Nanoscale

ROYAL SOCIETY
OF CHEMISTRY

NCNST

PAPER
Dae Sung Yoon *et al.*
Identifying DNA mismatches at single-nucleotide resolution by probing
individual surface potentials of DNA-capped nanoparticles

Identifying DNA mismatches at single-nucleotide resolution by probing individual surface potentials of DNA-capped nanoparticles†

Hyungbeen Lee,‡[a] Sang Won Lee,‡[b] Gyudo Lee, [ID] ‡[b,c] Wonseok Lee,[a]
Kihwan Nam,[d] Jeong Hoon Lee,[e] Kyo Seon Hwang,[f] Jaemoon Yang,[g]
Hyeyoung Lee,[h] Sangsig Kim,[i] Sang Woo Lee[a] and Dae Sung Yoon [ID] *[b]

Here, we demonstrate a powerful method to discriminate DNA mismatches at single-nucleotide resolution from 0 to 5 mismatches (χ_0 to χ_5) using Kelvin probe force microscopy (KPFM). Using our previously developed method, we quantified the surface potentials (SPs) of individual DNA-capped nanoparticles (DCNPs, ~100 nm). On each DCNP, DNA hybridization occurs between ~2200 immobilized probe DNA (*p*DNA) and target DNA with mismatches (*t*DNA, ~80 nM). Thus, each DCNP used in the bioassay (each *p*DNA–*t*DNA interaction) corresponds to a single ensemble in which a large number of *p*DNA–*t*DNA interactions take place. Moreover, one KPFM image can scan at least dozens of ensembles, which allows statistical analysis (*i.e.*, an ensemble average) of many bioassay cases (ensembles) under the same conditions. We found that as the χ_n increased from χ_0 to χ_5 in the tDNA, the average SP of dozens of ensembles (DCNPs) was attenuated owing to fewer hybridization events between the *p*DNA and the *t*DNA. Remarkably, the SP attenuation *vs.* the χ_n showed an inverse-linear correlation, albeit the equilibrium constant for DNA hybridization exponentially decreased asymptotically as the χ_n increased. In addition, we observed a cascade reaction at a 100-fold lower concentration of *t*DNA (~0.8 nM); the average SP of DCNPs exhibited no significant decrease but rather split into two separate states (no-hybridization *vs.* full-hybridization). Compared to complementary *t*DNA (*i.e.*, χ_0), the ratio of no-hybridization/full-hybridization within a given set of DCNPs became ~1.6 times higher in the presence of tDNA with single mismatches (*i.e.*, χ_1). The results imply that our method opens new avenues not only in the research on the DNA hybridization mechanism in the presence of DNA mismatches but also in the development of a robust technology for DNA mismatch detection.

[a]Department of Biomedical Engineering, Yonsei University, Wonju 26493, Korea
[b]School of Biomedical Engineering, Korea University, Seoul 02841, Korea.
E-mail: dsyoon@korea.ac.kr
[c]Harvard T.H. Chan School of Public Health, Boston, Massachusetts 02115, USA
[d]Research & Development Team, Korea Medical Devices Industrial Coop. Association (KMDICA), Seoul 04578, Korea
[e]Department of Electrical Engineering, Kwangwoon University, Seoul 01897, Korea
[f]Department of Clinical Pharmacology and Therapeutics, College of Medicine, Kyung Hee University, Seoul 02447, Korea
[g]Department of Radiology, Yonsei University College of Medicine, Seoul 03722, Korea
[h]Department of Biomedical Laboratory Science, Yonsei University, Wonju 26493, Korea
[i]Department of Electrical Engineering, Korea University, Seoul 02841, Korea
†Electronic supplementary information (ESI) available. See DOI: 10.1039/c7nr05250b
‡These authors contributed equally to this work.

Introduction

Genetic diversity, which arises from genetic variation, is the most important factor in the evolution of life on Earth.[1,2] In principle, genetic variations are caused by DNA mismatches in the human genome. DNA mismatches are often associated with various diseases, such as diabetes,[3] cancer,[4] and numerous genetic disorders or diseases (*e.g.*, Huntington, myotonic dystrophy, and fragile-X syndrome).[5] Accordingly, precise identification of the number and position of nucleotide mismatches in DNA is fundamental to determine the disease type of a patient and anticipate the time of onset. In general, the mutations of more than two adjacent nucleotides are more aggressive than those caused by a single-nucleotide polymorphism, because the more the consecutive position changes in a codon, the greater the risk of mutations of multiple amino acids.[6] This implies that DNA mismatch detection at single-nucleotide resolution is necessary to more accurately

predict the disease expression[7] and to develop clinical diagnostic methods.[8]

Given these considerations, there is a growing need for DNA mismatch detection. For example, polymerase chain reaction,[9] microcantilevers,[10,11] surface-enhanced Raman spectroscopy,[12] microarray patterns,[13,14] Kelvin probe force microscopy (KPFM),[15,16] and a nanopore system[17–19] have been developed and used for highly efficient detection of DNA mismatches. Despite such efforts, DNA mismatch detection at single-nucleotide resolution with respect to a wide range of genetic mutations remains a challenge.[11,20–23] More importantly, there is a difficulty in that the outcome of genotyping is significantly dependent not only on the number of mismatches in target DNA (tDNA) but also on the concentration of tDNA in a sample, since conventional methods for DNA mismatch detection rely strongly on a change in the binding affinity between probe DNA (pDNA) and tDNA. This makes it difficult to disentangle the concentration effect of tDNA from an experimental result searching for a wide range of mutations; it is a fatal drawback in sensing mutations in an unknown sample (*i.e.*, when there is no information about the sample concentration and the number of mismatches).

To overcome these hurdles, we have adopted DNA-capped gold nanoparticles (DCNPs) as a sensing platform in which pDNA was immobilized onto uniformly-sized nanoparticles to capture tDNA. In addition to DCNPs, KPFM[24] was used for high-resolution detection of individual DCNPs.[16] As a model, we selected tDNA with *BRCA1* gene mutations that are related to breast or ovarian cancer.[10,25] The official name of *BRCA1* is known to be 'breast cancer 1'.[26] According to a previous study,[26] several thousands of pDNA strands can be immobilized on a single nanoparticle, depending on its size. Therefore, each DCNP used in the bioassay (each pDNA–tDNA interaction) corresponds to a single DNA ensemble in which a large number of pDNA–tDNA interactions take place with reduced electrostatic and steric barriers. One KPFM image can provide both topography and surface potential values of every DCNP on the substrate (at least dozens of ensembles), allowing for statistical analysis (*i.e.*, an ensemble average) of many bioassay cases (ensembles) under the same conditions.[15] This suggests that efficient detection of DNA mismatches can be accomplished by using KPFM.

Herein, using the above-mentioned concept, we demonstrate that various mismatches (1 to 5 mismatched nucleotides) in the *BRCA1* gene can be rapidly determined at single-nucleotide resolution using KPFM of DCNPs. In comparison with the other conventional methods mentioned above, KPFM analysis of DCNPs can provide precise identification of DNA mismatches in a simple, label-free manner, without any labelling or amplification processes. In addition, our findings regarding how the concentration of tDNA affects the surface potentials of DCNPs are discussed, and its peculiar characteristics are presented in thermodynamic terms. We believe that our technology will be a significant breakthrough in the detection of DNA mismatches.

Experimental

Preparation of gold nanoparticles (GNPs)

GNP (~100 nm in diameter, molecular weight = 196.97 g mol^{-1}) solution was purchased from Sigma-Aldrich (St Louis, MO, USA). GNPs were suspended in 0.1 mM PBS at ~4 × 10^9 particles per milliliter.

