Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA

Chenxiang Lin[1,2,3], Ralf Jungmann[1,4], Andrew M. Leifer[5†], Chao Li[1,6], Daniel Levner[1,6], George M. Church[1,6], William M. Shih[1,2,3]★ and Peng Yin[1,4]★

The identification and differentiation of a large number of distinct molecular species with high temporal and spatial resolution is a major challenge in biomedical science. Fluorescence microscopy is a powerful tool, but its multiplexing ability is limited by the number of spectrally distinguishable fluorophores. Here, we used (deoxy)ribonucleic acid (DNA)-origami technology to construct submicrometre nanorods that act as fluorescent barcodes. We demonstrate that spatial control over the positioning of fluorophores on the surface of a stiff DNA nanorod can produce 216 distinct barcodes that can be decoded unambiguously using epifluorescence or total internal reflection fluorescence microscopy. Barcodes with higher spatial information density were demonstrated via the construction of super-resolution barcodes with features spaced by ∼40 nm. One species of the barcodes was used to tag yeast surface receptors, which suggests their potential applications as *in situ* imaging probes for diverse biomolecular and cellular entities in their native environments.

luorescence microscopy is a versatile tool used to visualize nanometre- to micrometre-sized entities. To study multiple species of interest, it is essential to develop multiplexed fluorescent tags (barcodes). Most existing fluorescent barcodes are constructed using either intensity encoding (for example, microbeads with anisotropically embedded fluorophores[1–5], cells with combinatorially expressed fluorescent proteins[6], nanoarrays that consist of different fluorescent (deoxy)ribonucleic acid (DNA) tiles[7] and messenger ribonucleic acid hybridized to fluorescent probes[8]) or geometrical encoding (for example, inorganic particles with optical features[9–14] and nucleic acid double helices with tandem fluorescent labels[15,16]). Intensity encoding relies on the combination of multiple spectrally differentiable fluorophores in a controlled molar ratio. Geometrical encoding, however, is obtained by separating optical features beyond the microscope's resolution limit (typically ∼250 nm for diffraction-limited imaging and ∼25–40 nm for current super-resolution imaging[17]) and arranging them in a specific geometric pattern. The multiplexing capability of geometrically encoded barcodes increases exponentially as additional spatially distinguishable fluorophores are incorporated. Thus, with a structurally stiff scaffold capable of defining the spatial arrangement of the fluorescent molecules, combinatorially large barcode libraries may be constructed.

Despite the remarkable success in synthesizing fluorescent barcodes for *in vitro* multiplexed detection[1–5,9–16], little effort has been made to create robust single-molecule barcodes suitable as *in situ* imaging probes. Additionally, most existing fluorescent barcodes are in the range 2–100 μm in size and the construction of barcodes with smaller dimensions remains challenging, with only a few reports[3,7,8,12] of submicrometre barcodes and no more than 11 distinct barcodes demonstrated experimentally[8]. Here, we report a group of geometrically encoded fluorescent barcodes self-assembled from DNA. These barcodes are 400–800 nm in
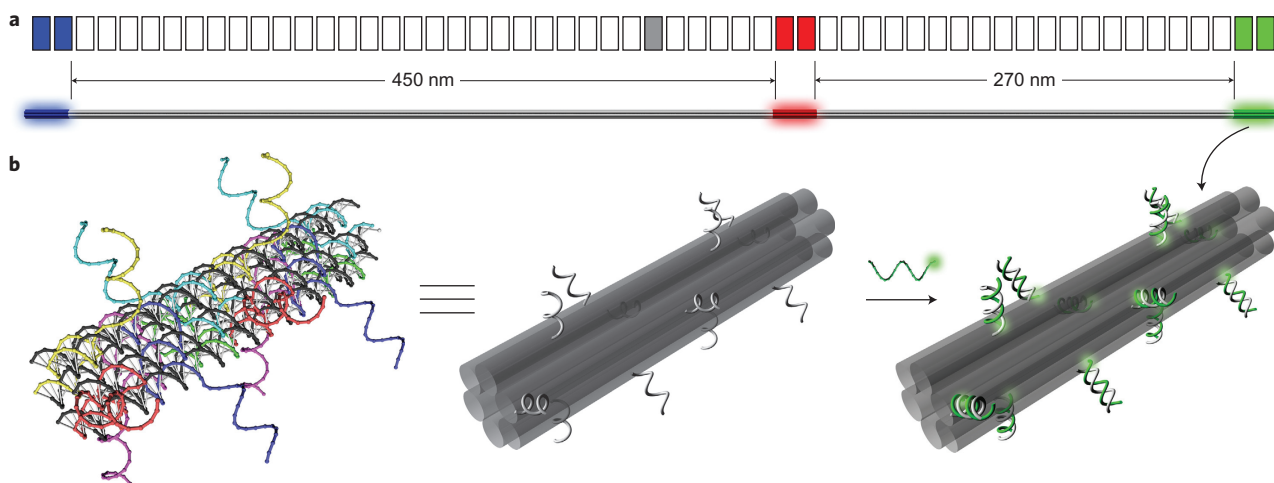
length, structurally stiff, reprogrammable in a modular fashion and easy to decode using epifluorescence or total internal reflection fluorescence (TIRF) microscopy. As evidence of their multiplexing power, 216 distinct barcode species were constructed and resolved using diffraction-limited TIRF microscopy Unlike the NanoString nCounter[15], no stretching step is necessary. The separation between the fluorescent spots was created exclusively by the stiffness of the DNA nanorod. Furthermore, barcodes with fluorescent features spaced below the diffraction limit were resolved using super-resolution microscopy. Finally, one species of the barcodes was used to tag yeast surface receptors, which suggests their potential application as *in situ* imaging probes.

Structural DNA nanotechnology exploits the well-defined double-helical structure of DNA and the predictable Watson–Crick base-paring rules to self-assemble designer nano-objects and devices[18–23]. A particularly effective method is DNA origami[24–30]. By folding a long, single-stranded 'scaffold strand' using many short synthetic 'staple strands', this approach generates complex, shape-controlled, fully addressable nanostructures with sizes of up to hundreds of nanometres. By functionalizing selected staple strands, such nanostructures can be used to organize fluorescent guest molecules spatially, including small organic molecules[31–33], and metallic[34] and semiconductive nanoparticles[35]. These properties make DNA origami a promising platform on which to build robust fluorescent barcodes, as the control over the exact ratio of different fluorophores allows intensity encoding and the spatial positioning of the fluorophores facilitates geometrical encoding, which can help to minimize undesired interfluorophore quenching.

