BMC Genomics

BioMed Central

Methodology article

Open Access

High accuracy genotyping directly from genomic DNA using a rolling circle amplification based assay

Osama A Alsmadi†, Carole J Bornarth†, Wanmin Song, Michele Wisniewski, Jing Du, Joel P Brockman, A Fawad Faruqi, Seiyu Hosono, Zhenyu Sun, Yuefen Du, Xiaohong Wu, Michael Egholm, Patricio Abarzúa, Roger S Lasken and Mark D Driscoll*

Address: Molecular Staging, Inc. 300 George Street, Suite 701, New Haven, CT 06511, USA

Email: Osama A Alsmadi - osamaa@molecularstaging.com; Carole J Bornarth - carole@molecularstaging.com; Wanmin Song - wanmins@molecularstaging.com; Michele Wisniewski - michelew@molecularstaging.com; Jing Du - jingd@molecularstaging.com; Joel P Brockman - joelb@molecularstaging.com; A Fawad Faruqi - fawadf@molecularstaging.com; Seiyu Hosono - seiyuh@molecularstaging.com; Zhenyu Sun - zhenyus@molecularstaging.com; Yuefen Du - yuefend@molecularstaging.com; Xiaohong Wu - xiaohongw@molecularstaging.com; Michael Egholm - michaele@molecularstaging.com; Patricio Abarzúa - patoa@molecularstaging.com; Roger S Lasken - rogerl@molecularstaging.com; Mark D Driscoll* - markd@molecularstaging.com

* Corresponding author †Equal contributors

Published: 30 May 2003

BMC Genomics 2003, **4**:21

Received: 4 February 2003
Accepted: 30 May 2003

This article is available from: http://www.biomedcentral.com/1471-2164/4/21

Abstract

Background: Rolling circle amplification of ligated probes is a simple and sensitive means for genotyping directly from genomic DNA. SNPs and mutations are interrogated with open circle probes (OCP) that can be circularized by DNA ligase when the probe matches the genotype. An amplified detection signal is generated by exponential rolling circle amplification (ERCA) of the circularized probe. The low cost and scalability of ligation/ERCA genotyping makes it ideally suited for automated, high throughput methods.

Results: A retrospective study using human genomic DNA samples of known genotype was performed for four different clinically relevant mutations: Factor V Leiden, Factor II prothrombin, and two hemochromatosis mutations, C282Y and H63D. Greater than 99% accuracy was obtained genotyping genomic DNA samples from hundreds of different individuals. The combined process of ligation/ERCA was performed in a single tube and produced fluorescent signal directly from genomic DNA in less than an hour. In each assay, the probes for both normal and mutant alleles were combined in a single reaction. Multiple ERCA primers combined with a quenched-peptide nucleic acid (Q-PNA) fluorescent detection system greatly accellerated the appearance of signal. Probes designed with hairpin structures reduced misamplification. Genotyping accuracy was identical from either purified genomic DNA or genomic DNA generated using whole genome amplification (WGA). Fluorescent signal output was measured in real time and as an end point.

Conclusions: Combining the optimal elements for ligation/ERCA genotyping has resulted in a highly accurate single tube assay for genotyping directly from genomic DNA samples. Accuracy exceeded 99 % for four probe sets targeting clinically relevant mutations. No genotypes were called incorrectly using either genomic DNA or whole genome amplified sample.

Background

Sequencing of the human genome has led to the identification of mutations and single nucleotide polymorphisms (SNPs) that can be linked with specific phenotypes. A number of assays have been developed to genotype known SNPs and mutations. However, the majority of these methods have many steps and are difficult to automate. Currently, most genotyping methods require that the target region of the genomic DNA be amplified using PCR prior to the genotyping assay (reviewed in [1,2]). Following amplification, the PCR product is either sequenced directly, or probed for the mutation of interest. Similarly, OCP ligation/ERCA can be used to accurately genotype PCR products in a biallelic format, where the probes for both alleles are present in a single reaction [3]. The OCP ligation/ERCA method is based on ligation dependent circularization of allele specific DNA probes followed by exponential rolling circle amplification of the circularized probes [4,5]. ERCA probes have also been variously described as 'padlock probes' [6,7] or 'circularizable probes' [4]. Although OCP ligation/ERCA can be used to genotype PCR products, it also provides a means for directly genotyping genomic DNA without requiring an initial amplification of the genomic DNA locus. For example, we recently described a high throughput ligation/ERCA protocol for genotyping genomic DNA samples in microtiter plates [8].

