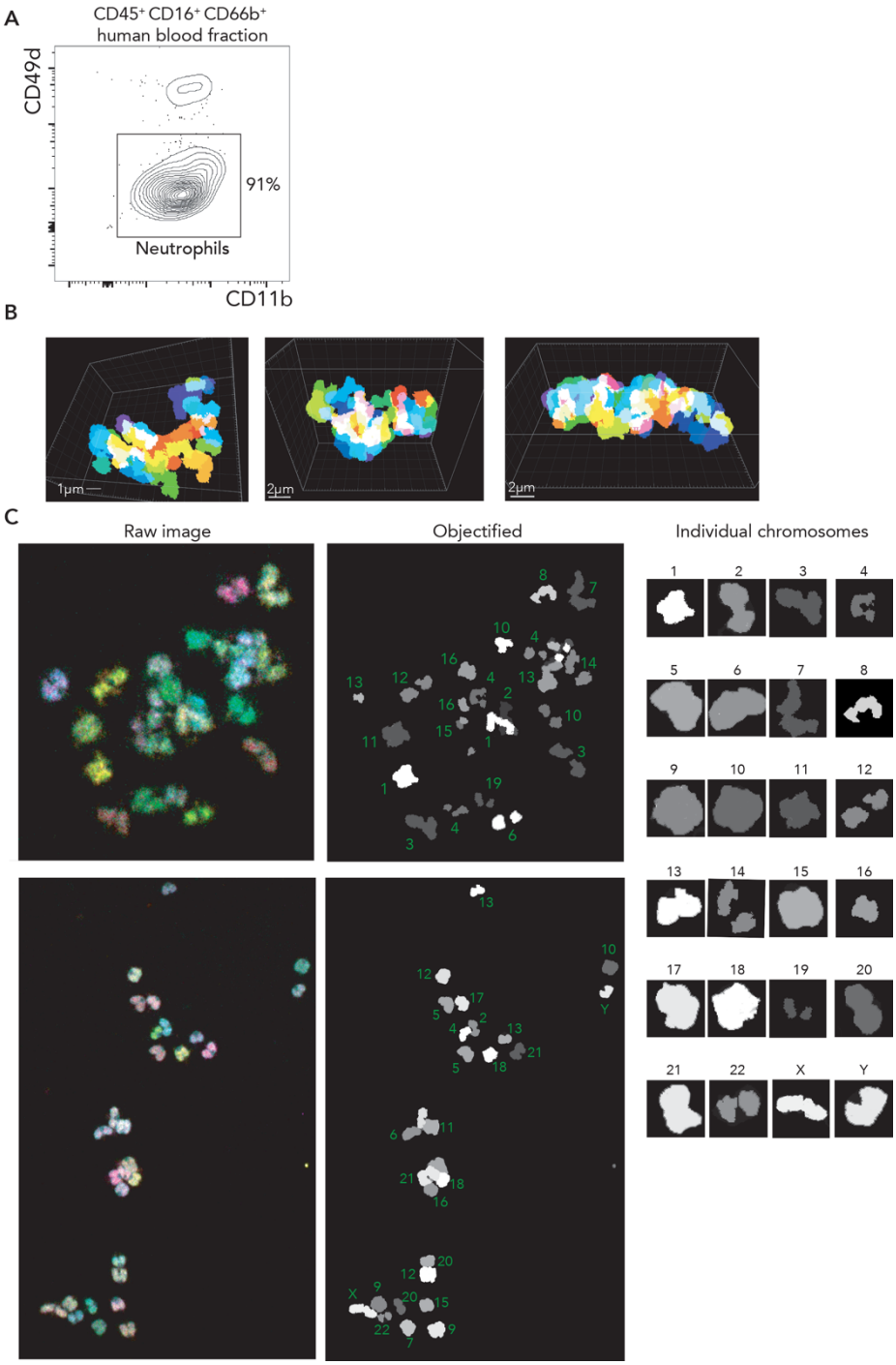
iScience, Volume *24*

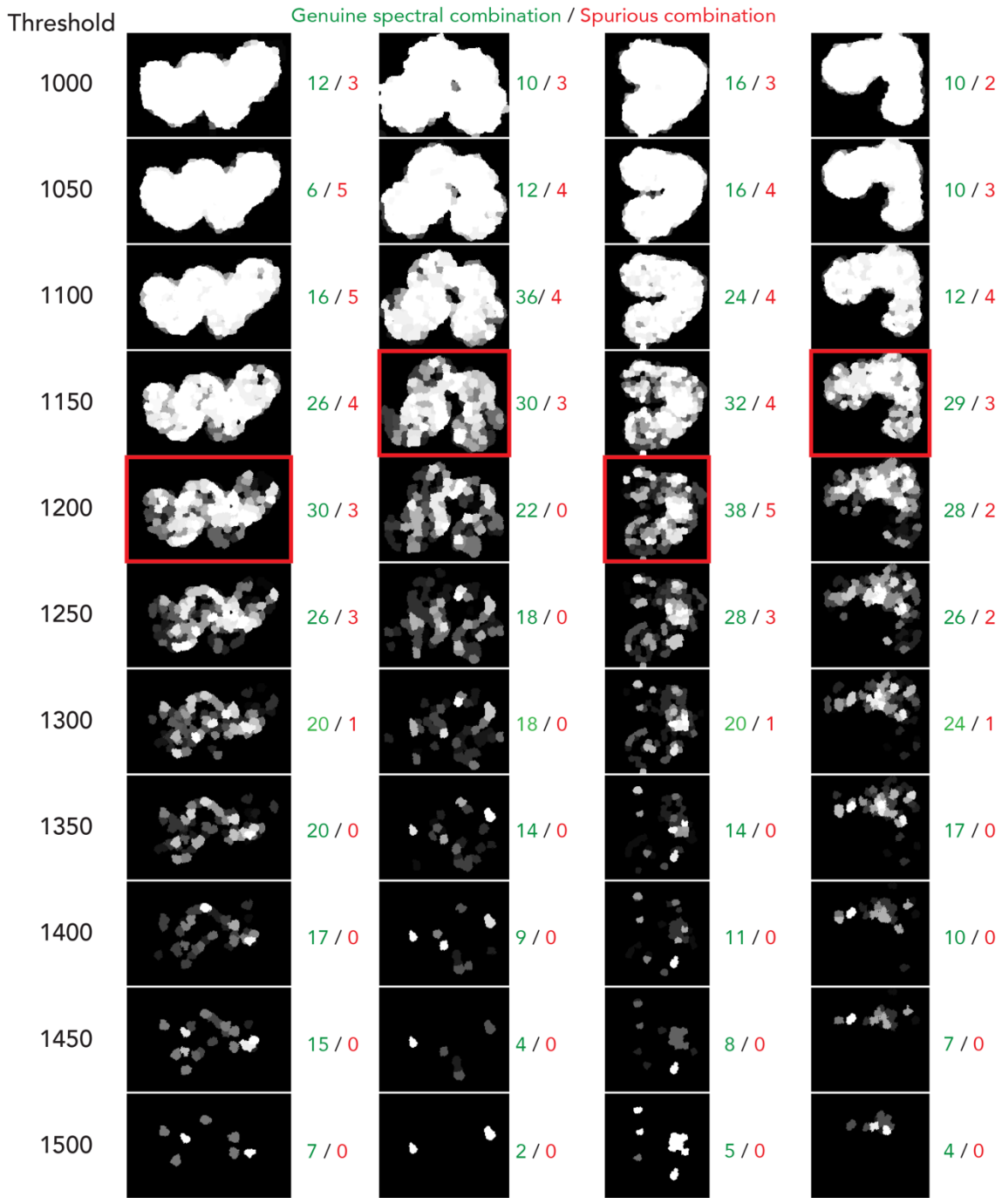
Supplemental information

Chromosomes distribute randomly to, but not within, human neutrophil nuclear lobes

