Clinical Validation of Quantum Dot Barcode Diagnostic Technology

Jisung Kim, Mia J. Biondi, Jordan J. Feld, and Warren C. W. Chan

Version  Post-Print/Accepted Manuscript


Publisher’s Statement  This document is the Accepted Manuscript version of a Published Work that appeared in final form in ACS Nano, copyright ©American Chemical Society after peer review and technical editing by the publisher. To access the final edited and published work see http://dx.doi.org/10.1021/acsnano.6b01254.

How to cite TSpace items

Always cite the published version, so the author(s) will receive recognition through services that track citation counts, e.g. Scopus. If you need to cite the page number of the TSpace version (original manuscript or accepted manuscript) because you cannot access the published version, then cite the TSpace version in addition to the published version using the permanent URI (handle) found on the record page.
Clinical Validation of Quantum Dot Barcode Diagnostic Technology

Jisung. Kim,\textsuperscript{a,b} Mia. J. Biondi,\textsuperscript{c} Jordan. J. Feld,\textsuperscript{c,*} Warren. C. W. Chan\textsuperscript{a,b,d,e,f,*}

\textsuperscript{a}Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Ontario M5S 3G9, Canada
\textsuperscript{b}Terrence Donnelly Centre for Cellular and Bimolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada
\textsuperscript{c}Sandra Rotman Centre for Global Health, University Health Network, Toronto, Ontario M5G 1L7, Canada
\textsuperscript{d}Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada
\textsuperscript{e}Department of Chemical Engineering, University of Toronto, Toronto, Ontario M5S 3E5, Canada
\textsuperscript{f}Department of Materials Science and Engineering, University of Toronto, Toronto, Ontario M5S 3E4, Canada

*Corresponding Authors:

Warren. C. W. Chan, 164 College Street, Toronto, ON, Canada, M5S 3G9, 416-946-8416,

Jordan. J. Feld, 101 College Street, Toronto, ON, Canada, M5G 1L7, 416-603-6230,

Keywords: Multiplex, Quantum Dot, Quantum Dot Barcodes, Sensitivity, Specificity, Infectious Diseases, Nanotechnology, Hepatitis B Virus, Diagnosis, \textit{in vitro} diagnostic assay

Abstract: There has been a major focus on the clinical translation of emerging technologies for diagnosing patients with infectious diseases, cancer, heart disease, and diabetes. However, most developments still remain at the academic stage where researchers use spiked target molecules to demonstrate the utility of a technology and to assess the analytical performance. This approach does not account for the biological complexities and variabilities of human patient samples. As a technology matures and potentially becomes clinically viable, one important intermediate step in the translation process is to conduct a full clinical validation of the technology using a large number of patient samples. Here we present a full detailed clinical validation of Quantum Dot (QD) barcode technology for diagnosing patients infected with Hepatitis B Virus (HBV). We further demonstrate that the detection of multiple regions of the viral genome using multiplexed QD barcodes improved clinical sensitivity from 54.9-66.7\% to 80.4-90.5\%, and describe how to use QD barcodes for optimal clinical diagnosis of patients. The use of QDs in biology and medicine was first introduced in 1998 but has not reached clinical care. This study describes our long-term systematic development strategy to advance QD technology to a clinically feasible product for diagnosing patients. Our “blueprint” for translating the QD barcode research concept could be adapted for other nanotechnologies, to efficiently advance diagnostic techniques discovered in the academic laboratory to patient care.
There has been a significant number of discussions on the need to translate academic discoveries into clinical utility for improving patient care.\textsuperscript{1,2} Despite this focus, there are only a few nanotechnology-based diagnostic devices that advanced to use in clinical settings (e.g. gold nanoparticle bio-barcode assay developed by Mirkin’s group).\textsuperscript{3,4} In a typical translation process, the first step is to use laboratory-prepared mock samples to evaluate the technology’s ability to detect the biomarkers in the spiked sample via measurements of the analytical sensitivity, limits of detection, dynamic range, and cross-reactivity. This is followed by the analysis of a few clinical samples (~10 patient samples) to show feasibility of diagnostics with real-world patient samples. The final step before advancing a technology beyond the academic research objective is to conduct a full clinical evaluation (>50 patient samples) to obtain clinical sensitivity and specificity values. These numbers indicate the degree of false positive and negative detections of a diagnostic device. Interestingly, there is a large number of published studies from the initial step, less publications in the subsequent step, and to a handful of studies in the clinical validation step. As a result, the clinical feasibility of many nanotechnology-based diagnostic devices is unclear because only few published studies have explored the clinical sensitivity and specificity of these systems using patient samples.

Over the last 15 years, we have followed the three development steps to advance QD barcoding technology to diagnose patients. The concept of QD barcoding was first presented by Nie and co-workers where they demonstrated the use of these barcodes to detect spiked DNA sequences in buffer.\textsuperscript{5} They reported that QD barcodes are advantageous for multiplexed diagnostic detection over the traditional barcodes doped with organic fluorophores (e.g. Luminex barcodes) because one single light source can excite multiple barcode emissions,\textsuperscript{6,7} and QDs can achieve better discrimination and identification of barcodes due to narrower spectral line width and longer photostability respectively.\textsuperscript{8-10} There is also a greater capacity to design more barcode signals by using QDs because the fluorescence of the QDs can be tuned by manipulating the size, shape, and chemical composition.\textsuperscript{11} Although these QD properties would allow simplified readout and cost-effective diagnostic procedure for multiplex detection of diseases in patients, the QD barcoding technology has not advanced to the clinic. In this study, we advance QD barcode technology one step closer to use in patient care by specifically evaluating the feasibility of QD barcodes to diagnose over 70 clinical samples.

