PNAS
www.pnas.org

Supplementary Information for

Unbiased classification of mosquito blood cells by single-cell genomics and high-content imaging

Maiara S. Severo, Jonathan J.M. Landry, Randall L. Lindquist, Christian Goosmann, Volker Brinkmann, Paul Collier, Anja E. Hauser, Vladimir Benes, Johan Henriksson, Sarah A. Teichmann, Elena A. Levashina

Elena A. Levashina
Email: levashina@mpiib-berlin.mpg.de

This PDF file includes:

Figs. S1 to S6
Tables S1 and S2

Other supplementary materials for this manuscript include the following:

Datasets S1 to S7

www.pnas.org/cgi/doi/10.1073/pnas.1803062115

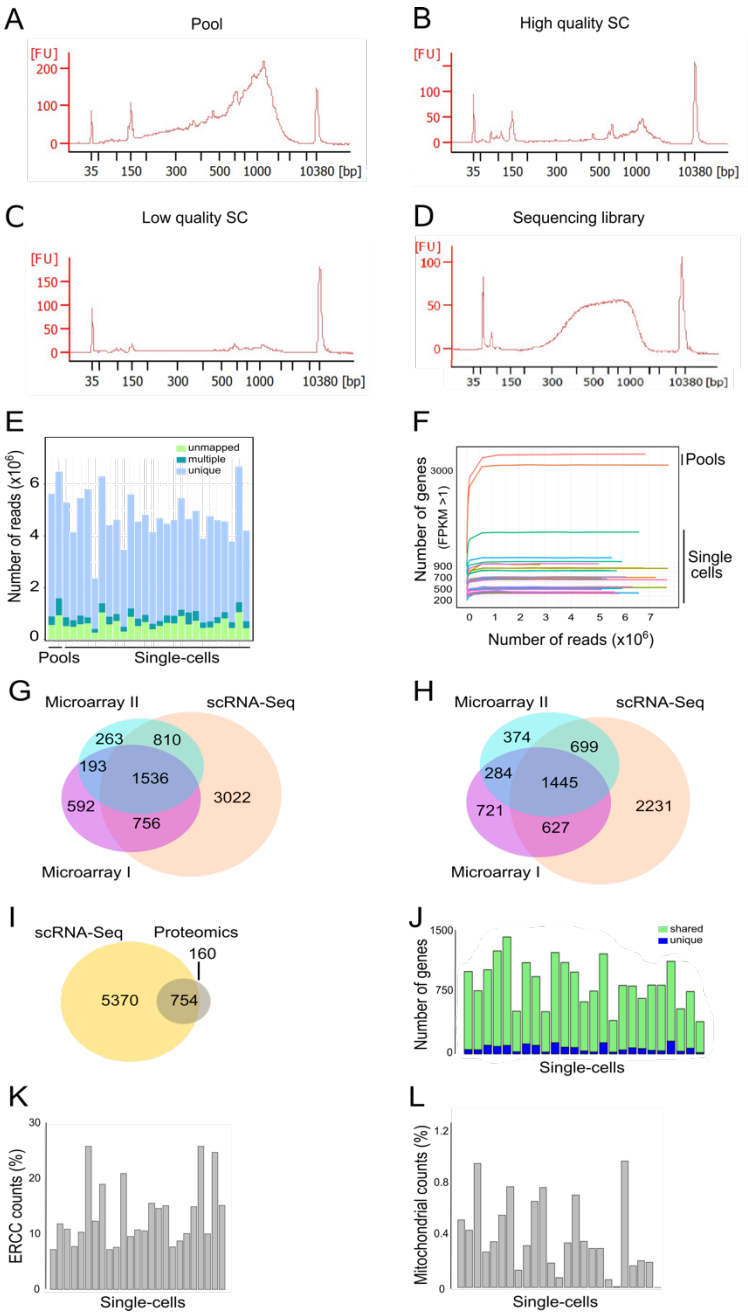

Figure S1: Quality assessment of scRNA-seq of mosquito blood cells. Related to Figure 1. (A-D) Representative electropherograms (Bioanalyzer) of pre-amplified cDNAs and cDNA library from *A. gambiae* RFP-positive hemocytes. (A) Representative graph of the cDNA size distribution obtained from a successful cDNA preamplification from 30 RFP-positive hemocytes (pool). (B-C) Representative graphs of a (B) high and (C) low quality cDNA profile from a single RFP-positive hemocyte. (D) Representative DNA library. (E) Read mapping distribution of sequences obtained with our scRNA-seq approach. (F) Saturation curves generated by randomly selecting a given number of raw reads from each library. Comparison of genes detected (normalized count ≥1) in our scRNA-seq and previous (G-H) microarray and (I)

