Supplemental Methods:

ChIP-seq peak calling:

Peak calling was performed using MACS2 [1,2]. Broad peaks were called to obtain consensus peaks per cell type. QPD and control ChIP datasets were combined per histone mark, per cell-type, and paired with cell-type specific inputs, as below:

```
macs2 callpeak –t QPD_cellT_chip_pool.bam CTL_cellT_chip_pool.bam\
–c QPD_cellT_input_pool.bam CTL_cellT_input_pool.bam -g hs \
-q 0.005 --fe-cutoff 2.0 –broad –max-gap 250 –min-length 300
```

Narrow peaks were called for ranking comparison of our H3K27ac datasets with those from Roadmap Epigenomics (described below). QPD, control, and consensus peak sets were generated per cell type:

```
macs2 callpeak –t QPD_cellT_H3K27ac_pool.bam –c
QPD_cellT_input_pool.bam\
-q 0.005 --fe-cutoff 4.0
```

Fold-enrichment signal tracks were generated from pooled biological replicates in bedgraph format with the MACS2 bdgcmp utility using the lambda.bdg file generated per pooled replicate as a control.

```
macs2 callpeak –t QPD_cellT_chip_pool.bam –c \
QPD_cellT_input_pool.bam --bdg
macs2 bdgcmp –t treat_pileup.bdg –c control_lambda.bdg -m FE
```

Tracks were visualized using the WashU epigenome browser (http://epigenomegateway.wustl.edu/legacy/).

To compare H3K27ac peaks between granulocytes and megakaryocytes, we first obtained a consensus peak set by merging granulocyte and megakaryocyte H3K27ac broad peaks using bedtools merge. Counts per peak were obtained for each individual sample and differential analysis was performed for peaks between megakaryocytes and granulocytes using DESeq2 v1.24 [3]. Counts from QPD samples were divided by 1.5 to account for the extra copy-gain for peaks overlapping the duplicated region (chr10:75659017-75736956). Raw p-values (Wald test) were corrected for the number of peaks within the duplicated region to obtain a locus p-value (reported in text) using the p.adjust function in R with "method=fdr".

H3K27ac ranking analysis

Comparison of H3K27ac peak rankings was performed using MACS2 narrowPeak files either generated as described above (for study samples) or downloaded from Roadmap Epigenomics [4]. Roadmap peaks were further filtered for criteria q-value < 0.005 and fold-enrichment value > 4 to match our called peaks. Peaks within each file were ranked based on the reported fold-enrichment value, which corresponds to the 7th column of the narrowPeak file. The equivalent of ENH_{QPD} was selected as the peak

overlapping the megakaryocyte and granulocyte consensus broad region (chr10:7571516-75719950). In case of multiple overlapping peaks, the highest fold-enrichment value was used; in case of no overlapping peaks, a value of 0 was used.

Accessions for published datasets used in this study

H3K27ac peaks in narrowPeak format were downloaded from Roadmap Epigenomics: https://egg2.wustl.edu/roadmap/data/byFileType/peaks/consolidated/narrowPeak/

ENCODE K562 CTCF ChIP-seq and CTCF ChIA-PET datasets utilized those pre-loaded in the WashU epigenome browser. They correspond to accessions ENCFF933ZLL and ENCFF000KYD, respectively.

K562 and IMR90 Hi-C datasets reported by [5] were visualized as contact matrices using Juicebox [6]:

Human megakaryocyte transcription factor ChIP-seq datasets reported by [7] were obtained from GEO: GSE24674

Mouse CD41 progenitor ChIP-seq datasets reported by [8] were obtained from GEO: GSE69101

Human megakaryocyte promoter capture Hi-C reported by [9] were obtained as pre-processed interactions from Supplemental data files (Data S1).
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5123897/bin/mmc4.zip

Primer Sequences:

4C-seq primers:

pPLAU_reading TCCCTACACGACGCTCTTCCGATCT **CCTCTCCCCTGGTGCTGATC**

pPLAU_nonread GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC **TTTCTCCTGGCTGGAAACCC**

ENH_{QPD}_reading TCCCTACACGACGCTCTTCCGATCT **AGAGAGGTACTATTAGGATC**

ENH_{QPD}_nonread GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC **TTATGTCAGGTCCTTAGGT**