Preparation of DNA solution

All oligonucleotides (Bio Basic Canada, Markham, ON, Canada) were purified by high-performance liquid chromatography (HPLC). They were dissolved in the same quantity of phosphate buffer (pH 7.0) so that all DNA solutions had an identical concentration (*i.e.*, 76.8 nM) in all experiments, except for the tDNA concentration tests (Fig. 5).

GNP functionalization with pDNA

To immobilize pDNA on the GNPs, we used a thiol-modified DNA (sequence: 5′-HS-CTA CCT TTT TTT TCT G-3′). An aliquot of 2.5 ml pDNA solution was mixed with 2.5 ml colloidal GNPs at room temperature. The mixture was gently vortexed for 16 h to accomplish full immobilization, then 15 ml of phosphate buffer was added to the mixture, and it was gently vortexed for 40 h. After chemical functionalization, the mixture was centrifuged for 25 min at 14 000 rpm at room temperature to remove excess reagents. The supernatant (16 ml) was removed, and 1 ml of 0.1 mM phosphate buffer was added to the tube containing the pDCNPs.

Hybridization of pDNA and tDNA

A mixture of 2.5 ml tDNA solution (76.8 nM) and 5 ml pDCNP solution was gently vortexed for 3 h at room temperature to produce p-tDCNPs. After the hybridization process, 12.5 ml of phosphate buffer was added to the mixture. Unhybridized tDNA was removed using the same method (*via* removal of the supernatant) as described above.

Preparation of the gold substrate

An SiO$_2$ (5000 Å)/Si (p-type) wafer (Silicon Technology, Tokyo, Japan) was cleaned with piranha solution (a mixture of 1 : 1 (v/v) H$_2$SO$_4$ and H$_2$O$_2$) at room temperature for at least 10 min. To obtain Au (200 Å)/Cr (50 Å)/SiO$_2$ (5000 Å)/Si substrates, chromium and gold were subsequently evaporated on the SiO$_2$/Si substrates using a thermal evaporator. After thermal evaporation, the wafer was diced into 10 × 10 mm^2 samples.

Preparation of KPFM samples

For KPFM imaging, 100 μl of the DCNPs were dropped onto each gold substrate, and the DCNPs were adsorbed on the gold substrate for 1 h. The gold substrate was then rinsed with de-ionized H$_2$O and gently blow-dried with nitrogen to avoid aggregation of the DCNPs before AFM imaging or surface potential measurements.

Surface potential measurements of DCNPs

The topography and the surface potential measurements of all DCNP specimens were performed using a commercial AFM (Multimode V, Veeco, CA, USA) at room temperature. A conducting cantilever tip (SCM-PIT, Bruker, CA, USA) was mounted in a tip holder (MMEFCH, Veeco, CA, USA), which is capable of controlling the voltage of the tip. Surface potential measurements were performed using a lift-mode KPFM.[15] The standard method of optimizing KPFM conditions is detailed in the ESI.†
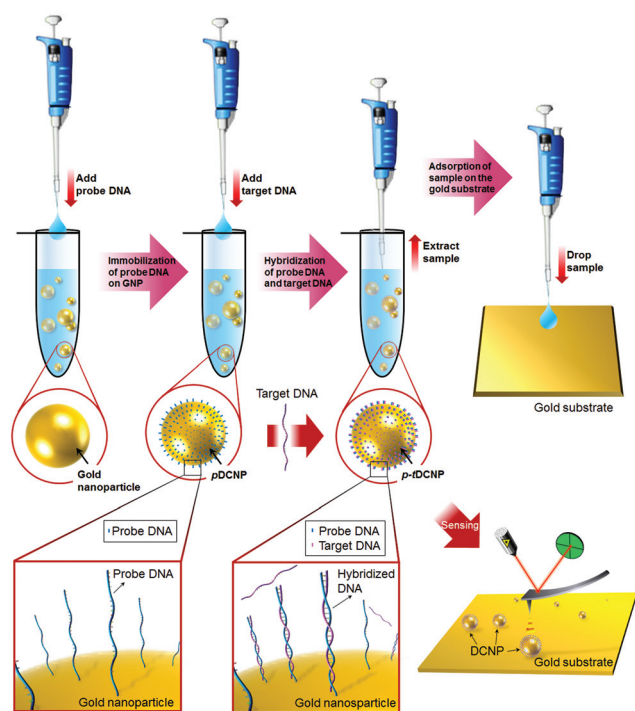

Fig. 1 Schematic representation of the assembly of DCNPs and their surface potential characterization by KPFM for detection of point mutations. The GNPs were functionalized with *p*DNA against target *t*DNA. In detail, bare GNPs were diluted in PBS and the *p*DNA was immobilized on the bare GNPs, forming *p*DCNPs. The *t*DNA, with 1 to 5 nucleotide mismatches, hybridized with the *p*DCNP, creating the *p*-*t*DCNP. To detect point mutations, the topography and surface potential images of each sample were obtained by KPFM.

Results and discussion

Our approach was aimed at efficient detection of point mutations by measuring the surface potentials of DCNPs consisting of a GNP (~100 nm in diameter) (ESI Fig. S1a†) and a hybridized form of DNA (Fig. 1). The DNA sequence (16 mer; Table 1) used is a core part of the human *BRCA1* gene.[10,25] It is important that the analytical tool or platform is capable of highly sensitive detection of a single-nucleotide polymorphism. To test how precisely our tool can identify a point mutation, we artificially created oligonucleotides having 1 to 5 mismatched nucleotides in *BRCA1* (Table 1). The *p*DNA, consisting of a sequence complementary to *BRCA1* without any mutation, was modified by a thiol group (-HS) in order to immobilize it on the GNP.[27,28] After immobilization, *p*DCNPs were obtained. More information about the materials and chemical functionalization is detailed in the Methods section. When the *p*DCNPs interacted with different *t*DNAs having 1 to 5 mismatched nucleotides, they formed double-stranded DCNPs (*i.e.*, *p*-*t*DCNP) with different binding affinities depending on the number of mismatched nucleotides. The DCNPs were then spread onto an Au/Si substrate freshly cleaned with piranha solution, gently dried with pure nitrogen at 25 °C, and then quantitatively probed by KPFM. For efficient detection of point mutations, we presumed that a single DCNP represents a single biological interaction (*i.e.*, DNA hybridization), where ~2200 of the *p*DNAs[26,29] are involved. The probability of the interaction occurring can be measured as the surface potential of the *p*-*t*DCNP using KPFM. This is because the binding of *t*DNA to *p*DNA results in an alteration in the surface potential of the *p*-*t*DCNP due to an increase in negative charge. Using our setup, a number of DCNPs spread on the gold surface were easily examined with a single KPFM image. This implies that the statistical analysis (ensemble average) of several dozen bioassays with the same conditions can be accomplished within a short time frame (*ca.* 30 min). Accordingly, this approach represents a breakthrough necessary to overcome a long analytical time, a critical drawback of atomic force microscopy (AFM) and KPFM. In this work, both the high-resolution topography and the surface potential images were analysed, as shown in