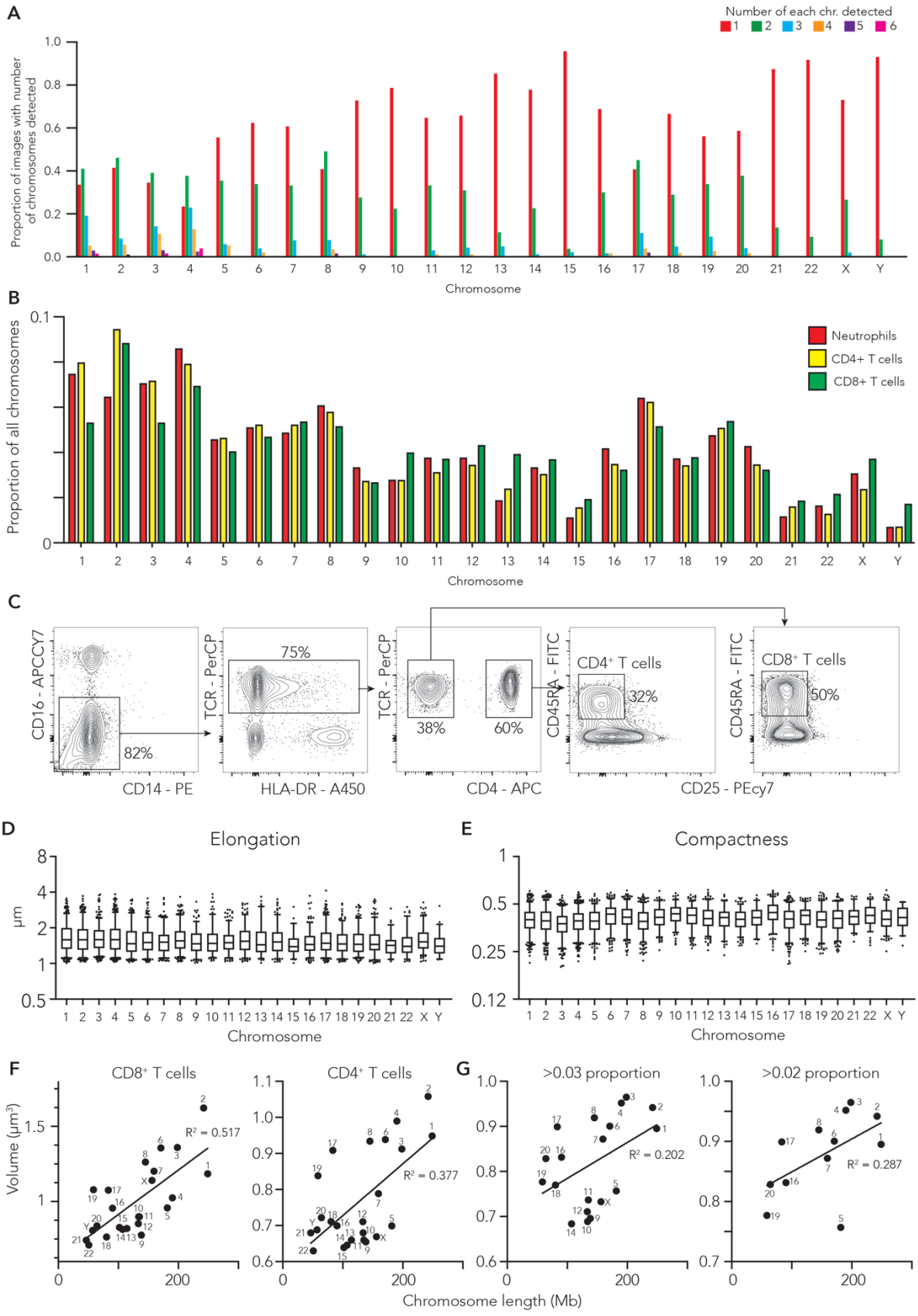
Christine R. Keenan, Michael J. Mlodzianoski, Hannah D. Coughlan, Naiara G. Bediaga, Gaetano Naselli, Erin C. Lucas, Qike Wang, Carolyn A. de Graaf, Douglas J. Hilton, Leonard C. Harrison, Gordon K. Smyth, Kelly L. Rogers, Thomas Boudier, Rhys S. Allan, and Timothy M. Johanson