4C-Sanger primers:

rs1916341_fwd: AAGGAAGGAGAAGTCAGGGCA

rs1916341_rev: AGAGAGGTACTATTAGGATCCAGGC

rs2227574_fwd: CATGACCTGTGACCAGCACT

rs2227574_rev: AGAGAGGTACTATTAGGATCACTTTTATT

Cloning primers:

ENH_{QPD_CONS}_FP: CCCCCGGGTTATGTCAGGTCC

ENH_{QPD_CONS}_RP: CCTCTGAGTGGAAGGGACGG

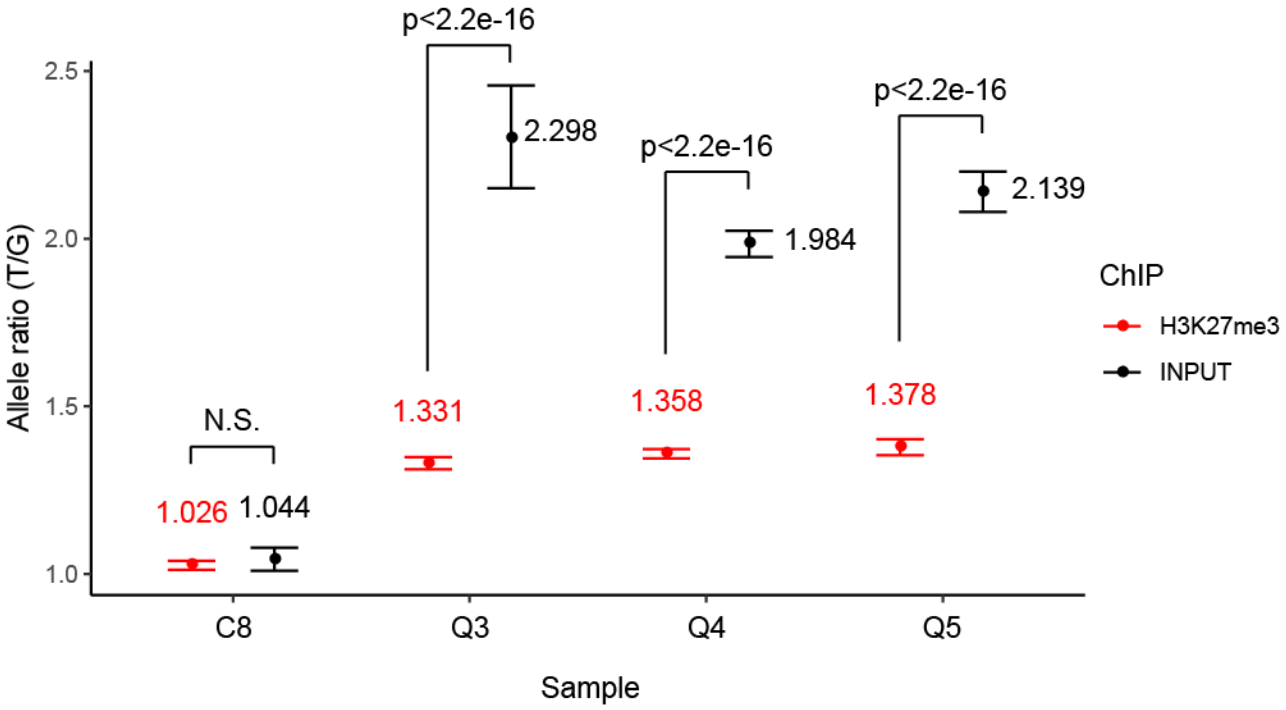
pPLAU_gDNA_F: AGTCTTAGGCAAGTTGGGGC

pPLAU_gDNA_R: ATCTCAGGACCGCGGCA

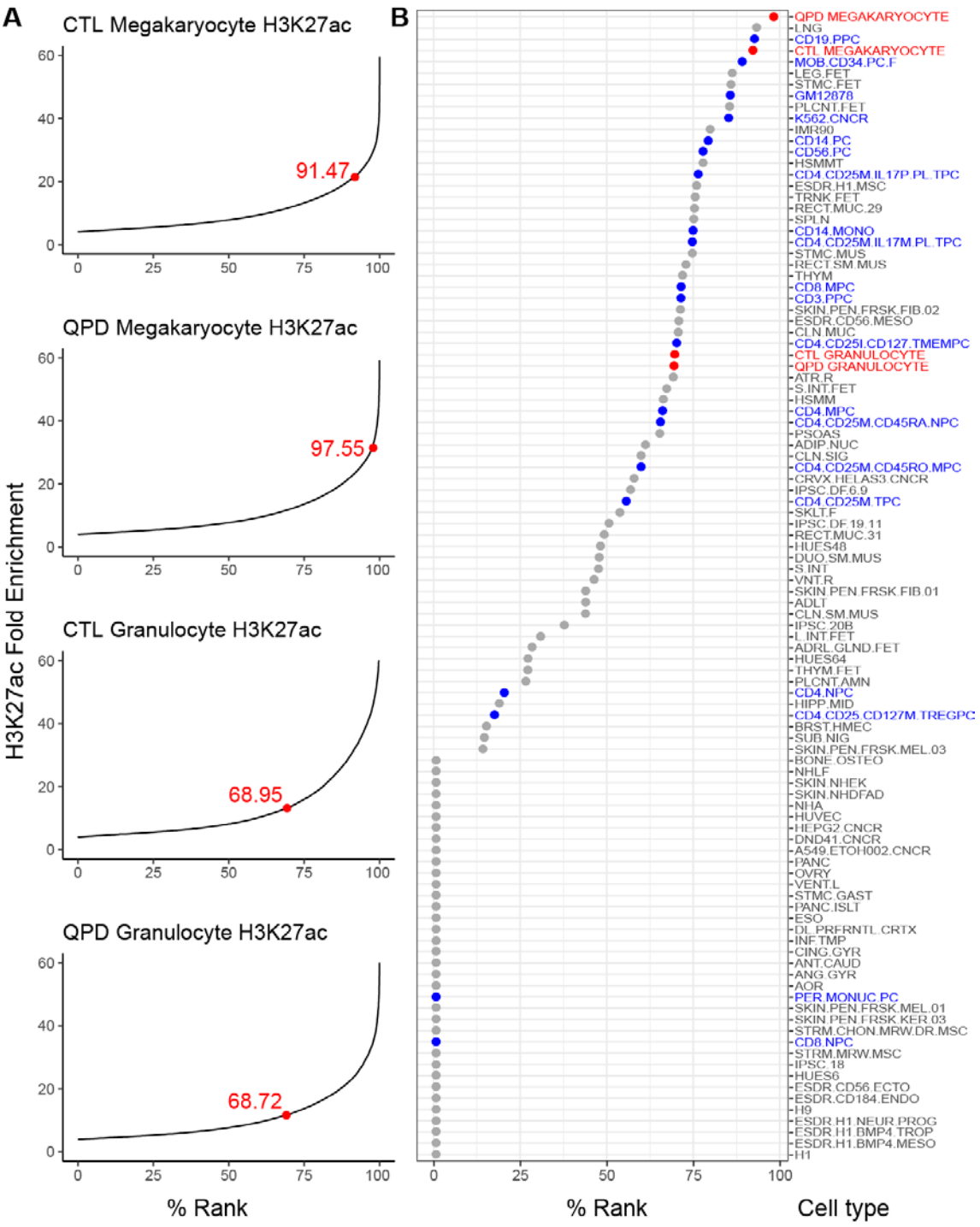
pVCL_gDNA_F: TGAGGGTTTCGTGTGAAGGG

pVCL_gDNA_R: GTGCGCGTATGAAACACTGG

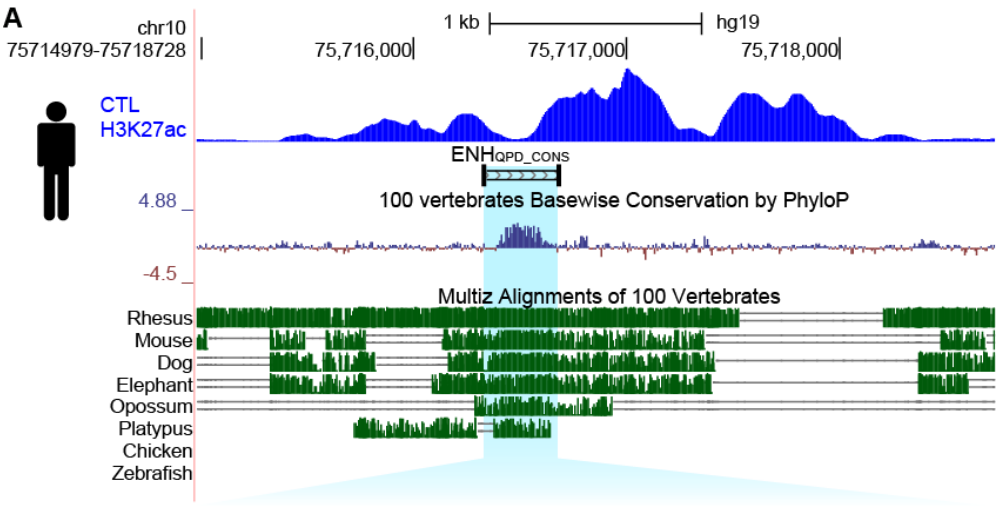
Supplemental Figures:



Supplemental Figure 1: Digital Droplet PCR analysis of rs1916341 allele ratios in H3K27me3 ChIP and sonication Input libraries. The analysis was undertaken to determine if there was allele