S1 Figure

A) Venn diagrams showing the overlap of mESC, Day 7 EB CD71+, and Day 7 EB CD71+ Ter119+ with E10.5 blood.

B) Heatmaps showing the expression levels of α-globin and β-globin in mESC, EB day 7 CD71+, and EB day 7 CD71+ Ter119+.

C) ATAC-seq PCA plot showing the distribution of Day 7 EB CD71+ and E10.5 blood.

D) Venn diagrams showing the overlap of CTCF, Day 7 EB CD71+, E10.5 blood, and APH spleen Ter119+.

| Condition          | Day 7 EB CD71+ | E10.5 blood | APH spleen Ter119+ |
|--------------------|----------------|-------------|--------------------|
| mESC               | 38,856         | 29,100      | 27,562             |
| Fetal Liver E12.5  | 20,100         | 32,435      | 25,475             |
| E9.5               |                |             |                    |
| E10.5              |                |             |                    |
| E11.5              |                |             |                    |
| Total EB CD71+     |                |             |                    |
| E10.5 blood        | 27,562         | 25,475      | 38,383             |
| APH spleen Ter119+ |                |             |                    |
| EB-only            | 6,658          | 2,324       | 9,810              |
| E10.5-only         |                | 24,464      |                    |
| Spleen-only        |                |             |                    |
| Intersection       |                |             |                    |

| Condition          | Day 7 EB CD71+ | E10.5 blood | APH spleen Ter119+ |
|--------------------|----------------|-------------|--------------------|
| mESC               | 11,147         | 8,591       | 9,430              |
| Fetal Liver E12.5  |                | 8,591       | 8,394              |
| E9.5               |                |             |                    |
| E10.5              |                |             |                    |
| E11.5              |                |             |                    |
| Total EB CD71+     |                |             |                    |
| E10.5 blood        | 9,430          | 8,394       | 18,366             |
| APH spleen Ter119+ |                |             |                    |
| EB-only            | 1,717          | 14          | 8,660              |
| E10.5-only         |                | 8,329       |                    |
| Spleen-only        |                |             |                    |
| Intersection       |                |             |                    |
S1 Fig: More detailed ATAC-Seq analyses of CD71 sub-populations (CD71+ S1, CD71+/Ter119+ S3) and CTCF and H3K4me3 ChIP-Seq data confirm the primitive nature of these populations.

A) Venn diagrams of overlapping accessible regions between tissue datasets. Peaks were included in the analysis (if present) in data from three biological replicate tracks. Left: mESCs vs EB-derived cells; Right: EB-derived, primitive and definitive erythroid cells.

B) RPKM-normalised ATAC-seq track averaged for three replicates of double positive CD71+Ter119+ cells compared to data from total CD71+ populations, both derived from day 7 EBs.

C) PCA plot of genome-wide ATAC-seq peaks, as in Figure 2B, with addition of primitive E9.5, E10.5 and E10.5 blood, the EB-derived CD71+ (S1) and CD71+/Ter119+ (S3) populations.

D) Venn diagrams of overlapping ChIP-seq peaks between tissue datasets. Peaks included in the analysis are from three biological replicate tracks for APH Ter119+ spleen and E10.5 blood and duplicates from EB day 7 CD71+ population. Upper diagram: CTCF ChIP-seq peaks with tables containing the detailed dataset from all three cell types. Lower diagram: H3K4me3 ChIP-seq peaks presented similarly as above.