Table 1 Probe and target DNA sequences used in the experiment

| Abbreviation of DNA	Full name of DNA	DNA sequence
Probe		
*p*DNA	Probe DNA	5′-HS-CTACCTTTTTTTTCTG-3′
Target		
*c*DNA	Complementary DNA	5′-CAGAAAAAAAGGTAG-3′
χ_1DNA	1 point mismatched DNA	5′-CAGAAAA**T**AAAGGTAG-3′
χ_2DNA	2 point mismatched DNA	5′-CAGAAA**TT**AAAGGTAG-3′
χ_3DNA	3 point mismatched DNA	5′-CAGAAA**TTT**AAGGTAG-3′
χ_4DNA	4 point mismatched DNA	5′-CAGAA**TTTT**AAGGTAG-3′
χ_5DNA	5 point mismatched DNA	5′-CAGAA**TTTTT**AGGTAG-3′
*nc*DNA	Non-complementary DNA	5′-**AGCTTCCGTACTCGTT**-3′
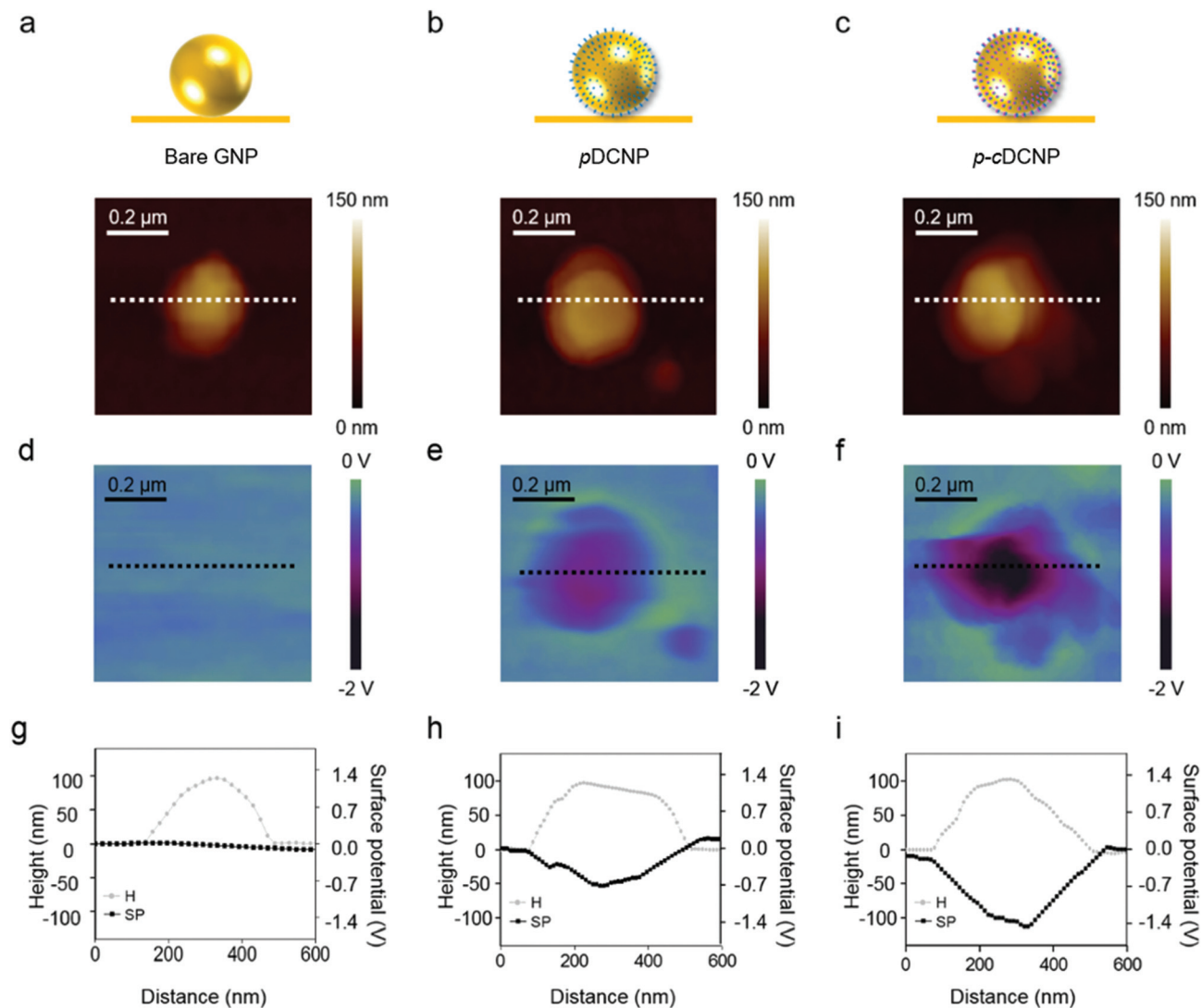
Fig. 2 Topography and surface potential images of GNP, *p*DCNP, and *p-c*DCNP. (a–c) Height map images observed by TMAFM: (a) bare GNP (100 nm), (b) *p*DCNP, and (c) *p-c*DCNP. (d–f) Magnified surface potential map images produced by KPFM: (d) bare GNP, (e) *p*DCNP, and (f) *p-c*DCNP. Black or white dotted lines in each image represent the trajectory of the line scan for the following cross-sectional views: (g–i) corresponding cross-sectional views taken through the height map images and surface potential map images from (d–f): (g) bare GNP, (h) *p*DCNP, and (i) *p-c*DCNP. The gray line in each graph represents a topographic cross-section of a single complex, whereas a black line depicts the cross-sectional surface potential.

Fig. 2. The size of the GNP was determined to be ∼90 nm by tapping mode AFM (ESI Fig. S1b and c†), which is attributable to a glitch arising from the amplitude damping effect taking place between the AFM tip and the sample surface,[30] and the wobbling effect of GNPs during AFM imaging.[31] For precise measurements, we optimized the conditions of the lift scan height and scan speed of a conductive cantilever tip in KPFM on the basis of our previous studies.[16,32–35] Because the GNP consists of the same material as the substrate (*i.e.*, a gold surface), it is reasonable to assume that the surface potential of a bare GNP would not be different from that of the substrate (*i.e.*, ∼0 V) (Fig. 2d and g). Briefly, we found a critical condition where the surface potentials (SPs) of bare GNPs placed on the Au substrate became zero, and thereby identical to that of the substrate. It was expected that under this condition we can avoid noise signals coming from various sources, which makes the measured surface potentials attributed to only the DNA–DNA interactions. Using heuristic evaluation of the size-dependent work function,[32,36] we optimized this measurement condition so that there was no change in the surface potential near the border between the GNP (∼100 nm) and the gold surface, and a flat surface potential was produced over the entire imaged area (ESI Fig. S3†). This process is essential in DNA bioassays using KPFM, because it enables one to measure the surface potential change due to the DNA interaction without interference from environmental factors. Fig. 2h and i show DNA (*i.e.*, *p-t*DNA). This is plausible because DNA is negatively charged owing to ionized phosphates in its back-

bone. In the case of the topographic height, however, it seems that there was little height difference between all, revealing that the measured surface potential of a DCNP was affected only by the associated DNA, including both single-stranded DNA (*i.e.*, *p*DNA) and double-stranded samples. The topographic heights of bare GNPs, *p*DCNP, and *p-c*DCNP were 89.0 ± 7.3 nm, 107.0 ± 6.6 nm, and 108.5 ± 8.0 nm, respectively. The slightly thicker topography of DCNPs may be attributable to the addition of DNA. In more detail, the adsorption of DCNPs on the substrate is likely to be stronger than that of bare GNPs. This leads to the slightly greater height of the DCNPs in comparison with bare GNPs because of the elimination of AFM artifacts, such as the wobbling effect[37] mentioned above. Accordingly, the DCNPs always exceeded ~100 nm (bare GNP diameter) in height.