Our main objective was to develop and translate QD barcode technologies for diagnosis of patients. We overcame the first barrier to using QD barcodes for clinical analysis by developing and optimizing methods to synthesize stable QD barcodes with consistent fluorescence emissions and monodispersed sizes in 2008.\textsuperscript{12} As we started to study and evaluate different methods to synthesize QD barcodes, we found that past methods for generating barcodes such as swelling,\textsuperscript{5} polymerization,\textsuperscript{13} and QD entrapment inside mesoporous silica beads\textsuperscript{14} were problematic for broad use of the technology in patients. Barcodes that were synthesized using the swelling method led to misdetection because QDs were loaded onto the surface of the bead and became exposed to small ions, which led to varied fluorescence signals in differing buffer conditions. Moreover, QD signals were typically quenched or altered by initiator using the polymerization method. Finally, there was a high degree of non-specific binding for barcodes prepared using a silica approach, which prevented the ability to detect biological targets with a detection limit that is competitive to conventional analytical techniques.\textsuperscript{15} To mitigate these problems, we developed a fluidic-flow focusing approach to reproducibly synthesize monodispersed QD barcodes,\textsuperscript{12,16} which automated the manufacturing of the barcodes. This strategy has been further developed by other researchers for large scale preparation of QD barcodes.\textsuperscript{17} Subsequently, we optimized protocols and methods
to conjugate bio-recognition molecules on the surface of QD barcodes\textsuperscript{18} and demonstrated the ability to detect genetic and protein targets spiked in buffer.\textsuperscript{5,19} These spiked buffer samples were used to characterize the analytical performance of the QD barcode assay.

In a typical single-plexed QD barcode assay (\textit{i.e.} detection of one molecule type), the barcode is mixed with a buffer solution containing fluorescently labelled detection probe. If this mixed solution also contains the target analyte that recognizes the bio-recognition molecule conjugated to the barcode surface and detection probe, a sandwich structure is formed between these two molecules. The barcode’s optical signal now includes the fluorescence emitted from the detection probe in addition to the original fluorescence generated from QDs, which signifies a positive detection. In the absence of the target analyte, a negative detection is signified by the barcode signal that remains identical to its initial state. For detection of multiple molecule types in solution (\textit{i.e.} multiplexed assay, Figure 1), a library of different emitting barcodes is prepared, conjugated with different bio-recognition molecules, and added to a single vial. The barcodes conjugated with bio-recognition molecules identify the specific molecule types to be detected. Optical read-out (\textit{e.g.} flow cytometer) will identify which barcodes contain the secondary probe signals, and determine the presence of the molecule in solution. Using this multiplexing strategy, different types of diseases can be detected by designing various panels of barcodes.\textsuperscript{6,7,19} Positive and negative control barcodes are also added in a typical multiplexed QD barcode assay, where the target DNA corresponding to positive control barcode is always spiked in to validate cross-reactivity and ensure that there is no significant false-negative nor false-positive signals arising from multiplexing.\textsuperscript{20,21}

A conventional QD barcode assay has a LOD ranging from atto- to femtomole DNA per microliter,\textsuperscript{6,7,20} which limits the ability to detect pathogens that are present at lower concentrations, but still require clinical intervention. To overcome this challenge, we added an amplification step to improve the analytical sensitivity. We used a Recombinase Polymerase Amplification (RPA) method, which is an isothermal nucleic acid amplification technique that can achieve high analytical sensitivity by using recombinase-driven nucleoprotein complex to facilitate strand exchange at the homologous sequence of template DNA without the need for temperature cycling required by standard polymerase chain reaction (PCR).\textsuperscript{22} This enabled us to test QD barcodes for diagnosing a small number of patient samples (step 2), which demonstrated the ability of QD barcodes to detect 7 HBV and 10 Human Immunodeficiency Virus (HIV) patient samples.\textsuperscript{21}

Presently, there are two major milestones remaining in the development of QD barcode technology for diagnostic applications. The first is to develop a handheld device that can read and identify QD barcodes. This was demonstrated recently when we developed a smartphone device that can image the barcodes deposited on glass slides and utilized in-house built algorithm to deconvolve the corresponding signals.\textsuperscript{21} This device still requires optimization for practical use in patients but we were able to demonstrate detection of a small number of patient samples (step 2 in the development process). This device is currently being scaled-up prior to full clinical testing. Secondly, there is a need to clinically evaluate the diagnostic feasibility of the QD barcode technology, which is the main focus of the current study. This study would allow us to determine how QD barcode technology compares to current diagnostic technologies (\textit{e.g.} PCR). Thus, the main objective of this study is to quantify the clinical sensitivity and specificity and to evaluate the applicability of these barcodes in a clinical laboratory. This is a final key step, from the academic perspective, for translation of a technology. This clinical evaluation strategy will be adapted to assessing the feasibility of our smartphone device in future studies (step 3).
We chose HBV as the model pathogen, which was used in our previous study, because of the global impact of the disease, the need for improved diagnostic tests and its ability to mutate, thus allowing us to confirm the stability of our platform. HBV-related liver diseases account for 0.5-1 million deaths per year, and approximately 350-400 million people are estimated to be chronic carriers for HBV surface antigen (sAg), although most are unaware of their infection. Current diagnostics for HBV involve the use of standard serological tests followed by quantification of HBV DNA by PCR. DNA detection is an essential component of HBV diagnosis as it is frequently used as an indicator of the need for treatment and is used to monitor patients during antiviral therapy. However, viruses exist as quasi-species, representing a major challenge to the development of nucleic acid-based diagnostic tests. Sequence variations accumulated as a result of high mutation rates in the HBV genome pose a risk of producing false-negative results reducing diagnostic sensitivity. In the absence of therapy, between $10^{11}$ and $10^{13}$ virions are produced per day, with an error rate of $10^{-4}$ to $10^{-5}$ at each base, per round of replication.

HBV is also categorized into genotypes A-J, which can differ by more than 8% and are concentrated in different geographical regions globally. Such variability among DNA sequences can be unfavourable to the RPA process as the use of RPA has been shown to be hampered by mismatches near the 3’-end of primers, and mismatches of greater than 7-9 base pairs. In this study, we overcome these current challenges of diagnosing HBV by (1) amplifying multiple sites within the HBV genome, and (2) using multiple QD barcodes to detect the various sequences simultaneously in a single reaction vessel (Figure 2). The viral DNA is first extracted from patient serum using magnetic microbeads, various regions of the extracted genome are amplified by RPA, amplified products are detected by multiplexed QD barcode assay, and finally fluorescence signals are measured via flow cytometry. Using this strategy, we show that the detection of multiple targets leads to a significant enhancement of the clinical sensitivity for the diagnosis of HBV. Our study fosters the clinical translation of QD barcodes, and presents the clinical feasibility of QD barcodes for hospital use in the near future.

