disorders,” *Leukemia*, vol. 22,
802 no. 5, pp. 905–914, 2008.

803 [52] J. C. Aster, W. S. Pear, and S. C. Blacklow, “The varied
804 roles of notch in cancer,” *Annual Review of Pathology:*
805 *Mechanisms of Disease*, vol. 12, pp. 245–275, 2017.

806 [53] F. P. Lampreia, J. G. Carmelo, and F. Anjos-Afonso,
807 “Notch signaling in the regulation of hematopoietic stem cell,”
808 *Current stem cell reports*, vol. 3, no. 3, pp. 202–209, 2017.

809

Figures and Table

Figure 1. *HOXA9* low and high expression stratifies patients in AML. (A) *HOXA9* expression in AML patients is significantly bimodal (unimodality test rejected with $p < 2.2 \times 10^{-16}$), suggesting a role as a genetic switch. (B) *HOXA9* high and low cohorts have divergent AML prognosis, consistent with known *HOXA9* biology in AML. *HOXA9* expression also partially explains (C) FAB and (D) molecular classifications of AML patients. Using Sankey diagrams, we find that the M3 FAB subtype as well as the PML-RAR α and CBF (CBFB-MYH11 and RUNX1-RUNXT1) translocations are solely linked to low expression of *HOXA9*. Similarly, the high cohort possesses specific AML subclasses: M0, M5, MLL/NUP98 translocations and complex cytogenetics. (E) *HOXA9* cohorts show distinct clinical characteristics: high cohort patients are older, display a higher white blood cell counts (WBC) and tend to have higher percentage of blasts in the bone marrow.

829

830 **Figure 2. *HOXA9* determines cancer cell fates in MPN**

831 **diseases with *JAK2* and *TET2* mutations. (A) MPN patients**

832 **with *JAK2/TET2* mutations showing different clinical**

833 **characteristics can be explained by a simple gene motif**

834 **including a switching property. The model starts from a healthy**

835 **state on the left (wild type) and sequentially acquires mutations**

836 in *JAK2* and *TET2* genes. The first mutation affects the gene
837 target expression (middle networks) which remains stable
838 when the second mutation appears. The order in which
839 mutations occur impacts on the gene target expression but
840 also the phenotypes, CMP expansion and erythroid
841 differentiation (networks on the right). Despite the identical final
842 mutational state, the cell shows different hematopoietic
843 behaviours. **(B)** *HOXA9* is downstream of both *TET2* and *JAK2*
844 and acts as a developmental switch through its self-positive
845 feedback loop. Direct interaction drawn as continuous arrows
846 and post translation modification as dotted arrows. **(C)** The
847 *JAK2/TET2/HOXA9* molecular network is built with the
848 BioModelAnalyzer (BMA) tool and integrates six phenotypes
849 considered as our model outputs. Using the Linear Temporal
850 Logic model checking tool available in BMA, we observe a
851 bifurcation after simulation of the double mutant state. The
852 bifurcation analysis identifies two stable states with different
853 phenotype values that fit the mutation order characteristics
854 observed in MPN patients with *JAK2* and *TET2* mutations. **(D)**
855 Ranking of major cancer pathways in *JAK2* analyses
856 determine the RTK-RAS pathway as the most correlated to
857 *JAK2*. This pathway contains *JAK2* and so this result is as
858 expected. The second pathway is NOTCH. The overall
859 accuracy of the pathway is computed with the Matthews
860 Correlation Coefficient. **(E)** We plot SHAP scores for the
861 NOTCH pathway to determine which genes of the pathway
862 have been important classifiers, that is genes with an important
863 expression correlation with *JAK2*. Among them, *PSEN1* and

864 *ADAM10* are involved in the cleavage of NOTCH membrane
 865 receptors. *NUMB* activates *ITCH* which degrades NOTCH. **(F)**
 866 The heatmap of the NOTCH pathway and HOXA family
 867 generated using MPN microarray datasets from [23] validate
 868 NOTCH expression in our model highlighting the importance of
 869 this pathway in MPN disease progression. HOXA heatmap
 870 confirms HOXA bimodality but show different levels of
 871 expression to what we find in our model. "JAK2" and "TET2"
 872 refer to the single mutant mouse models, and "DM" is the
 873 double mutant with *JAK2* mutated first. "WT" designates the
 874 wild type (no mutation) genotype.
 875
 876

877

	WT	<i>TET2</i>	<i>JAK2</i>	<i>TET2</i> first	<i>JAK2</i> first
Stem Cell Renewal	1	2	1	2	2
CMP expansion	1	2	1	2	1
GMP expansion	1	2	2	2	2
GMP differentiation	1	0	1	1	1
Erythroid differentiation	1	0	2	1	2
MEP expansion	1	1	2	2	2

878

879

880 **Table 1. Specification table for *JAK2/TET2* BMA model.** In

881 order from the left to the right columns, the specifications are

882 the wild type state, the *TET2* single mutant, the *JAK2* single

883 mutant and finally the double mutants, which consists of a

884 bifurcation with two state attractors that represent the case

885 where *TET2* is mutated before *JAK2* (*TET2* first) and the

886 alternative case where *JAK2* is mutated first. We determine

887 phenotype values using literature for the single mutants and

888 Ortmann et al [13] for the double mutants. The value 1

889 represents the healthy state, 0 the lowered/inactive state and 2

890 the overactive state.

891