nature chemistry

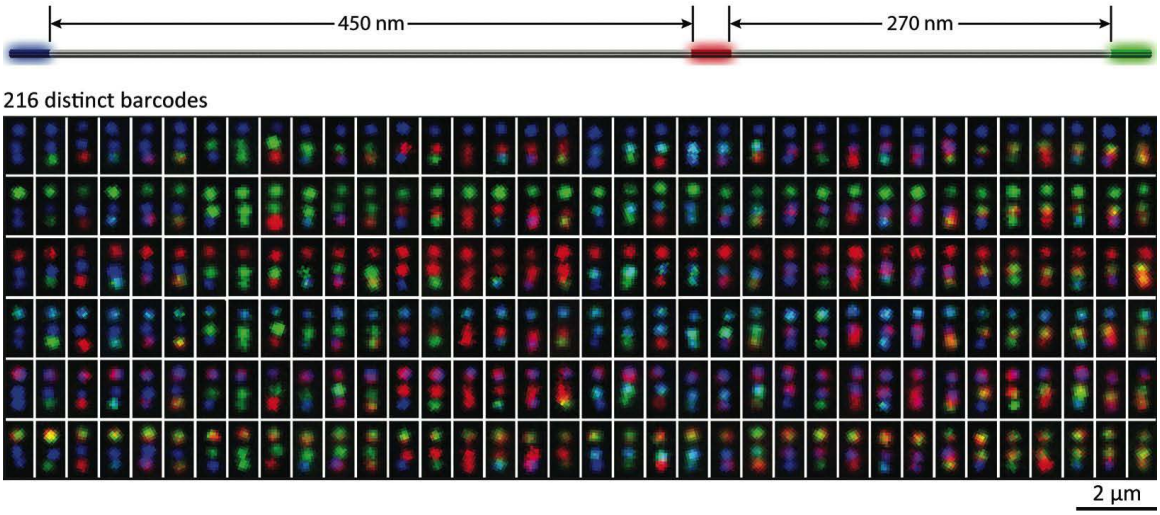
SUPPLEMENTARY INFORMATION

DOI: 10.1038/NCHEM.1451

Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA

Chenxiang Lin, Ralf Jungmann, Andrew M. Leifer, Chao Li, Daniel Levner, George M. Church, William M. Shih and Peng Yin

TABLE OF CONTENT



Summary figure. Life-science research and biomedical diagnostics call for robust fluorescence barcodes of compact size and high multiplexing capability. Here DNA-origami technology was used to construct a new kind of geometrically encoded barcodes with excellent structural stiffness. They hold promise for *ex* and *in situ* imaging of diverse biologically relevant entities.

Materials and Methods

Materials

Non-modified DNA oligonucleotides were purchased from Bioneer Inc. (Alameda, CA). Fluorescently modified DNA oligonucleotides were purchased from Integrated DNA Technology (Coralville, IA) and Biosynthesis (Lewisville, TX). Polyclonal antibodies to *C. albicans* were purchased from Thermo Scientific (Rockford, IL). Streptavidin was purchased from Invitrogen (Carlsbad, CA). Glass slides and coverslips were purchased from VWR (Radnor, PA). All other reagents were purchased from Sigma-Aldrich (St. Louis, MO). M13mp18 scaffold strand for the DNA-origami drift markers was obtained from New England Biolabs (Ipswich, MA).

Design of the nano-barcodes

The main-body of the linear nano-barcode is a DNA six-helix-bundle (6hb) nanorod dimer designed using the caDNAno software (cadnano.org). The structure of the barcode is illustrated as a strand diagram in Figure S1. Staples in the fluorescent labeling zones were extended at the 3'-end with 21-base long single-stranded overhangs (handles) for fluorescent labeling (see Table S1 for staple extension sequences). In addition, ten staples (five per monomer, shown as magenta strands in the diagram) were extended at the 5'-end with biotin. The three-point-star like barcode is self-assembled from three identical DNA 6hb nanorod monomers and a DNA 6hb ring. Here, 12 handles protrude out from a helix on the outer edge of the ring. Three of them serve as sticky-ends to link with the 6hb nanorods while the other nine are used for fluorescent labeling. The 6hb nanorod carries a handle at the front end to serve as the complementary sticky-end and 24 handles at the rear end for fluorescent labeling. The design is illustrated as a strand diagram in Figure S1 and the staple extension sequences used in this design are listed in Table S2.

Self-assembly of the nano-barcodes

To fold the ~800 nm DNA-nanorod based barcode, the monomer nanorods were each assembled and fluorescently labeled in a separate test tube and then mixed together to form the dimer. The assembly of the monomer was accomplished in one-pot reaction by mixing 100 nM scaffold strands (7,308-base long, termed p7308, see Table S4 for sequence) derived from M13 bacteriophage with a pool of oligonucleotide staple strands (600 nM of each; reverse-phase cartridge purified.) in folding buffer containing 5 mM Tris, 1 mM EDTA, 20 mM $MgCl_2$, 50 mM NaCl (pH 8) and subjecting the mixture to a thermal-annealing ramp that cooled from 80 ºC to 60 ºC over the course of 80 minutes and then cooled from 60 ºC to 24 ºC over 15 hours. Excessive staples were removed from the folded nanorods by polyethylene glycol fractionation.[1] To fluorescently label the nanorods, desired DNA oligonucleotides with complementary sequence to the handles (termed anti-handles, see Table S3 for sequences) were added to the monomers (1.2:1 molar ratio between anti-handles and handles) and the hybridization was carried out at 37 ºC for 2 hours. For dimerization, stoichiometric amount of the two fluorescently labeled monomers were mixed and incubated at 16 ˚C overnight. The final product was loaded on a non-denaturing, 1.5% agarose gel and electrophoresed at 3V/cm in ice-cold 0.5× TBE buffer containing 10 mM $MgCl_2$ for 3.5 hours. It is important not to expose the fluorescently labeled nanorod dimers to elevated temperature (>25˚C) to avoid the cross-contamination because some handles in different fluorescent labeling zones share the

same oligonucleotide sequence (Table S1). The desired band was excised and the nanorod dimers were recovered from the agarose gel using a pellet-pestle method[2]. To prepare an equimolar mixture of 2, 27, 72 or 216 barcodes, all the barcode species involved are folded, labeled and dimerized separately and then mixed at equal molar ratio for agarose gel electrophoresis purification. To produce the three-point-star like barcode, the DNA nanorod was folded and labeled as described above and the DNA ring was prepared under similar conditions, with the p3024 scaffold (3,024-base long circular ssDNA derived from pBluescript vector, see Table S4 for sequence), staple and $MgCl_2$ concentration changed to 25 nM, 125 nM (each) and 10 mM, respectively. The labeled ring and the nanorod were then purified separately through agarose gel electrophoresis, mixed together at 1:3 molar ratio and incubated at 37 ºC overnight. The final product was purified via gel electrophoresis as described above.

Tagging yeast cells with nano-barcodes

Approximately 3.8×10^7 *C. albicans* yeast cells were suspended in 150 µL of PBS buffer (2.5×10^5 cells per microliter). 1 µL of 4 mg/mL polyclonal antibody (with or without biotin label) was added to this suspension and allowed to bind to the cell surface for 15 minutes. The yeasts were then spun in a desktop centrifuge at 4000 rcf for 15 minutes to form a pellet and washed with 150 µL of PBS buffer twice. The antibody-coated cells was then resuspended in 150 µL of PBS buffer and mixed with 2 µL of 10 mg/mL streptavidin. After incubating for 10 minutes, the cells were again spun and washed as described above. The cell pellet was then resuspended in 150 µL of PBS buffer containing 10 mM $MgCl_2$ and 0.05% Tween-20 (termed PBS+ buffer here) and treated with 2 µL of 2 nM biotinylated or non-biotinylated GRG nano-barcode. After 30 minutes incubation on a desktop shaker, the yeast cells were pelleted and washed three times with 150 µL of PBS+ buffer.

Sample preparation for fluorescence imaging

A piece of coverslip (No. 1.5, 18×18 mm^2, ~0.17 mm thick) and a glass slide (3×1 inch^2, 1 mm thick) were sandwiched together by two strips of double-side tapes to form a flow chamber with inner volume of ~15 µL. First, 20 µL of biotin labeled bovine albumin (1 mg/mL, dissolved in Buffer A (10 mM Tris-HCl, 100 mM NaCl, 0.05% Tween-20, pH 7.5)) was flown into the chamber and incubated for 2 minutes. The chamber was then washed using 20 µL of Buffer A twice. 20 µL of streptavidin (0.5 mg/mL, dissolved in Buffer A) was then flown through the chamber and allowed to bind for 2 minutes. After washing twice with 20 µL of Buffer B (5 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM EDTA, 0.05% Tween-20 pH 8, supplemented with 2.5 mM protocatechuic acid, 10 nM protocatechuate-3,4-dioxygenase and 1 mM Trolox), 20 µL of biotin labeled nano-barcodes (10 pM in Buffer B) were finally flown into the chamber and incubated for 2 minutes. The chamber was washed using 20 µL of buffer B for three times, sealed with nail polish and mounted on the sample stage of a microscope with the coverslip facing towards the objective. For DNA-PAINT imaging, samples were prepared as described above. The final imaging buffer solution contained 20 nM ATTO488, Cy3b and ATTO655 labeled imager strands in Buffer B respectively. For the three-point-star shape barcode and the yeast cells, 5 µL of specimen was directly deposited on the glass slide, covered immediately with a piece of coverslip, sealed with nail polish and incubated for 5 min before mounting onto the sample stage.

Diffraction-limited fluorescence imaging

Fluorescence imaging was carried out on an inverted Leica DM1600B microscope (Buffalo Grove, IL) applying an objective-type TIRF configuration using a Leica TIRF illuminator with an oil-immersion objective (HCX PL APO 100×/1.47 oil CORR TIRF). Additional lenses were used to achieve a final imaging magnification of ~112 fold, corresponding to a pixel size of ~71 nm. A 488 nm solid state laser (JDS Uniphase, FCD 488-10), a 561 nm solid state laser (LASOS YLK6110) and a 635 nm diode laser (BSR, ChromaLase 635) were used for TIRF excitation. Laser beams were filtered with clean up filters (zet488x, zet561x, 635x) and coupled into the microscope objective using a multi-band beamsplitter (zt488, zt561, zt635). Fluorescence light was spectrally filtered with emission filters (ET525/36, ET600/32, ET705/72) and imaged on an EMCCD camera (Hamamatsu C9100-02, Hamamatsu, Japan). Exposure times were 1 second for 488 nm excitation and 700 ms for 561 and 635 nm excitations. Two consecutive images were taken for the same area of sample and then averaged to yield the final image. The laser beam alignment and imaging process were automatically controlled by the Leica LAS software. The TIRF images were processed using ImageJ and a custom written software for decoding (see the decoding software section).

For epi-fluorescence imaging, the full-spectrum light generated from a mercury lamp was filtered through desired filters (470/40, 560/40 and 645/30 for blue, green and red channel, respectively) to serve as excitation source. The same beam splitter and emission filters were used as in the TIRF mode. The exposure time for all channels was 2 seconds. For the barcode treated yeast sample, an additional bright field image was taken to show the profile of the cells.

DNA-PAINT imaging

Super-resolution fluorescence imaging using DNA-PAINT was carried out on three different microscope systems. Figure 4b and c and Figure S14a were acquired on an inverted Zeiss LSM 710 ELYRA PS.1 microscope (Carl Zeiss, Germany) applying an objective-type total internal reflection fluorescence (TIRF) configuration using an oil immersion objective (100x alpha Plan Apo, NA 1.46, Oil, Zeiss). An additional 1.6 magnification was used to obtain a final imaging magnification of ~160 fold, corresponding to a pixel size of 100 nm. Samples were imaged using custom made microscopic chambers. Three lasers (488 nm (100 mW), 561 nm (100 mW) and 642 nm (100 mW), power levels are nominal laser output power before fiber coupling) were used for TIRF excitation. The laser beam was coupled into the microscope objective using a multi-band beamsplitter. Fluorescence light was spectrally filtered with an emission filter (BP 495-575/LP 750 for ATTO488 imaging, BP 570-65-/LP 750 for Cy3b imaging and LP 655 for ATTO655 imaging) and imaged on an EMCCD camera (Andor iXon 3, Andor Technologies, North Ireland). 5000 frames were recorded at a frame rate of 50 Hz for each color channel sequentially. Using this microscope setup, DNA-PAINT provides a resolution of ~27 nm (FWHM of a 2D Gaussian fit to the reconstructed point spread function (PSF)) in the red, ~22 nm in the green and ~23 nm in the blue channel.

Figure S13 was acquired on an inverted Nikon Ti-E microscope (Nikon Instruments, Melville, NY) with the Perfect Focus System, applying an objective-type TIRF configuration using a Nikon TIRF illuminator with an oil-immersion objective (100× Plan Apo, NA 1.49, Oil, Nikon). An additional 1.5 magnification was used to

obtain a final imaging magnification of ~150 fold, corresponding to a pixel size of 107 nm. Three lasers (491 nm (27 mW), 561 nm (16 mW) and 640 nm (6 mW), power levels are measured at the objective) were used for TIRF excitation. The laser beam was filtered with a clean up filter (488/20 for ATTO488 imaging, 561/20 for Cy3b imaging and 642/20 for ATTO655 imaging, all Chroma Technologies) and coupled into the microscope objective using a multi-band beamsplitter (zt405/488-491/561/638rpc, Chroma Technologies). Fluorescence light was spectrally filtered with an emission filter (525/50 for ATTO488 imaging, 600/50 for Cy3b imaging and 700/75 for ATTO655 imaging, all Chroma Technologies) and imaged on an EMCCD camera (Hamamatsu ImagEM C-9100-13, Hamamatsu, Japan). 5000 frames were recorded at a frame rate of 10 Hz for each color channel sequentially.