Results and Discussion

Design of QD Barcodes, RPA, and Probe Sequences.

Six QD barcodes were prepared in total by infusing various color and intensity ratios of QDs inside polymeric microbeads using concentration-controlled flow-focusing (CCFF) method (Table S1). To avoid overlap in the optical signals, we prepared our barcode panel by using different combinations of QD concentrations and emission wavelengths. Specifically, we utilized at least 10 times difference in QD concentration between the barcodes (e.g. B5, B2 and B3, Table S1) which can be observed as 10 times incremental increase in the FL2 signal on flow cytometry (Fig. S1). Also, QDs of significantly different emission wavelengths (e.g. B4 and B6, Table S1), and a combination of the various emission wavelengths (e.g. B1, Table S1) were used.

For the single-plexed assay, barcode B1 was conjugated to four capture probes designed to be complementary to the first half of amplicons, and used to detect four conserved regions (R1-R4) within the HBV genome (Fig. S2). For multiplexed assay, four barcodes were conjugated with corresponding capture probes (B1-CP1, B2-CP2, B3-CP3 and B4-CP4, Table S2). B5 and B6 were used as positive and negative controls during multiplexed detection and conjugated with capture probes CPP and CPN respectively. CPP and CPN sequences were designed to be complementary to genes of other sexually transmitted infectious pathogens (Treponema pallidum and Neisseria
gonorrhoeae bacteria respectively), which are known to contain no similarity in their DNA sequences with respect to HBV genome, and therefore assure high specificity for the controls during our multiplexed detection.

The average surface density of capture probe DNA ranged from approximately $5 \times 10^3$ to $11 \times 10^3$ molecules/µm² and $3 \times 10^3$ to $10 \times 10^3$ molecules/µm² for single-plexed and multiplexed detections, respectively (Fig. S3). The six-barcode signals were then confirmed to have unique and distinctive signatures, and to be differentiable from each other on a flow cytometer (Fig. S1). These unique barcode signals were subsequently used for decoding multiple signals.

In order to ensure that our test could be used worldwide, we conducted sequence alignments to assess which areas of the genome would be ideal for primer binding both with respect to conservation and properties for the amplification strategy. Ideally our products would be approximately 100 nucleotides in length as longer target strands have previously been shown to lower hybridization efficiency, and were chosen to minimize the formation of capture and detection probe secondary structures (Table S2). In total, four regions within the circular HBV genome were identified as meeting all criteria for the assay (Fig. S2). QD barcodes can then detect multiple sequences simultaneously in a single reaction, improving the probability of true-positive detection.

**Characterization of Clinical Samples.**

We collected 72 clinical samples from Toronto Western Hospital Liver Clinic, and used them for this study to ensure sufficient statistical power for assessing our diagnostic sensitivity (Figure 3). All samples were first tested by standard clinical tools to identify hepatitis B surface antigen (HBsAg), anti-hepatitis B core antibodies (anti-HBc) and HBV DNA (tested by gold-standard instrument assessing HBV viral load, COBAS Amplicor HBV Monitor Test or the COBAS® AmpliPrep/COBAS® TaqMan® HBV Test, v2.0) with a lower limit of detection of 20-60 IU/mL. Healthy controls were negative for all markers (n=5), those with resolved past infection were negative for HBsAg and HBV DNA with detectable anti-HBc antibodies (n=5) and those with chronic infection were positive for HBsAg and anti-HBc antibodies with varying levels of HBV DNA (n=62). HBV DNA titres were undetectable in 11 patients on suppressive antiviral therapy and ranged from $10^1$ to greater than $10^9$ IU/mL in those with untreated infection (n=51). These 51 samples were identified as patients positive for HBV covering genotypes A-E, accounting for the five most common genotypes globally. Serological and clinical data are shown in Table S3.

**Clinical Sensitivity of a Single QD Barcode Assay.**

A single-plex assay refers to the detection of a single genetic target. We determined the clinical sensitivity and specificity of the four selected genetic HBV targets by comparing results from the gold-standard instrument with a single-plexed QD barcode assay. Extracted viral DNA was directly used for amplification in RPA, where for each sample, the amplification was performed four times using the four primer pairs (FP1-FP4 and RP1-RP4 for forward and reverse primers respectively, Table S2) to produce four amplicons (T1-T4, Table S2). Purified and denatured amplicons were directly incubated with corresponding detection probes that are labeled with Alexa Fluor 647 and designed to be complementary to the second half of amplicons (DP1-
DP4, Table S2), hybridization buffer, and QD barcodes at 37°C for 30 minutes, washed three times with a washing buffer, and stored in PBST buffer for flow cytometry.

Single-plexed detection of four regions is first presented as normalized detection probe intensity plots (Fig. S4). We first determined the cutoff intensity levels \(i.e. 3\) standard deviations above the background signal) by analyzing water control samples. The results are shown in Figure 3A. Signal-to-cutoff value greater than or equal to 1 was considered as positive detection, whereas signal-to-cutoff value lower than 1 was considered as negative detection. Our results show that detection of each individual genomic region achieved an overall clinical sensitivity ranging from 58.8% (95% CI: 44.2-72.1%) to 64.7% (95% CI: 50.0-77.2%), and a specificity of 100% (95% CI: 80.8%-100%) (Figure 4A). The difference in sensitivity among the four amplification regions highlights sequence-dependent amplification performance behavior, which directly affected the assay. Receiver operating characteristic (ROC) curves were also developed from the normalized intensity plots (Figure 4B). These ROC curves describe the tradeoff or relationship between sensitivity and specificity as we vary the cutoff intensity level from 0 to 1 on the normalized detection probe intensity plots,\(^9\) and can also be used to preset cutoff intensity values if one desires to achieve a certain target sensitivity or specificity. Lines that are closer to the top left corner on this plot indicate that they are approaching higher sensitivity and specificity levels, and any point along the diagonal line \(i.e. Line of No Discrimination\) means that the diagnostic result is no better than random guess estimations. In general, sensitivity and specificity levels were higher with Region #1 and #2 compared to Region #3 and #4. For instance, at a specificity level of 0.8, the sensitivities of Region #1 and #2 were ~0.77 and 0.84 respectively, whereas for Region #3 and #4 sensitivities were ~0.69 and 0.67 respectively. Similarly, at a sensitivity level of 0.8, the specificities of Region #1 and #2 were ~0.67 and 0.81 respectively, whereas for Region #3 and #4 specificities were ~0.1 and 0.24 respectively. Such discrepancies suggest that RPA is favored in some regions due to more efficient primer-template binding, and the level of non-specific binding by the detection probes varies from one sequence to another. Additionally, we analyzed whether the assay performance varies among two different patient groups: 1) patients who are currently on-treatment or have previous treatment record, and 2) treatment-naïve patients (Fig. S5). Interestingly, ROC curves are much closer to the top left corner of the plot for treatment-naïve group, suggesting that our diagnostic technique provides a better prediction of infectivity when the patients have not received treatment before.