Fluorescent barcodes with single-labelled zones
We first designed a family of 27 barcodes based on six-helix bundle DNA nanorods[36] that were ∼800 nm long (Fig. 1a). Three zones of

[1]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, Massachusetts 02115, USA, [2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA, [3]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA, [4]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA, [5]Program in Biophysics, Harvard University, Cambridge, Massachusetts 02138, USA, [6]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA; [†]Present address: Lewis Sigler Institute for Integrative Genomics, Princeton University, Princeton, New Jersey 08544, USA. *e-mail: py@hms.harvard.edu; william_shih@dfci.harvard.edu

Figure 1 | Design of a barcode based on a DNA nanorod. a, Two schematics of the blue//red/green (BRG) barcode with a segment diagram on the top and a three-dimensional (3D) view at the bottom ('//' and '/' denote longer and shorter interzone distances in the barcode, respectively). The main body of the barcode is a DNA nanorod formed by dimerizing two origami monomers, each consisting of 28 segments of 42 bp (~14 nm in length). The grey segment in the middle represents the junction where the two monomers join together. Three zones of the nanorod of 84 bp were labelled fluorescently (shown as blue, red and green segments) to produce the BRG barcode. **b**, 3D illustrations to show the details of one fluorescently labelled zone. The left diagram shows a strand model of an 84 bp zone before labelling. Each of the twelve 63-base staples (rainbow colours) contains two parts: the 42-base region at the 5′-end to fold the scaffold (black) into a six-helix bundle nanorod and the 21-base extension at the 3′-end protrudes out for fluorescent labelling. The middle diagram shows a simplified model to emphasize the six-helix bundle structure (each helix is shown as a semitransparent grey cylinder) and the positioning of the 12 staple extensions (light-grey curls). The diagram on the right shows the zone with 'green' fluorescent labelling. The labelling is achieved by hybridizing the Cy3 (glowing green spheres)-modified strands to the staple extensions.
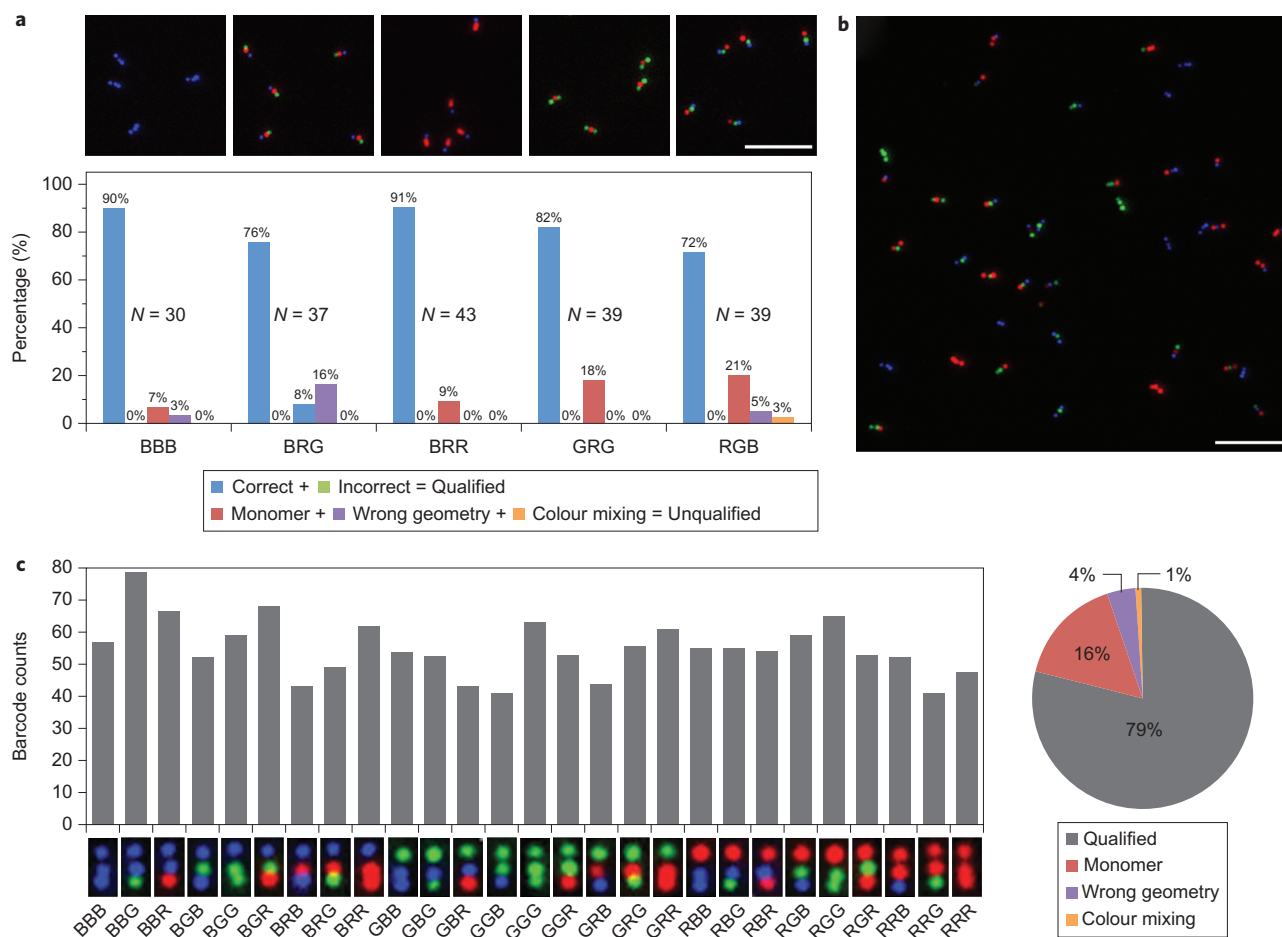
84 base pairs (bp; ~28 nm) on the nanorod were selected for fluorescent labelling, with interzone distances of 450 nm and 270 nm between the first and the last two zones, respectively. The fluorescently labelled zones were arranged spatially to generate an asymmetric fluorescent pattern that is decodable by diffraction-limited microscopy. For example, labelling the three zones (from left to right in Fig. 1a) with 'blue' (B, Alexa Fluor 488), 'red' (R, Alexa Fluor 647) and 'green' (G, Cy3) fluorophores (the pseudocolours reflect the excitation wavelength of the fluorophores) resulted in a BRG barcode that should be distinguishable from a GRB barcode. Therefore, $3^3 = 27$ different barcodes can be made from three spectrally distinguishable fluorophores. Within each zone (Fig. 1b), fluorescently modified oligonucleotides were hybridized to the twelve 21-base staple extensions (~6 nm between adjacent fluorophores) that protruded out from the main body of the nanorod. To facilitate imaging, ten additional staple strands were designed with 5′-biotinylated extensions to enable surface attachment (Supplementary Fig. S1, Tables S1–S4).