This report describes modifications to OCP ligation/ERCA genotyping that reduce reaction time and improve accuracy to above 99%. Methods for improving accuracy include simultaneous detection of two alleles per well, probes designed to contain hairpin sequences that regulate ligation discrimination and prevent nonspecific amplification, and introduction of a Q-PNA fluorescent detection system[9]. Improvements in probe design, rolling circle reaction design, and assay methods were integrated into an improved high accuracy genotyping assay, and the utility of the method was demonstrated for four clinically important mutations. The Q-PNA fluorescent detection system produced a six-fold reduction in time to result.

In many instances only limited quantities of genomic DNA are available from clinical samples for genetic testing. DNA can also be limiting when large numbers of assays must be performed on a single sample. Recently, a whole genome amplification method (WGA) has been developed that is capable of accurately producing large quantities of genomic DNA from limited sample [10,11]. In cases where genomic DNA was limiting, WGA genomic DNA was substituted for genomic DNA. Complete and faithful amplification of the entire genome resulted in WGA product that could be substituted for genomic DNA without affecting the accuracy of genotyping. WGA DNA amplification and OCP ligation/ERCA can be combined to create an automated high throughput mutation assay that includes DNA amplification from limiting samples and allele detection in a single tube.

Results

Probe Design

The OCP ligation/ERCA assay was improved to meet the requirements for fast, and accurate, genotyping. Probes for both alleles were combined in a single reaction tube (Figure 1). One OCP annealed perfectly to the normal sequence of interest, and the other targeted the mutant. The 3' terminal nucleotide of the OCP annealed opposite the variant nucleotide. If the OCP and the target matched at the mutation site, the OCP was circularized by ligase. If the OCP was mismatched at the 3' end, it was not ligated. There are a number of advantages to having both probes present in the same reaction, as opposed to having each probe in a separate reaction [3]. For instance, if amplification fails in a biallelic (ie., two allele) reaction, signal is not generated for either allele, and the result can be discarded. However, if only one of the two reactions fail in a two reaction/ single allele per well format, a heterozygote could be miscalled as a homozygote. Furthermore, the two allele per well format reduced non-specific amplification. With both mutant and normal OCP in the reaction, either one or the other or both OCPs were circularized, depending on the genotype. Therefore, there was a specific rolling circle reaction in every assay, regardless of the genotype. The kinetics of the specific rolling circle reaction were sufficiently rapid so that non-specific amplification was greatly reduced or eliminated. Details of the mechanism of the Exponential Rolling Circle Amplification (ERCA) reaction are shown in Figure 2.

The exponential amplification of the ligated probe sequence generates double stranded DNA products that are unit lengths of the probe [4]. Electrophoretic separation on an agarose gel generates a ladder of DNA bands as seen in Figure 3. Typically, the products ranged in size from 80 bases to 1.2 kb, with most products less than 500 bases long. The ERCA products for normal and mutant alleles incorporate a different fluorescent P1 primer. In Figure 3, signal from the normal allele results in FAM (false color green) fluorescence, while signal from the mutant allele results in Cy3 (false color red) fluorescence. A sample with heterozygous genotype produces both colors.

The OCPs were redesigned to incorporate a hairpin structure (Figure 4) that improved target specificity analogous to beacon [12] probe design. Parameters affecting the stability of the hairpin were tailored to increase the specificity of target recognition. Target directed hybridization using a hairpin is more specific than hybridization of a