Figure 4d and S14b and c were acquired using the same setup used for Figure S13 except for different excitation lasers (Agilent MLC400B laser launch for 488nm (28 mW), 561nm (28 mW) & 647nm (80 mW), power levels are measured at the objective) and a different EMCCD camera (Andor iXon 3, Andor Technologies, North Ireland).

Super-resolution images were reconstructed using spot-finding and 2D Gaussian fitting algorithms programmed in LabVIEW[3]. A simplified version of this software is available for download at www.e14.ph.tum.de. Images were drift corrected and the different channels aligned using DNA-origami drift markers with binding sites for all three imager strands (cf. Figure S1d). The origami structure is based on Rothemund's rectangular shaped origami[4]. The black staple strands in the strand diagram were randomly extended with 2 out of three docking/handle strands for DNA-PAINT imaging. The resulting structure displays ~80 binding sites on average for the red, green and blue imager strands respectively. This high density of binding sites yields a constantly "occupied" i.e. bound state of one or more imager strands at every given point in time. This makes the structures ideal for systematic drift correction in each channel and because the structures are imaged in each color channel at the same positions, they can be used to correct for offsets and align the red, green and blue channel accordingly.

The calculated image resolution (FWHM of 2D Gaussian fit to the reconstructed PSF, N=16) for Figure 4b, c and Figure S14a are ~27 nm in the red, ~22 nm in the green and ~23 nm in the blue channel. The calculated image resolution for Figure S14b and c is ~17 nm.

Automated characterization of the geometry of BRG & GRG barcodes

For the purpose of assessing the barcode geometry, we chose BRG and GRG barcodes as examples. The software first determined the centroids of all visible spots in the diffraction limited TIRF images based on their fluorescence intensities. The centroid of a spot was defined as the brightest central pixel in a 3-pixel by 3-pixel region. The software processed blue, green and red fluorescence channels independently to identify all the centroids in each channel (Figure S4). It then superimposed the three channels and found the correct barcodes based on the design blueprint. The distances between the neighboring fluorescently labeled zones were calculated as the distances between the centroids of the adjacent spots on a barcode. The barcode bending angle (α) was calculated using the cosine formula as shown below, where a and b are the distances between neighboring centroids and c is the distance between the farthest two centroids within a barcode. The measured parameters are listed in Table S5.

$$\cos\alpha = \frac{a^2 + b^2 - c^2}{2ab}$$

Decoding software for diffraction-limited TIRF images

Custom MATLAB scripts were used to decode barcodes from diffraction-limited TIRF images. The software can run in either a supervised or an unsupervised mode. In the supervised mode, each decoded barcode is presented to the user to accept or reject. In unsupervised mode, the software decodes all barcodes without any user input.

The software operates in two steps: In the first step the software identifies the location and position of barcodes. In the second step the software actually decodes the barcode (See Figure S10). For step one, the software preprocesses the three-channel TIRF image containing the barcodes (one channel each for fluorescence in the red, green and blue). The image is spatially high-pass filtered to remove slowly spatially varying illumination background. Each color channel is then thresholded independently so that only pixels corresponding to putative barcodes remain. The three-channel thresholded image is then merged to generate a single channel binary image. The software identifies shapes from the binary image that match the length and width criteria of a barcode. The position and orientation of these shapes are recorded. In step two, the software reads each barcode identified in step one. For each barcode, the software extracts the corresponding region of the original three-channel image. This constitutes the observed barcode. The software then performs a matched filter to compare the observed barcode against a library of all possible reference barcodes. For each reference, three simultaneous correlations are performed, one in each color channel, while keeping all channels aligned. To reduce computations, these correlations are actually performed on one-dimensional projections of the 2D image of the observed barcode in each channel, so that each fluorescent spot appears as a peak in a 1D trace. The observed barcode is assigned the identity of the reference barcode with the highest combined correlation across all three color channels. When run in supervised mode, the computer will present the user with the image of the barcode and its most likely identity. The user can either accept the computer's proposal or manually enter the correct identity of the barcode. In cases where the software erroneously decodes misshapen or malformed barcodes, the user can reject the barcode entirely. The software finally outputs the counts of all barcode species as well as the coordinates and identity of each individual barcode. When run in supervised mode, it also registers the number of occasions when user approves the computer's proposal. For the 72-barcode pool, this agreement rate is approximately 80%.

Supplementary Tables

Table S1 Handle sequences for rod-like barcodes

Staple extension sequences used for constructing the 800-nm nano-barcode and 400 nm super-resolution barcode. The staple extension names are colored to match the staples in the strand diagram shown in Figure S1. Handles 1 and 2 are located at the front monomer and handles 4–6 are at the rear monomer (Figure S1a). The 3'-ends of biotinylated handles and the 5'-ends of handles 1–6 are linked to the staples.

Staple extension name	Sequence
Biotinylated extensions	5'-/biotin/TTTTT-3'
For Dual-labeling:	
Handle **1** & **3**	5'-TTCCTCTACCACCTACATCAC-3'
Handle **2** & **5**	5'-TAACATTCCTAACTTCTCATA-3'
Handle **4**	5'-CTACCATCTCTCCTAAACTCA-3'
Handle **6**	5'-ACTCACCACCACAATATTATA-3'
For single-labeling:	
Handle **1** & **2** & **3** & **4**	5'-TTCCTCTACCACCTACATCAC-3'
Handle **5** & **6**	5'-TAACATTCCTAACTTCTCATA-3'
For DNA-PAINT:	
Handle **P1.1**, **P2.1**, **P3.1** & **P4.1**	5'-TTATACATCTA-3'
Handle **P1.2**, **P2.2**, **P3.2** & **P4.2**	5'-TTATCTACATA-3'
Handle **P1.3**, **P2.3**, **P3.3** & **P4.3**	5'-TTTCTTCATTA-3'

Table S2 Handle sequences for three-point-star-like barcodes

Staple extension sequences used for constructing the three-point-star like nano-barcode. The staple extension names are colored to match the staples in the strand diagram shown in Figure S1b. Handles 1 and 2 are located at the ring and handles 3 and 4 are on the nanorod (Figure S1b). The 5'-ends of the handles are linked to the staples.

Staple extension name	Sequence
For fluorescent labeling:	
Handle **1**	5'-TTCCTCTACCACCTACATCAC-3'
Handle **3**	5'-TAACATTCCTAACTTCTCATA-3'
For linking nanorods to the ring:	
Handle **2**	5'-CTTCACACCACACTCCATCTA-3'
Handle **4**	5'-TAGATGGAGTGTGGTGTGAAG-3'

Table S3 Anti-handle sequences for rod-like barcodes

Anti-handle sequences used for labeling rod-like barcodes. /AlexF488/ = Alexa Fluor 488; /AlexF647/ = Alexa Fluor 647. Colors refer to the pseudo-colors used in fluorescence imaging.

Anti-handle name	Color	Sequence
For Dual-labeling:		
Anti-handle **1** & **3**	Blue	5'-GTGATGTAGGTGGTAGAGGAA/AlexF488/-3'
	Green	5'-GTGATGTAGGTGGTAGAGGAA/Cy3/-3'
	Red	5'-GTGATGTAGGTGGTAGAGGAA/AlexF647/-3'
Anti-handle **2** & **5**	None	5'-TATGAGAAGTTAGGAATGTTA-3'
	Green	5'-TATGAGAAGTTAGGAATGTTA/Cy3/-3'
	Red	5'-TATGAGAAGTTAGGAATGTTA/AlexF647/-3'
Anti-handle **4**	None	5'-TGAGTTTAGGAGAGATGGTAG-3'
	Green	5'-TGAGTTTAGGAGAGATGGTAG/Cy3/-3'
	Red	5'-TGAGTTTAGGAGAGATGGTAG/AlexF647/-3'
Anti-handle **6**	Blue	5'-TATAATATTGTGGTGGTGAGT/AlexF488/-3'
	Green	5'-TATAATATTGTGGTGGTGAGT/Cy3/-3'
	Red	5'-TATAATATTGTGGTGGTGAGT/AlexF647/-3'
For single-labeling:		
Anti-handle **1** & **2** & **3** & **4**	Blue	5'-GTGATGTAGGTGGTAGAGGAA/AlexF488/-3'
	Green	5'-GTGATGTAGGTGGTAGAGGAA/Cy3/-3'
	Red	5'-GTGATGTAGGTGGTAGAGGAA/AlexF647/-3'
Anti-handle **5** & **6**	None	5'-TGAGTTTAGGAGAGATGGTAG-3'
	Green	5'-TATGAGAAGTTAGGAATGTTA/Cy3/-3'
	Red	5'-TATGAGAAGTTAGGAATGTTA /AlexF647/-3'
For DNA-PAINT:		
Anti-handle **P1.1, P2.1, P3.1 & P4.1**	Red	5'-CTAGATGTAT/ATTO655/-3'
Anti-handle **P1.2, P2.2, P3.2 & P4.2**	Green	5'-TATGTAGATC/Cy3b/-3'
Anti-handle **P1.3, P2.3, P3.3 & P4.3**	Blue	5'-GTAATGAAGA/ATTO488/-3'

Table S4 Scaffold sequences
Scaffold sequences for six-helix-bundle ring (p3024) and nanorod (p7308).

P3024 sequence
CCCGGTACCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTT
TTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCA
CATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCC
CAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAG
CGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGC
GCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTC
AAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGA
CCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGAC
GGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAA
ACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGC
CGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATT
TTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTTCGGGGAAATGTGCGCGG
AACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAA
TAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACAT
TTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCA
GAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTA
CATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACG
TTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATT
GACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTT
GAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATT
ATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAAC
GATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAAC
TCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTG
ACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAAC
TACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTG
CAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGG
AGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCC
CTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAA
TAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCA
AGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCT
AGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTT
CCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTT
CTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGT
TTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCG
CAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAAC
TCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCA
GTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGG
CGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACG
ACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCC
GAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGC
GCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTC
GCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTAT
GGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGC
TCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTG
AGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGC
GAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGAT
TCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCA

ACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCT
TCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAG
CTATGACCATGATTACGCCAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCT
GGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCGTAAATCAATGACT
TACGCGCACCGAAAGGTGCGTATTGTCTATAGCCCCCTCAGCCACGAATTCGTCTGA
CGACGACAAGACAAGCTTGCGTGTGAATTCCCTGGCTTCTCCTGAGAAA