proteomics studies. In (H), only genes identified in the pool samples were used for comparison. (J) Distribution of unique and shared genes across single-cell samples. Percentages of (I) ERCCs and (J) mitochondrial counts in single-cells. Microarray I: Dataset from Baton et al, 2009. Microarray II: Dataset from Pinto et al, 2009.

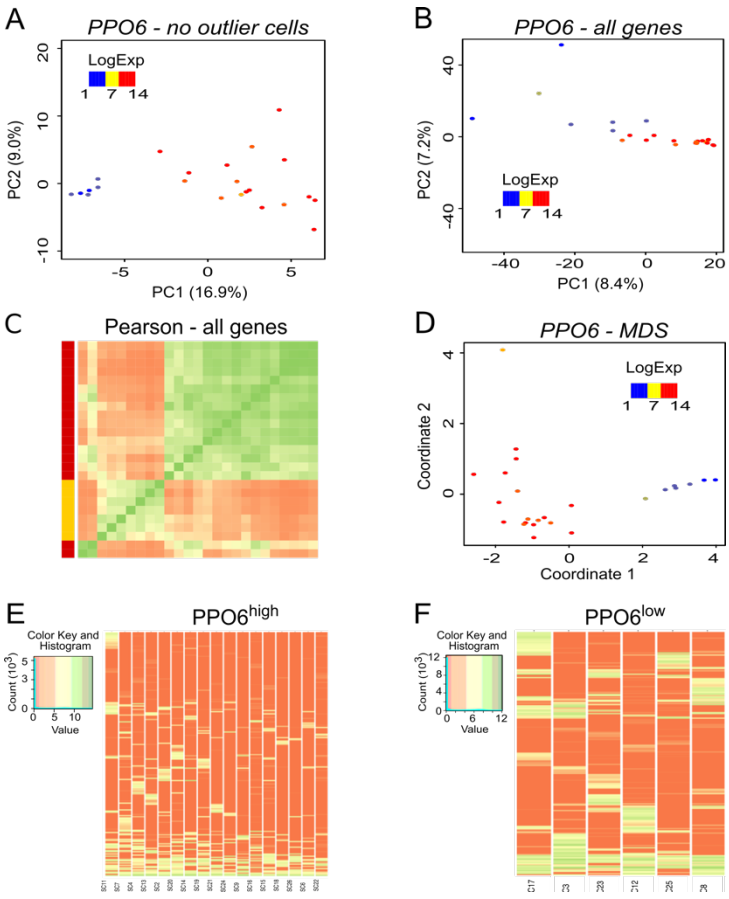

Figure S2: PPO6 cell populations exhibit intra-population heterogeneity. PCA plot based on the expression of (A) highly variable genes in the absence of the outlier cells identified in Figure 2C. A PCA plot (B) and a Pearson correlation heatmap (C) based on all identified genes are shown. Color bar indicates the groups identified in Figure 2. (D) A MDS plot based on the expression of the *PPO* genes identified in the scRNA-Seq. The first two principal components or coordinates are shown, and each point represents one single hemocyte. *PPO6* expression, as log10 (normalized counts +1), is overlaid onto the PCA and MDS plots. Heatmap of genes identified in at least one cell in either PPO6^{high} (E) or PPO6^{low} (F) cell populations.