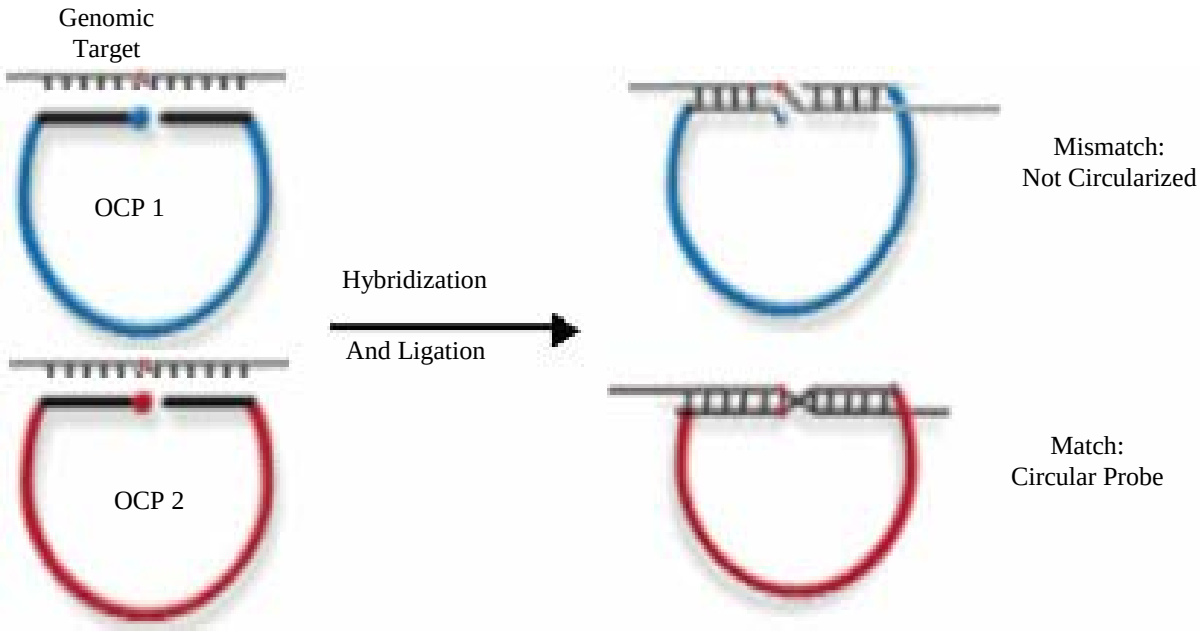

Figure 1
Ligation of OCP to genomic DNA target Open circle probe (OCP) anneals to genomic DNA target, and the 3' base either base pairs with the target (red) or creates a mismatch (blue). The matched OCP is circularized by ligase, and becomes topologically linked to the target.

linear probe because the structure of the hairpin probe is maintained in a folded state unless the hairpin sequence encounters its preferred target sequence. If the probe does not encounter the correct target, the hairpin will remain self-annealed, reducing the possibility of nonspecific or transient annealing events. Only the extensive base pairing between the probe and target sequence is sufficient to disrupt the hairpin.

The new probe design also provides a way to reduce background by rendering unused probe inert. (Figure 5). During ERCA, the 3'-end of the hairpin of unused probe is extended by DNA polymerase to produce a non-reactive double-stranded product (data not shown). In the absence of a method such as this for either removing or inactivating unused OCP, nonspecific background amplification can occur [13]. Sequencing of misamplification products (data not shown, [14]) has revealed that misamplification products produced in the absence of ligase usually consist of a short OCP-derived sequence immediately flanked by primer sequences. Each 'primer-OCP-primer' unit was repeated dozens or hundreds of times in tandem. The unit sequence varied randomly from reaction to reaction, and sometimes more than one sequence was present in a reaction. The exact mechanism for generating and expanding misamplification products with this overall sequence motif is not completely understood. However, improved probe design that results in enzymatic elimination of unused OCPs, as demonstrated with the 3' hairpin OCP, can reduce or eliminate unwanted nonspecific amplifications.

To determine the effect of the 3' hairpin on genotyping accuracy, probes with and without 3' hairpins were designed for both the Factor V Leiden and Hemochromatosis H63D mutations (Table 1). Control probes lacking hairpins were constructed in two different ways: 1) Hairpins in the FVL probes were removed by simply deleting the non-target directed OCP sequence involved in hairpin formation. This did not change priming sites or the base pairing between the probe and target, but non-hairpin probes were reduced in length by 10 nucleotides compared to the hairpin probes. 2) To eliminate the 3' hairpin without changing the length of the OCP, the hairpins in the H63D probes were removed by changing the sequence of the OCP such that the length of the probes was conserved but the base pairing in the hairpin was disrupted. Real time ERCA genotyping reactions were performed