P7308 sequence

AATGCTACTACTATTAGTAGAATTGATGCCACCTTTTCAGCTCGCGCCCCAAATGAA
AATATAGCTAAACAGGTTATTGACCATTTGCGAAATGTATCTAATGGTCAAACTAAA
TCTACTCGTTCGCAGAATTGGGAATCAACTGTTATATGGAATGAAACTTCCAGACAC
CGTACTTTAGTTGCATATTTAAAACATGTTGAGCTACAGCATTATATTCAGCAATTAA
GCTCTAAGCCATCCGCAAAAATGACCTCTTATCAAAAGGAGCAATTAAAGGTACTCT
CTAATCCTGACCTGTTGGAGTTTGCTTCCGGTCTGGTTCGCTTTGAAGCTCGAATTAA
AACGCGATATTTGAAGTCTTTCGGGCTTCCTCTTAATCTTTTTGATGCAATCCGCTTTG
CTTCTGACTATAATAGTCAGGGTAAAGACCTGATTTTTGATTTATGGTCATTCTCGTT
TTCTGAACTGTTTAAAGCATTTGAGGGGGATTCAATGAATATTTATGACGATTCCGCA
GTATTGGACGCTATCCAGTCTAAACATTTTACTATTACCCCCTCTGGCAAAACTTCTT
TTGCAAAAGCCTCTCGCTATTTTGGTTTTTATCGTCGTCTGGTAAACGAGGGTTATGA
TAGTGTTGCTCTTACTATGCCTCGTAATTCCTTTTGGCGTTATGTATCTGCATTAGTTG
AATGTGGTATTCCTAAATCTCAACTGATGAATCTTTCTACCTGTAATAATGTTGTTCC
GTTAGTTCGTTTTATTAACGTAGATTTTTCTTCCCAACGTCCTGACTGGTATAATGAG
CCAGTTCTTAAAATCGCATAAGGTAATTCACAATGATTAAAGTTGAAATTAAACCAT
CTCAAGCCCAATTTACTACTCGTTCTGGTGTTTCTCGTCAGGGCAAGCCTTATTCACT
GAATGAGCAGCTTTGTTACGTTGATTTGGGTAATGAATATCCGGTTCTTGTCAAGATT
ACTCTTGATGAAGGTCAGCCAGCCTATGCGCCTGGTCTGTACACCGTTCATCTGTCCT
CTTTCAAAGTTGGTCAGTTCGGTTCCCTTATGATTGACCGTCTGCGCCTCGTTCCGGC
TAAGTAACATGGAGCAGGTCGCGGATTTCGACACAATTTATCAGGCGATGATACAAA
TCTCCGTTGTACTTTGTTTCGCGCTTGGTATAATCGCTGGGGGTCAAAGATGAGTGTT
TTAGTGTATTCTTTTGCCTCTTTCGTTTTAGGTTGGTGCCTTCGTAGTGGCATTACGTA
TTTTACCCGTTTAATGGAAACTTCCTCATGAAAAAGTCTTTAGTCCTCAAAGCCTCTG
TAGCCGTTGCTACCCTCGTTCCGATGCTGTCTTTCGCTGCTGAGGGTGACGATCCCGC
AAAAGCGGCCTTTAACTCCCTGCAAGCCTCAGCGACCGAATATATCGGTTATGCGTG
GGCGATGGTTGTTGTCATTGTCGGCGCAACTATCGGTATCAAGCTGTTTAAGAAATTC
ACCTCGAAAGCAAGCTGATAAACCGATACAATTAAAGGCTCCTTTTGGAGCCTTTTT
TTTGGAGATTTTCAACGTGAAAAAATTATTATTCGCAATTCCTTTAGTTGTTCCTTTCT
ATTCTCACTCCGCTGAAACTGTTGAAAGTTGTTTAGCAAAATCCCATACAGAAAATT
CATTTACTAACGTCTGGAAAGACGACAAAACTTTAGATCGTTACGCTAACTATGAGG
GCTGTCTGTGGAATGCTACAGGCGTTGTAGTTTGTACTGGTGACGAAACTCAGTGTT
ACGGTACATGGGTTCCTATTGGGCTTGCTATCCCTGAAAATGAGGGTGGTGGCTCTG
AGGGTGGCGGTTCTGAGGGTGGCGGTTCTGAGGGTGGCGGTACTAAACCTCCTGAGT
ACGGTGATACACCTATTCCGGGCTATACTTATATCAACCCTCTCGACGGCACTTATCC
GCCTGGTACTGAGCAAAACCCCGCTAATCCTAATCCTTCTCTTGAGGAGTCTCAGCCT
CTTAATACTTTCATGTTTCAGAATAATAGGTTCCGAAATAGGCAGGGGGCATTAACT
GTTTATACGGGCACTGTTACTCAAGGCACTGACCCCGTTAAAACTTATTACCAGTAC
ACTCCTGTATCATCAAAAGCCATGTATGACGCTTACTGGAACGGTAAATTCAGAGAC
TGCGCTTTCCATTCTGGCTTTAATGAGGATTTATTTGTTTGTGAATATCAAGGCCAAT
CGTCTGACCTGCCTCAACCTCCTGTCAATGCTGGCGGCGGCTCTGGTGGTGGTTCTGG
TGGCGGCTCTGAGGGTGGTGGCTCTGAGGGTGGCGGTTCTGAGGGTGGCGGCTCTGA
GGGAGGCGGTTCCGGTGGTGGCTCTGGTTCCGGTGATTTTGATTATGAAAAGATGGC
AAACGCTAATAAGGGGGCTATGACCGAAAATGCCGATGAAAACGCGCTACAGTCTG

ACGCTAAAGGCAAACTTGATTCTGTCGCTACTGATTACGGTGCTGCTATCGATGGTTT
CATTGGTGACGTTTCCGGCCTTGCTAATGGTAATGGTGCTACTGGTGATTTTGCTGGC
TCTAATTCCCAAATGGCTCAAGTCGGTGACGGTGATAATTCACCTTTAATGAATAATT
TCCGTCAATATTTACCTTCCCTCCCTCAATCGGTTGAATGTCGCCCTTTTGTCTTTGGC
GCTGGTAAACCATATGAATTTTCTATTGATTGTGACAAAATAAACTTATTCCGTGGTG
TCTTTGCGTTTCTTTTATATGTTGCCACCTTTATGTATGTATTTTCTACGTTTGCTAACA
TACTGCGTAATAAGGAGTCTTAATCATGCCAGTTCTTTTGGGTATTCCGTTATTATTG
CGTTTCCTCGGTTTCCTTCTGGTAACTTTGTTCGGCTATCTGCTTACTTTTCTTAAAAA
GGGCTTCGGTAAGATAGCTATTGCTATTTCATTGTTTCTTGCTCTTATTATTGGGCTTA
ACTCAATTCTTGTGGGTTATCTCTCTGATATTAGCGCTCAATTACCCTCTGACTTTGTT
CAGGGTGTTCAGTTAATTCTCCCGTCTAATGCGCTTCCCTGTTTTTATGTTATTCTCTC
TGTAAAGGCTGCTATTTTCATTTTTGACGTTAAACAAAAAATCGTTTCTTATTTGGAT
TGGGATAAATAATATGGCTGTTTATTTTGTAACTGGCAAATTAGGCTCTGGAAAGAC
GCTCGTTAGCGTTGGTAAGATTCAGGATAAAATTGTAGCTGGGTGCAAAATAGCAAC
TAATCTTGATTTAAGGCTTCAAAACCTCCCGCAAGTCGGGAGGTTCGCTAAAACGCC
TCGCGTTCTTAGAATACCGGATAAGCCTTCTATATCTGATTTGCTTGCTATTGGGCGC
GGTAATGATTCCTACGATGAAAATAAAAACGGCTTGCTTGTTCTCGATGAGTGCGGT
ACTTGGTTTAATACCCGTTCTTGGAATGATAAGGAAAGACAGCCGATTATTGATTGG
TTTCTACATGCTCGTAAATTAGGATGGGATATTATTTTTCTTGTTCAGGACTTATCTAT
TGTTGATAAACAGGCGCGTTCTGCATTAGCTGAACATGTTGTTTATTGTCGTCGTCTG
GACAGAATTACTTTACCTTTTGTCGGTACTTTATATTCTCTTATTACTGGCTCGAAAAT
GCCTCTGCCTAAATTACATGTTGGCGTTGTTAAATATGGCGATTCTCAATTAAGCCCT
ACTGTTGAGCGTTGGCTTTATACTGGTAAGAATTTGTATAACGCATATGATACTAAAC
AGGCTTTTTCTAGTAATTATGATTCCGGTGTTTATTCTTATTTAACGCCTTATTTATCA
CACGGTCGGTATTTCAAACCATTAAATTTAGGTCAGAAGATGAAATTAACTAAAATA
TATTTGAAAAAGTTTTCTCGCGTTCTTTGTCTTGCGATTGGATTTGCATCAGCATTTAC
ATATAGTTATATAACCCAACCTAAGCCGGAGGTTAAAAAGGTAGTCTCTCAGACCTA
TGATTTTGATAAATTCACTATTGACTCTTCTCAGCGTCTTAATCTAAGCTATCGCTAT
GTTTTCAAGGATTCTAAGGGAAAATTAATTAATAGCGACGATTTACAGAAGCAAGGT
TATTCACTCACATATATTGATTTATGTACTGTTTCCATTAAAAAAGGTAATTCAAATG
AAATTGTTAAATGTAATTAATTTTGTTTTCTTGATGTTTGTTTCATCATCTTCTTTTGCT
CAGGTAATTGAAATGAATAATTCGCCTCTGCGCGATTTTGTAACTTGGTATTCAAAGC
AATCAGGCGAATCCGTTATTGTTTCTCCCGATGTAAAAGGTACTGTTACTGTATATTC
ATCTGACGTTAAACCTGAAAATCTACGCAATTTCTTTATTTCTGTTTTACGTGCAAAT
AATTTTGATATGGTAGGTTCTAACCCTTCCATTATTCAGAAGTATAATCCAAACAATC
AGGATTATATTGATGAATTGCCATCATCTGATAATCAGGAATATGATGATAATTCCG
CTCCTTCTGGTGGTTTCTTTGTTCCGCAAAATGATAATGTTACTCAAACTTTTAAAATT
AATAACGTTCGGGCAAAGGATTTAATACGAGTTGTCGAATTGTTTGTAAAGTCTAAT
ACTTCTAAATCCTCAAATGTATTATCTATTGACGGCTCTAATCTATTAGTTGTTAGTG
CTCCTAAAGATATTTTAGATAACCTTCCTCAATTCCTTTCAACTGTTGATTTGCCAACT
GACCAGATATTGATTGAGGGTTTGATATTTGAGGTTCAGCAAGGTGATGCTTTAGATT
TTTCATTTGCTGCTGGCTCTCAGCGTGGCACTGTTGCAGGCGGTGTTAATACTGACCG
CCTCACCTCTGTTTTATCTTCTGCTGGTGGTTCGTTCGGTATTTTTAATGGCGATGTTT
TAGGGCTATCAGTTCGCGCATTAAAGACTAATAGCCATTCAAAAATATTGTCTGTGC
CACGTATTCTTACGCTTTCAGGTCAGAAGGGTTCTATCTCTGTTGGCCAGAATGTCCC
TTTTATTACTGGTCGTGTGACTGGTGAATCTGCCAATGTAAATAATCCATTTCAGACG
ATTGAGCGTCAAAATGTAGGTATTTCCATGAGCGTTTTTCCTGTTGCAATGGCTGGCG
GTAATATTGTTCTGGATATTACCAGCAAGGCCGATAGTTTGAGTTCTTCTACTCAGGC
AAGTGATGTTATTACTAATCAAAGAAGTATTGCTACAACGGTTAATTTGCGTGATGG
ACAGACTCTTTTACTCGGTGGCCTCACTGATTATAAAAACACTTCTCAGGATTCTGGC
GTACCGTTCCTGTCTAAAATCCCTTTAATCGGCCTCCTGTTTAGCTCCCGCTCTGATTC

TAACGAGGAAAGCACGTTATACGTGCTCGTCAAAGCAACCATAGTACGCGCCCTGTA
GCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTG
CCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCC
GGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTT
TACGGCACCTCGACCCCAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGCCAT
CGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGG
ACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGGCTATTCTTTTGATTTA
TAAGGGATTTTGCCGATTTCGGAACCACCATCAAACAGGATTTTCGCCTGCTGGGGC
AAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATC
AGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAA
CCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCG
ACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGG
CACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGG
ATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGAATTCGAGCTCGGTAC
CCGGGGATCCTTATACGGGTACTAGCCATGCGTATACGGTCGCTAGCGGACTTGCCT
CGCTATCAAAGGTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCACTGGCCGTCGT
TTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGC
ACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTC
CCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGA
AGCGGTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGT
CCCCTCAAACTGGCAGATGCACGGTTACGATGCGCCCATCTACACCAACGTGACCTA
TCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGTTACTCG
CTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTT
GATGGCGTTCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAATGCGAATT
TTAACAAAATATTAACGTTTACAATTTAAATATTTGCTTATACAATCTTCCTGTTTTTG
GGGCTTTTCTGATTATCAACCGGGGTACATATGATTGACATGCTAGTTTTACGATTAC
CGTTCATCGATTCTCTTGTTTGCTCCAGACTCTCAGGCAATGACCTGATAGCCTTTGT
AGATCTCTCAAAAATAGCTACCCTCTCCGGCATTAATTTATCAGCTAGAACGGTTGA
ATATCATATTGATGGTGATTTGACTGTCTCCGGCCTTTCTCACCCTTTTGAATCTTTAC
CTACACATTACTCAGGCATTGCATTTAAAATATATGAGGGTTCTAAAAATTTTTATCC
TTGCGTTGAAATAAAGGCTTCTCCCGCAAAAGTATTACAGGGTCATAATGTTTTTGGT
ACAACCGATTTAGCTTTATGCTCTGAGGCTTTATTGCTTAATTTTGCTAATTCTTTGCC
TTGCCTGTATGATTTATTGGATGTT

Table S5 Barcode geometry measurement results

Results of automated barcode geometry measurement by analyzing one $71 \times 71\ \mu m^2$ image for each barcode species. The geometrical parameters are listed as $(mean \pm stdev)$. Each pixel is equal to 71 nm.

Barcode type	Barcode identified	Bending angle α (degree)	Larger distance (pixel)	Smaller distance (pixel)
BRG	70	165.11±10.08	6.18±0.76	3.77±0.74
GRG	81	161.51±11.73	6.12±0.95	3.95±0.98

Table S6 Decoding result of the 72-barcode pool

Computer aided barcode-counting result for the 72-barcode-species mixture. The barcodes are sorted by counts (from highest to lowest). The barcode species that were expected in the mixture are highlighted. In total 2,556 of the 2,617 registered barcodes are the expected species. On average there are 35.5 ± 8.2 $(mean \pm stdev)$ barcodes per expected species and only 1.4 ± 0.7 $(mean \pm stdev)$ barcodes per false positive species. B = Blue, G = Green, R = Red; '//' and '/' denote longer and shorter interzone distances in the barcode, respectively.