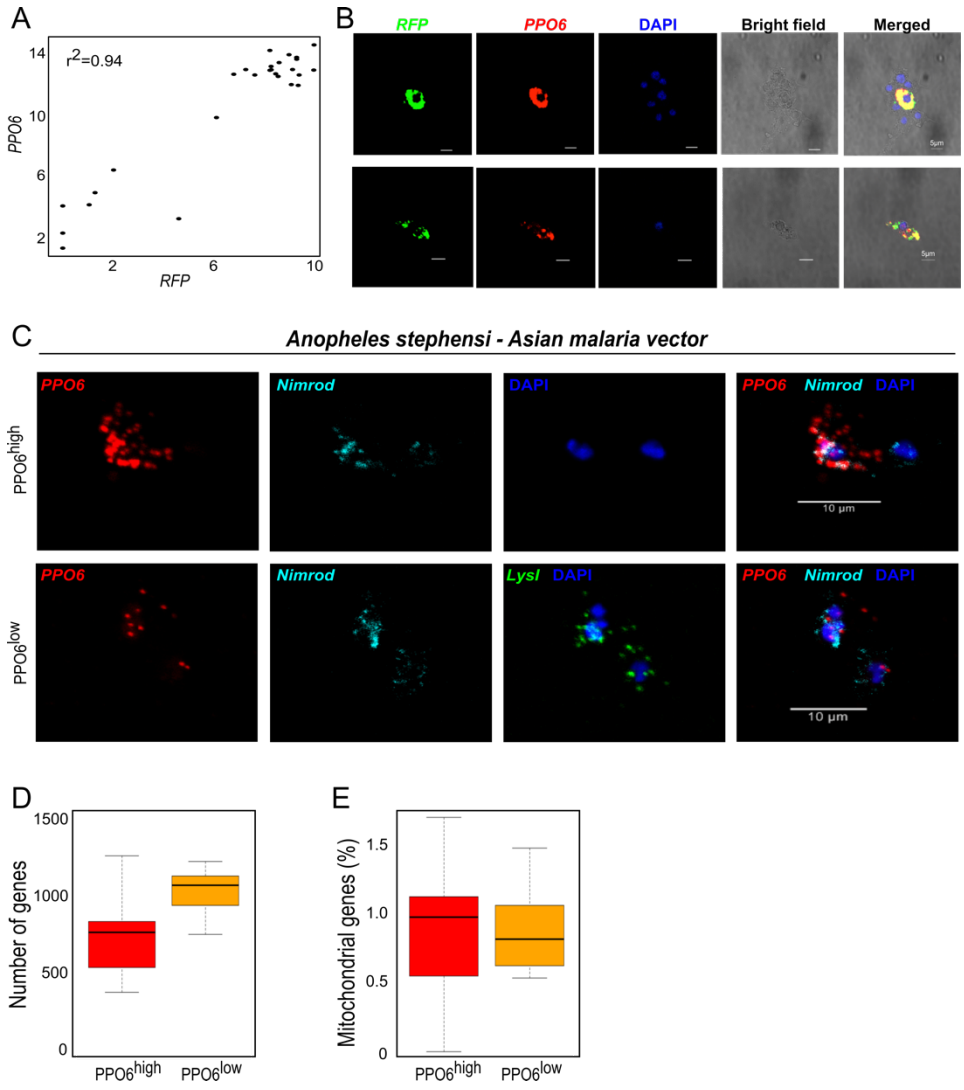

Figure S3: Gene expression analyses of blood cells from *PPO6::RFP* and *Anopheles stephensi* mosquitoes, and transcriptional characteristics of PPO6 cell subpopulations. Related to Figures 2 and 3. (A) Scatter plot of expression of *RFP* and *PPO6* (log normalized counts + 1) in single-cells. r^2 indicates the Pearson correlation. (B) RNA-FISH validation of *RFP* expression in *PPO6::RFP* mosquitoes. (C) Similar patterns of expression of the identified cell-type markers for PPO6 populations are observed in the Asian mosquito vector, *A. stephensi.* Scale bars: (B) 5 µm, (C) 10 µm. DNA is stained with DAPI. (D) Total number of genes and (E) percentage of mitochondrial genes in $PPO6^{high}$ and $PPO6^{low}$ populations. A gene was considered present when at least one count was measured in at least one cell. Boxplots indicate the median, first and third quartile, and min and max values.