We assembled DNA-origami nanorods following the published protocol[36] with slight modifications (Supplementary Methods and Materials). To validate our system, we randomly chose five distinct barcodes for quality-control experiments. Two distinct features of these barcodes were clearly visible from the TIRF and epifluorescence images (Fig. 2a (top) and Supplementary Figs S2,S3): first, each fluorescent zone was resolved as a single-colour spot and each complete barcode consisted of three such spots; second, two of the neighbouring spots were separated by a small gap with the other two neighbours sat closely together. Therefore, we can recognize visually the geometrically encoded barcodes based on the colour identity of the spots and their relative spatial positions, even without the aid of any specialized decoding software. Computer-aided analysis of the BRG barcodes (Supplementary Fig. S4, Table S5) measured the average centre-to-centre distance between the neighbouring spots to be 433 ± 53 nm (mean $\pm$ standard deviation (s.d.), $N = 70$, longer distance) and 264 ± 52 nm ($N = 70$, shorter distance), which confirmed the

correct formation of the barcodes. The small discrepancy between these experimentally measured distances and the designed values (478 nm and 295 nm, respectively) may be attributed to random thermal bending of the nanorods (persistence length of ~1–2 μm (ref. 29); see Supplementary Fig. S5). Also, the spot intensities were not perfectly uniform across the whole image, which may be the result of factors such as uneven illumination of the sample and differences in labelling efficiency. Nevertheless, the TIRF images showed that the barcodes assembled as designed and were resolved unambiguously.

We then manually investigated TIRF images with an area of 50×50 μm^2 for five selected barcode species (Fig. 2a, bottom). The objects in the images were first sorted into qualified (that is, three single-colour spots arranged in a nearly linear (with bending angle $\geq 120°$, see Supplementary Fig. S4) and asymmetric fashion, as designed) and unqualified (that is, all other objects) barcodes. A qualified barcode was further categorized as correct or incorrect (false positive) based on whether the geometric pattern of its constituent fluorescent spots corresponded to the designed type. The unqualified barcodes were sorted further into (1) monomer nanorods (single spot or two touching spots), (2) barcodes with the 'wrong' geometry (that is, extreme bending) and (3) barcodes that contained at least one spot with multiple colours. On average, >80% of the visible objects were determined as qualified barcodes (that is, $\leq$20% false negative out of 188 observed objects) and these qualified barcodes were all correct (that is, 0% false positive out of 154 qualified barcodes (see the Supplementary Technical Notes)).

For many applications it is necessary for several different barcode species to coexist in one pool. Thus, it is important to examine the performance of our system when different barcode species are mixed. In an initial test, BRG and RGB barcodes were synthesized separately, mixed together at an equal molar ratio and co-purified via gel electrophoresis. The TIRF analysis (Supplementary Fig. S6) confirmed the 1:1 stoichiometry of the two barcodes and an overall assembly success rate (qualified barcode/all objects) of ~80%, which suggests that both barcodes maintained their integrity

Figure 2 | Fluorescent barcodes with single-labelled zones. a, Superimposed TIRF microscopy images of five barcode species (top) and the statistics from manual counting (bottom). The barcode types are given below the horizontal axis of the diagram. Each bar graph was generated by manually counting the objects in a $50 \times 50\ \mu m^2$ image (~40 barcodes; the exact sample size N is given next to the corresponding bar graph). **b**, A representative image of the equimolar mixture of 27 barcode species. **c**, Statistics obtained by analysing 27 images ($50 \times 50\ \mu m^2$) of the 27-barcode mixture (1,485 barcode molecules). The bar chart shows counts of the 27 species (55 ± 9, mean $\pm$ s.d.). A representative TIRF image ($1.4 \times 0.7\ \mu m^2$) of each barcode type is placed below the corresponding bar. The pie chart shows the object-sorting result. The colour scheme used for the bar graphs in (**a**) and the pie chart in (**c**) is: blue, correct barcodes (qualified barcode with expected identity); green, incorrect barcodes (qualified barcode with unexpected identity); red, monomer nanorods (one spot or two touching spots); purple, barcodes with the wrong geometry (bending angle $<120°$, see Supplementary Methods); orange, barcodes that contain at least one spot with two colours. In the 27-barcode pool, correct versus incorrect barcodes were not distinguishable because all barcode types were expected. As a result, the bars and pie slices that represent the qualified barcodes in (**c**) are shown in grey. Scale bars, 5 μm.
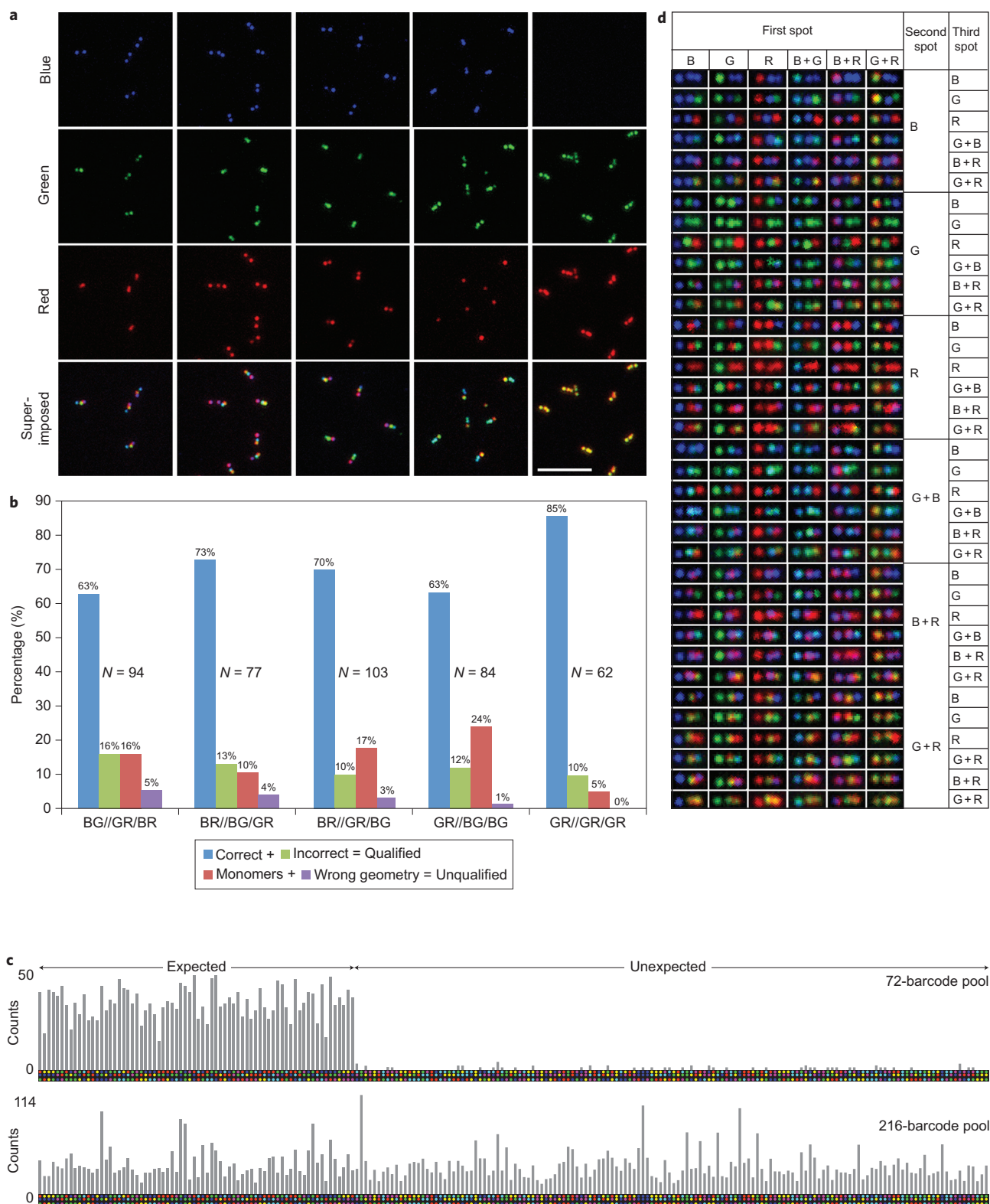
during the mixing and co-purification processes. In addition, over 98% of the qualified barcodes were either BRG or RGB barcodes. The 2% false-positive rate resulted from an unexpected barcode type, BGB (Supplementary Fig. S6), which could be attributed to the rare occasion on which the front monomer of the BRG barcode lay in close proximity to the rear monomer of the RGB barcode.