**Multiplexed Detection of Four Amplification Regions**

Next, we performed multiplexed detection of four amplification regions. Table S2 summarizes the six barcodes and corresponding DNA sequences used for this assay. For each sample, amplicons produced from the four regions (T1-T4) were pre-mixed, denatured and incubated with the six barcodes (B1-CP1, B2-CP2, B3-CP3, B4-CP4, B5-CP P and B6-CPN), positive control DNA (TP), detection probes (DP1-DP4 and DPC), and hybridization buffer at 37°C for 30 minutes, washed three times with a washing buffer, and stored in PBST buffer for flow cytometry. The barcode signals were then deconvoluted as shown in Fig. S6. First, the entire barcode population was gated from debris, aggregates or artifacts using forward vs. side scatter (FSC vs. SSC) plots followed by the gating of specific barcode populations in FL1 vs. FL2 plot. The histogram of FL4 signal corresponding to Alexa Fluor 647 signal from the detection probe is plotted to calculate the median intensity level. The presence of corresponding amplicons enables the formation of a sandwich structure between capture probe functionalized microbeads and
detection probes, thereby producing strong FL4 signal, whereas in the absence of amplicons, there should be minimal FL4 signal detected. For all samples tested, positive control barcodes produced signals that are at least 3-fold higher than the corresponding signals produced from negative control barcodes confirming that our assay worked correctly in a 6-plexed format (Figure 5). As expected from the single-plexed assay, each amplification region produced intensity profiles unique to itself across 72 samples that are screened (Figure 5). Similar to the single-plexed detection, we used five water controls for each region to calculate cutoff intensity levels (i.e. 3 standard deviations above the average water signals).

As illustrated in the heat map diagram (Figure 5A), we confirmed that multiple barcodes that detect different HBV genomic regions could be performed in a single reaction vial with minimal cross-reactivity between barcodes. The genomic regions were identified by the emission of the original barcodes, and the presence of the genetic targets was identified by the emission from the secondary probes. As hypothesized, we detected unique signals across the 72 samples. There were regions that produced higher intensity signals compared to the other regions of the same HBV genome even at high viral loads, and the presence of several barcodes enhanced the rate of detection (e.g. samples 19, 47 and 58, Table S4). This finding signifies the importance of using multiple barcodes to reduce false-negative detections, and to increase the likelihood of producing true-positive diagnostic results. There are in total 4 false-positive across the four regions and out of the 21 negative samples tested. Interestingly, all false-positive were on suppressive antiviral therapy with undetectable viral loads on PCR. Also, 3 of the false-positive are from the same patient (#67), suggesting further confirmatory testing may be useful for this patient. On the other hand, all of the healthy volunteers (n=5), and resolved infection group (n=5) were correctly identified as true-negatives on our assay. Regarding the false-negatives, much of the false-negative signals arose from the on-treatment patient group or patients with previous treatment experiences (55/92 = 60%, 23 samples in this group x 4 regions = 92 in total), whereas treatment-naïve group produced significantly lower false-negative rate (17/112 = 15%, 28 samples in this group x 4 regions = 112 in total). We speculate that this is primarily due to suppressed viral loads for on-treatment patients, and development of drug-resistant mutations after the treatment.

Combinatorial Analysis

As a next step, we evaluated whether the combinatorial signal arising from the detection of the different genomic regions can lead to an increase in clinical sensitivity. We performed combinatorial analysis of clinical sensitivity by calculating sensitivity for various combinations of barcodes. Positive detection was determined if one or more barcodes produced a signal above the cutoff intensity as defined in Figure 5. As shown in Figure 6A, there was a systematic increase in clinical sensitivity up to combinations of two barcodes and saturation in sensitivity was observed with three and four barcodes. We also observed an increase in the sensitivity as we applied various sample inclusion criteria (i.e. samples that have viral loads greater than 200, 2,000, and 20,000 IU/mL). We chose an analytical threshold of 2,000 IU/mL because this corresponds to the level of viremia for which antiviral treatment is recommended by international treatment guidelines.23 Individuals with viral loads greater than or equal to 2,000 IU/mL, and with elevated alanine transaminases (ALTs) require further assessment and potentially initiation of antiviral treatment (Table S3).23 For all criteria tested, a combination of B4 and B2 (Region #4 and 2 in Figure 5) was enough to saturate the sensitivity curves and therefore used to calculate clinical sensitivity and specificity for samples containing viral loads greater than the clinically relevant threshold levels.
of 2,000 and 20,000 IU/mL. The signal-to-cutoff values shown in Figure 6B demonstrate that the clinical sensitivities of 90.5% (95% CI: 76.5-96.9%) and 100% (95% CI: 88.3-100%) are achieved when 2,000 and 20,000 IU/mL thresholds are applied respectively, and clinical specificity of 95.2% (95% CI: 74.1-99.8%) is achieved. More detailed analysis of samples can be seen on Fig. S7. The reduction in specificity compared to single-plexed detection is speculated to be caused by a higher level of non-specific binding of detection probes to the barcode surface as there are 3 times more detection probes added for multiplexed assay compared to a single-plexed assay (15 pmol/μL versus 5 pmol/μL). In our case, the presence of more than one positive barcode improved specificity, particularly for samples with high levels of viremia. Only one false-positive showed positive signals arising from multiple barcodes (sample 69) and all samples with only a single positive barcode were either false positives (n=1, sample 66) or had low levels of HBV DNA (n=5, samples 2, 3, 8, 12, and 50), suggesting that further confirmatory testing may be useful for samples yielding a single positive barcode. Furthermore, the level of non-specific binding can be reduced with various blocking strategies (e.g. BSA or digested BSA surface modification, Fig. S8). An alternative method to mitigate this reduction in specificity is to preset cutoff threshold intensity levels by developing ROC curves as shown in Fig. S9. Here we developed ROC curves for each region based on Figure 5, and adjusted cutoff intensity level to achieve 100% specificity for all four regions. Applying this new threshold gave 100% specificity (95% CI: 80.8-100%) while maintaining the sensitivity level (90.5%, CI: 76.5-96.9%). Thus, these results suggest we can achieve near perfect clinical diagnosis of patients infected with HBV by using multiple QD barcodes in the detection process.