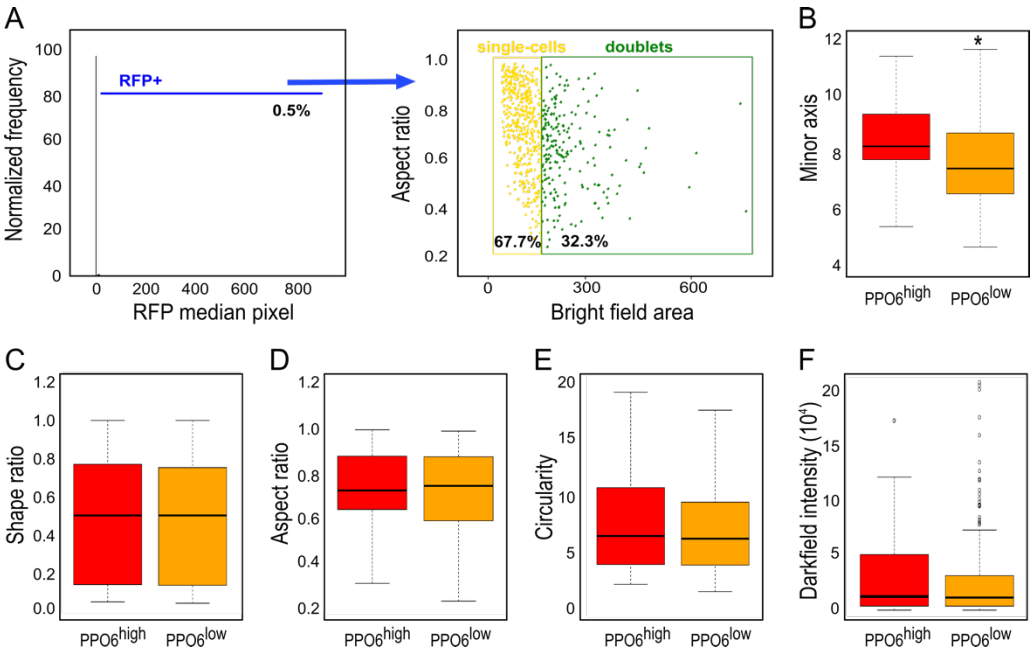

Figure S4: Imaging flow cytometry analyses of PPO6^{high} and PPO6^{low} cell subpopulations. Related to Figure 4. (A) Flow cytometry gating approach to identify RFP-positive single blood cells. Numbers indicate the percentages of events per gate (n=130,000). Single-cells and doublets were identified after initial gating based on RFP median pixel (blue arrow). (B-F) Boxplots showing distribution of morphological measurements according to the groups. Asterisks represent p <0.01 based on Mann–Whitney–Wilcoxon comparisons. Boxplots indicate the median, first and third quartile, and min and max values.

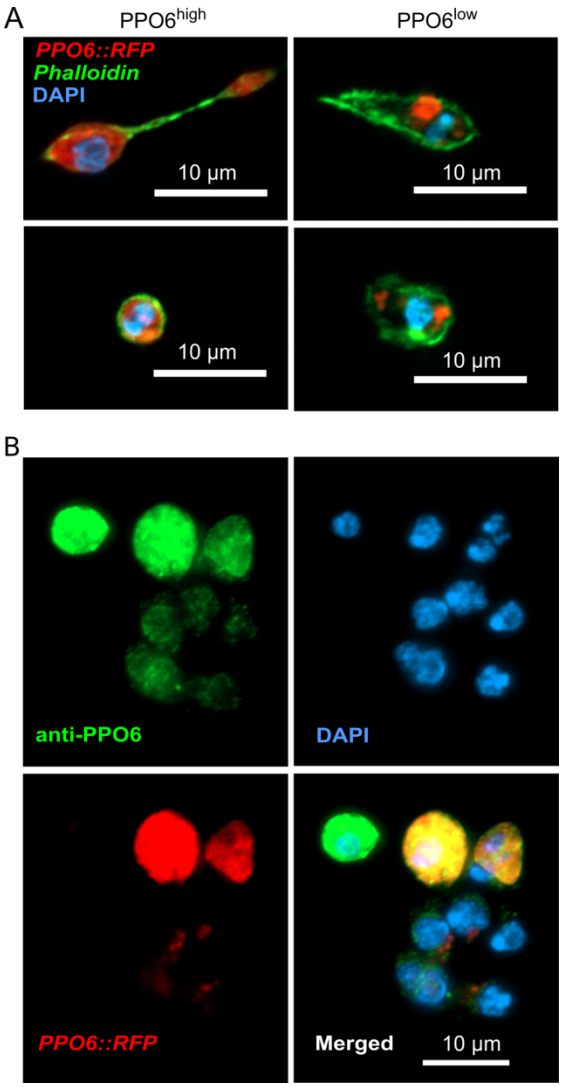

Figure S5: Characterization of PPO6 subpopulations by fluorescence microscopy upon cell attachment. RFP-expressing hemocytes (red) were obtained by perfusion of female adults and allowed to attach onto microscopy glasses. Cells were classified as PPO6^{high} (left panel) or PPO6^{low} (right panel) according to the expression of RFP (red). (A) PPO6^{high} and PPO6^{low} cell subpopulations exhibit different morphologies across the groups. Phalloidin staining indicates attachment foci and pseudopodia. (B) Cells were stained with anti-PPO6 antibodies for comparison to the RFP expression. A cell can display PPO6 staining but lack RFP expression. Scale bars: 10 µm. DNA is stained with DAPI. Representative images of 2 independent experiments are shown.