Next we imaged a pool of all 27 members of the barcode family in which all the species were mixed at an equimolar ratio. TIRF images (Fig. 2b and Supplementary Fig. S7) showed that all the types of barcodes were resolved. Statistical analyses of ~1,500 barcodes across 27 images ($50 \times 50\ \mu m^2$) revealed an average count of 55 barcodes per type with a s.d. of nine (Fig. 2c), in rough agreement with the expected stoichiometry, given pipetting and sampling errors. The distribution of observed objects over the four categories was consistent with the values measured from single-type barcode samples (here, correct versus incorrect barcodes were not distinguishable as all 27 types were included). The above observations suggest that the submicrometre-long DNA nanorod represents a reliable platform on which to construct geometrically encoded barcodes with built-in structural stiffness.
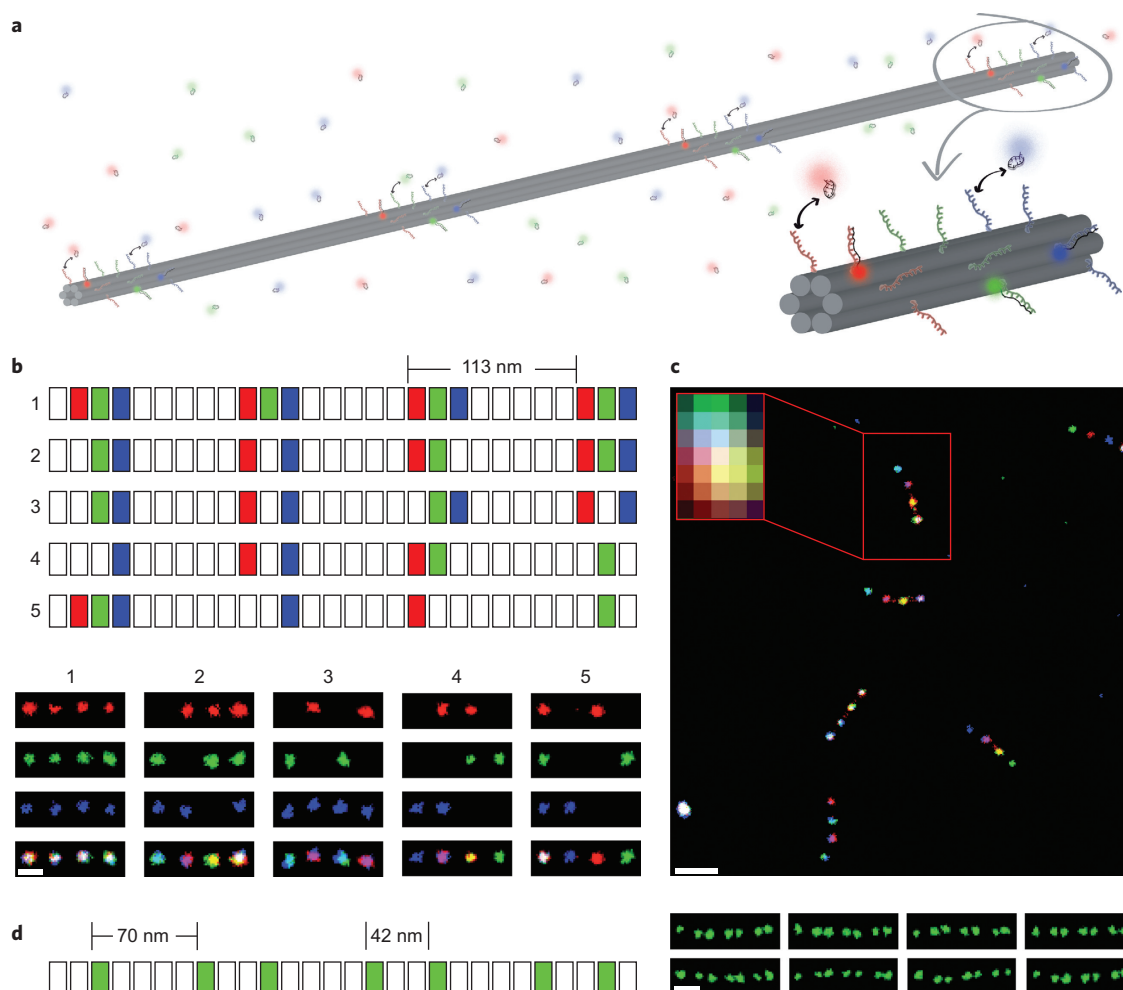
Fluorescent barcodes with dual-labelled zones