Supp Fig 1. Neutrophil isolation and chromosome paint quality control. Related to Fig 1. (A) Example purity check of enriched human blood neutrophils. (B) Example 3D renders of human neutrophil nuclei. Colours are arbitrary but represent individual chromosomes. (C) Example of chromosome calling using our analysis pipeline on human metaphase chromosome spreads. Raw images (left), objectified images (centre) with called chromosome number, and select individual chromosomes (right) are shown.

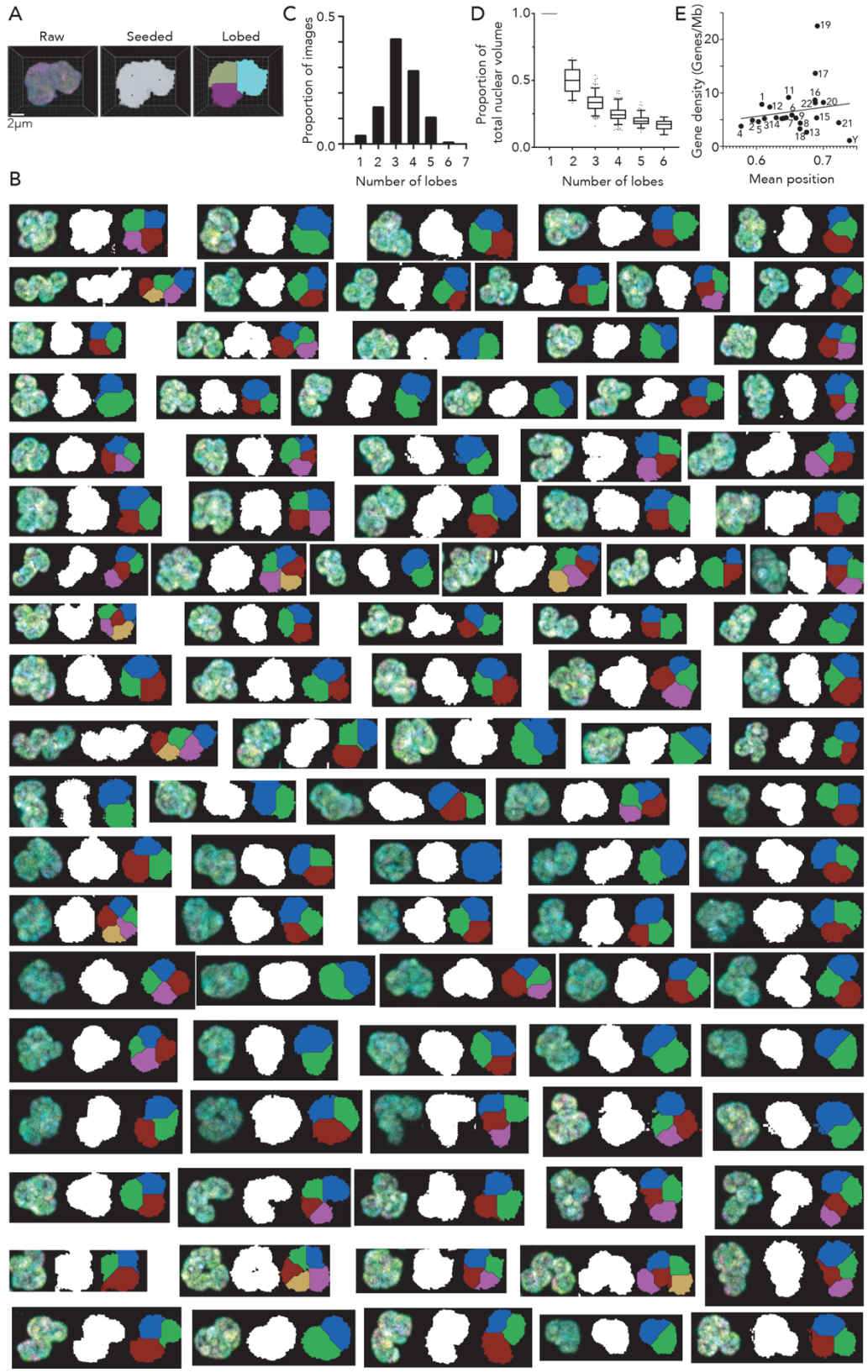


Supp Fig 2 Serial thresholding. Related to Fig 1. Images showing the number of objects called with genuine spectral combinations (green) and spurious combinations (red) in four select neutrophil nuclei images at thresholds between 1000 and 1500. Red boxed image shows the threshold closest to that called by the algorithm. For clarity, examples of images thresholded in increments of 50 are shown. In fact, increments of 10 are used by the algorithm.



Supp Fig 3 Detecting and measuring chromosomes. Related to Fig 1. (A) Number of chromosomes defined within a nucleus for all neutrophil images (proportion shown). (B) Proportion

of total chromosomes detected made up by each chromosome in human blood neutrophils, $CD4^+$ T cells and $CD8^+$ T cells. Data from 171 $CD4^+$ T cells and 30 $CD8^+$ T cells. (C) Example sort profile of human blood CD4+ and CD8+ T cells. (D, E) Box and whisker plot (5^{th}-95^{th} percentile) showing the elongation (D) and compactness (E) of chromosomes detected in human blood neutrophils. Elongation is the ratio between the largest axis of a fitted 3D ellipsoid and the second largest. Compactness is the normalized ratio between volume and surface. (F) Scatterplot