bias in repression of *PLAU* in QPD by evaluating H3K27me3 enrichment at *PLAU*. Data compare the allele ratios in sonication input libraries compared to H3K27me3 ChIP libraries for QPD participants that were heterozygous at rs1916341 (each with two copies of the T allele on the QPD chromosome and one copy of the G allele on the other chromosome) and one heterozygous control participant (one copy of each of the T and G alleles). Allele ratios were calculated by dividing the number of copies of the T allele by number of copies of the G allele. Error bars show 95% Poisson confidence intervals, as calculated using the poisson.test() function in R. Fisher's exact test was used to assess significance between ratio of counts observed in H3K27me3 ChIP versus Input libraries per sample. Raw count data are provided in **Supplemental Table 1.**



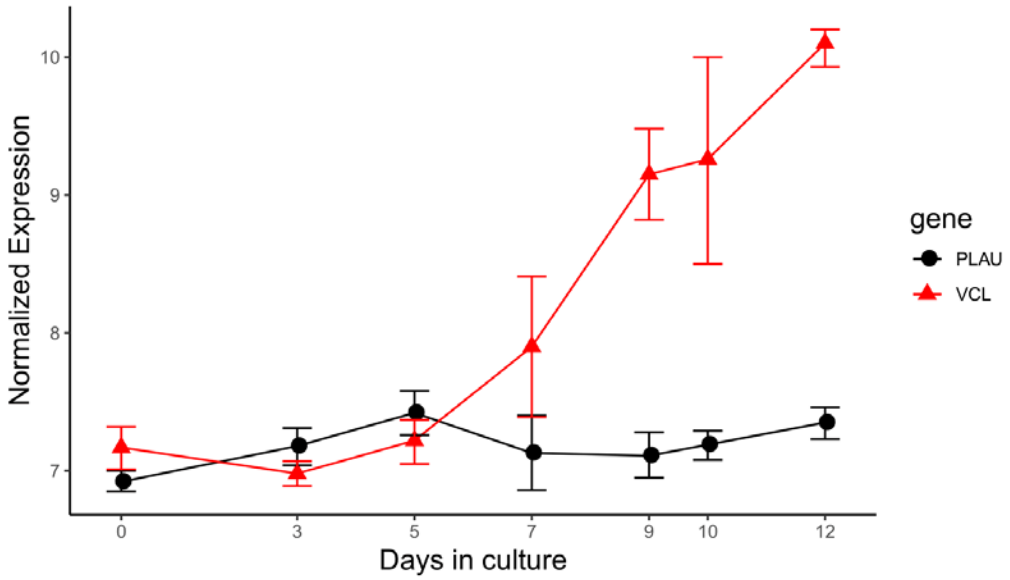
Supplemental Figure 2: Ranking analysis of H3K27ac enrichment at ENH_{QPD}. a, Ranking of H3K27ac peaks (a marker of active enhancers) in QPD and control granulocytes and megakaryocytes. Y-axis shows percent ranking of H3K27ac peaks, ranked by H3K27ac fold-enrichment score. ENH_{QPD} is depicted by the red dot. **b,** Comparison of H3K27ac percent rank score at ENH_{QPD} in QPD and control megakaryocytes and tissues profiled in Roadmap Epigenomics. Cells of hematopoietic origin are colored blue. This figure illustrates the high enrichment of H3K27ac at ENH_{QPD} in megakaryocytes relative to other cells and tissues.