The multiplexing capability of the barcodes was enhanced by increasing the number of fluorophore species allowed per zone. We changed the sequence of six staple extensions per zone so that instead of using 12 identical fluorescent oligonucleotides for labelling, we used a combination of up to two fluorophore species, which allows more-distinct fluorescence signatures (pseudocolours) for each zone. This 'dual-labelling' strategy generates six pseudocolours (B, R, G, BG, BR and GR) from three spectrally differentiable fluorophores. Consequently, the total number of distinct barcodes was raised from $3^3 = 27$ to $6^3 = 216$. Five members from the barcode family with dual-labelled zones were chosen for quality-control testing (Fig. 3a). The barcodes can be decoded visually either solely from the superimposed TIRF image or by examining the three separate channels simultaneously. For example, as shown in the first column of Fig. 3a, the barcode 'BG//GR/BR' exhibited two spots each in the blue, green and red channels, but with descending gaps between them. In the superimposed image, the barcodes have the expected colour-mixing signature, cyan//yellow/pink. The correct formation of
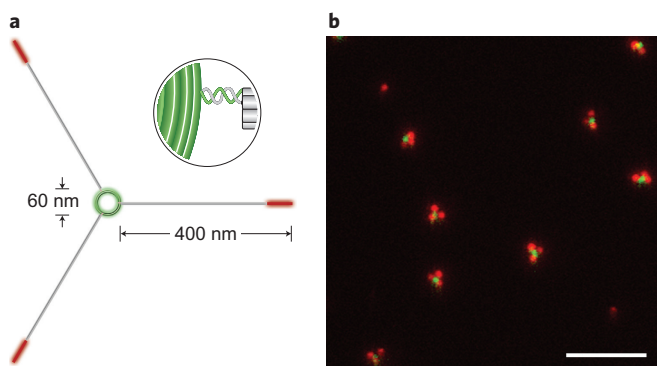
Figure 3 | Fluorescent barcodes with dual-labelled zones. a, Typical TIRF microscopy images of five selected barcode species, shown both in separate channels and as superimposed images. Scale bar, 5 μm. **b**, Statistics obtained by analysing two 50 × 50 μm^2 images of each barcode species (∼85 barcodes, the exact sample size N is given beside the corresponding bar graph). The barcode types are given below the horizontal axis of the diagram. Colour scheme (unrelated to the pseudocolours of the fluorophores): blue, correct barcodes (correct geometry and colour identity); green, incorrect barcodes (correct geometry, but incorrect colour identity); red, monomer nanorods (one spot or two touching spots); purple, barcodes with the wrong geometry (bending angle <120°, see Supplementary Methods). **c**, Computer-aided barcode-counting results of the 72-barcode pool ($N = 2,617$) and the 216-barcode pool ($N = 7,243$), where the number of qualified barcodes are plotted as bar graphs (see Supplementary Tables S6,S7 for the numbers). A computer-generated reference barcode image is placed below the corresponding bar. The same horizontal axis is used for both bar-graph diagrams. **d**, One representative TIRF image (1.4 × 0.7 μm^2) for each of the 216 barcode species with dual-labelled zones.

Figure 4 | Super-resolution fluorescent barcodes. a, Schematic of barcodes for DNA-PAINT super-resolution imaging. The 400 nm DNA nanorod consists of four binding zones evenly spaced by ~113 nm. Each zone can be decorated with the desired combination of 'docking' sequences for red, green or blue imager strands. The orthogonal imager strands bind transiently to their respective 'docking' sites on the nanorod, which creates the necessary 'blinking' for super-resolution reconstruction. **b**, The top panel is a segment diagram (similar to that in Fig. 1a) of the nanorod monomers used to create five barcodes. The lower panel comprises super-resolution images of the five barcodes shown in separate channels and as superimposed images. Scale bar = 100 nm. **c**, Super-resolution image showing all five barcodes from (**b**) in one mixture. The inset shows the diffraction-limited image of a barcode. Scale bar = 250 nm. **d**, An asymmetric barcode that consists of seven binding zones for green imager strands spaced at ~70 nm and ~42 nm for longer and shorter distances, respectively. Scale bar, 100 nm.

the other four barcode species was verified similarly (Fig. 3a and Supplementary Fig. S8). Although the pseudocolours of the dual-labelled zones were not always uniform (for example, some yellow spots were green tinted and others were red tinted) because of inconsistent labelling efficiency, the fluorescence signature of any given spot could be identified by checking the raw images acquired from the three imaging channels. Similar to barcodes with single-labelled zones, manual analysis of two $50 \times 50 \, \mu m^2$ images of each of the five barcode species with dual-labelled zones revealed that 75–95% of the objects were qualified barcodes, of which 80–90% were of the correct type (Fig. 3b). Compared to barcodes in the single-labelled zone, the percentage of qualified barcodes remained the same, although the false-positive rate increased from zero (Fig. 2a) to 10–20%, which reflects the expected decrease in the robustness of the dual-labelling strategy (see Supplementary Technical Notes). All five barcodes examined here have each of their zones labelled with two distinct fluorophore species, which probably makes them among the most error-prone members of the barcode family with dual-labelled zones. Therefore, we would expect a lower average false-positive rate from the whole family.

We next imaged a mixture that contained 72 barcode species with dual-labelled zones that were individually assembled and co-purified (Supplementary Fig. S9). Custom MATLAB scripts were used to assist the decoding process (Supplementary Fig. S10) in either a fully automated unsupervised or in a supervised mode in which the software's best estimate of the barcode's identity was presented to the user for approval. Comparison between supervised and unsupervised decoding results showed >80% agreement between the computer software and the user (see Supplementary Methods and Fig. S11). The supervised analysis of 36 TIRF images ($64 \times 64 \, \mu m^2$) registered 2,617 qualified barcodes that belonged to 116 different species (Fig. 3c, top panel, and Supplementary Table S6). The expected 72 species constituted ~98% of the total barcode population with an average barcode count of 36 per species and a s.d. of eight. In contrast, the unexpected species averaged only ~1.4 barcodes per species (maximum four counts). Finally, we analysed a mixture that contained all 216 members of the barcode family with dual-labelled zones (Supplementary Fig. S12). Unsupervised software analysis of 60 images ($64 \times 64 \, \mu m^2$) registered an average count of 34 ± 17 (mean ± s.d., $N = 7,243$; see Fig. 3c, bottom panel, and

Figure 5 | Fluorescent barcode with nonlinear geometry. a, Schematic of three identical ~400 nm long DNA nanorods linked to the outer edge of a DNA ring (diameter ~60 nm) through the hybridization between staple extensions (inset). The ring and the end of the rod are labelled with Cy3 (green) and Cy5 (red), respectively. **b**, A representative TIRF microscopy image. Scale bar, 5 μm.