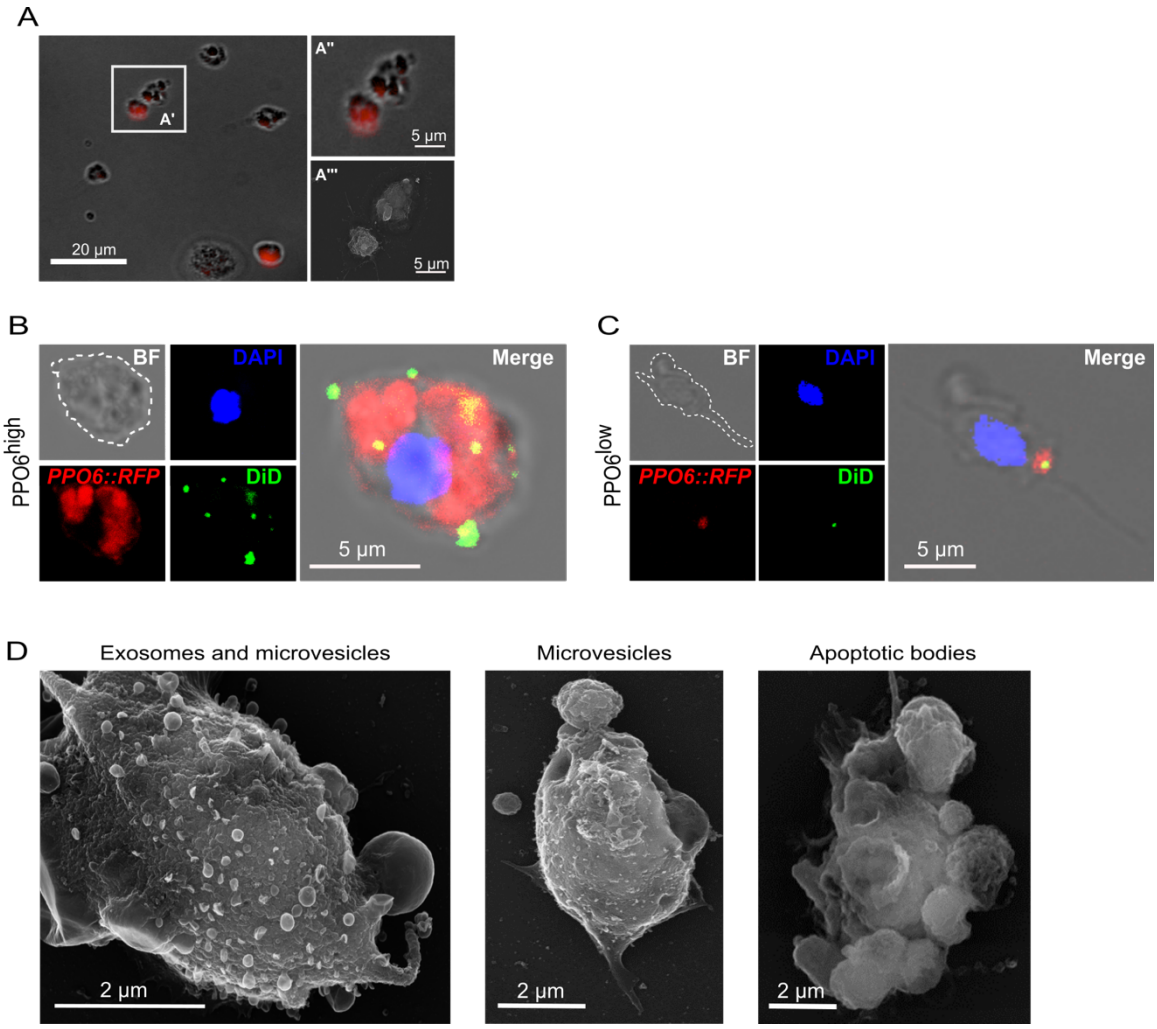

Figure S6: Vesicular structures in mosquito hemocytes. Related to Figure 5. (A-A'') Correlative fluorescence and scanning electron microscopy (SEM) of RFP-expressing cells obtained by hemolymph perfusion. (B) PPO6^{high} and (C) PPO6^{low} cell subpopulations display DiD fluorescence suggestive of vesicles. (D) Representative SEM images of mosquito blood cells displaying exosomes, microvesicles and apoptotic bodies. Scale bars: (A) 20 µm, (A''-C) 5 µm, (D) 2 µm. DNA is stained with DAPI.

Table S1: Alignment statistics for all sequenced libraries.

Well	Sample	Total number of reads	Number of mapped reads	Unique mapping reads	Multiple mapping reads	Unmapped reads
A1	**Pool 1**	5615688	5060145	4729439	330706	555543
A2	**Pool 2**	6481103	5553377	4897255	656122	927726
A4	**SC1**	5274357	4761977	4421679	340298	512380
A5	**SC2**	4138770	3653831	3400144	253687	484939
A6	**SC3**	5448767	4863893	4528385	335508	584874
A7	**SC4**	5784480	5165299	4933024	232275	619181
B8	**SC5**	2357650	2099445	1954616	144829	258205
C1	**SC6**	6285191	5246147	4894194	351953	1039044
C3	**SC7**	4401499	3822267	3519586	302681	579232
C7	**SC8**	4620786	3909509	3641622	267887	711277
C8	**SC9**	3459264	3159000	2974689	184311	300264
D1	**SC10**	5595460	4723182	4395667	327515	872278
D3	**SC11**	4546394	4032611	3744325	288286	513783
D4	**SC12**	4814842	4147753	3866959	280794	667089
D5	**SC13**	4129338	3740817	3526535	214282	388521
D9	**SC14**	4664492	4144974	3955985	188989	519518
E4	**SC15**	4456352	3820728	3531385	289343	635624
F10	**SC16**	5444609	4545799	4297554	248245	898810
F1	**SC17**	4606035	3991175	3693107	298068	614860
F3	**SC18**	4647533	4070486	3607177	463309	577047