B

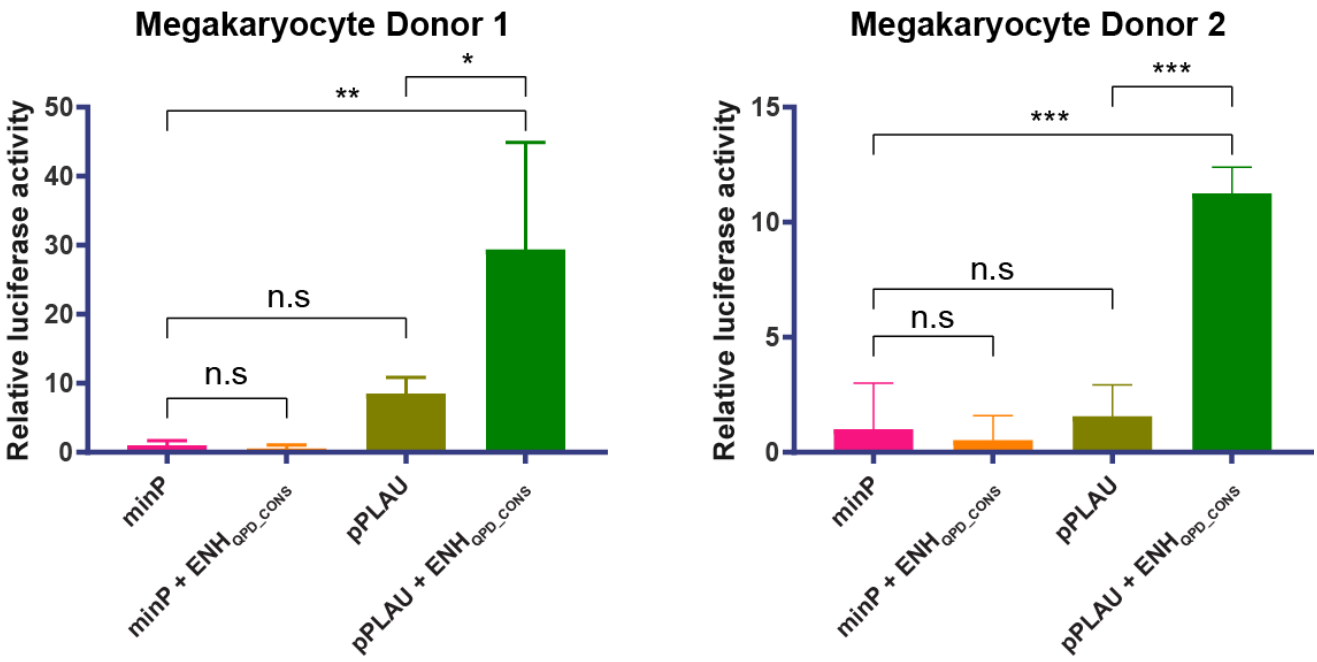
```
Human       cccccgggttatgtcaggtccttaggtggaactctgagtcccacatcactaaatgctgac
Mouse       cctctgggctaagccagagccccaggcagaactctgaagcctatagcact---------c
Dog         cctctgg---------------ttagcagcacttccaagctcacacgactaaatgctatc
            **.*.**               . .*..* ***.. *. *..*.*  ***         *

Human       atgcccgcagaatccaggaggctcc-agtgactagaaagaagtttcatcaaatt-tgaac
Mouse       accctcacagtacccacaaggctc--agtgactagaa---agtttcatcgaattctgaac
Dog         a--cttgcagaatccatgaagctccaagtgactagaaaggagtttcatcgaattctgaac
            *  *...*** *.*** .*.****  ***********   *********.**** *****

Human       aggaagcagcagtggcagctgataatctgtgacctgaatatagctatccattatctcact
Mouse       aggaagcagcagtggcagctgataatctgtgacctgaatatagctatccattatctcact
Dog         aggaagcagcaatggcagctgataatctgtgacctgaatatagccatccattatctcact
            ***********.*********************************.***************

Human       gtggtttctggtttccctgacagagaaaacgcccacaactttcttaaagttattctaaac
Mouse       gtggtttcttgtttccctgacagaggaaacgcccacgactttcttgaagttattctaaac
Dog         gtggtttcttgtttctctgacagaggaaacacccacgactttcttgaagttattctaaac
            ********* *****.*********.****.*****.********.**************

Human       actcagttccttgtttggctttggcaggataatgttacccctgggtgcccctgtgctttg
Mouse       actcagttccttgttcggctctggcaggataatgtttcccctgggtgccccggtgttttg
Dog         acccagtttcttgtttagctctggcaggataatgttacccctgggcactcctgtgctttg
            **.*****.******..***.*************** ********..*.** ***.****

Human       ggctgtggcctccctgccagcctgaggctccagctcttcagccaatctgtgggg-caacc
Mouse       agctgcagcctccctgccagcctgaggctcaagctcagcggtcaacctgtggggccaact
Dog         ggctatggcctccttgccttcctgggctcccgggtgagcagccaacctgtgggg-taact
            .***...******.****  ****.* ..* .* *   *.*.***.******** .***.

Human       gtcccttccactcagagg
Mouse       gtcccctccacggaggag
Dog         gtcctctcatcttggtgg
            ****..**  *  .* .*
```