Supplementary Table S7) for each barcode species. This relatively large s.d. may be attributed to the increase in the number of distinct barcode species and to the decoding error in the fully automated data analysis. Although this system might be improved further to achieve better assembly and decoding accuracy, our study demonstrates that 216 barcode species were constructed and resolved successfully (Fig. 3d).
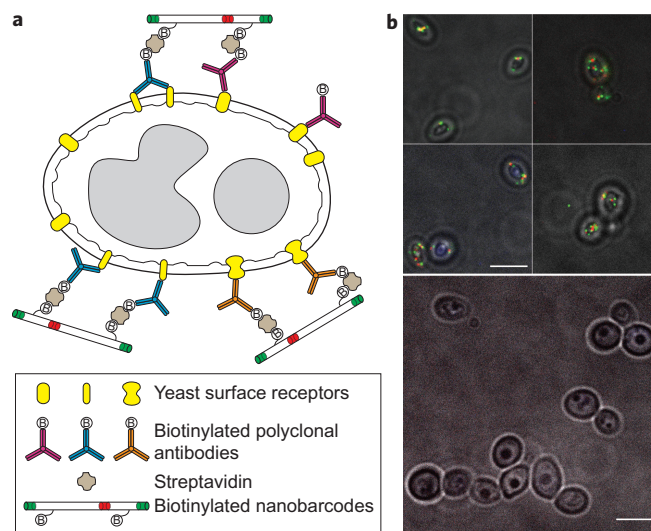
Subdiffraction fluorescent barcodes

The multiplexing capability of the barcodes can also be enhanced by increasing the number of fluorescent zones and resolving them using super-resolution techniques[37–40]. We here used DNA-PAINT[31] to obtain super-resolved images of barcodes with fluorescent zones spaced below the diffraction limit. In stochastic reconstruction microscopy[41,42] most molecules are switched to a fluorescent dark (OFF) state, and only a few emit fluorescence (ON state). Each molecule is localized with nanometre precision[43] by fitting its emission to a 2D Gaussian. DNA-PAINT exploits repetitive, transient binding of fluorescently labelled oligonucleotides ('imager' strands) to complementary 'docking' strands to obtain stochastic switching between fluorescence ON and OFF states (Fig. 4a). In the unbound state, only background fluorescence is observed. On binding of an imager strand, its fluorescence emission is detected using TIRF microscopy.

Previously published work demonstrated single-colour imaging with DNA-PAINT[31]. Here, we extended the technique to three-colour imaging by using orthogonal imager strand sequences coupled to three spectrally distinct dyes (Atto488, Cy3b and Atto655). To demonstrate the feasibility of the three-colour super-resolution barcode system, we designed a DNA nanorod monomer with four binding zones in a symmetric arrangement. The neighbouring zones were separated by ~113 nm, well below the diffraction limit. Each binding zone consisted of 18 staple strands, which can display three groups of orthogonal docking sequences (six per group) to which the blue, green or red imager strands can bind. As a proof-of-principle experiment, we designed five different barcodes (Fig. 4a and the top panel of Fig. 4b). The bottom panel of Fig. 4b shows the super-resolution reconstruction for each channel separately, as well as an overlay of all the channels. Figure 4c shows a larger area that contains all five barcodes. The unique pattern of the barcodes in all three channels can be resolved. The transient, repetitive binding of imager strands to docking sequences on the nanorod not only creates the necessary 'blinking' behaviour for localization, but also makes the imaging more robust, as DNA-PAINT is not prone to

photobleaching (Supplementary Fig. S13). We also constructed an asymmetric barcode version that consisted of seven binding zones for green imager strands with longer and shorter spacings of ~70 nm and ~42 nm, respectively (Fig. 4d, overview given in Supplementary Fig. S14). If three colours in this arrangement are used, $(2^3 - 1)^7 = 823{,}543$ different barcodes can be constructed. In principle, we could also construct shorter barcodes and maintain the multiplexing capability of the diffraction-limited barcode system. For example, ~100 nm long barcodes with three asymmetric binding zones would allow for 343 different signatures. The inherent modularity of our design allows the interzone distances to be tailored to a wide range of microscopes and applications.

Nonlinear fluorescent barcodes

More sophisticated barcodes can be generated using DNA nanostructures with nonlinear geometry. Figure 5 shows an example in which three ~400 nm DNA rods were linked to the outer edge of a ~60 nm DNA ring through hybridization between staple extensions (Fig. 5a, inset). Fluorescently labelling the ring and the far end of the nanorods generated a pattern that resembled a three-point star and was clearly resolvable under fluorescence microscopy. TIRF microscopy and transmission electron microscopy (Fig. 5b and Supplementary Figs S15,S16) revealed that about 50% of successfully assembled barcodes featured three nanorods that surrounded the ring with a roughly 120° angle between each other, as designed, although many other barcodes had significantly biased angles between neighbouring nanorods because of the semi-flexible double-stranded DNA linker between the ring and the nanorods. It is conceivable that we could increase the system's accuracy and robustness using a similar design to connect three identical 'satellite' linear barcodes to a central hub (here the three satellite barcodes may share the hub as a common fluorescently labelled zone). Such a design would triplicate the encoding redundancy. In addition, stiffer linkers between the ring and the protrusions (for

Figure 6 | Tagging yeast cells with the GRG barcodes as *in situ* imaging probes. a, In this illustration of the tagging mechanism, the biotinylated barcodes are anchored on the yeast cell through streptavidin molecules bound to biotinylated polyclonal antibodies coated on the yeast surface. For clarity, only two of the ten biotinylated staples on the barcode are shown. **b**, Superimposed microscopy images (acquired from bright-field and TIRF microscopy) of the yeast cells treated with the barcodes. The top four images show yeast cells treated as illustrated in (**a**). The bottom image is a negative control in which yeast cells were treated with non-biotinylated barcodes. Scale bars, 5 μm.

example, multihelix DNA with strand cross-overs) could be used to achieve better-defined barcode geometry.