On the other hand, there was a negligible height difference between the *p*DCNP and the *p-c*DCNP ($\Delta h \approx 1.5$ nm), despite the addition of complimentary *t*DNA. This is because *t*DNA does not bind to the terminal end of *p*DNA, but instead binds to its complementary bases in parallel. However, the surface

potential measurement exhibited a manifest difference between *p*DCNP and *p-c*DCNP. As expected, we observed a doubling of the surface potential[16] when *p*DCNPs ($\phi_{pDNA} = -0.743$ V) were exposed to their complementary *t*DNA to make *p-c*DCNPs ($\phi_{p-cDNA} = -1.455$ V) (ESI Table S1†). This implies that almost all of the *p*DNA immobilized on the GNP hybridized with the complementary *t*DNA. It also indicates that the *p*DCNPs used in this work had the capacity for efficient DNA hybridization despite the steric hindrance of DNA–DNA interactions. Notably, the observed doubling of the surface potential provides us with an important clue indicating that the quantity of hybridized *t*DNA is linearly proportional to the surface potential of *p-t*DCNP, which suggests the potential for highly accurate detection of point mutations for the early diagnosis of cancer. To investigate whether sequence-specific detection of gene mutations can be achieved using our platform, we performed KPFM imaging of the *p-t*DCNPs in which *p*DNA interacted with *t*DNA possessing 1 to 5 mismatched nucleotides (Table 1). Fig. 3 depicts the representative three-dimensional (3D) topography and surface potential images of
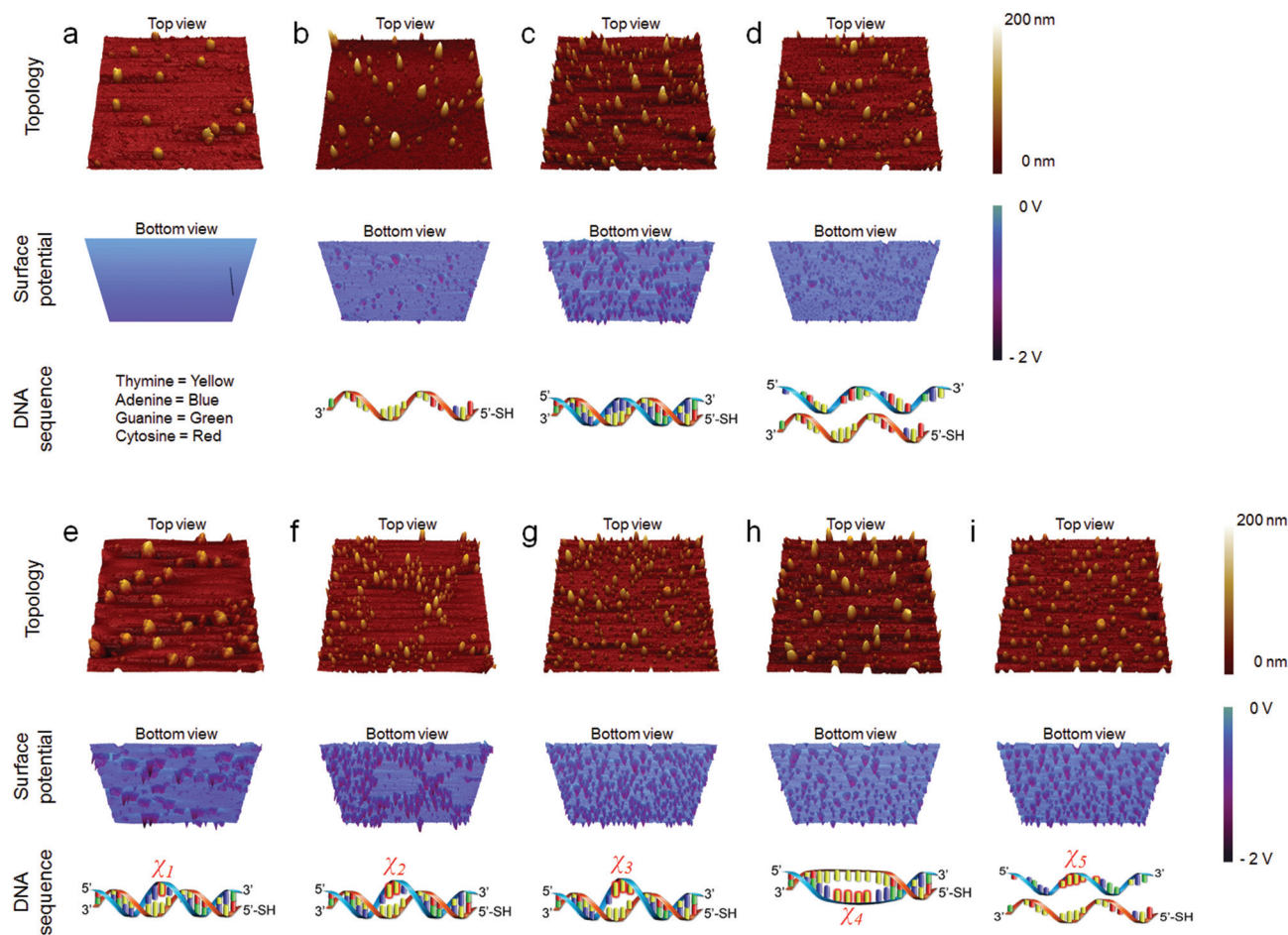

Fig. 3 Bioassay for identification of gene mutations by KPFM. (a−i) 3D images (10 × 10 μm²) of the DCNPs (top view images for topography and bottom view images for surface potential): (a) bare GNP (100 nm), (b) *p*DCNP, (c) *p-c*DCNP, (d) *p-nc*DCNP, (e) *p-χ₁*DCNP, (f) *p-χ₂*DCNP, (g) *p-χ₃*DCNP, (h) *p-χ₄*DCNP, and (i) *p-χ₅*DCNP. Schematic illustrations of the DNA sequences represent the hybridization states between *p*DNA and *t*DNA (with *t* = *c*, *χ₁*−*χ₅*, and *nc*). Every image contains dozens of DCNPs, the surface potential distribution of which was acquired through a single KPFM image.

the p-tDCNPs obtained by KPFM. All images were acquired at the optimal scan speed of 10 μm s^{-1} under ambient conditions (ESI Fig. S4†). The 3D images show that the DCNPs protruded into positive topographic values, whereas they exhibited negative potential values with hollow cone-shaped patterns in surface potential images. This trend can be seen more clearly in the bottom view of the 3D surface potential images compared with the top view images (ESI Fig. S5†). For precise point mutation detection, it was essential to conduct a negative control experiment in which we could check whether the pDCNPs interact with the non-complementary tDNA (ncDNA). Fig. 3d depicts the KPFM result of the p-ncDCNPs, showing ϕ = ~0.747 V, which is equivalent to the surface potential distribution (ϕ = ~0.743 V) of the original pDCNPs that were not exposed to any tDNA (see ESI Table S1†). This indicates that the pDCNPs rarely interacted with the ncDNA.

Fig. 3e–i reveal a consistent trend of surface potential distributions of the p-tDCNPs with respect to the number of mismatched nucleotides, from 1 to 5 (χ_{1-5}). We found that the surface potential was likely to decrease as the number of mismatches increased ($\chi_1 \rightarrow \chi_5$). To evaluate whether our approach is applicable to clinical diagnostics, we conducted quantitative and statistical analyses of the surface potential changes of all the p-tDCNPs, including p-c (*i.e.*, p-χ_0), p-χ_1, p-χ_2, p-χ_3, p-χ_4, p-χ_5, and p-ncDCNP. Here, χ represents a point mutation, and the subscript indicates the number of mismatched nucleotides: accordingly, $t = c$ (complementary; χ_0), χ_1–χ_5, or nc (non-complementary).