Conclusions

The work presented herein describes a full clinical validation of QD barcode technology for the diagnosis of a large number of clinical samples that represent the spectrum of disease of HBV with differing viral loads, relevant viral genotypes and both treatment-experienced and treatment-naïve populations. Our combinatorial analysis demonstrated that the QD barcode diagnostic sensitivity is 54.9% (95% CI: 40.5-68.6%) with a single barcode but can be improved to 80.4% (95% CI: 66.5-89.7%) with two barcodes combination when all samples are included, and from 66.7% (95% CI: 50.4-80.0%) with a single barcode to 90.5% (95% CI: 76.5-96.9%) with two barcodes combination if the clinically relevant threshold level (2000 IU/mL) is applied (Figure 6). We also demonstrated that ROC curves can be developed to preset cutoff intensity levels to achieve desired specificity (Fig. S9). With respect to those samples that are not detected by all four barcodes, we do see a relationship between viral load and number of barcodes, where higher viral loads are more likely to be detected by two or more barcodes (Table S4). For instance, 75% (3/4) of the discordant samples with viral loads less than $10^4$ IU/mL were detected by only one barcode, whereas 90% (9/10) of the discordant samples with viral loads greater than $10^4$ IU/mL were detected by two or more barcodes. Thus, while our data suggest that sensitivity of a single barcode increases as the viral load increases, we have identified that for a subset of samples, multiple barcodes are required even in the presence of high viral load.

In theory, any false-negative samples identified in our assay that were above the clinically relevant threshold of 2,000 IU/mL could fall within the error of the diagnostic assay, and therefore if repeated, the clinical sensitivity of our assay may be closer to 100%. This is evidenced by the discrepancy in clinical sensitivities reported by independent groups for the gold-standard automated HBV detection system, which are affected by HBV e antigen (eAg) status, HBV...
genotype, and by intra-assay variability of serum samples (reported coefficients of variation ranged from 1.22 to 8.22% for TaqMan assay). Based on the clinical data we had available, we did not observe a difference in detection of different HBV genotypes, or samples with differing eAg status, even when these samples had viral loads near our defined analytical cut-off. Furthermore, we were able to detect samples with polymerase drug resistance mutations selected by available treatment, (Table S3) which is an important finding, as one of our detection regions included these mutations (R4, Fig. S2). This confirms that even single nucleotide substitutions in our primer regions will not affect the diagnostic accuracy of the test, an important consideration for constantly evolving viral pathogens.

Additionally, depending on the desired clinical need, if our test were only to be utilized to determine patients with obviously active chronic hepatitis (i.e. high ALT levels and serum DNA levels greater than 20,000 IU/mL), the sensitivity would increase to 100%. These patients require therapy immediately. This would be particularly applicable in settings that lack the equipment and training to perform liver biopsies and/or other investigations used to make treatment decisions.

Although we used HBV, a DNA virus, as the model for our study, this technique can also be applied to improve the diagnosis of RNA viruses like Hepatitis C virus (HCV) or HIV by using a reverse-transcription RPA to amplify the signal. In fact, for viruses that either have similar presentations, or often no clinical presentation at all (HBV, HCV, HIV), several barcodes for each virus could be combined allowing for multiplexing of a single virus with improved sensitivity, as well as multiplexing for several viruses simultaneously. Because of the size and shape-dependent optical properties of QDs, using different combinations of QDs and varying the intensity level can theoretically design over 1 million barcodes. Thus, the use of our combinatorial strategy could enable the diagnosis of a multitude of infectious pathogens with high clinical sensitivity.

Finally, the presented procedure utilizes techniques that are ideal for application in point-of-care (POC) settings. RPA was used to replace conventional PCR for nucleic acid amplification because RPA can be performed in 30 minutes, is simple to use, and operates at a low constant temperature (37°C); hence, eliminating the complexity of using an expensive thermocycler for application in low resource or remote settings. RPA is also advantageous over other well-known isothermal amplification techniques like rolling circle amplification (RCA), loop-mediated isothermal amplification (LAMP), or helicase-dependent amplification (HDA) because it is cheaper, extremely quick, power saving, and does not require complex primer design. Furthermore, it has been recently demonstrated that human body heat may be employed for incubation of RPA reactions suggesting an extremely low-cost solution for operation in resource-limited settings. Also, various microfluidic approaches have been shown to automate the RPA procedure and simplify sample-handling steps. Likewise, the QD barcode assay can be easily automated, and the signals can be detected using a smartphone camera as demonstrated in our previous work. In the near future, we envision a final black box device with various chambers to encompass all necessary components and automate our diagnostic procedure. Each compartment will contain lyophilized reagents to be dissolved by corresponding buffers and samples will be transferred from one compartment to another using capillaries or electrically driven flow. The barcodes will then be deposited on a substrate, imaged using the smartphone camera accompanied with various optics and excitation source, and signals will be deconvoluted via our custom software algorithm. Such smartphone-based diagnostic system can facilitate wireless transmission of diagnostic data enabling surveillance, mapping and prediction of diseases.
One of the limitations of the current design is that RPA for multiple amplification regions was performed in a parallel setup requiring more operational time than the detection of a single amplification region. We envision that this challenge can be addressed in the near future by implementing multiplexed RPA, where multiple regions are simultaneously amplified, as demonstrated by Nickisch-Rosenegk and colleagues.\textsuperscript{34} Also, quantification of initial viral loads will improve monitoring of disease progression and treatment response, which should be further examined by following quantitative RPA (qRPA) approach suggested by Richards-Kortum and colleagues.\textsuperscript{35} The influence of sequence mismatches on RPA specificity and barcode detection can be further analyzed by high-throughput sequencing, and can be used for not only detecting pathogens, but also for identification of various genotypes and subspecies. Although we acknowledge that there was a small reduction in clinical specificity with the multiplexed detection scheme, we speculate that this is primarily due to increase in the level of non-specific binding, which can be mitigated by various surface passivation strategies, or by simply presetting cutoff intensity values that would result in desired specificity from previously developed ROC curves. In a complete POC system, extraction of genetic targets of interest need to be employed in a more streamlined format either by integrating microcapillary with FTA membrane or microfluidic approach.\textsuperscript{36-38} Finally, the integration of extraction, amplification and detection steps into a single device unit will accelerate the transition to the POC application. Our longer-term systematic strategy was to advance powerful QD properties into a multiplex diagnostic platform that can be used to diagnose real infected patients. The adaptation of this strategy will increase the clinical translation of nanotechnology-based devices.