of median chromosome volume in $CD8^+$ T cells (left) and $CD4^+$ T cells (right) against chromosome linear length. Data fitted with a straight line (y=2.94e-3x+0.621, $R^2 = 0.517$ and p-value 7.5e-5 and y=1.41e-3x+0.588, $R^2 = 0.377$ and p-value 0.00141, respectively). (G) Scatterplot of median chromosome volume in neutrophils excluding chromosomes called at a proportion of total lower than 0.03 (left) and 0.02 (right) against chromosome linear length. Data fitted with a straight line (y=7.93e-4x+0.704, $R^2 = 0.202$ and p-value 0.0753 and y=5.48e-3x+0.794, $R^2 = 0.287$ and p-value 0.072, respectively).



Supp Fig 4 Defining nuclear lobes. Related to Fig 2. (A) Slice of a neutrophil image moving through the lobe calling pipeline, from raw image to the final lobed nucleus defined using watershed analysis. (B) Further examples of lobe calling pipeline results. (C) Bar plot showing the proportion of

human blood neutrophils containing between 1-7 nuclear lobes. (D) Box and whisker plot (5^{th}-95^{th} percentile) showing the volume of lobes as a proportion of the total nuclear volume in human blood neutrophils containing between 1-6 lobes. (E) Scatterplot of the median of the mean chromosome volume radial position within human neutrophil nuclear lobes plotted as a function of gene density (genes/Mb). Data fitted with a straight line (y=17.29e-4x-4.77, $R^2 = 0.028$ and p-value 0.433).

Transparent Methods

Ethics Statement

Collection of human blood for research studies was approved by Human Research Ethics Committees of Melbourne Health and Walter and Eliza Hall Institute of Medical Research (application 88/03). Written consent was obtained from donors. Donors were all males aged between 25-40 years of age.