Barcode name	Counts	Barcode name	Counts
B//B/R	50	B//R/G	50
R//G/B	49	B//GB/R	48
GB//B/G	48	GB//R/GB	48
GB//RB/G	47	B//B/G	46
GB//G/B	45	R//B/G	45
B//G/B	44	G//G/B	44
RB//R/G	44	GB//B/RG	43
RG//R/B	43	G//GB/R	42
RB//RB/G	42	RG//B/RG	42
G//B/R	41	G//B/RG	41
RB//G/RB	41	RG//G/B	41
RG//G/RB	41	B//R/GB	40
G//B/G	40	R//G/RB	40
B//RB/G	39	B//R/B	38
RB//B/R	38	RG//RB/G	38
RB//B/G	37	RB//RG/B	37
RG//B/R	36	B//G/R	35
B//G/RB	35	G//G/RB	35
G//R/G	35	G//RB/G	35
GB//R/B	35	R//RB/G	35
RG//R/GB	35	G//G/R	34
RB//B/RG	34	RB//G/B	34
G//R/B	33	G//R/GB	33
R//G/R	33	R//RG/B	33
RG//GB/R	33	GB//G/RB	32
B//RG/B	31	GB//GB/R	31
GB//R/G	31	RB//R/GB	31
RG//RG/B	31	GB//RG/B	29
RG//B/G	29	R//GB/R	28
R//R/B	28	GB//B/R	27
R//R/G	27	R//R/GB	27
B//B/RG	26	RG//G/R	26
GB//G/R	24	R//B/R	24

RB//GB/R	24	RG//R/G	23
G//RG/B	21	R//B/RG	19
RB//R/B	17	RB//G/R	15
G//RB/R	4	B//RB/GB	3
RB//RB/B	3	B//GB/RG	2
B//RG/G	2	B//RG/GB	2
G//RG/RB	2	GB//GB/RG	2
GB//RG/GB	2	R//GB/G	2
RB//GB/RG	2	RG//RB/B	2
RG//RG/RB	2	B//B/B	1
B//G/GB	1	B//G/RG	1
B//GB/B	1	B//GB/RB	1
B//RB/B	1	G//GB/RG	1
G//R/RB	1	G//RG/GB	1
G//RG/R	1	GB//G/GB	1
GB//R/RB	1	GB//RB/B	1
GB//RB/RG	1	GB//RG/RB	1
R//B/GB	1	R//G/RG	1
R//RB/GB	1	R//RG/RB	1
RB//B/GB	1	RB//G/GB	1
RB//GB/B	1	RB//GB/G	1
RB//R/RB	1	RB//R/RG	1
RB//RB/RG	1	RB//RG/RB	1
RG//GB/B	1	RG//GB/G	1
RG//GB/RB	1	RG//RG/R	1

Table S7 Decoding result of the 216-barcode pool

Computer aided barcode-counting result for the 216-barcode-species mixture. The barcodes are sorted by counts (from highest to lowest). In total 7,243 qualified barcodes were registered. On average there are 33.5 ± 16.7 $(mean \pm stdev)$ barcodes per species. B = Blue, G = Green, R = Red; '//' and '/' denote longer and shorter interzone distances in the barcode, respectively.

Barcode name	Counts	Barcode name	Counts
B//G/G	114	B//B/B	102
G//B/B	99	B//G/B	95
B//B/G	86	G//G/B	81
G//B/G	81	G//G/G	71
B//GB/G	71	GB//G/G	70
GB//B/B	70	G//RB/R	69
R//G/B	63	B//RG/RG	62
B//RG/R	61	B//GB/B	61
GB//B/G	59	G//R/B	57
RG//B/B	57	R//G/G	55
RB//B/G	55	RB//G/G	54
R//B/B	53	B//R/R	52
B//GB/R	51	B//R/G	50
B//RG/G	50	RG//G/B	50
B//G/R	49	RB//GB/G	48
G//RG/B	47	RB//RB/G	46
GB//G/B	46	GB//GB/G	45
R//B/G	43	B//B/R	43
GB//GB/R	43	G//GB/R	42
RG//RB/R	42	RB//B/B	42
GB//RB/RG	42	G//G/R	41
G//RB/B	41	G//GB/G	41
B//R/RG	41	B//RB/R	41
B//RB/RG	41	RG//GB/B	41
RB//G/B	41	RB//RG/R	41
RB//RB/R	41	GB//B/R	41
G//GB/B	40	G//B/R	39
B//RG/B	39	RG//GB/RB	39
R//RG/B	38	R//GB/G	38
G//RG/RB	38	RB//GB/B	38
GB//RG/G	38	R//RB/R	37
R//GB/B	37	G//B/RG	37
B//GB/GB	37	GB//G/RG	37
G//R/R	36	RG//G/G	36
RB//GB/R	36	GB//R/R	36

GB//R/RG	36	R//RB/G	35
B//G/RG	35	B//RG/RB	35
RG//RB/RB	35	GB//RG/R	35
G//RG/R	34	G//GB/RB	34
G//RG/G	33	B//GB/RG	33
B//GB/RB	33	RG//B/G	33
RG//RB/B	33	RG//GB/G	33
GB//RB/B	33	GB//GB/RG	33
G//R/RB	32	G//B/RB	32
B//R/B	32	B//RB/G	32
RB//RG/G	32	GB//RG/RG	32
B//R/GB	31	RB//B/RG	31
GB//R/RB	31	GB//B/RG	31
R//G/GB	30	G//G/RB	30
RG//R/B	30	RG//R/RB	30
RG//B/R	30	RB//R/G	30
RB//G/R	30	GB//GB/B	30
R//R/G	29	R//B/R	29
B//B/RG	29	B//RG/GB	29
B//RB/GB	29	RG//R/R	29
RG//G/R	29	RB//RB/RG	29
RB//GB/RB	29	GB//RG/B	29
GB//RG/RB	29	GB//RG/GB	29
GB//GB/RB	29	R//RG/RB	28
G//RG/RG	28	G//RB/G	28
G//RB/RB	28	RB//R/R	28
GB//GB/GB	28	G//RB/RG	27
B//G/GB	27	B//B/GB	27
RG//B/GB	27	RG//RG/RB	27
RB//G/GB	27	RB//B/R	27
RB//B/GB	27	RB//RG/B	27
RB//RB/B	27	GB//R/GB	27
R//R/RG	26	R//G/R	26
R//RG/GB	26	R//GB/GB	26
G//G/RG	26	G//GB/GB	26
B//RB/RB	26	RG//GB/R	26
RB//GB/GB	26	GB//R/G	26
R//R/R	25	R//R/B	25
R//B/GB	25	R//RB/GB	25
G//R/G	25	RG//B/RB	25
RG//RG/B	25	RG//GB/GB	25

GB//R/B	25	GB//G/R	25
GB//G/RB	25	GB//B/GB	25
GB//RB/R	25	GB//RB/G	25
R//RG/G	24	G//G/GB	24
GB//RB/RB	24	R//RB/B	23
R//RB/RG	23	R//GB/R	23
RG//RG/G	23	RB//RG/RB	23
GB//G/GB	23	R//G/RB	22
G//R/GB	22	B//G/RB	22
RG//RB/GB	22	RB//G/RG	22
RB//RG/RG	22	GB//B/RB	22
GB//RB/GB	22	R//B/RG	21
G//B/GB	21	G//GB/RG	21
RG//R/GB	21	RG//B/RG	21
RG//RG/R	20	RG//RB/G	20
RB//RB/GB	20	RB//GB/RG	20
R//R/GB	19	R//RG/RG	19
B//RB/B	19	RG//GB/RG	19
RB//R/RG	19	RB//R/RB	19
RG//RB/RG	18	RB//B/RB	18
R//B/RB	17	B//B/RB	17
RG//G/RB	17	RG//RG/GB	17
RB//RG/GB	17	R//G/RG	16
R//RG/R	16	R//RB/RB	16
G//R/RG	16	G//RB/GB	16
RB//G/RB	16	RB//RB/RB	16
R//R/RB	15	R//GB/RB	15
G//RG/GB	15	B//R/RB	15
RG//R/RG	15	RG//G/GB	15
RB//R/B	15	R//GB/RG	14
RG//R/G	13	RG//G/RG	13
RB//R/GB	13	RG//RG/RG	12

Technical Notes

Barcode defects and possible improvement

The unqualified (both single- and dual-labeled-zone case) and incorrect (dual-labeled-zone case only) barcodes likely arose from assembly defects, sample damage during handling, and the overlapping of nanorods on the surface, which can be largely reduced by optimizing the sample preparation and imaging protocol. For example, fluorophores with higher quantum efficiency and less susceptibility to photobleaching could be used to improve imaging quality. Higher quality staples (e.g., enzymatically produced and HPLC purified) could be used to construct the barcode to enhance fluorescent labeling efficiency by minimizing the number of missing or truncated staple extensions in the binding zones. Barcodes could be purified under more gentle conditions (e.g., recover DNA from agarose gel through electro-elution instead of crush-and-soak[5] to reduce the possibility of breaking a dimer nanorod into two monomers. The purified barcodes could be chemically cross-linked to improve their thermal and mechanical stability. Redundant coding could be built into the barcode to generate more accurate and robust readouts.

False positive rate and unqualified barcodes

It is common practice in single-molecule detection techniques to discard unqualified data points in order to minimize false positive rate. For example, in flow cytometry, aggregated or dead cells are often times sorted out based on their light-scattering profile and are generally not included in the final data analysis. In comparison with Nanostring, where only "~20% of the detected molecules were counted", ~80% of the DNA-origami-based barcodes are qualified ones.

While the unqualified barcode may bind to the target and thus introduce false-negative signal (by max. 20% in our case), it is important to note that no barcode species gives rise to unqualified barcodes systematically more than other species do (cf. Fig 2 & 3). Therefore, although we cannot directly use the barcode counts as the number of target molecules, we should be able to compare the abundance of multiple targets without bias caused by our imaging probes.

Single- *vs.* dual-labeled-zone barcodes in robustness

In our current design, a dual-labeled zone carried 6 staple extensions for each fluorophore species. In contrast, 12 staples per fluorophore were used in a single-labeled zone. As a result, the dual-labeled-zone barcode consisted of dimmer spots that were more susceptible to damages such as photo bleaching. In this sense, one can conceptualize a single-labeled-zone barcode as a redundantly encoded dual-labeled-zone barcode. For instance, a single-labeled-zone barcode can still be recognized as the correct type when six fluorophores were missing from each zone, a scenario in which the dual-labeled-zone barcode could be disqualified or categorized as incorrect. In the latter case, it will increase the false positive rate. In principle, the false positive rate can be decreased by increasing the copy number of each fluorophore species in a dual-labeled zone.

Barcode cost

To make one type of diffraction-limited barcode (10 μL of 10 nM = 0.1 pmol), the material cost is the following (in US$):

Calculation is based on the listed price provide by the vendors: Bioneer, Inc; Integrated DNA Technologies, Inc and New England Biolabs, Inc.

Label-free oligonucleotides: \$0.1 per base, 4 nmol final product per oligo.
Dye-labeled 21-mer oligonucleotides: \$300 per oligonucleotide, 10 nmol final product per oligo.
M13mp18: \$31 for 10 μg (~ 4 pmol).

- Label-free oligonucleotides:
$(\text{number of bases}) \times (\text{synthesis cost per base}) \times (\text{DNA amount in pmol})/(\text{synthesized amount})$
which is $(7308 \times 2 + 21 \times 12 \times 3) \times 0.1 \times 0.5/4000 \approx 0.19$
- Fluorescently labeled oligonucleotides:
$(\text{average cost per oligo}) \times (\text{oligos per barcode}) \times (\text{DNA amount in pmol})/(\text{synthesized amount})$
which is $300 \times 6 \times 0.12/10000 \approx 0.02$
- Cost of scaffold if purchased (negligible if made in house):
$(\text{cost of m13 scaffold}) \times (\text{DNA amount in pmol})/(\text{purchased amount})$
which is $31 \times 0.1/4 \approx 0.78$

- In total: \$0.22 (if scaffold is made in house) or \$1.0 (if scaffold is purchased)

Therefore the fluorescently labeled oligonucleotides do not constitute the major part of the total cost. Note that only 10 distinct species of fluorescently labeled oligonucleotide is necessary to construct a library of 216 barcodes.