**Methods**

*QD Synthesis*

CdSe alloyed ZnS capped QDs with peak emission wavelengths of 506nm (“QD506”), 547nm (“QD547”), 560nm (“QD560”), 580nm (“QD580”), 596nm (“QD596”), and 615nm (“QD615”) were synthesized and characterized according to published procedures,\textsuperscript{39,40} and stored in chloroform at room temperature in dark until later use.

*QD Microbeads Synthesis*

QD microbeads were synthesized via concentration-controlled flow-focusing technique as outlined in our previous work.\textsuperscript{12} Polymer solution (4 wt%) was first prepared by dissolving 400 mg of poly(styrene-co-maleic anhydride) (32%, cumene-terminated, Sigma-Aldrich) in 10 mL chloroform, and filtered using a 0.2 µm PTFE syringe filter (Nalgene). Various QD color and concentration ratios plus fixed concentration of FeO magnetic nanoparticles (150 µL of 36.5 ng/mL) were mixed with the polymer solution to make a final volume of 1 mL QD polymer solution as outlined in Table S1. The QD polymer solution and double-distilled (DD) water were then introduced into a customized nozzle system (Ingeniatrics) submerged inside a beaker partially filled with DD water at a focused and focusing flow rates of 0.9 mL/hr and 180 mL/hr respectively using a syringe pump (Harvard Apparatus). The synthesized microbeads were stabilized by overnight stirring, filtered using a 35 µm BD falcon nylon mesh strainer cap, concentrated into a single tube, and stored at 4°C in dark until later use. The size distribution (Table S1) and concentration of microbeads were characterized using Beckman Coulter Vi-Cell counter.
Conjugation of Capture Probe to Microbead Surface

Conjugation of capture DNA strands to microbead surface was done through reaction with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) and N-hydroxysulfosuccinimid (sulfo-NHS) chemistry. DNA capture strands from IDT DNA Technologies (CPP and CPN, Table S2) and Bio Basic Inc. (CP1-CP4, Table S2), purchased HPLC-purified, were designed with an amine group and C12 (CPP and CPN, Table S2) and C6 (CP1-CP4, Table S2) spacer on the 5’ end, prepared at a concentration of 10 pmol/µL in TE buffer, and stored at 4°C until later use. For singleplexed assay, capture probes CP1, CP2, CP3 or CP4 were conjugated with B1, and for multiplexed assay, capture probes CP1, CP2, CP3, CP4, CPP and CPN were conjugated with B1, B2, B3, B4, B5 and B6 respectively. To conjugate, EDC and sulfo-NHS were first dissolved in MES buffer (pH 5, 100 mM) in concentrations of 0.0192g/64µL and 0.01g/100µL respectively. Then, approximately 10^6 microbeads were mixed with 55 µL of MES buffer (pH 5, 100 mM), 10 µL of the sulfo-NHS solution, 32 µL of the EDC solution, and 2.88 µL of the 10 pmol/µL capture DNA stock solution. The reaction was allowed to take place overnight, at which point additional 0.01 g of EDC was added and the mixture was allowed to incubate for another four hours.

To calculate surface density of capture probe DNA for each barcode, 1 µL of 5% Tween was added to the microbead solution, centrifuged at 3500g for 5 minutes to form bead pellet at the bottom of microcentrifuge tube, and 50 µL of the supernatant containing unbound DNA was extracted. The same conjugation procedure described above was performed for no conjugate control cases for each barcode, where DD water was added in place of the microbeads. SYBR gold (Invitrogen), dissolved in DMSO, was first diluted to 0.5:10000 ratio by adding 0.5 µL of SYBR gold to 10 mL of TE buffer, and standard curves (Fig. S3C) were developed by activating 2.5, 5.0, 7.5 and 10.0 µL of no conjugation control samples with 197.5, 195.0, 192.5, and 190 µL of SYBR Gold (0.5X) respectively in a black 96-well plate for 5 minutes before being read by BMG Labtech plate reader. In a black 96-well plate, 10 µL of supernatants from all conjugation cases, as well as 10 µL of the four blank cases containing only DD water, were each added to individual wells, 190 µL of the diluted SYBR gold was added to each well in the plate, and incubated at room temperature for 5 minutes before being read by plate reader. The fluorescence signal from the unbound DNA was used to estimate conjugation for each barcode by comparing the fluorescence of the conjugation cases with their respective controls containing no microbeads. That is, the lower fluorescence signal indicates higher amount of conjugation. Results were first converted to efficiency percentage and normalized to surface area to determine surface density (Fig. S3A and S2B). To complete conjugation, conjugated microbeads were washed twice by removing rest of the supernatant, resuspending microbead pellet in 100 µL of 0.05% Tween, and centrifuging at 3500g for 5 minutes. Washed microbeads were resuspended in 100 µL of 0.05% Tween (~10, 000 beads/µL) and stored at 4°C in dark until later use.

Sample Selection, Collection and Viral DNA Extraction

The de-identified clinical samples were obtained from the Toronto Western Hospital Liver Clinic (protocol approved by the Research Ethics Board of the University Health Network, affiliate of the University of Toronto). All patients provided written informed consent for storage and use of their specimens for research. Patient serum was collected by venipuncture in a Vacutainer, and stood upright for 30-60 minutes. Samples were spun in a refrigerated centrifuge, and serum was
aliquoted and stored at -80 °C. Viral HBV DNA was extracted using the Chemagic Viral DNA/RNA Kit (PerkinElmer).