F5	**SC19**	4961559	4244663	3866505	378158	716896
F6	**SC20**	3880894	3440424	3228198	212226	440470
F7	**SC21**	4754095	4297398	4059186	238212	456697
F9	**SC22**	4602225	3978043	3745810	232233	624182
G1	**SC23**	4553910	3744282	3501359	242923	809628
G3	**SC24**	3776579	3339027	3115424	223603	437552
G4	**SC25**	6662345	5611870	5226449	385421	1050475
H4	**SC26**	4202565	3771295	3502259	269036	431270

Table S2: Fisher's discriminant factor results obtained from the comparison between all parameters analyzed for PPO6^{high} and PPO6^{low} blood cell populations.

Features	RD Mean (ppo_high,ppo_low)
Width RFP	1.7
Minor Axis RFP	1.69
Thickness Maximum RFP	1.63
Modulation RFP	1.4
Median Pixel RFP	1.24
Mean Pixel RFP	1.14
Intensity RFP	1.11
Bright Detail Intensity R7 RFP	1.08
Standard Deviation RFP	1.03
Circularity RFP	0.92
Symmetry 3 RFP	0.41
Symmetry 4 RFP	0.45
Bright Detail Similarity R3 RFP	0.18
Delta Centroid XY RFP Intensity Weighted	0.16
Saturation Count BF	0
Saturation Percent BF	0
Diameter Adaptive Erode BF	0
Spot Area Minimum Adaptive Erode BF	0

Additional datasets (separate Excel files)

Dataset S1: Mosquito blood cell transcriptome based on scRNA-seq of pools and single hemocytes.

Dataset S2: Genes expressed in at least 90% of the single-cells.

Dataset S3: Gene ontologies for genes identified in all samples (Dataset S2).

Dataset S4: Highly variable genes identified in the single-cells based on technical noise estimation.

Dataset S5: Gene ontologies for PPO6^{high} blood cell population.

Dataset S6: Gene ontologies for PPO6^{low} blood cell population.

Dataset S7: Morphological parameters analyzed for the 319 cells using imaging flow cytometry.