Cell isolation

Neutrophils were isolated from 2ml of healthy male donor whole blood following the EasySep Human Neutrophil Enrichment kit manufacturer's protocol). Purity was ~90%. Human T cells were isolated from the PBMC layer of the remaining whole blood after density separation in Leucosep™ tubes containing 15 mL Ficoll-Paque, following the manufacturer's protocol. These cells were stained with TCRab-PerCP-eFluor 710 (eBioscience Cat.No. 46-9986-42), CD4-APC (BD Pharmingen Cat.No. 555349), CD25-PECy7 (BD Bioscience Cat.No. 557741), CD45RA-FITC (eBioscience Cat.No. 556626), CD14-PE (BioLegend Cat.No. 367104), CD16- APC-Cy7 (BD Bioscience Cat.No. 557758), HLA-DR-eFluor450 (eBioscience Cat.No. 48-9952-42) and CD19-BV650 (BioLegend Cat.No. 302238). CD4+ T cells (CD16- CD14- TCRab+ CD4+ CD45RA+ CD25-) and CD8+ T cells (CD16- CD14- TCRab+ CD4- CD45RA+ CD25-) were sorted to a purity >97%.

Chromosome paint

Sixty thousand cells were settled on a poly-L-Lysine (Sigma Aldrich Cat no. P4707) coated cover slip at 37^{o}C for 20 minutes, washed with phosphate buffer saline, before fixing with fresh 3:1 methanol:acetic acid (glacial) for 15 minutes at 22^{o}C. The cells were then rinsed twice with water before being incubated in 2x saline-sodium citrate buffer (SSC) for 2 minutes. The cells were then dehydrated in an ethanol series (75%, 85% and 100%) for 2 minutes each. 7uL of Metasystems

24XCyte Human multicolour FISH probes (Metasystems Cat no. D-0125-060-DI) was pre-warmed on a glass slide at 37°C for 5 minutes before a further 2 minutes at 75°C with the cells added. The cells were then sealed and incubated at 37°C for 18 hours in a dark, humid chamber. After hybridisation was complete the cells were immersed in 72°C 0.4x SSC for 2 minutes, 2x SSC with 0.05% Tween-20 for 30 seconds before being sealed on a glass slide in 85% glycerol.

Confocal Microscopy

Imaging experiments were performed on a Zeiss 880 confocal microscope using a 63x 1.4 NA objective lens. The system was run in lambda mode, recording fluorescence signal in 32 spectral channels over a spectral range of 410 nm to 690 nm (in 8.9 nm increments). Samples were imaged in 3D using Nyquist sampling of 70 nm pixel size and z-steps of 200 nm using the 405 nm, 488 nm, 561 nm and 633 nm lasers to excite the samples consisting of the fluorescent labels DEAC, FITC, Spectrum Orange, Texas Red and Cy5. Single colour control experiments were performed to determine the spectral signatures using the above microscope settings.

Tetraspeck beads (ThermoFisher) adhered to a coverslip and mounted to a microscope slide were imaged in 3D. Custom written Matlab (Mathworks, Natick, MA, USA) scripts measured the axial positions of the Tetraspeck beads in each spectral channel and removed the chromatic aberrations from the sample images. The five spectral signatures from the single colour controls were used to linearly unmix the 32-channel spectral fluorescence signal in each voxel into the five constituent fluorophores.

Chromosome assignment and measurement

After channel unmixing and normalisation, voxels with similar values in the 5 channels are clustered together using a 3D Simple Linear Iterative Clustering (SLIC) algorithm (Tran Thi Nhu et al., 2017). Thresholding is then applied to the 5 channels and chromosome identity is assigned according to known channel combinations. The threshold is set automatically by examining a range of thresholds

and select that which yields the optimal number of chromosome objects (2) and the minimum number of false combination objects. Chromosome measurements are performed on assigned chromosomes using algorithms and tools from the 3D ImageJ suite (Ollion et al., 2013). 3D Eroded Volume Fraction (Ballester et al., 2008) is also performed to compute the position of chromosomes within the nucleus.

Nuclear lobe calling

Nuclei boundaries are detected by summing all channels and global thresholding. The approximate 3D positions of the lobe centres are manually marked before a watershed method separates the nucleus into lobes.

Chromosome shuffling

While considering the chromosomes three-dimensional character, the centre positions of all chromosomes are randomly distributed within the nuclear space until all chromosomes fit and no overlapping is observed between chromosomes. Due to space constraints shuffling was not possible for every nucleus attempted.