Supplemental Figure 3: Phylogenetic analysis of ENH_{QPD}. a, Genome browser view of ENH_{QPD} in human. Basewise PhyloP conservation scores[10] and Multiz multiple sequence alignments of 100 vertebrates [11] are shown for select species. Striped bar marks the position of ENH_{QPD_CONS}. **b,** multiple sequence alignment of ENH_{QPD_CONS} between human, mouse, and dog. MSA sequences from the Multiz multiple sequence alignments of 100 vertebrates were downloaded with the UCSC table browser using the hg19 coordinates chr10:75716316 -75716690. Extracted sequences were realigned using MAFFT with method = L-INS-i [12,13]. Pairwise percent sequence identity (reported in main text) was calculated using BLASTn [14].

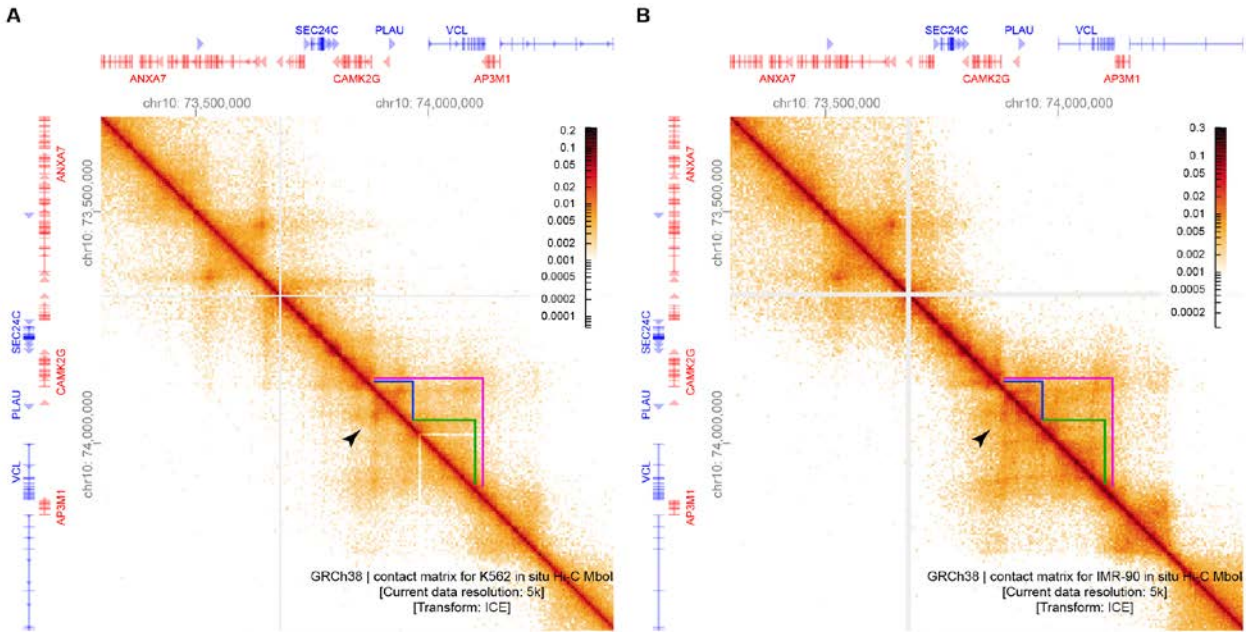


Supplemental Figure 4: *PLAU* and *VCL* expression during normal megakaryopoiesis.
Normalized microarray gene expression of *PLAU* (black) and *VCL* (red) in developing cord blood-derived cultured megakaryocytes at days 0 (CD34+ progenitors), 3, 5, 7, 9, 10, and 12 from n=3 samples per time point. Data was retrieved using the "Expressed!" tool from the Haemgen Tools portal: https://haemgen.haem.cam.ac.uk/expressed/

The figure illustrates that *PLAU* and *VCL* expression are uncoupled during normal megakaryopoiesis. *VCL* expression normally increases during megakaryopoiesis whereas *PLAU* expression does not.



Supplemental Figure 5: Luciferase reporter assay for ENH_{QPD} in megakaryocyte cultures. Relative luciferase activity for megakaryocyte cultures from n=2 biological replicates (separate donors and transfections) for minimal (minP) and *PLAU* (p*PLAU*) promoter constructs, with or without ENH_{QPD_CONS}. Each panel shows measurements from 3 pipetting replicates per construct from a single biological replicate. All values are normalized relative to minP in their respective cell type. Statistical analysis was performed using ANOVA one-way with Tukey correction. Asterisks immediately above data points denote significance compared to minP. Error bars show standard deviation of mean. Asterisks above denote significance of select pairwise comparisons. ***, p>0.001; **, p>0.01; n.s., not significant.