Fig. 4a depicts the histograms of the topographic height distributions of all the types of p-tDCNPs, each of which was obtained from 15–40 individual complexes captured in a single image under each condition. The topographic heights were 107.0 ± 6.6, 103.8 ± 5.3, 104.4 ± 4.5, 105.3 ± 6.5, 105.2 ± 3.9, 106.3 ± 7.7, and 103.5 ± 6.1 nm for p-c (*i.e.*, p-χ_0), p-χ_1, p-χ_2, p-χ_3, p-χ_4, p-χ_5, and p-ncDCNP, respectively. Our observations reveal that the interaction of a pDCNP with any tDNA did not severely affect its topographic height, as expected. In contrast to the topographic height, the surface potential distribution was severely altered owing to the addition of negative charges with tDNA (Fig. 4b). As anticipated, the surface potential of each p-tDCNP strongly depended on χ. Both the mean value and standard deviation for each case were extracted by the Gaussian fit: p-cDNA (−1.455 ± 0.026 V), p-χ_1DNA (−1.379 ± 0.036 V), p-χ_2DNA (−1.334 ± 0.041 V), p-χ_3DNA (−1.156 ± 0.044 V), p-χ_4DNA (−0.912 ± 0.055 V), p-χ_5DNA (−0.748 ± 0.034 V), and p-ncDNA (−0.747 ± 0.041 V). Given these results, it is obvious that the absolute value of the surface potential decreases as χ increases. This implies that the existence of mismatched nucleotides in tDNA weakens the binding affinity of tDNA for pDNA.

We plotted a box chart and performed t-tests using the surface potential data of both neighboring conditions (*i.e.*, p-c/p-χ_1, p-χ_1/p-χ_2, p-χ_2/p-χ_3, p-χ_3/p-χ_4, and p-χ_4/p-χ_5), the results of which are shown in Fig. 4c and d. In the t-tests, all p-values were much less than 0.05. As a result, we show that the surface potential measurement of the p-tDCNPs enabled reliable dis-
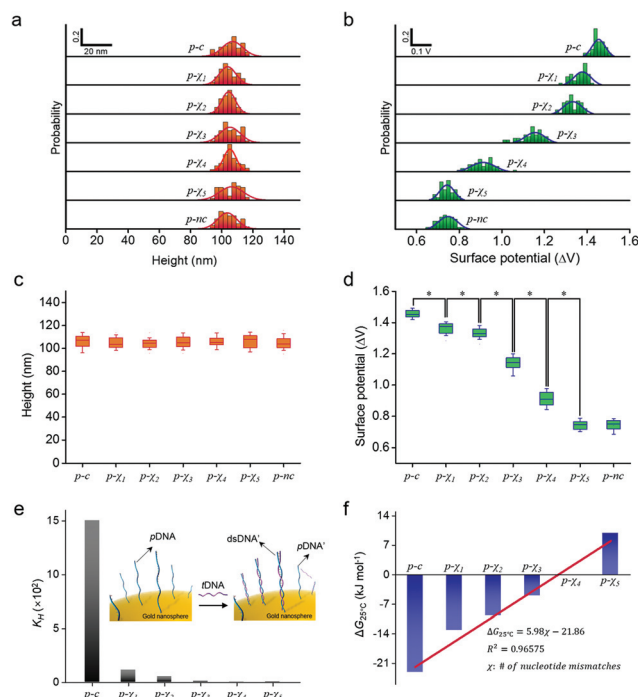

Fig. 4 Statistical analysis of DCNPs for the precise identification of gene mutations. (a and b) Histograms of the height and surface potential of the p-tDCNPs with $t = c$, χ_1–χ_5, and nc. A Gaussian fit was added to each histogram, which helped to determine the trend of surface potential distribution. (c and d) Box plots of the height and the surface potential distributions. In all cases, the center line represents the median, and the lower and upper limits of the box correspond to the 25th and 75th percentiles, respectively. To confirm whether our approach can discriminate between two neighboring conditions, t-tests were performed to compare the results between each group. P-Values were calculated using a t-test (*P < 0.05). (e) The calculated equilibrium constant (K_H) of each DNA hybridization reaction is shown (Fig. 4b and Table S2†). Inset: Schematic illustration of a DNA hybridization reaction on a GNP surface. (f) Gibbs free energy (ΔG) of each DNA hybridization reaction extracted from (b).

crimination of single-nucleotide level variations in the gene. This also suggests that a single KPFM image is sufficient for an accurate discrimination of single and/or multiple nucleotide mutations in disease-specific genes.

To quantify the binding affinity between pDNA and tDNA with different numbers of nucleotide mutations (χ_{1-5}), we used statistical analysis to evaluate the kinetics of DNA hybridization. Using a stoichiometric approach,[38,39] we extracted the equilibrium constants from the chemical interaction between oligonucleotides with different point mutations, (χ). Specifically, we formulated a statistical analysis of the DNA hybridization between pDNA and tDNA according to the following equation:

$$p\text{DNA} + t\text{DNA} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \text{dsDNA} \qquad (1)$$
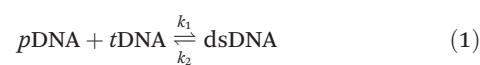
where k_1 is the rate constant for the forward reaction, and k_2 is the rate constant for the reverse reaction. Thus, the equili-

brium constant for DNA hybridization (K_H) between pDNA and tDNA is determined by:[38,39]

$$k_H = \frac{k_1}{k_2} = \frac{[\text{dsDNA}']}{[p\text{DNA}'][t\text{DNA}']} = \frac{|\phi_{\text{dsDNA}'}|}{|\phi_{p\text{DNA}'}||\phi_{t\text{DNA}'}|} \quad (2)$$

where the prime (′) denotes the remaining components in the equilibrium state after the chemical reaction (inset in Fig. 4e). ϕ indicates the surface potential of each type of DNA. Note that ϕ does not represent the surface potential of a single molecule of DNA, but the average potential of dozens of DCNP ensembles spread on the gold surface. Under the optimal KPFM measurement conditions (ϕ_{Au} substrate $- \phi_{\text{GNP}} \approx 0$ V) (Fig. 2g) it enables the estimation of the pure contribution of any DNA to the surface potential, which makes the calculation of K_H successful. When pDNA and tDNA are fully hybridized and form double strands, it is assumed that the surface potential of a DCNP is acquired by KPFM measurements, Φ is twice that of the pDNA surface potential ($\Phi = \phi_{\text{dsDNA}'} = 2\phi_{p\text{DNA}}$). Of course, this assumption was confirmed by the previous experiment (Fig. 2h and i). However, in other cases where an insufficient amount of tDNA was used, dsDNA and the remaining pDNA coexisted on GNPs after hybridization. Accordingly, Φ is the total potential by individual contributions of the remaining pDNA ($\phi_{p\text{DNA}'}$) and dsDNA ($\phi_{\text{dsDNA}'}$) after the hybridization reaction ($\Phi = \phi_{\text{dsDNA}'} + \phi_{p\text{DNA}'} = 2\phi_{p\text{DNA}} - \phi_{t\text{DNA}'}$). Substituting this condition into eqn (2) provides the following K_H equation consisting of Φ:

$$K_H = \frac{|\Phi - \phi_{p\text{DNA}'}|}{|\phi_{p\text{DNA}'}|(2\phi_{p\text{DNA}} - \Phi)} \quad (3)$$