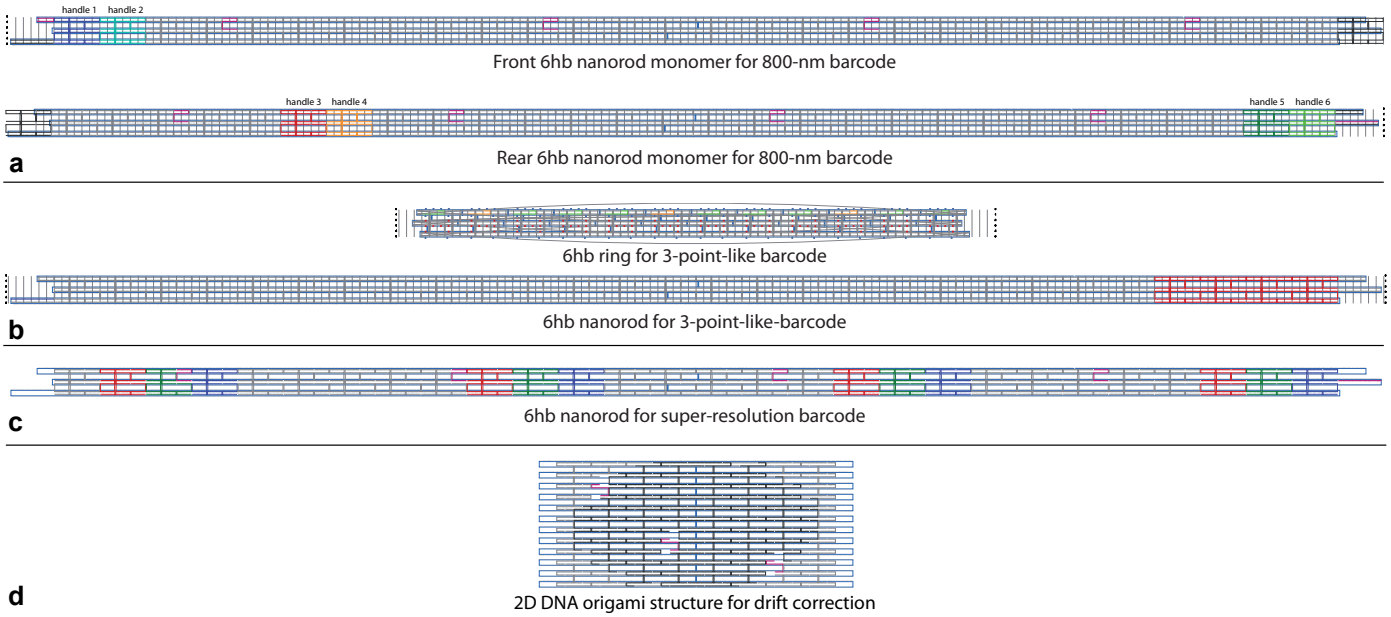


Fig. S1. Strand diagram of barcode design. The strands are colored according to the staple extensions in Table S1. (a) 800-nm-rod-based barcode (b) Three-point-star-like barcode. (c) 400 nm super-resolution barcode. (d) Single-layer DNA origami marker for super-resolution drift correction and alignment. (Zoom in to see details)

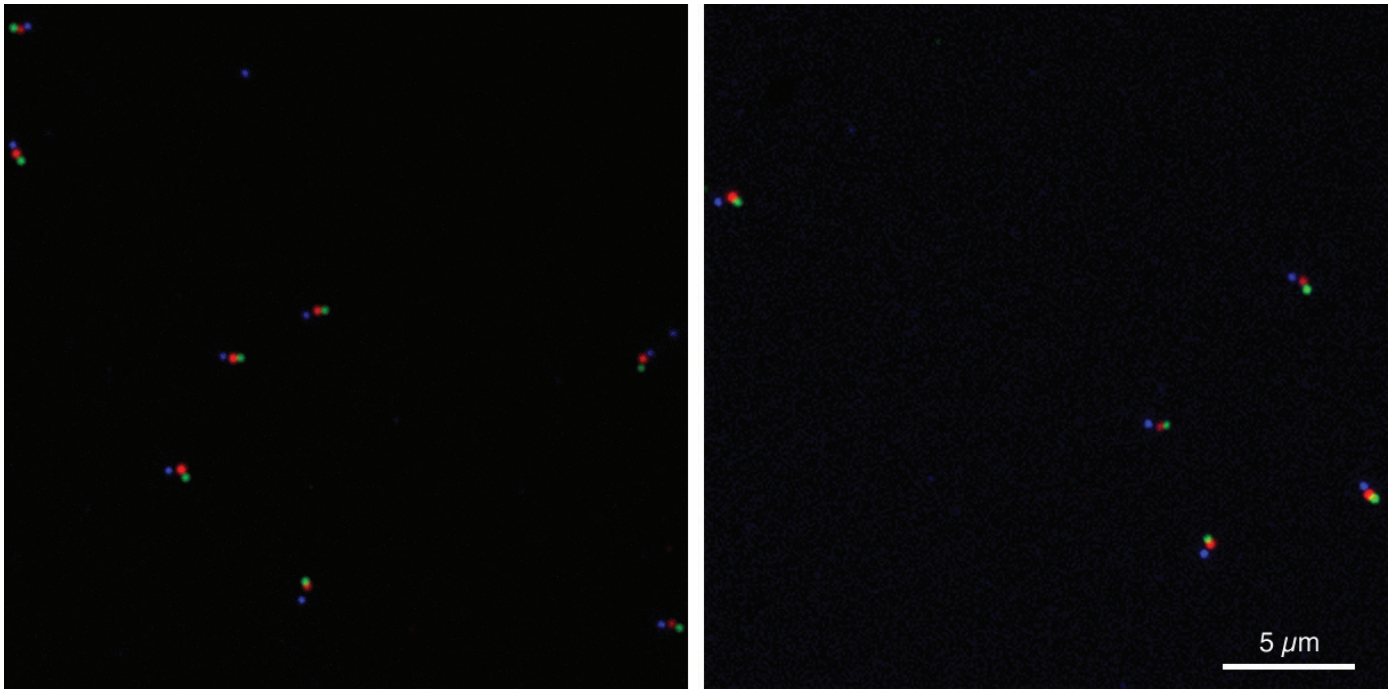


Fig. S2. Representative epi-fluorescence microscope image of BRG nano-barcode. Scale bar: 5 μm.

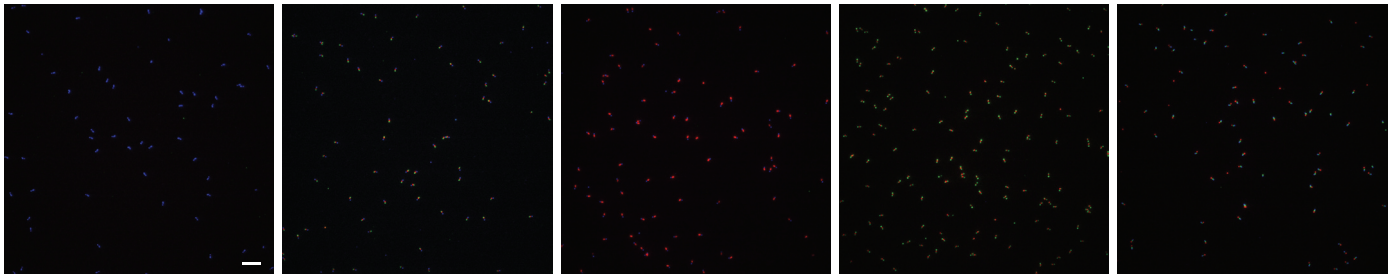


Fig. S3. Representative TIRF microscope images of the selected 5 single-labeled-zone nano-barcode. Scale bar: 5 μm. (This image can be downloaded separately in higher resolution)

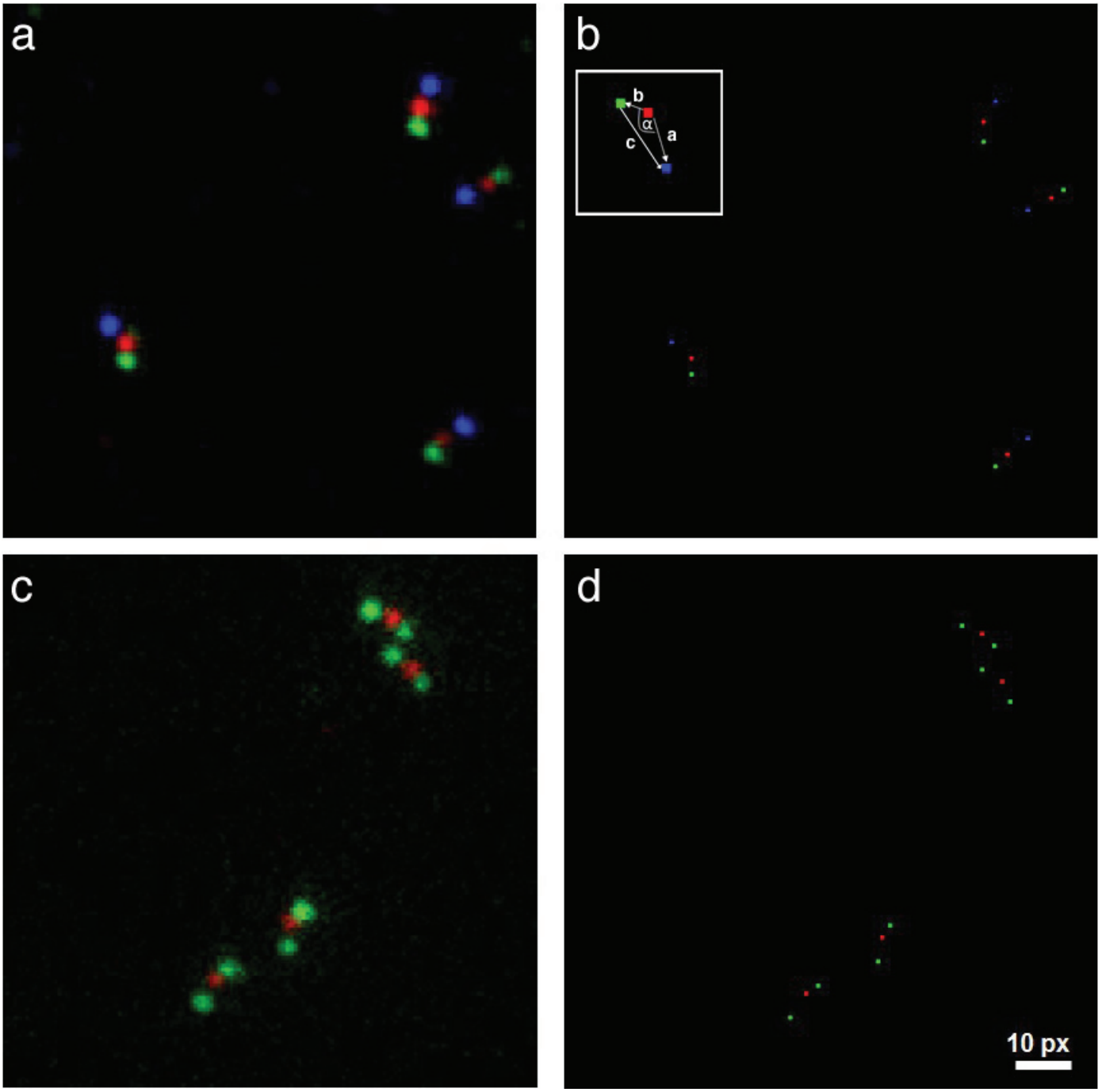


Fig. S4. Diagrams showing the automated barcode geometry measurement. (a) TIRF microscope image of BRG barcode. (b) The same image after computer processing. The centroid of each fluorescent spot is shown as one square (pixel). Inset: a schematic diagram to illustrate the distances and angles measured by the software to characterize the barcode geometry. (c) TIRF microscope image of GRG barcode. (d) The same image after computer processing. Scale bar: 10 pixel = 710 nm.

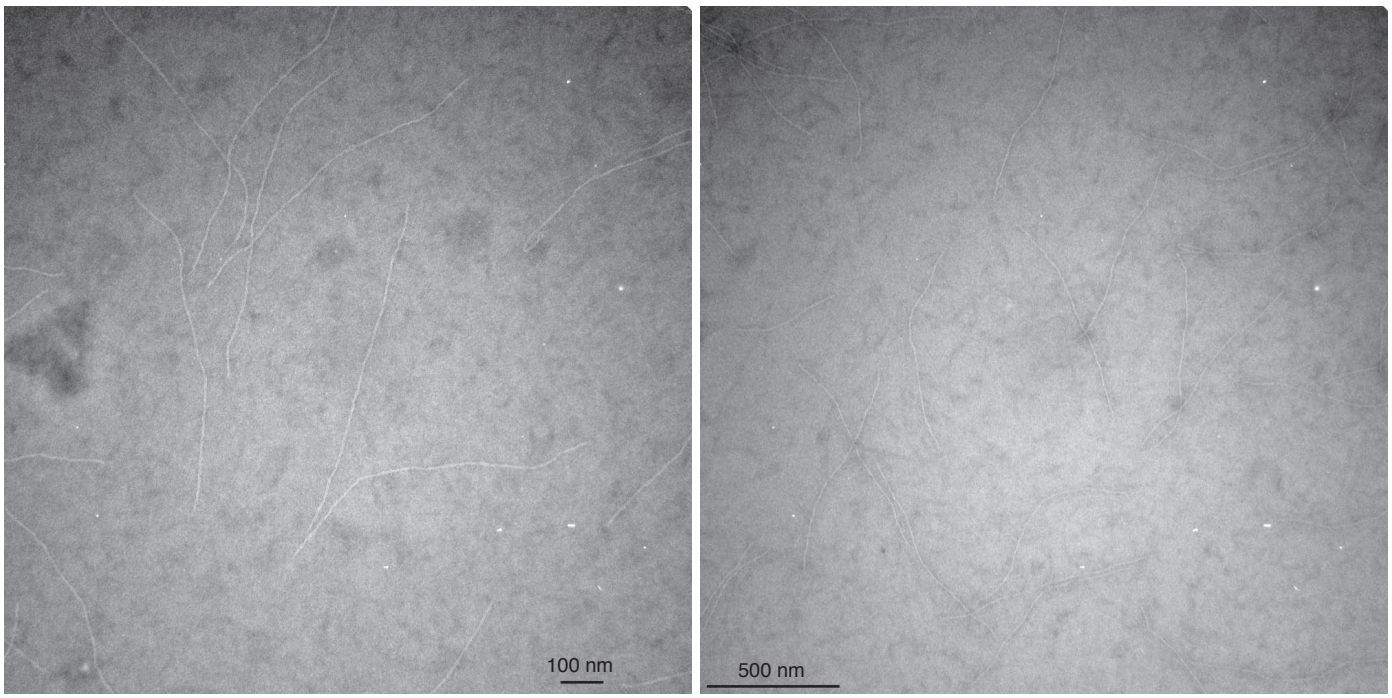


Fig. S5. Representative TEM image of the 800-nm-rod-based barcodes.