In order to demonstrate the feasibility of using the proposed assay in a variety of settings globally, we chose samples that represented a diverse population. Males and females were represented, with the majority of individuals being between 31-60 years of age at the time of sample collection (Figure 3A). As HBV is a blood-borne infection, we wanted to categorize the various risk factors that may have contributed to HBV infection in a clinic at a major health center in Toronto, Canada (Figure 3B). Samples were selected to reflect diverse HBV serology, virology, and treatment. Individuals were categorized as uninfected healthy controls, CoreAb-Positive/sAg-Negative (resolved infection), and various phases of chronic HBV infection including the immune tolerant phase, the immune active phase, the immune control phase, and chronic HBV on therapy. At each stage, different combinations of viral proteins may be present or absent, which also influences the level of HBV DNA. Although our assay detects viral DNA as opposed to viral proteins or ALT levels, it was essential to validate and determine whether our analytical LOD would correspond to disease stages that require immediate attention. Viral loads from $10^3$ to greater than $10^9$ IU/mL were included, as well as genotypes A-E, both treatment-naïve and experienced, and those with drug-resistance mutations, notably present in one of the regions used for our detection (Table S3).

**RPA and Purification**

RPA was performed using extracted DNA, primer pairs (Bio Basic Inc.), and TwistAmp® Basic kit (TwistDx). Primer pairs (FP1-FP4, and RP1-RP4, Table S2) were purchased HPLC-purified, and first prepared at a concentration of 100 pmol/µL in TE buffer. The stock primer solution was then diluted to 10 pmol/µL aliquots and stored at 4°C for later use. For each amplification reaction, a premix solution containing 2.4 µL of each forward and reverse primers (10 pmol/µL), 9.2 µL of nuclease-free water, 29.5 µL of rehydration buffer, 2.5 µL of magnesium acetate (280 mM), and 4 µL of the extracted DNA was prepared in a total volume of 50 µL. For no template controls, 4 µL of nuclease-free water was added instead of extracted DNA. This solution was then transferred to a tube containing the lyophilized enzyme pellet, mixed, and incubated at 37°C for 30 minutes. For all 72 samples and no template controls, RPA was performed four times using the four primer pairs (FP1-FP4, RP1-RP4, Table S2) corresponding to four amplification regions (T1-T4, Table S2). RPA products were then purified using EZ-10 spin column DNA gel extraction kit (Bio Basic Inc.), eluted into 50 µL, visualized by agarose gel electrophoresis, and stored at 4°C until later use.

**Single-plexed Assay**

Detection probes (DP1-DP4 and DPC, Table S2) from IDT DNA technology were purchased HPLC-purified, modified with Alexa Fluor 647 on 3’ end (DP1-DP4, Table S2) or 5’ end (DPC, Table S2), prepared at a concentration of 100 pmol/µL, and stored at 4°C in dark until later use. For each assay reaction, a premix solution containing 1 µL of conjugated microbeads (~10, 000 beads, B1-CP1, B1-CP2, B1-CP3, or B1-CP4), 1 µL of detection probes (100 pmol/µL, DP1, DP2, DP3, or DP4), 7 µL of DD water, and 10 µL of hybridization buffer (10x SSC, 0.1% SDS, heated to 60°C) making up a total volume of 19 µL was prepared. Purified RPA products were then denatured by heating at 100°C for 10 minutes and stored in ice immediately. 1 µL of
denatured RPA product was mixed with 19 µL of premix solution and incubated at 37°C for 30 minutes. The assay solution was washed three times by mixing with 200 µL of washing buffer (0.5x SSC, 0.1% SDS), placing in a magnetic rack (MagnaRack, Life Technologies) for 10 minutes to allow magnetically encoded microbeads to settle to the wall of microcentrifuge tube, and removing supernatant. Washed product was resuspended in 200 µL of PBST buffer for flow cytometry (BD FACSCalibur). The assay was performed individually for all four regions across 72 samples plus no template controls using the corresponding barcode, detection probe and amplicon matches.

**Multiplexed Assay**

For all 72 samples plus 5 no template controls tested, 5 µL of RPA products from four regions were mixed into a single tube to make a final volume of 20 µL. For each assay reaction, a premix solution consisting 0.5 µL of barcode B1-CP1, 0.5 µL of barcode B2-CP2, 0.5 µL of barcode B3-CP3, 0.5 µL of barcode B4-CP4, 0.5 µL of barcode B5-CPP, 0.5 µL of barcode B6-CPN (~5,000 beads per barcode), 0.5 µL of detection probe DP1 (50 pmol), 0.5 µL of detection probe DP2 (50 pmol), 0.5 µL of detection probe DP3 (50 pmol), 0.5 µL of detection probe DP4 (50 pmol), 1 µL of detection probe DPC (100 pmol), 0.5 µL of positive control DNA (TP, 0.5 pmol), 1.5 µL of water, and 10 µL of hybridization buffer (10x SSC, 0.1% SDS, heated to 60°C) making up a total volume of 18 µL was prepared. Mixed RPA products were then denatured by heating at 100°C for 10 minutes and stored in ice immediately. 2 µL of denatured RPA product was mixed with 18 µL of premix solution and incubated at 37°C for 30 minutes. The assay solution was washed three times by mixing with 200 µL of washing buffer (0.5x SSC, 0.1% SDS), placing in a magnetic rack (MagnaRack, Life Technologies) for 10 minutes to allow magnetically encoded microbeads to settle to the wall of microcentrifuge tube, and removing supernatant. Washed product was resuspended in 200 µL of PBST buffer for flow cytometry (BD FACSCalibur).