In the above equation, measurable values among the surface potentials are $\phi_{p\text{DNA}}$ and Φ, which is shown in Fig. 2 and 3. This means that transforming eqn (3) into a function of $\phi_{p\text{DNA}}$ and Φ is necessary for the convenient calculation of K_H. Thus, we have assumed that, after hybridization, both amounts of the remaining pDNA ($\phi_{p\text{DNA}'}$) on GNPs and the remaining tDNA ($\phi_{t\text{DNA}'}$) in solution are equal ($\phi_{p\text{DNA}'} = \phi_{t\text{DNA}'}$). For an ideal case of $\phi_{p\text{DNA}'} = \phi_{t\text{DNA}'}$, eqn (3) can be further simplified to:

$$K_H = \frac{|\Phi - (2\phi_{p\text{DNA}} - \Phi)|}{|(2\phi_{p\text{DNA}} - \Phi)|^2} \quad (4)$$

Based on the model (eqn (4)), we calculated the K_H of the hybridization reaction between pDNA and tDNA, the result of which is shown in Fig. 4e. The result reveals that the K_H of the p-tDCNP exponentially diminished asymptotically as χ increased. The K_H values of the p-tDCNPs with χ_{1-3} exceeded 1, implying that the forward reaction (hybridization) was predominant. On the other hand, we found that $K_H = 1.031$ in the case of χ_4 (Fig. 4f), which indicates that the rate of the forward reaction was almost equal to the rate of the reverse reaction. Moreover, the K_H in the case of χ_5 was much less than 1 (~0.019), indicating that tDNAs with χ_5 were easily detached from the pDCNP by thermal drift because the binding affinity for pDNA was extremely low.

Indeed, we found that the K_H of the p-χ_1DCNP was reduced by 86.612% in comparison with that of the p-cDCNP (ESI Fig. S6†). For quantitative comparison and a further understanding of the thermodynamics given χ, we extracted the Gibbs free energy (ΔG), which estimates the chemical potential during DNA–DNA interactions,[40] from the relationship between ΔG and K_H using the following equation: $\Delta G = -RT \ln K_H$, where R is the molar gas constant (8.314 J mol K^{-1}) and T is the absolute temperature in Kelvin. The ΔGs of all the DNA–DNA interactions are plotted in Fig. 4f. The ΔG of DNA hybridization linearly increased as a function of χ, $i.e.$, it was well-expressed as $\Delta G_{25°C} = 5.98\chi - 21.86$. This result not only determines the spontaneity of the chemical reaction in the thermodynamics of DNA hybridization taking place on the GNP but also shows the ability of our sensor system to discriminate nucleotide mutations at the single-nucleotide level. For example, the negative value of ΔG supports that the hybridization processes of all p-tDCNPs occurred spontaneously. By contrast, for p-χ_5 and p-ncDNA, the hybridization processes were not favorable, as indicated by the positive values of ΔG. ΔG is useful not only for an intuitive understanding of the point mutations but also for determining the ability to capture tDNA. For instance, we realized that, in our system, it was meaningless to conduct a hybridization experiment using tDNA with more than χ_5. In other words, the precise detection of point mutations under χ_4 is possible, at least for the $BRCA1$ sequence.

For further verification of our sensor system, a full understanding of how the quantity of the tDNA added affects the results is of essential importance. It is possible that, in our system, an insufficient quantity of tDNA will produce misleading results, comparable to the results obtained using samples with a large number of point mutations (χ). We performed the surface potential analysis of p-cDNA, p-χ_1DNA, and p-ncDNA at 3 different concentrations: 768 pM, 7.68 nM, and 76.8 nM (Fig. 5, ESI Fig. S7, and Table S2†). At 76.8 nM tDNA, both the p-cDNA and the p-χ_1DCNPs went through full hybridization, exhibiting a single narrow Gaussian distribution in surface potential. Remarkably, at lower tDNA concentrations (7.68 nM and 768 pM), we found that both the p-cDNA and the p-χ_1DCNPs exhibited bimodal Gaussian distributions in surface potential with a large gap between the two peaks. All left peaks in the bimodal distributions in Fig. 5d, e, g, and h are 0.741–0.765 V in average potential, which is seemingly identical to that ($\phi = $ ~0.743 V) of the original pDCNP before hybridization (ESI Table S1†). Therefore, this peak indicates that these complexes experienced no hybridization with the tDNA, and were left unbound. Conversely, each right peak in Fig. 5d, e, g, and h shows a much higher average potential than that of the corresponding left peak, implying that the hybridization reaction of this group (the right peaks) of the complexes was nearly complete. It is noteworthy that the amount of DCNPs left unbound decreased as the tDNA concentration increased. Lastly, for a negative control, we confirmed that the pDCNP never interacted with the ncDNA at any concentration, instead exhibiting a single Gaussian distribution with $\phi = 0.741$–0.752 V ($i.e.$, <1.5%).
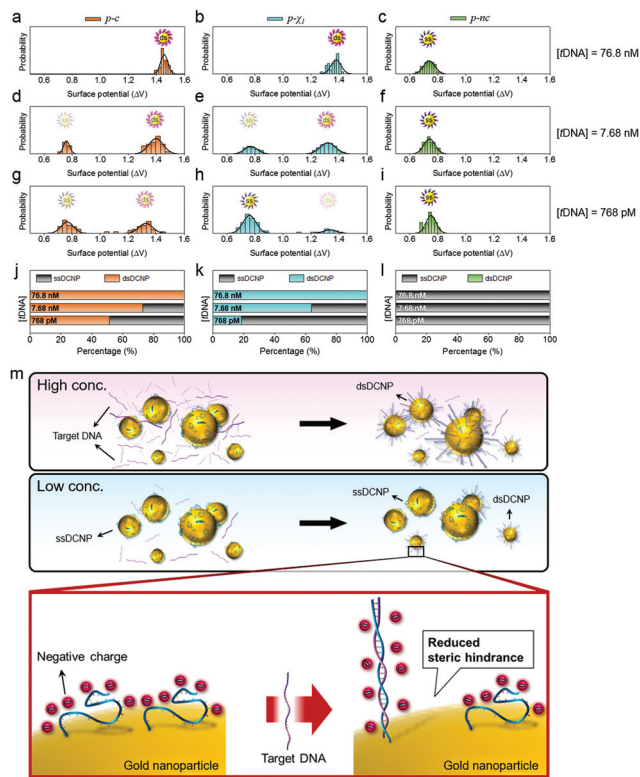
Fig. 5 Statistical analysis of the surface potential distributions of the *p*-*t*DCNPs with different *t*DNA concentrations. (a–i) Surface potential distributions of *p*-*t*DCNPs with different concentrations of *t*DNA, including (a, d, and g) *c*DNA, (b, e, and h) χ_1DNA, and (c, f, and i) *nc*DNA. (j–l) Quantification of the ratio of the two different states (*i.e.*, hybridized and unbound states) of the DCNPs with different concentrations of *t*DNA, including (j) *c*DNA, (k) χ_1DNA, and (l) *nc*DNA. (m) The schematic illustration shows the conformational change of the DNA molecule, which provides the remnant *t*DNA enough room to easily bind with the neighboring *p*DNA around dsDNA; this may occur in a cascading fashion.

The above result has never been obtained using any other detection methods, such as colorimetric, fluorescence, electrochemical, or nanomechanical routes. Accordingly, we speculate that the existence of two split Gaussian distributions can be used as a measure to estimate whether an unknown sample (with no concentration information) has sufficient *t*DNA concentration for a reliable bioassay. The following description (Fig. 5m) may explain why this peculiar behavior of the surface potential distribution takes place at low *t*DNA concentrations. DNA hybridization on gold surfaces has been known to occur randomly but in a cascading fashion.[41] The hybridization of the first single *t*DNA (*t* = c or χ_1) molecule onto the *p*DCNP requires high activation energy, because of steric hindrance owing to its neighboring *p*DNA molecules on the GNP. Once the first *t*DNA is hybridized with a single flexible *p*DNA on the GNP, a conformational change takes place, as the flexible ssDNA becomes a stiff double helix dsDNA.[42] This conformational change provides the remnant *t*DNA enough room to easily bind with the *p*DNA surrounding the dsDNA, remarkably decreasing the activation energy for hybridization.