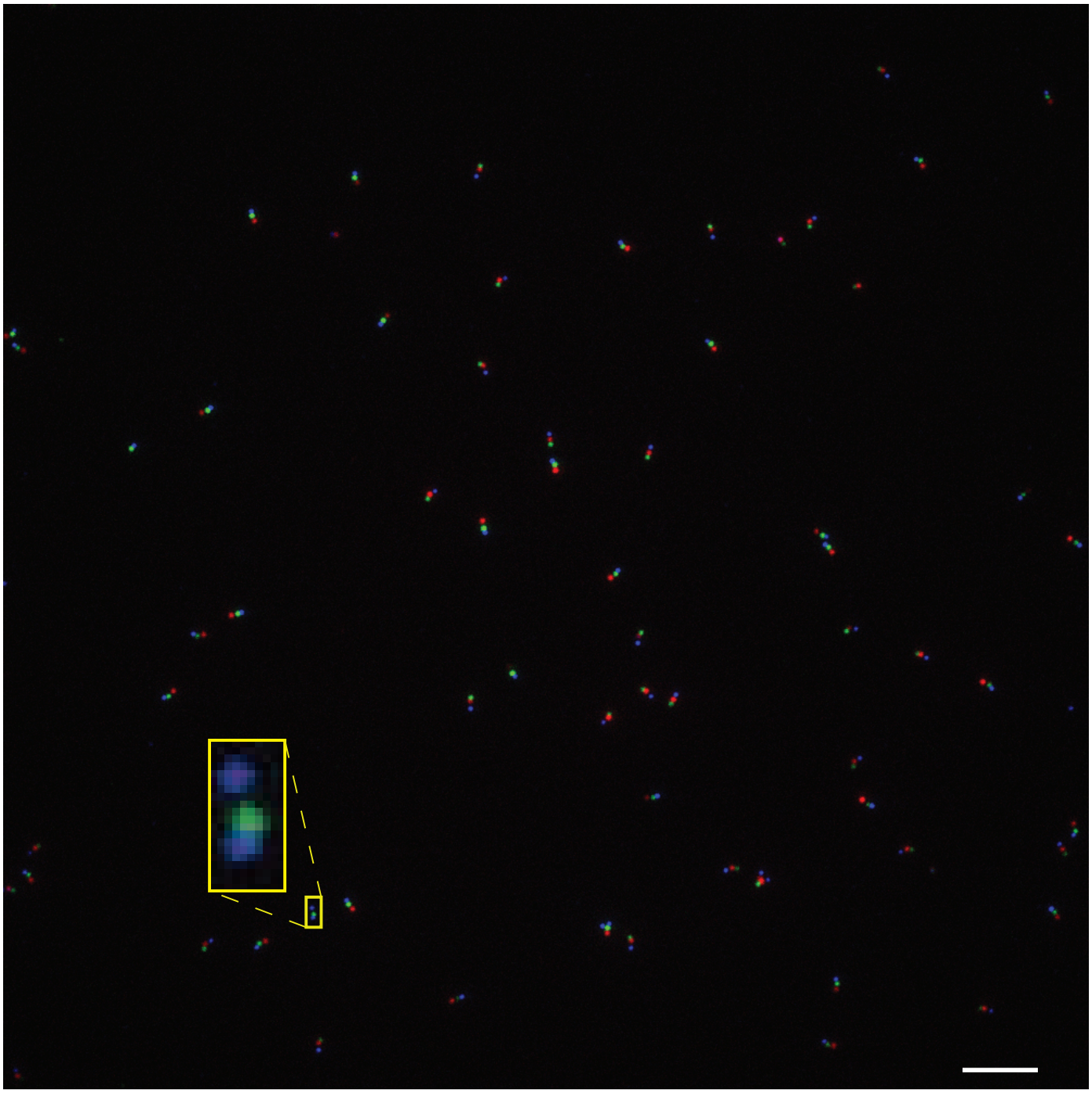


Fig. S6. Representative TIRF microscope image of 1:1 mixture of BRG and RGB barcode. Note the false-positive barcode highlighted by a yellow 1.4×0.7 μm^2 rectangle and zoomed-in $5\times$ in the inset. Scale bar: 5 μm.

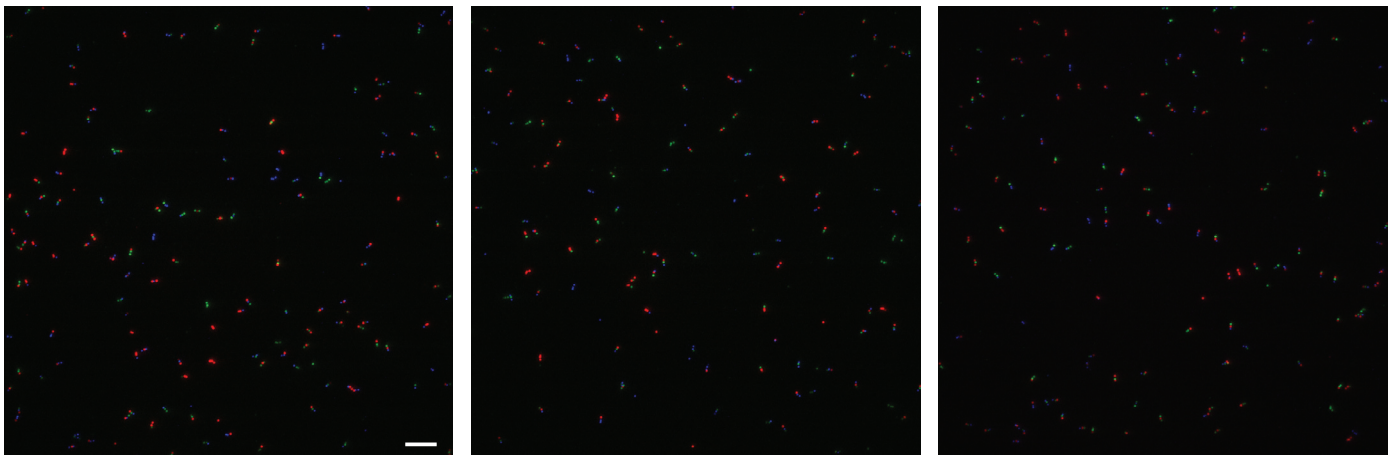


Fig. S7. Representative TIRF microscope images of an equimolar mixture of all 27 single-labeled-zone barcode family members. Scale bar: 5 μm. (This image can be downloaded separately in higher resolution)

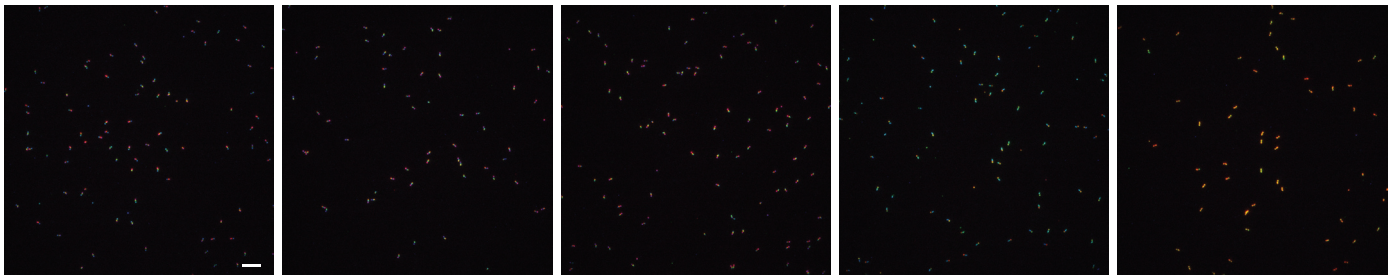


Fig. S8. Representative TIRF microscope images of the selected 5 dual-labeled-zone nano-barcode. Scale bar: 5 μm. (This image can be downloaded separately in higher resolution)

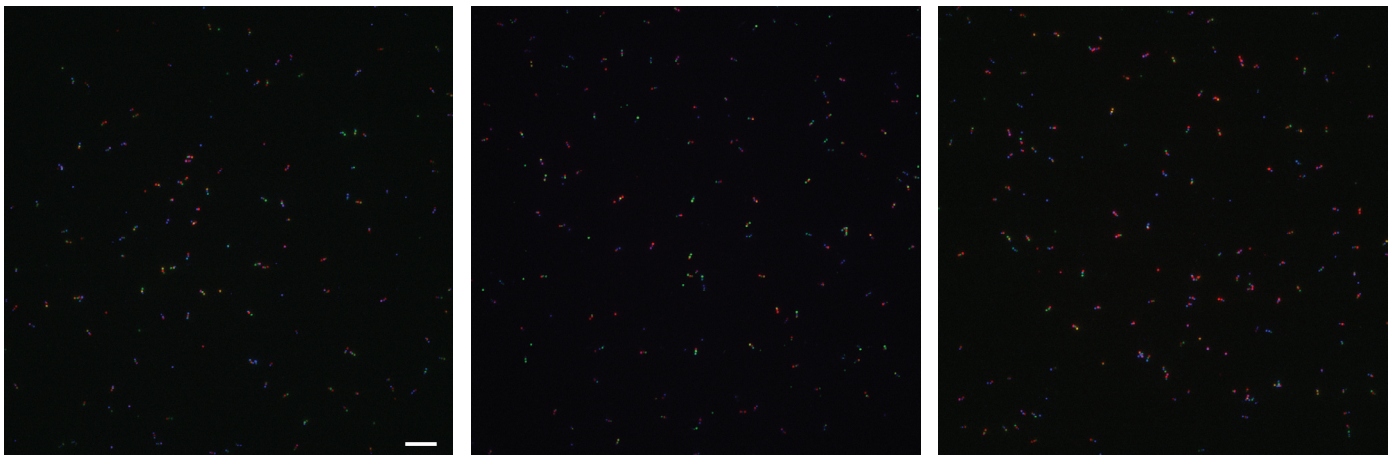


Fig. S9. Representative TIRF microscope images of an equimolar mixture of the selected 72 dual-labeled-zone barcode family members. Scale bar: 5 μm. (This image can be downloaded separately in higher resolution)

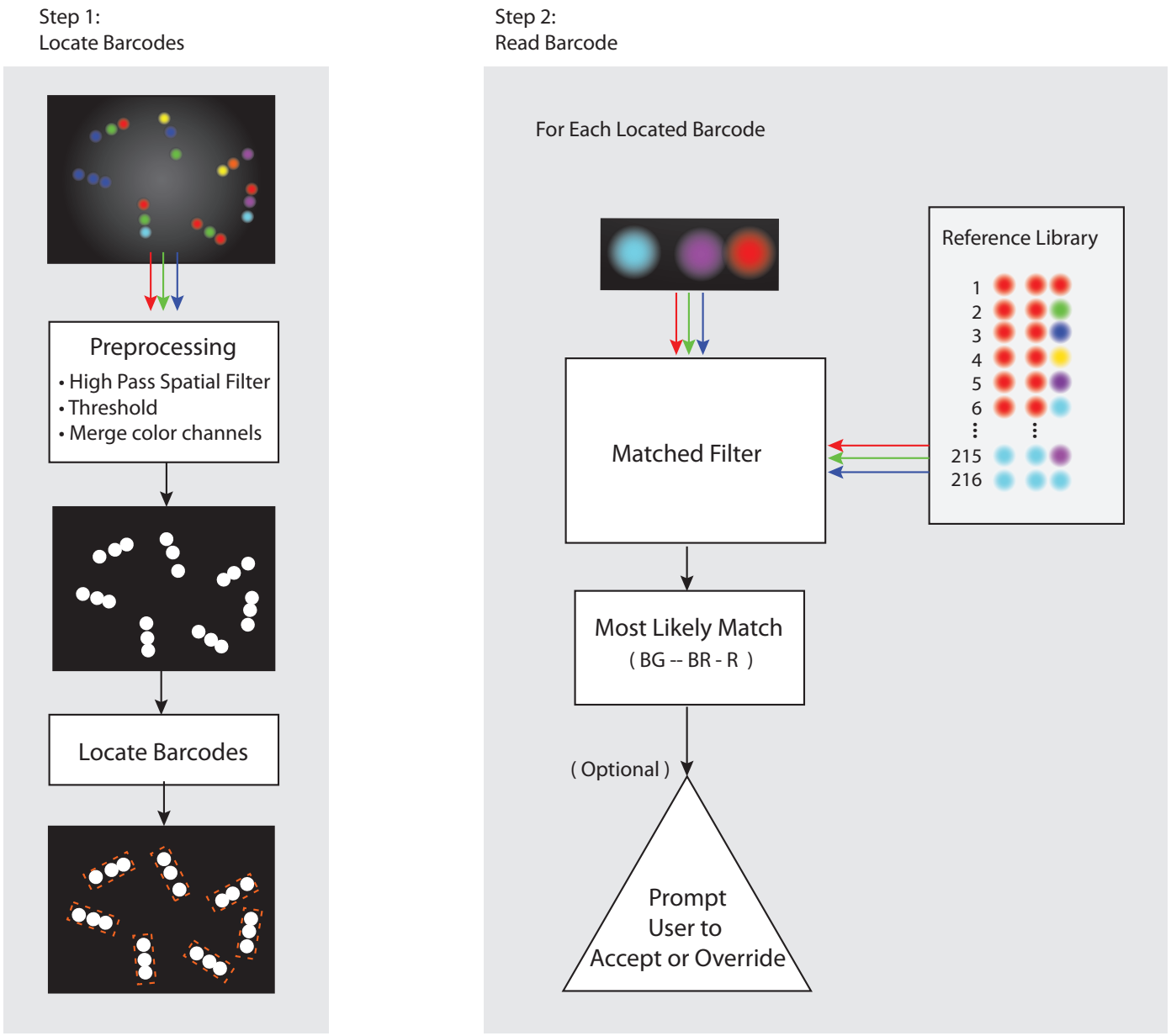


Fig. S10. Block diagram showing the principle of automated decoding software.