All images were stored within an OMERO Database (Allan et al., 2012), processing and analysis were then automated using the TAPAS home system (Whitehead, 2018).

Chromosome measurements analysis

Chromosome measurements were analysed with R using packages dplyr and purr. Objects not assigned to a known combination of values were removed for analysis and plotting. Straight lines were fitted to the data with function lm() from R.

Chromosome co-localisation in lobes

Nuclei with less than ten chromosomes detected or just one lobe were excluded from the analysis. For the remaining nuclei, the volume of each chromosome in each lobe was recorded. The co-localisation

score of any two chromosomes in a nucleus was quantified by $Sum(p_{il} * p_{jl})$, where p_{il} is the proportion by volume of chromosome i in lobe l , p_{jl} is the proportion by volume of chromosome j in lobe l , and the sum is over all lobes in the nucleus. Co-localization p-values were obtained as follows. Each chromosome was treated in turn as the reference chromosome. The co-localisation scores for the reference chromosome with each other chromosome were ranked within each nucleus. The ranks were summed across nuclei and converted to z-scores assuming uniformly distributed ranks for each nucleus. The p-values were adjusted for multiple testing using the Bonferroni correction. Heatmaps of the $-\log_{10}$(p-value) were generated using the R package gplots with the function heatmap.2.

***In situ* HiC**

As described by Rao et al. (Rao et al., 2014), $2x10^6$ human blood neutrophils were resuspended with culture media at $1x10^6$ cells/ml and fixed with 1% v/v formaldehyde (Sigma). Crosslinked cells were lysed with 10mM Tris-HCl pH8.0, 10 mM NaCl, 0.2% Igepal CA630 (Sigma), and protease inhibitors (Sigma). Pelleted nuclei were then digested with 100U of MboI (NEB) overnight and subsequently biotin-labelled with Klenow fragment (NEB) and biotin-dATP (Invitrogen). Filled ends were ligated with T4 DNA ligase (NEB) and sonicated (Covaris). The resulting DNA fragments were biotin-pulled down and end-repaired with T4 polynucleotide kinase (NEB), T4 DNA polymerase (NEB) and Klenow fragment, followed by A-tailing with 3'-5' Klenow (exo-) fragment (NEB), and adaptor ligation using Quick ligase (NEB). The resultant Hi-C library was amplified with Phusion Polymerase (Thermo), size-selected and purified with AMPure XP magnetic beads (Beckman) and sequenced on an Illumina NextSeq 500 to produce 81-bp paired-end reads. Approximately 200 million read pairs were generated for one biological replicate.. HiC libraries for $CD4^+$ T cells, $CD8^+$ T cells and B cells are from GSE105776. The data pre-processing and analysis was performed with the *diffHic* pipeline (Lun and Smyth, 2015) in R with changes in parameters (Johanson et al., 2018a). Where biological replicates were available, the libraries were summed after pre-processing.

Chromosomal looping interactions were detected using a method described by Rao et al. (Rao et al., 2014) and in (Johanson et al., 2018b). In brief, read pairs were counted in bin pairs of 50 kbp anchors for all libraries. For each bin pair, the log-fold change over the average abundance of each of several neighbouring regions was computed. Neighbouring regions in the interaction space included a square quadrant of sides 'x+1' that was closest to the diagonal and contained the target bin pair in its corner; a horizontal stripe of length '2x+1' centred on the target bin pair; a vertical stripe of '2x+1', similarly centred; and a square of sides '2x+1', also containing the target bin pair in the centre. The enrichment value for each bin pair was defined as the minimum of these log-fold changes, i.e., the bin pair had to have intensities higher than all neighbouring regions to obtain a large enrichment value. The neighbourhood counts for the libraries at 50 kbp bin size were computed with the neighborCounts function from the *diffHic* package with flank (x) 5 bin sizes (i.e., 250 kbp) and enrichment values were determined with filterPeaks function with get.enrich=TRUE. Looping interactions were then filtered with filterPeaks. Loops were defined as those with enrichment values above 1, were more than 100 kbp from the diagonal and with minimum count greater than 10 for all libraries except the neutrophils which used a minimum count of 5 to accounts for differences in library size. Directly adjacent loops in the interaction space were aggregated into clusters to a maximum cluster size of 500 kbp using the clusterPairs function from the *csaw* package v1.18.0 (Lun and Smyth, 2016). Blacklisted genomic regions were obtained from ENCODE for hg38 (Consortium, 2012). Loops that that had at least one anchor in a blacklisted genomic region were removed.