Accordingly, subsequent hybridizations after the first occur rapidly. This explains why there are bimodal Gaussian distributions without any intermediate states in the surface potential of the *p*-*t*DCNPs at low *t*DNA concentrations.

As described above, we observed a bimodal Gaussian distribution in the surface potential of the *p*-*t*DCNPs at low *t*DNA concentrations, exhibiting a different population ratio between the two peaks under each condition. Herein, the population ratio (P_u/P_h) of the unbound to the hybridized DCNPs under each condition is thought to be a significant factor that can be directly correlated with the amount of the *t*DNA that exists in the target sample. As the *t*DNA concentration increases, P_u/P_h gradually decreases, finally reaching zero. This trend is similar for both the cases of *p*-*c*DCNPs and *p*-*m*$_1$DCNPs, but the detailed characteristic is somewhat different (Fig. S8†). The ratio of P_u/P_h for the *p*-*m*$_1$DCNP varies steeper with the *t*DNA concentration than that of the *p*-*c*DCNP. When an unknown sample is given, measuring the surface potentials of both the excessive and the insufficient *t*DNA cases will help one accomplish the precise identification of its mismatch type. In addition, it is possible that this behavior will enable the development of other applications, including those useful for quantitative analysis of very small amounts of biological samples such as mRNA or microRNA extracted from disease-related rare cells,[43–45] or for analysis of viral genes.[46]

Conclusions

In this study, we confirmed that the combination of DCNPs and KPFM is a promising tool to identify the existence of DNA mutations with a different number of mismatched nucleotides, from 1 to 5, compared with a complementary sequence. To overcome the very long analytical time of KPFM, which is a critical drawback, we adopted *p*DNA immobilized on GNPs (*p*DCNPs) to serve as a nanoscale platform to capture target DNA (mutated *BRCA1*) during the hybridization reaction. Our results demonstrate that the average surface potential of *p*-*t*DCNPs steadily decreased as the number of mismatched nucleotides increased, exhibiting a manifest difference ($P <$ 0.05 in the *t*-test) in the surface potential between two neighboring conditions (*i.e.*, *p*-c/*p*-χ_1, *p*-χ_1/*p*-χ_2, *p*-χ_2/*p*-χ_3, *p*-χ_3/*p*-χ_4, and *p*-χ_4/*p*-χ_5). Thus, the surface potential ϕ of the DCNPs is a potentially useful factor that enables the accurate detection of nucleotide mutations in disease-related genes at the single-nucleotide level. The calculation of the equilibrium constant from the surface potential data (Φ and $\phi_{pDNA'}$) allows us to fully understand the binding affinity of *t*DNA in the hybridization reaction. Compared with other conventional optical and electrochemical detection methods,[18,47,48] our approach has several advantages, such as easy preparation, label-free detection, expeditious detection, high accuracy, and high reproducibility. Moreover, the statistical analysis of dozens of ensembles (*i.e.*, dozens of *p*-*t*DCNPs) could be performed in only one round of KPFM imaging, suggesting that our method enables efficient sequence-specific detection of point mutations. We

consider that this approach is applicable not only to the investigation into the relationships between nucleotide mutations and disease expression, but also to the detection of other disease-related biomolecules such as proteins, enzymes, peptides, carbohydrates, or even toxic ions.

Author contributions

H. L., G. L. and D. S. Y. conceived and designed this research. S. W. L., H. L. and W. L. performed the experiments. S. W. L., H. L., W. L., G. L., K. N., J. H. L., and K. S. H. analyzed the experimental data. H. L., G. L., J. Y., H. L., S. K., S. W. L. and D. S. Y. wrote the paper. D. S. Y. supervised the project.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 R. Dawkins, *The greatest show on earth: The evidence for evolution*, Simon and Schuster, 2009.
2 E. Nevo, A. Beiles and R. Ben-Shlomo, in *Evolutionary Dynamics of Genetic Diversity*, ed. G. S. Mani, Springer, Berlin, Heidelberg, 1984, pp. 13–213.
3 H. Raeder, S. Johansson, P. I. Holm, I. S. Haldorsen, E. Mas, V. Sbarra, I. Nermoen, S. A. Eide, L. Grevle, L. Bjorkhaug, J. V. Sagen, L. Aksnes, O. Sovik, D. Lombardo, A. Molven and P. R. Njolstad, *Nat. Genet.*, 2006, **38**, 54–62.
4 C. Greenman, P. Stephens, R. Smith, G. L. Dalgliesh, C. Hunter, G. Bignell, H. Davies, J. Teague, A. Butler, C. Stevens, S. Edkins, S. O'Meara, I. Vastrik, E. E. Schmidt, T. Avis, S. Barthorpe, G. Bhamra, G. Buck, B. Choudhury, J. Clements, J. Cole, E. Dicks, S. Forbes, K. Gray, K. Halliday, R. Harrison, K. Hills, J. Hinton, A. Jenkinson, D. Jones, A. Menzies, T. Mironenko, J. Perry, K. Raine, D. Richardson, R. Shepherd, A. Small, C. Tofts, J. Varian, T. Webb, S. West, S. Widaa, A. Yates, D. P. Cahill, D. N. Louis, P. Goldstraw, A. G. Nicholson, F. Brasseur, L. Looijenga, B. L. Weber, Y.-E. Chiew, A. deFazio, M. F. Greaves, A. R. Green, P. Campbell, E. Birney, D. F. Easton, G. Chenevix-Trench, M.-H. Tan, S. K. Khoo, B. T. Teh, S. T. Yuen, S. Y. Leung, R. Wooster, P. A. Futreal and M. R. Stratton, *Nature*, 2007, **446**, 153–158.
5 P. Hsieh and K. Yamane, *Mech. Ageing Dev.*, 2008, **129**, 391–407.
6 J. K. Pickrell, T. Berisa, J. Z. Liu, L. Segurel, J. Y. Tung and D. A. Hinds, *Nat. Genet.*, 2016, **48**, 709–717.
7 F. Di Silverio, P. Casale, D. Colella, L. Andrea, F. Seccareccia and A. Sciarra, *Cancer*, 2000, **88**, 835–843.
8 M. J. Garnett, E. J. Edelman, S. J. Heidorn, C. D. Greenman, A. Dastur, K. W. Lau, P. Greninger, I. R. Thompson, X. Luo, J. Soares, Q. Liu, F. Iorio, D. Surdez, L. Chen, R. J. Milano, G. R. Bignell, A. T. Tam, H. Davies, J. A. Stevenson, S. Barthorpe, S. R. Lutz, F. Kogera, K. Lawrence, A. McLaren-Douglas, X. Mitropoulos, T. Mironenko, H. Thi, L. Richardson, W. Zhou, F. Jewitt, T. Zhang, P. O'Brien, J. L. Boisvert, S. Price, W. Hur, W. Yang, X. Deng, A. Butler, H. G. Choi, J. W. Chang, J. Baselga, I. Stamenkovic, J. A. Engelman, S. V. Sharma, O. Delattre, J. Saez-Rodriguez, N. S. Gray, J. Settleman, P. A. Futreal, D. A. Haber, M. R. Stratton, S. Ramaswamy, U. McDermott and C. H. Benes, *Nature*, 2012, **483**, 570–575.
9 S. H. Lee, V. S. Vigliotti, J. S. Vigliotti, W. Jones and S. Pappu, *Am. J. Clin. Pathol.*, 2010, **133**, 569–576.
10 J. Mertens, C. Rogero, M. Calleja, D. Ramos, J. A. Martin-Gago, C. Briones and J. Tamayo, *Nat. Nanotechnol.*, 2008, **3**, 301–307.
11 F. Huber, H. P. Lang, N. Backmann, D. Rimoldi and C. Gerber, *Nat. Nanotechnol.*, 2013, **8**, 125–129.
12 S. J. Hurst, M. S. Han, A. K. Lytton-Jean and C. A. Mirkin, *Anal. Chem.*, 2007, **79**, 7201–7205.
13 H. Cho, J. Jung and B. H. Chung, *Chem. Commun.*, 2012, **48**, 7601–7603.
14 T. A. Taton, C. A. Mirkin and R. L. Letsinger, *Science*, 2000, **289**, 1757–1760.
15 H. Lee, S. W. Lee, G. Lee, W. Lee, J. H. Lee, K. S. Hwang, J. Yang, S. W. Lee and D. S. Yoon, *Nanoscale*, 2016, **8**, 13537–13544.
16 A. K. Sinensky and A. M. Belcher, *Nat. Nanotechnol.*, 2007, **2**, 653–659.
17 S. M. Iqbal, D. Akin and R. Bashir, *Nat. Nanotechnol.*, 2007, **2**, 243–248.
18 I. M. Derrington, T. Z. Butler, M. D. Collins, E. Manrao, M. Pavlenok, M. Niederweis and J. H. Gundlach, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 16060–16065.
19 B. M. Venkatesan and R. Bashir, *Nat. Nanotechnol.*, 2011, **6**, 615–624.
20 C. J. Yu, Y. Wan, H. Yowanto, J. Li, C. Tao, M. D. James, C. L. Tan, G. F. Blackburn and T. J. Meade, *J. Am. Chem. Soc.*, 2001, **123**, 11155–11161.
21 N. Moran, D. M. Bassani, J. P. Desvergne, S. Keiper, P. A. Lowden, J. S. Vyle and J. H. Tucker, *Chem. Commun.*, 2006, **48**, 5003–5005.