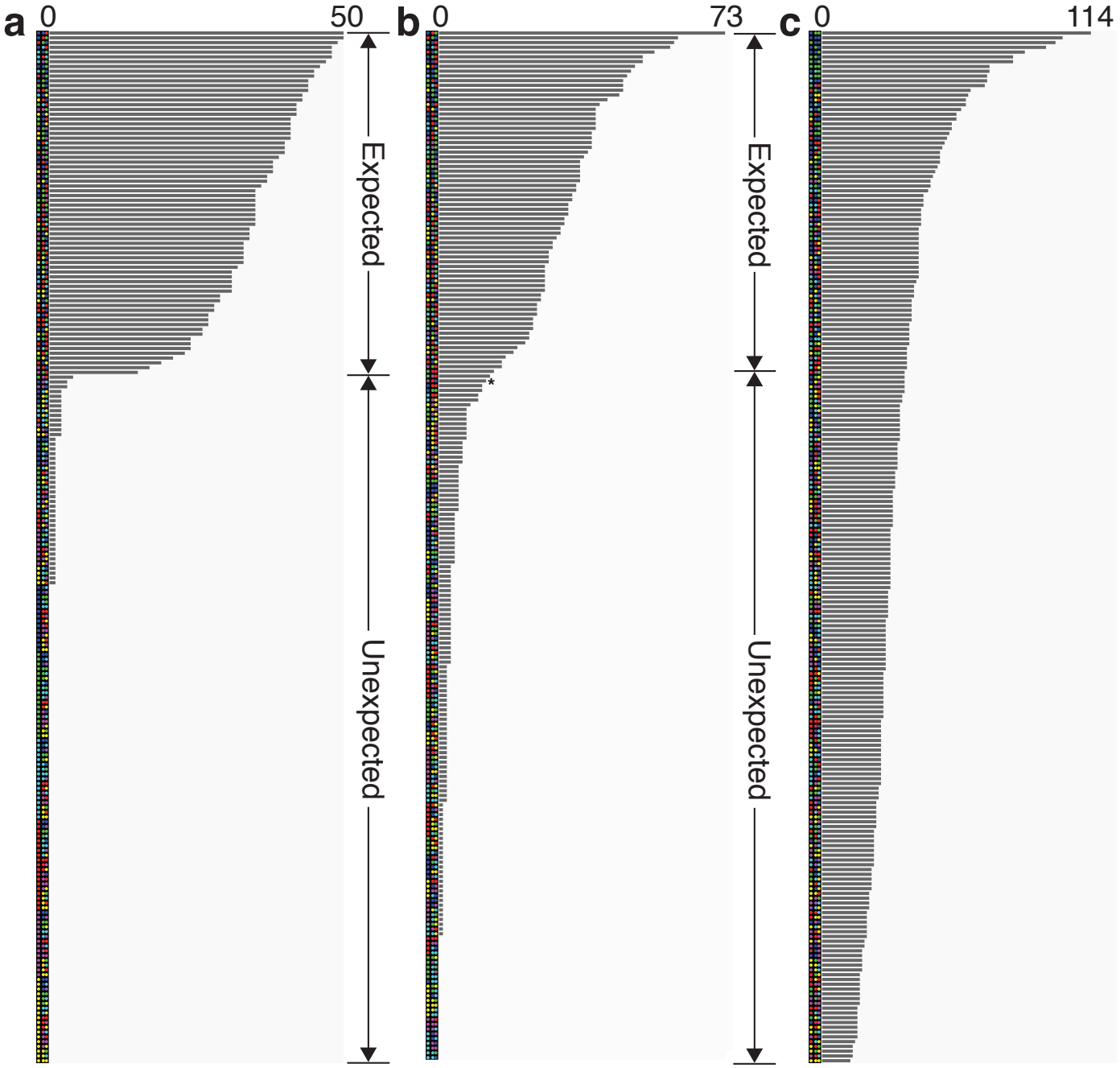


Fig. S11. Statistics obtained by analyzing mixtures of (a) 72 barcodes in supervised mode, (b) 72 barcodes in unsupervised mode, and (c) 216 barcodes in unsupervised mode. The barcode counts are listed in Table S6 and S7. The barcode types are shown as computer-generated ideal images and placed to the left of the corresponding bar-graphs. The barcodes are sorted in a descending order of counts. Annotations to the right of the bar-graphs were used to denote the expected and unexpected species in (a) and (b), with one exception in (b) where an expected barcode (BR--B-R) was denoted by *.

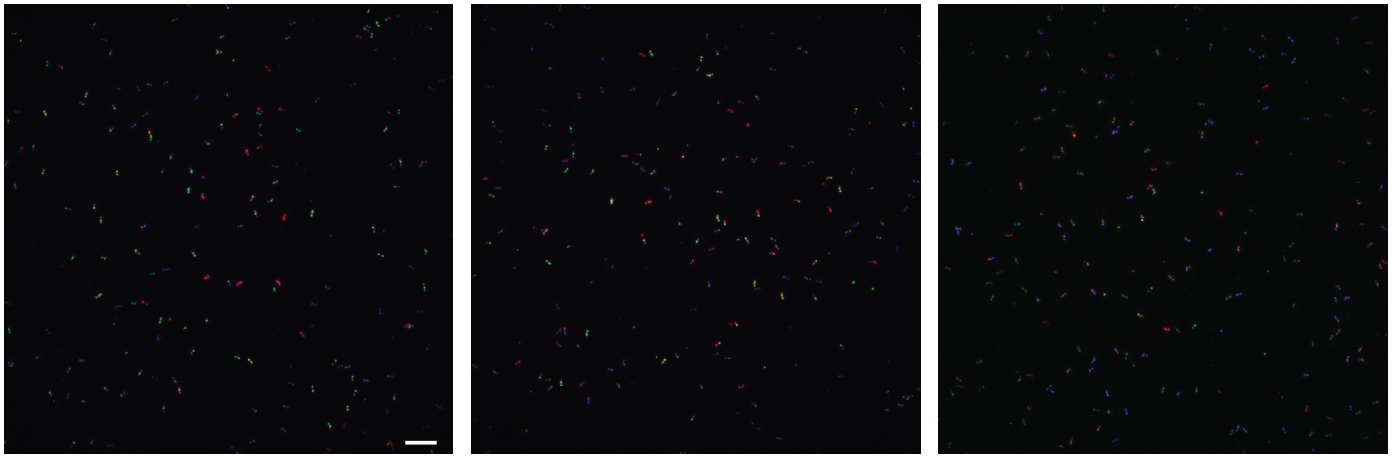


Fig. S12. Representative TIRF microscope images of an equimolar mixture of all 216 dual-labeled-zone barcode family members. Scale bar: 5 μm. (This image can be downloaded separately in higher resolution)

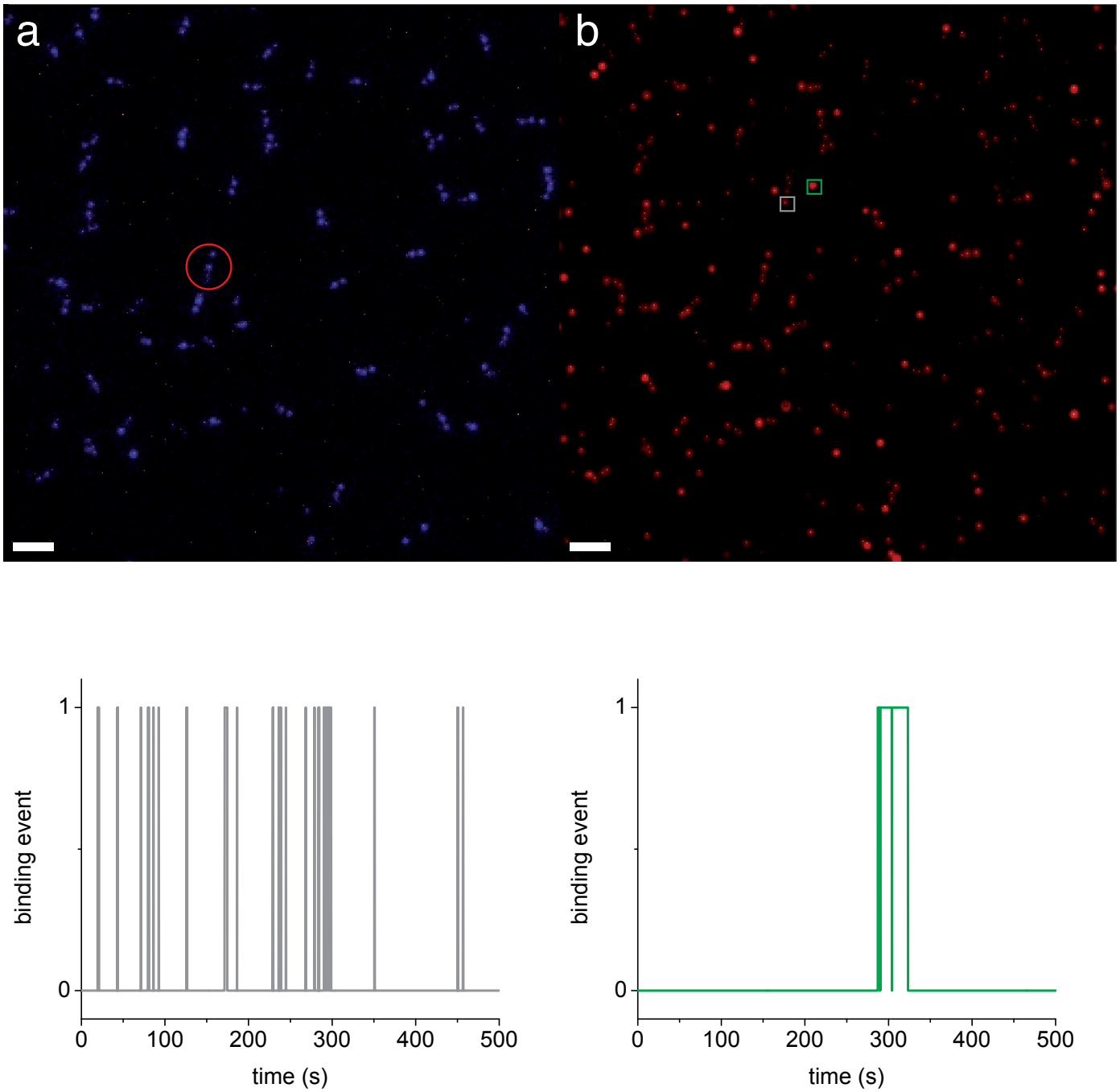


Fig. S13. Diffraction-limited and super-resolution overview images of 800 nm DNA origami-based barcodes. (a) Diffraction-limited image of nano-barcodes, statically labeled with Alexa488 together with a super-resolution reconstruction overlay obtained using DNA-PAINT. (b) Diffraction-limited image of the same sample region using DNA-PAINT and overlaid super-resolution reconstruction. Scale bars: 2 μm. The nano-barcodes are specifically imaged using DNA-PAINT. Some parts of the Alexa488 labeled nano-barcodes are already bleached or labeled with inactive fluorophores (see red highlighted area), but are correctly imaged using DNA-PAINT. The few non-specific spots in the DNA-PAINT image can be ruled out by investigating the binding frequency pattern, i.e. non-specific binding per area happens with a lower probability or even only once during imaging acquisition (cf. intensity vs. time plots for the grey - specific binding - and green - non-specific binding - highlighted regions).