**Data Analysis**

Data was first analyzed from FlowJo software by gating the entire microbead population in FSC vs. SSC plot. From this subpopulation, the specific barcode population was gated, which was finally used to plot histogram of FL4 signals and calculate median intensity (Fig. S6). For both single-plexed and multiplexed assays, FL4 signals corresponding to Alexa Fluor 647 from detection probes were subtracted with signals from blank samples containing only conjugated barcodes in PBST buffer. FL4 median intensity values for each amplification region were then normalized to its maximum value to plot bar graphs in Figure 5B and Fig. S4. For positive and negative controls in Figure 5B, intensity values were normalized to the maximum intensity of positive control samples. The cutoff intensity level was calculated for each region using the following equation:

\[ I_{cutoff} = \text{Avg}(I_{\text{no template controls}}) + 3\times \text{Stdev}(I_{\text{no template controls}}) \]

The heatmap diagram in Figure 5A was created using a custom-written Matlab script that accepts normalized intensity values as an input and outputs the heatmap diagram using `imagesc` function. ROC data in Figure 4B and Fig. S9A were computed using a custom-written Matlab script that accepts normalized intensity values as an input and outputs true and false positive rates in response to varying cutoff intensity level from 0 to 1 by an increment of 0.001. Adjusted cutoff
value in Fig. S9B was determined from this output by choosing the minimum intensity value that results in zero false positive rates. For the combinatorial analysis in Figure 6A, signal-to-cutoff values were first calculated for each region using the normalized intensity and $I_{cutoff}$ values calculated above, and assigned a value of 1 or 0 according the following equation:

If $(\text{signal}/\text{cutoff}) \geq 1$, then $Q_n = 1$, where $n = \text{barcode #}$
If $(\text{signal}/\text{cutoff}) < 1$, then $Q_n = 0$, where $n = \text{barcode #}$

From the above calculation, all combinations of barcodes were analyzed to determine positive or negative detection using the following equation:

If $[(Q_1 + Q_2), (Q_4 + Q_2), (Q_1 + Q_2 + Q_3), \text{etc.}] > 0$, then detected as positive
If $[(Q_1 + Q_2), (Q_4 + Q_2), (Q_1 + Q_2 + Q_3), \text{etc.}] = 0$, then detected as negative

Finally, the sensitivity and specificity are calculated using the following equation (95% confidence interval was determined using the online calculator, http://vassarstats.net/clin1.html):

Sensitivity = $\frac{\text{# of true positives}}{\text{(# of true positives + # of false negatives)}}$
Specificity = $\frac{\text{# of true negatives}}{\text{(# of true negatives + # of false positives)}}$

Supporting Information: HBV Genome, Capture Probe Surface Density, FL1 vs. FL2 Signals for Six QD Barcodes, Singleplexed Detection of Four Amplicons, Deconvolution of Barcode Signals, Adjusted Cutoff Intensity Level from ROC Curve, Samples Categorized by HBV Disease Markers and Viral Characteristics, List of Synthesized Microbeads, List of DNA Sequences and Corresponding Barcodes used for Multiplexed Assay, Summary of Discordant Sample Detection by Viral Load and Multiplexed Barcode Detection. This material is available free of charge via the Internet at http://pubs.acs.org.

Author Contributions: J.K., and W.C.W.C. conceived the concept, designed the study and wrote the paper. J.K. performed microbead synthesis, capture probe conjugation, RPA, single-plexed and multiplexed assays, flow cytometry, and data analysis. M.J.B. completed HBV sequence analysis, designed target regions, selected diverse patient samples, collected patient data, completed viral extraction, performed PCR, and assisted with manuscript preparation. J.J.F. provided clinical samples, and was involved in experimental design. All authors reviewed and edited the manuscript prior to submission.

Funding Sources: We would like to thank the Natural Sciences and Engineering Research council (RGPIN288231) and the Canadian Institute of Health Research (MOP130143, COP126588) and Canadian Research Chair (950-223824) for grant funds. J.K. acknowledges Natural Sciences and Engineering Research Council Postgraduate Scholarship for support.

Acknowledgment: We acknowledge B. Udagama and A. Lam for synthesizing barcodes B1 and B4 respectively.
Conflict of Interest: The authors declare no competing financial interest.

References


Figure 1. Multiplexed QD Barcode Assay. Multiple QD barcodes are incubated with target DNA and detection probes. Target DNA joins the barcode to corresponding detection probe, and unbound detection probes are washed away. Each barcode signal is first gated via flow cytometry, and detection probe signal intensity is calculated from the gated barcode population.
Figure 2. Integration of Multiplexed QD Barcode Assay with RPA. Multiple regions within HBV genome are amplified by RPA producing positive or negative amplification products. Amplicon-specific barcodes functionalized with capture DNA molecules are used to detect denatured amplicons in multiplexed QD barcode assay.
Figure 3. Participant Demographics. Participant demographics categorized by (A) Age and sex, and (B) Birth country and risk factors of HBV infection.
Figure 4. Singleplexed Detection of Four Amplification Regions. (A) Signal-to-cutoff values of four amplification regions relative to its reference test. (B) ROC curves developed for each region by varying cutoff intensity value from 0 to 1 by an increment of 0.001. Line of no discrimination predicts sensitivity and specificity by random guess estimations.
Figure 5. Multiplexed Detection of Four Amplification Regions. Normalized Alexa Fluor 647 (AF647) intensity across 72 samples screened with four barcodes corresponding to four amplification regions plus positive and negative control barcodes. (A) Heat map diagram (P: Positive control, N: Negative control). (B) Bar graphs. Dotted lines represent cutoff intensity levels calculated from five water controls (W1-W5, 3 standard deviations above average water intensity). Samples are ordered from left to right as the following: water controls, healthy subjects, CoreAb positive, undetectable by gold standard, and DNA positive samples. HBV positive samples (sample ID 1-20, 36-58, 60, 62-64, 68, 70-72) are ordered from the lowest to highest viral load in IU/mL.
Figure 6. Combinatorial Analysis. (A) Effect of various barcode/amplification region combinations on clinical sensitivity. Different sample inclusion criteria were applied (samples containing viral loads greater than 200, 2000, and 20,000 IU/mL). (B) Signal-to-cutoff values with combination of Region #4 and 2 (B4 and B2) relative to its reference test for samples containing different viral loads (All, >200 IU/mL, >2000 IU/mL, and >20,000 IU/mL).
Table of Contents Graphic

0 kb  Entire HBV Genome  3.2 kb

RPA  RPA  RPA  RPA  RPA

Negative Amplification  Negative Amplification

Multiplexed Detection

True Positive  False Negative  True Positive  False Negative  True Positive