22 B. Dubertret, M. Calame and A. J. Libchaber, *Nat. Biotechnol.*, 2001, **19**, 365–370.

23 E. Papadopoulou and S. E. Bell, *Angew. Chem., Int. Ed.*, 2011, **50**, 9058–9061.

24 H. Lee, W. Lee, J. H. Lee and D. S. Yoon, *J. Nanomater.*, 2016, **2016**, 21.

25 A. Antoniou, P. D. Pharoah, S. Narod, H. A. Risch, J. E. Eyfjord, J. L. Hopper, N. Loman, H. Olsson, O. Johannsson, A. Borg, B. Pasini, P. Radice, S. Manoukian, D. M. Eccles, N. Tang, E. Olah, H. Anton-Culver, E. Warner, J. Lubinski, J. Gronwald, B. Gorski, H. Tulinius, S. Thorlacius, H. Eerola, H. Nevanlinna, K. Syrjakoski, O. P. Kallioniemi, D. Thompson, C. Evans, J. Peto, F. Lalloo, D. G. Evans and D. F. Easton, *Am. J. Hum. Genet.*, 2003, **72**, 1117–1130.

26 H. D. Hill, J. E. Millstone, M. J. Banholzer and C. A. Mirkin, *ACS Nano*, 2009, **3**, 418–424.

27 S. J. Park, T. A. Taton and C. A. Mirkin, *Science*, 2002, **295**, 1503–1506.

28 C. M. Niemeyer, B. Ceyhan and P. Hazarika, *Angew. Chem., Int. Ed.*, 2003, **42**, 5766–5770.

29 F. Li, H. Zhang, B. Dever, X.-F. Li and X. C. Le, *Bioconjugate Chem.*, 2013, **24**, 1790–1797.

30 Y. Ebenstein, E. Nahum and U. Banin, *Nano Lett.*, 2002, **2**, 945–950.

31 G. Lee, H. Lee, K. Nam, J. H. Han, J. Yang, S. W. Lee, D. S. Yoon, K. Eom and T. Kwon, *Nanoscale Res. Lett.*, 2012, **7**, 608.

32 J. Park, J. Yang, G. Lee, C. Y. Lee, S. Na, S. W. Lee, S. Haam, Y.-M. Huh, D. S. Yoon, K. Eom and T. Kwon, *ACS Nano*, 2011, **5**, 6981–6990.

33 G. Lee, W. Lee, H. Lee, S. Woo Lee, D. Sung Yoon, K. Eom and T. Kwon, *Appl. Phys. Lett.*, 2012, **101**, 043703.

34 G. Lee, W. Lee, H. Lee, C. Y. Lee, K. Eom and T. Kwon, *Sci. Rep.*, 2015, **5**, 16220.

35 K. Nam, K. Eom, J. Yang, J. Park, G. Lee, K. Jang, H. Lee, S. W. Lee, D. S. Yoon, C. Y. Lee and T. Kwon, *J. Mater. Chem.*, 2012, **22**, 23348–23356.

36 N. T. Khoa, S. W. Kim, D.-H. Yoo, E. J. Kim and S. H. Hahn, *Appl. Catal., A*, 2014, **469**, 159–164.

37 S. Darwich, K. Mougin, A. Rao, E. Gnecco, S. Jayaraman and H. Haidara, *Beilstein J. Nanotechnol.*, 2011, **2**, 85–98.

38 P. W. Stevens, M. R. Henry and D. M. Kelso, *Nucleic Acids Res.*, 1999, **27**, 1719–1727.

39 D. Erickson, D. Li and U. J. Krull, *Anal. Biochem.*, 2003, **317**, 186–200.

40 R. McKendry, J. Zhang, Y. Arntz, T. Strunz, M. Hegner, H. P. Lang, M. K. Baller, U. Certa, E. Meyer, H.-J. Güntherodt and C. Gerber, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 9783–9788.

41 J. Zheng, G. Zhu, Y. Li, C. Li, M. You, T. Chen, E. Song, R. Yang and W. Tan, *ACS Nano*, 2013, **7**, 6545–6554.

42 J. Fritz, *Analyst*, 2008, **133**, 855–863.

43 S. Husale, H. H. Persson and O. Sahin, *Nature*, 2009, **462**, 1075–1078.

44 H. M. Heneghan, N. Miller, A. J. Lowery, K. J. Sweeney, J. Newell and M. J. Kerin, *Ann. Surg.*, 2010, **251**, 499–505.

45 G. K. Joshi, S. Deitz-McElyea, M. Johnson, S. Mali, M. Korc and R. Sardar, *Nano Lett.*, 2014, **14**, 6955–6963.

46 T. D. Harris, P. R. Buzby, H. Babcock, E. Beer, J. Bowers, I. Braslavsky, M. Causey, J. Colonell, J. DiMeo and J. W. Efcavitch, *Science*, 2008, **320**, 106–109.

47 S. M. Yoo, T. Kang, H. Kang, H. Lee, M. Kang, S. Y. Lee and B. Kim, *Small*, 2011, **7**, 3371–3376.

48 I. Singh, C. Wendeln, A. W. Clark, J. M. Cooper, B. J. Ravoo and G. A. Burley, *J. Am. Chem. Soc.*, 2013, **135**, 3449–3457.