Fluorescent barcodes as *in situ* imaging probes

Modifying the barcodes with functional ligands, such as antibodies and aptamers, would allow the barcodes to tag specific biological samples and serve as multiplexed *in situ* imaging probes. In a proof-of-principle experiment, the GRG barcode was used to tag wild-type *Candida albicans* yeast. The yeast cells were first mixed with a biotinylated polyclonal antibody specific to *C. albicans*, then coated with streptavidin and finally incubated with biotinylated GRG barcodes (Fig. 6a). TIRF microscopy revealed the barcodes attached to the bottom surface of the yeast cells (Fig. 6b, top panel, and Supplementary Fig. S17). Although some of the nanorods landed awkwardly on the uneven cell walls of the yeast cells, a number of GRG barcodes were resolved clearly. In contrast, no barcode tagging was observed when non-biotinylated antibodies or barcodes were used (Fig. 6b, bottom panel, and Supplementary Fig. S17), which suggests that no nonspecific interactions existed between the barcode and the cell surface. Here, only the bottom layer ($\sim$100 nm) of the 3D specimen was imaged to assure adequate lateral resolution and minimal background noise. With the development of 3D microscopy[44], protein conjugation chemistry[45–47] and DNA self-assembly techniques[21–23], we believe a quantitative and multiplexed *in situ* imaging system is within reach.

Discussion

In summary, we constructed a new kind of geometrically encoded fluorescent barcode using self-assembly based on DNA origami. Our approach differs from previous optically decodable barcoding systems[1–16] in three aspects.

First, our barcodes are 400−800 nm in length, substantially shorter than the existing geometrically encoded fluorescent barcodes. The submicrometre dimensions make them potentially useful as *in situ* single-molecule imaging probes. For example, we demonstrated the tagging of cell-surface proteins in yeast. The system also fulfils a technological challenge to build robust finite-size optical barcodes with the smallest feature approaching or even smaller than the diffraction limit of visible light.

Second, our barcodes are self-assembled from DNA. Unlike previous constructions, their synthesis and purification does not require enzymatic reaction[15], genomic engineering[6], photolithography[14], electrochemical etching[48] or microfluidic devices[1,14], and is easy to perform in a typical biochemistry laboratory (see Supplementary Technical Note for the barcode cost). The barcodes can be modified easily to display sequence-specific attachment sites simply through DNA-origami staple extensions.

Third, the robustness and programmability of the platform based on DNA origami enables a high multiplexing capability. Here, we constructed successfully 216 barcode species that were decoded in one pool (previous submicrometre-sized fluorescent barcodes contained no more than 11 distinct barcodes[3,7,8,12]). Even compared to a previous production of multicolour fluorescent barcodes up to 10 μm in size, our system is among the few examples[1,15] that demonstrate more than 100 distinct barcode species in practice. The NanoString nCounter demonstrated larger numbers of barcodes that coexisted in a homogenous solution[15]. However, the reporter probes of the NanoString ($\sim$7 kb or $\sim$2 μm double-stranded RNA–DNA hybrid molecules) require stretching to make the barcodes resolvable, which makes them difficult to use as *in situ* probes. In contrast, the stiff barcodes based on DNA origami maintain their structural integrity when applied to cell surfaces, which suggests their potential use as *in situ* probes. Furthermore, combination of the DNA-PAINT imaging technique and the nanorods based on DNA origami allowed us to create

nanoscopic 'super-resolution' barcodes, which suggests the potential to achieve >100,000 distinct barcodes. Such a large number of barcodes opens opportunities such as the measurement of human gene expression through *ex situ* imaging. In practice, it is probable that a small number of distinct codes could be sufficient to label the targets of interest in various application scenarios. For those applications, we could use selectively the best-performing subset of the barcode library and/or use barcodes of shorter sizes (for example, $\sim$100 nm barcode with three fluorescent spots).

In the future, we expect our barcodes to interface with the rapidly growing protein/antibody library and tagging techniques[44–47] to yield a versatile imaging toolbox for single-molecule studies of biological events and biomedical diagnostics. For example, we can envision the utility of a library of barcodes modified with monoclonal antibodies against the clusters of differentiation molecules[49] (>300 known) for immunophenotyping applications[50]. However, moving from *in vitro* imaging on a glass surface to *in situ* imaging of a cell membrane may present substantial technical challenges. For example, crowding of multiple barcodes may render an individual barcode unresolvable. Although our current barcodes are potentially suitable for *in situ* labelling of multiple cell types (each carrying distinct surface markers), they are too bulky for a multiplexed detection of many subcellular identities in a single cell. This challenge could be addressed by using barcodes of a more compact size and stiffer structures (for example, DNA tetrahedrons with $\sim$50 nm edges). Although such shorter barcodes may be used readily to tag the surface markers of fixed cells, imaging them in living cells will probably pose additional significant challenges. For example, to resolve barcodes on dynamic living cell membranes may require higher temporal resolution than that provided by current super-resolution microscopy techniques. The barcodes constructed here represent the first experimental demonstration of a large number of submicrometre geometrically encoded fluorescent barcodes that are structurally stiff and optically resolvable. The fast progress of nanotechnology's ability to engineer shape and microscopy's ability to resolve shape should enable the future construction of more versatile and powerful geometrically encoded barcodes for diverse biomedical applications.

References

1. Fournier Bidoz, S. *et al*. Facile and rapid one-step mass preparation of quantum-dot barcodes. *Angew. Chem. Int. Ed.* **47,** 5577–5581 (2008).
2. Han, M., Gao, X., Su, J. Z. & Nie, S. Quantum-dot-tagged microbeads for multiplexed optical coding of biomolecules. *Nature Biotechnol.* **19,** 631–635 (2001).
3. Li, Y., Cu, Y. T. H. & Luo, D. Multiplexed detection of pathogen DNA with DNA-based fluorescence nanobarcodes. *Nature Biotechnol.* **23,** 885–889 (2005).
4. Marcon, L. *et al*. 'On-the-fly' optical encoding of combinatorial peptide libraries for profiling of protease specificity. *Mol. BioSyst.* **6,** 225–233 (2010).
5. Xu, H. *et al*. Multiplexed SNP genotyping using the Qbead system: a quantum dot-encoded microsphere-based assay. *Nucleic Acids Res.* **31,** e43 (2003).
6. Livet, J. *et al*. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450,** 56–62 (2007).
7. Lin, C., Liu, Y. & Yan, H. Self-assembled combinatorial encoding nanoarrays for multiplexed biosensing. *Nano Lett.* **7,** 507–512 (2007).
8. Levsky, J. M., Shenoy, S. M., Pezo, R. C. & Singer, R. H. Single-cell gene expression profiling. *Science* **297,** 836–840 (2002).
9. Braeckmans, K. *et al*. Encoding microcarriers by spatial selective photobleaching. *Nature Mater.* **2,** 169–173 (2003).
10. Dejneka, M. J. *et al*. Rare earth-doped glass microbarcodes. *Proc. Natl Acad. Sci. USA* **100,** 389–393 (2003).
11. Gudiksen, M. S., Lauhon, L. J., Wang, J., Smith, D. C. & Lieber, C. M. Growth of nanowire superlattice structures for nanoscale photonics and electronics. *Nature* **415,** 617–620 (2002).
12. Li, X. *et al*. Controlled fabrication of fluorescent barcode nanorods. *ACS Nano* **4,** 4350–4360 (2010).
13. Nicewarner-Pena, S. R. Submicrometer metallic barcodes. *Science* **294,** 137–141 (2001).