Supplemental Figure 6: Hi-C interactions at the *PLAU* locus. *In-situ* Hi-C contact matrices are shown for published K562 (A) and IMR90 (B) datasets [5]. This figure illustrates TAD and nested subTAD structures at the *PLAU* locus. Triangles mark the approximate positions of subTAD$_{PLAU}$ (blue), subTAD$_{VCL}$ (green), and the larger *PLAU* locus TAD (magenta). Arrows mark the position of corner-dots (looping interactions) corresponding to the boundaries of subTAD$_{PLAU}$ that are present in K562 but diminished in IMR90. Preprocessed data were obtained and visualized from the 4D Nucleome webportal corresponding to experiments 4DNESI7DEJTM and 4DNES1ZEJNRU.

References

1. **Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9(9):R137.**
2. **Gaspar JM. Improved peak-calling with MACS2. *BioRxiv*. 2018;**
3. **Love MI, Huber W, Anders S. Moderated estimation of fold change and' ' dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.**
4. **Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518(7539):317–330.**
5. **Rao SSP, Huntley MH, Durand NC, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*. 2014;159(7):1665–1680.**
6. **Robinson JT, Turner D, Durand NC, et al. Juicebox.js Provides a Cloud-Based Visualization System for Hi-C Data. *Cell Syst.* 2018;6(2):256–258.e1.**
7. **Tijssen MR, Cvejic A, Joshi A, et al. Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Dev. Cell*. 2011;20(5):597–609.**
8. **Goode DK, Obier N, Vijayabaskar MS, et al. Dynamic gene regulatory networks drive hematopoietic specification and differentiation. *Dev. Cell*. 2016;36(5):572–587.**
9. **Javierre BM, Burren OS, Wilder SP, et al. Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell*. 2016;167(5):1369–1384.e19.**
10. **Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* 2010;20(1):110–121.**
11. **Blanchette M, Kent WJ, Riemer C, et al. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* 2004;14(4):708–715.**
12. **Katoh K, Kuma K, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 2005;33(2):511–518.**
13. **Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinformatics*. 2017;bbx108.**
14. **Zhang Z, Schwartz S, Wagner L, Miller W. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 2000;7(1-2):203–214.**