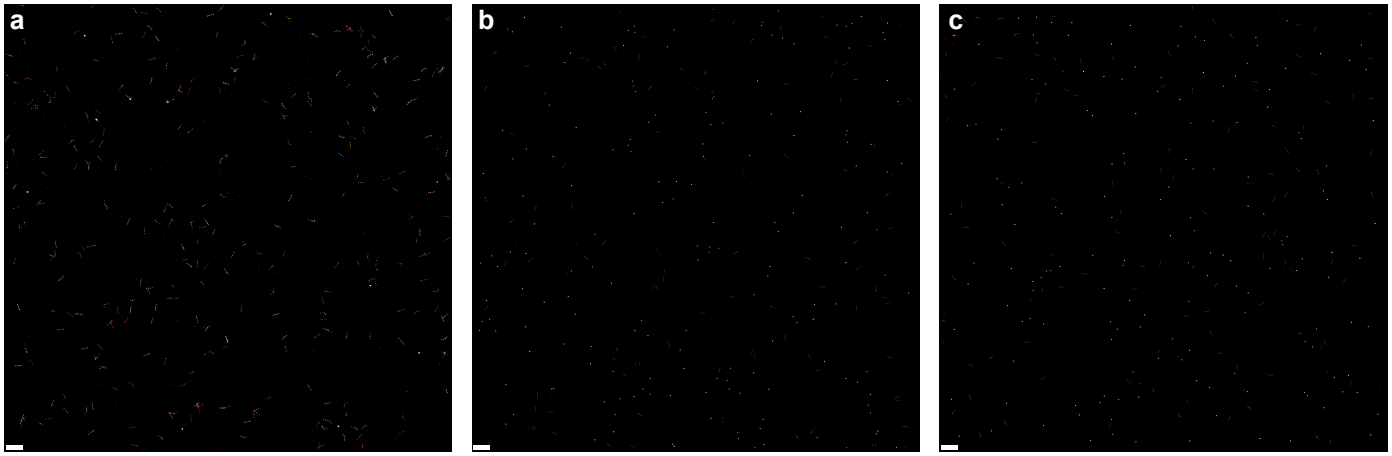


Fig. S14. Super-resolved DNA-PAINT zoomed-out images. Scale bars: 1 μm. (a) 3-color overview image for the 4-zone super-resolution barcode with 113 nm distance between binding zones. (b, c) Overview images for the 7-zone super-resolution barcodes with 70 nm and 42 nm distance between binding zones. (Zoom in to see details)

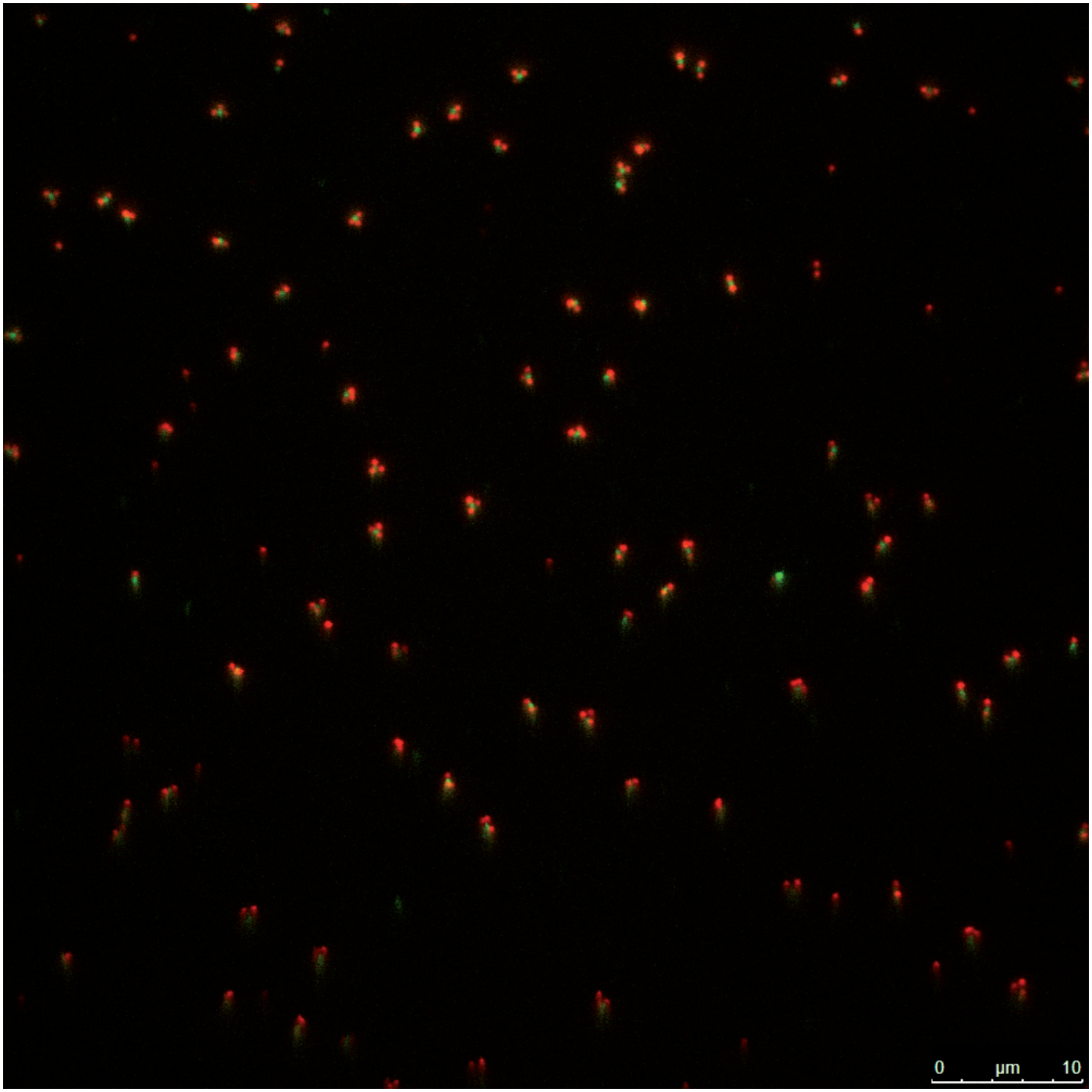


Fig. S15. Representative TIRF microscope images of the three-point-star-like barcodes. Scale bar: 10 μm.

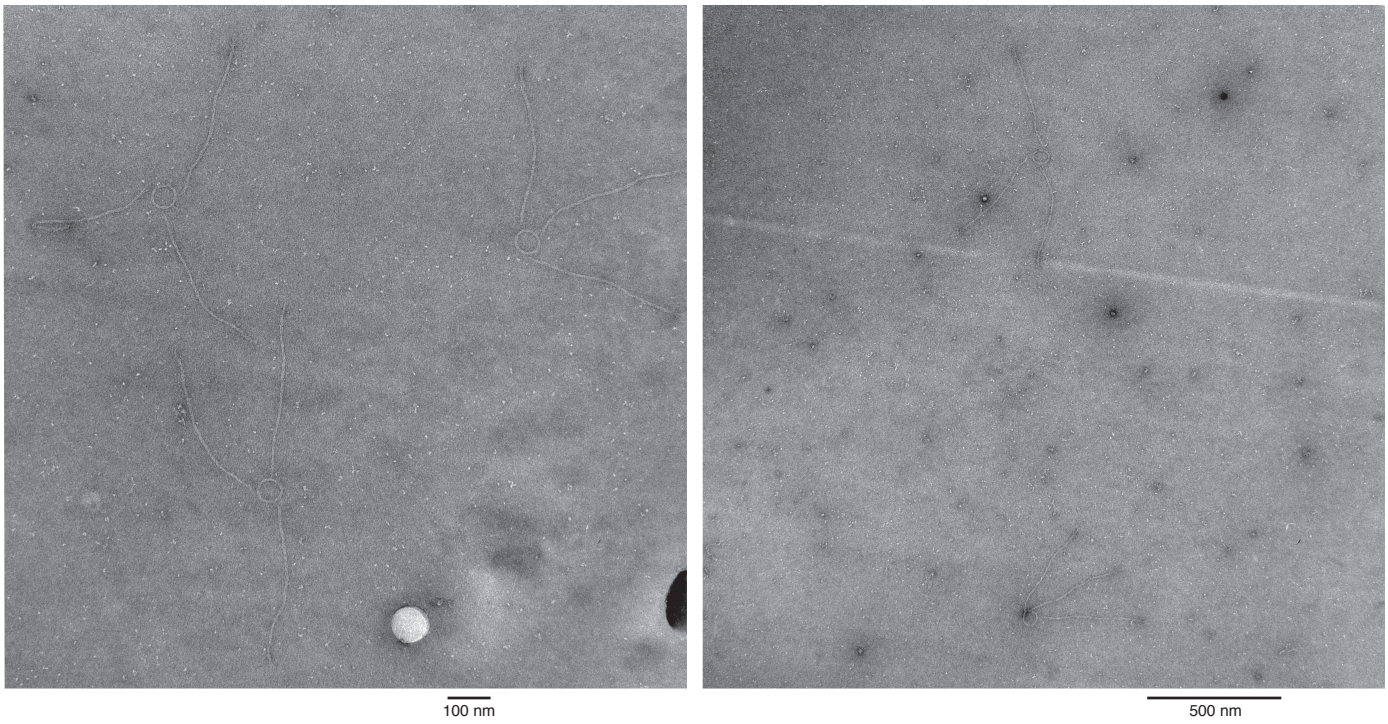


Fig. S16. Representative TEM images of the three-point-star-like barcodes.

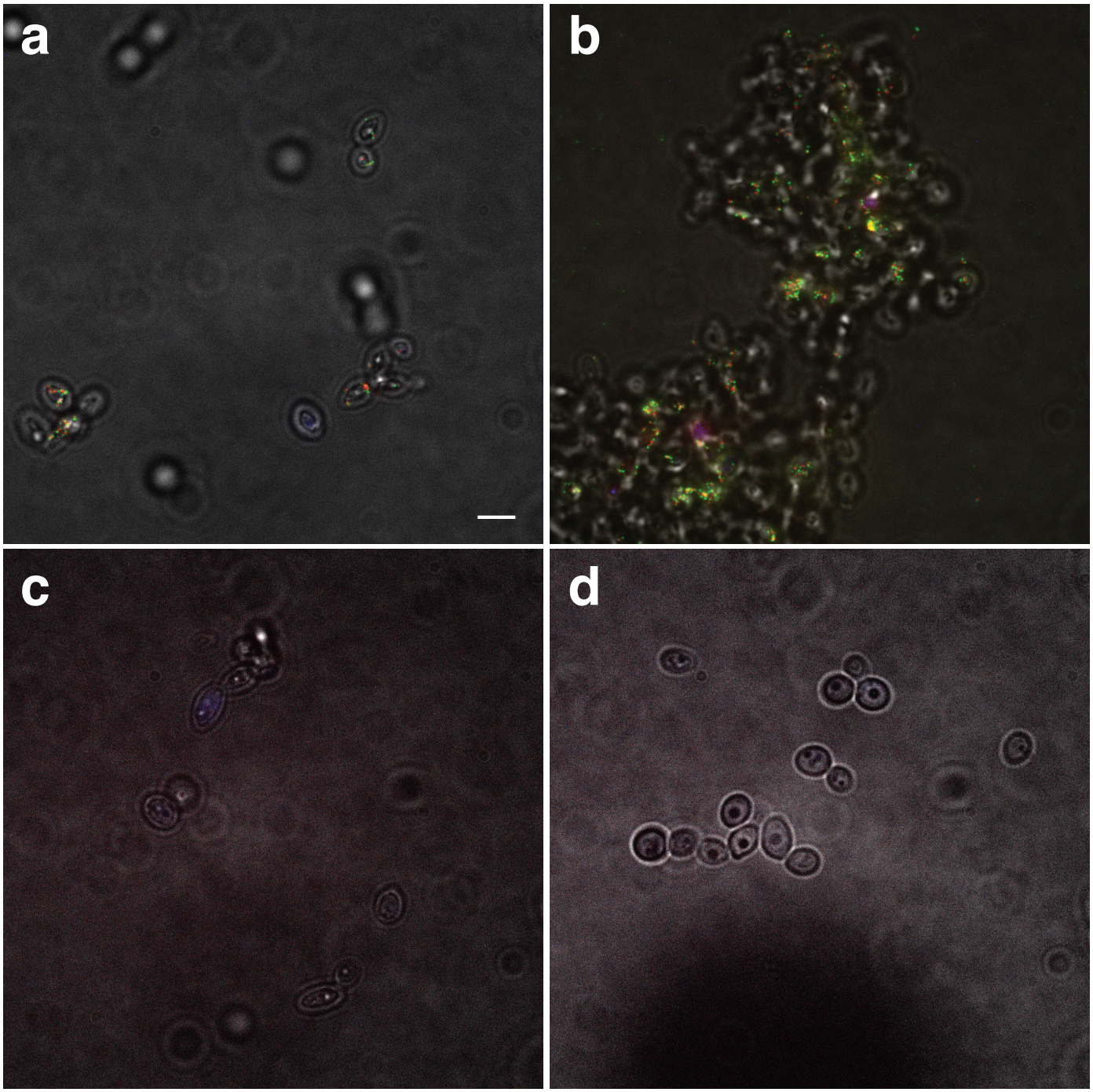


Fig. S17. Representative TIRF images showing the GRG-barcode-treated *C. albicans* cells. (a) GRG-barcode-tagged *C. albicans* cells (scattered). (b) GRG-barcode-tagged *C. albicans* cells (aggregated). (c) *C. albicans* cells treated with non-biotinylated antibodies. (d) *C. albicans* cells treated with non-biotinylated GRG barcodes. Scale bar: 5 μm.

References

1 Lis, J. T. & Schleif, R. Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucleic Acids Res.* **2**, 383-389 (1975).

2 Kurien, B. T. & Scofield, R. H. Extraction of nucleic acid fragments from gels. *Anal. Biochem.* **302**, 1-9 (2002).

3 Steinhauer, C., Forthmann, C., Vogelsang, J. & Tinnefeld, P. Superresolution Microscopy on the Basis of Engineered Dark States. *J. Am. Chem. Soc.* **130**, 16840-16841 (2008).

4 Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297-302 (2006).

5 Bellot, G., Mcclintock, M. A., Lin, C. & Shih, W. M. Recovery of intact DNA nanostructures after agarose gel-based separation. *Nature methods* **8**, 192-194 (2011).