14. Pregibon, D. C., Toner, M. & Doyle, P. S. Multifunctional encoded particles for high-throughput biomolecule analysis. *Science* **315,** 1393–1396 (2007).

15. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnol.* **26,** 317–325 (2008).

16. Xiao, M. *et al.* Direct determination of haplotypes from single DNA molecules. *Nature Methods* **6,** 199–201 (2009).

17. Toomre, D. & Bewersdorf, J. A new wave of cellular imaging. *Annu. Rev. Cell Dev. Biol.* **26,** 285–314 (2010).

18. Seeman, N. C. Nucleic acid junctions and lattices. *J. Theor. Biol.* **99,** 237–247 (1982).

19. Aldaye, F. A., Palmer, A. L. & Sleiman, H. F. Assembling materials with DNA as the guide. *Science* **321,** 1795–1799 (2008).

20. Lin, C., Liu, Y. & Yan, H. Designer DNA nanoarchitectures. *Biochemistry* **48,** 1663–1674 (2009).

21. Nangreave, J., Han, D., Liu, Y. & Yan, H. DNA origami: a history and current perspective. *Curr. Opin. Chem. Biol.* **14,** 608–615 (2010).

22. Shih, W. M. & Lin, C. Knitting complex weaves with DNA origami. *Curr. Opin. Struct. Biol.* **20,** 276–282 (2010).

23. Tørring, T., Voigt, N. V., Nangreave, J., Yan, H. & Gothelf, K. V. DNA origami: a quantum leap for self-assembly of complex structures. *Chem. Soc. Rev.* **40,** 5636–5646 (2011).

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

25. Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459,** 414–418 (2009).

26. Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **325,** 725–730 (2009).

27. Andersen, E. S. *et al.* Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **459,** 73–76 (2009).

28. Han, D., Pal, S., Liu, Y. & Yan, H. Folding and cutting DNA into reconfigurable topological nanostructures. *Nature Nanotechnol.* **5,** 712–717 (2010).

29. Liedl, T., Högberg, B., Tytell, J., Ingber, D. E. & Shih, W. M. Self-assembly of three-dimensional prestressed tensegrity structures from DNA. *Nature Nanotechnol.* **5,** 520–524 (2010).

30. Han, D. *et al.* DNA origami with complex curvatures in three-dimensional space. *Science* **332,** 342–346 (2011).

31. Jungmann, R. *et al.* Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano Lett.* **10,** 4756–4761 (2010).

32. Steinhauer, C., Jungmann, R., Sobey, T. L., Simmel, F. C. & Tinnefeld, P. DNA origami as a nanoscopic ruler for super-resolution microscopy. *Angew. Chem. Int. Ed.* **48,** 8870–8873 (2009).

33. Lund, K. *et al.* Molecular robots guided by prescriptive landscapes. *Nature* **465,** 206–210 (2010).

34. Pal, S., Deng, Z., Ding, B., Yan, H. & Liu, Y. DNA-origami-directed self-assembly of discrete silver-nanoparticle architectures. *Angew. Chem. Int. Ed.* **49,** 2700–2704 (2010).

35. Bui, H. *et al.* Programmable periodicity of quantum dot arrays with DNA origami nanotubes. *Nano Lett.* **10,** 3367–3372 (2010).

36. Douglas, S. M., Chou, J. J. & Shih, W. M. DNA–nanotube-induced alignment of membrane proteins for NMR structure determination. *Proc. Natl Acad. Sci. USA* **104,** 6644–6648 (2007).

37. Hell, S. W. Far-field optical nanoscopy. *Science* **316,** 1153–1158 (2007).

38. Huang, B., Babcock, H. & Zhuang, X. Breaking the diffraction barrier: super-resolution imaging of cells. *Cell* **143,** 1047–1058 (2010).

39. Vogelsang, J. *et al.* Make them blink: probes for super-resolution microscopy. *ChemPhysChem.* **11,** 2475–2490 (2010).

40. Walter, N. G., Huang, C. Y., Manzo, A. J. & Sobhy, M. A. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nature Methods* **5,** 475–489 (2008).

41. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313,** 1642–1645 (2006).

42. Rust, M. J., Bates, M. & Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* **3,** 793–795 (2006).

43. Yildiz, A. *et al.* Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science* **300,** 2061–2065 (2003).

44. Jones, S. A., Shim, S. H., He, J. & Zhuang, X. Fast, three-dimensional super-resolution imaging of live cells. *Nature Methods* **8,** 499–508 (2011).

45. Gautier, A. *et al.* An engineered protein tag for multiprotein labeling in living cells. *Chem. Biol.* **15,** 128–136 (2008).

46. Keppler, A. *et al.* A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature Biotechnol.* **21,** 86–89 (2003).

47. Klein, T. *et al.* Live-cell dSTORM with SNAP-tag fusion proteins. *Nature Methods* **8,** 7–9 (2011).

48. Cunin, F. *et al.* Biomolecular screening with encoded porous-silicon photonic crystals. *Nature Mater.* **1,** 39–41 (2002).

49. Matesanz-Isabel, J. *et al.* New B-cell CD molecules. *Immunol. Lett.* **134,** 104–112 (2011).

50. Maecker, H. T., McCoy, J. P. & Nussenblatt, R. Standardizing immunophenotyping for the Human Immunology Project. *Nature Rev. Immunol.* **12,** 191–200 (2012).

Author contributions

C. Lin conceived the project, designed and conducted the majority of the experiments, analysed the data and prepared the majority of the manuscript. R.J. conceived the super-resolution barcode study, designed and conducted experiments for this study, analysed the data and prepared the manuscript. A.M.L. wrote the MATLAB script for the automated barcode deciphering and prepared the manuscript. C. Li wrote the C script for the barcode geometry characterization. D.L. (with C. Lin) performed the yeast-tagging experiment. G.M.C. championed multiplexed *in situ*, supervised C. Li and D.L., and critiqued the data and the manuscript. W.M.S. conceived the project, discussed the results and prepared the manuscript. P.Y. conceived, designed and supervised the study, interpreted the data and prepared the manuscript. All authors reviewed and approved the manuscript.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permission information is available online at http://www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.Y. and W.M.S.

Competing financial interests

The authors declare competing financial interests: a provisional US patent application has been filed.