Heatmaps of the loops between chromosomes where generated using the R package gplots with the function heatmap.2. The frequency of transchromosomal interactions to the power of 0.25 was plotted as a function of the sum of the chromosome lengths. A linear model was fitted to data with the lm function.

Supplemental References

ALLAN, C., BUREL, J. M., MOORE, J., BLACKBURN, C., LINKERT, M., LOYNTON, S., MACDONALD, D., MOORE, W. J., NEVES, C., PATTERSON, A., PORTER, M., TARKOWSKA, A., LORANGER, B., AVONDO, J., LAGERSTEDT, I.,

LIANAS, L., LEO, S., HANDS, K., HAY, R. T., PATWARDHAN, A., BEST, C., KLEYWEGT, G. J., ZANETTI, G. & SWEDLOW, J. R. 2012. OMERO: flexible, model-driven data management for experimental biology. *Nat Methods,* 9**,** 245-53.

BALLESTER, M., KRESS, C., HUE-BEAUVAIS, C., KIEU, K., LEHMANN, G., ADENOT, P. & DEVINOY, E. 2008. The nuclear localization of WAP and CSN genes is modified by lactogenic hormones in HC11 cells. *J Cell Biochem,* 105**,** 262-70.

CONSORTIUM, E. P. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature,* 489**,** 57-74.

JOHANSON, T. M., COUGHLAN, H. D., LUN, A. T. L., BEDIAGA, N. G., NASELLI, G., GARNHAM, A. L., HARRISON, L. C., SMYTH, G. K. & ALLAN, R. S. 2018a. Genome-wide analysis reveals no evidence of trans chromosomal regulation of mammalian immune development. *PLoS Genet,* 14**,** e1007431.

JOHANSON, T. M., LUN, A. T. L., COUGHLAN, H. D., TAN, T., SMYTH, G. K., NUTT, S. L. & ALLAN, R. S. 2018b. Transcription-factor-mediated supervision of global genome architecture maintains B cell identity. *Nat Immunol,* 19**,** 1257-1264.

LUN, A. T. & SMYTH, G. K. 2015. diffHic: a Bioconductor package to detect differential genomic interactions in Hi-C data. *BMC Bioinformatics,* 16**,** 258.

LUN, A. T. & SMYTH, G. K. 2016. csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res,* 44**,** e45.

OLLION, J., COCHENNEC, J., LOLL, F., ESCUDE, C. & BOUDIER, T. 2013. TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinformatics,* 29**,** 1840-1.

RAO, S. S., HUNTLEY, M. H., DURAND, N. C., STAMENOVA, E. K., BOCHKOV, I. D., ROBINSON, J. T., SANBORN, A. L., MACHOL, I., OMER, A. D., LANDER, E. S. & AIDEN, E. L. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell,* 159**,** 1665-80.

TRAN THI NHU, H., ARROJO, E. D. R., BERGGREN, P. O. & BOUDIER, T. 2017. A novel toolbox to investigate tissue spatial organization applied to the study of the islets of Langerhans. *Sci Rep,* 7**,** 44261.

WHITEHEAD, L., WIMMER, V., LAFOURESSE, F., RATNAYAKE, D., CURRIE, P., GROOM, J., ROGERS, K. AND BOUDIER, T 2018. Towards an Automated Processing and Analysis System for multi-dimensional light-sheet microscopy big data using ImageJ and OMERO. *International Microscopy Congress IMC 19.*