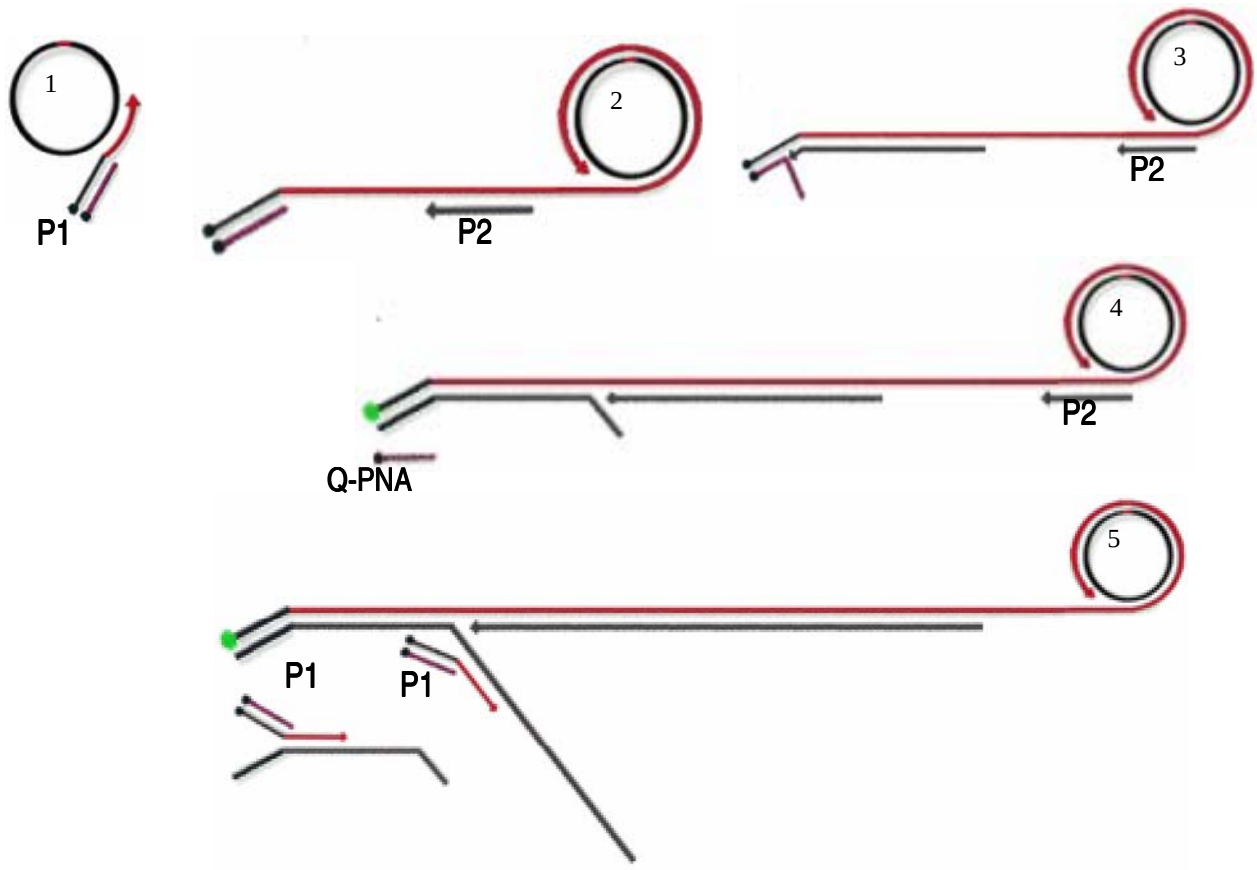

Figure 2
ERCA using Q-PNA, P1 and P2 Exponential amplification of circularized OCP was performed using primers P1 and P2. P1 is a composite of a Q-PNA annealed to a fluorescent DNA primer specific for the circle sequence directed against one allele. Q-PNA anneals to the fluorescent DNA primer such that the c-terminal quencher on the PNA quenches the 5' fluorescent moiety (either FAM or Cy3) on the DNA primer. P2 is a DNA primer with the same strand orientation as the circle, complementary to the product of the primary rolling circle reaction. Amplification of the circle specific for only one of the two alleles is illustrated. ERCA begins when P1 anneals to the circle (1), and polymerase copies the circle sequence, eventually copying the entire circle. The polymerase begins displacing the previously synthesized product, opening up single stranded P2 binding sites (2). Each turn of displacement synthesis around the circle exposes another P2 binding site, and the resulting synthesis from the annealed P2 primers results in the displacement of downstream primers and products, including the PNA (3). Displacement of the PNA separates the fluorescent label on the DNA from the quencher on the PNA, producing fluorescent signal specific for the P1/circle pair (4). Displacement of the downstream primers and products also opens up additional P1 binding sites, and the process continues in an exponential cascade.

using FVL and H63D probes with and without 3' hairpins (Figures 6, 7, 8 and 9). The fluorescence versus time was plotted for each of 48 samples, grouped according to genotype. The genotyping signals appeared in as little as 10 minutes using the Q-PNA detection system. Genotyping calls were made as described in Methods, (Determination of Genotypes, Real Time OCP ligation/ERCA). For both FVL and H63D, nonspecific amplification occurred when OCPs lacked the hairpin structure. Strong nonspecific signal appeared for samples known to be either homozygous mutant or normal, frequently generating a heterozygous genotyping miscall. 48 samples were genotyped for FVL and H63. FVL genotyping with the non-hairpin probes resulted in unacceptably high error, with only 65% of the samples called correctly (Figure 6). Results for non-hairpin H63D probes were similar, with only 65% called cor-

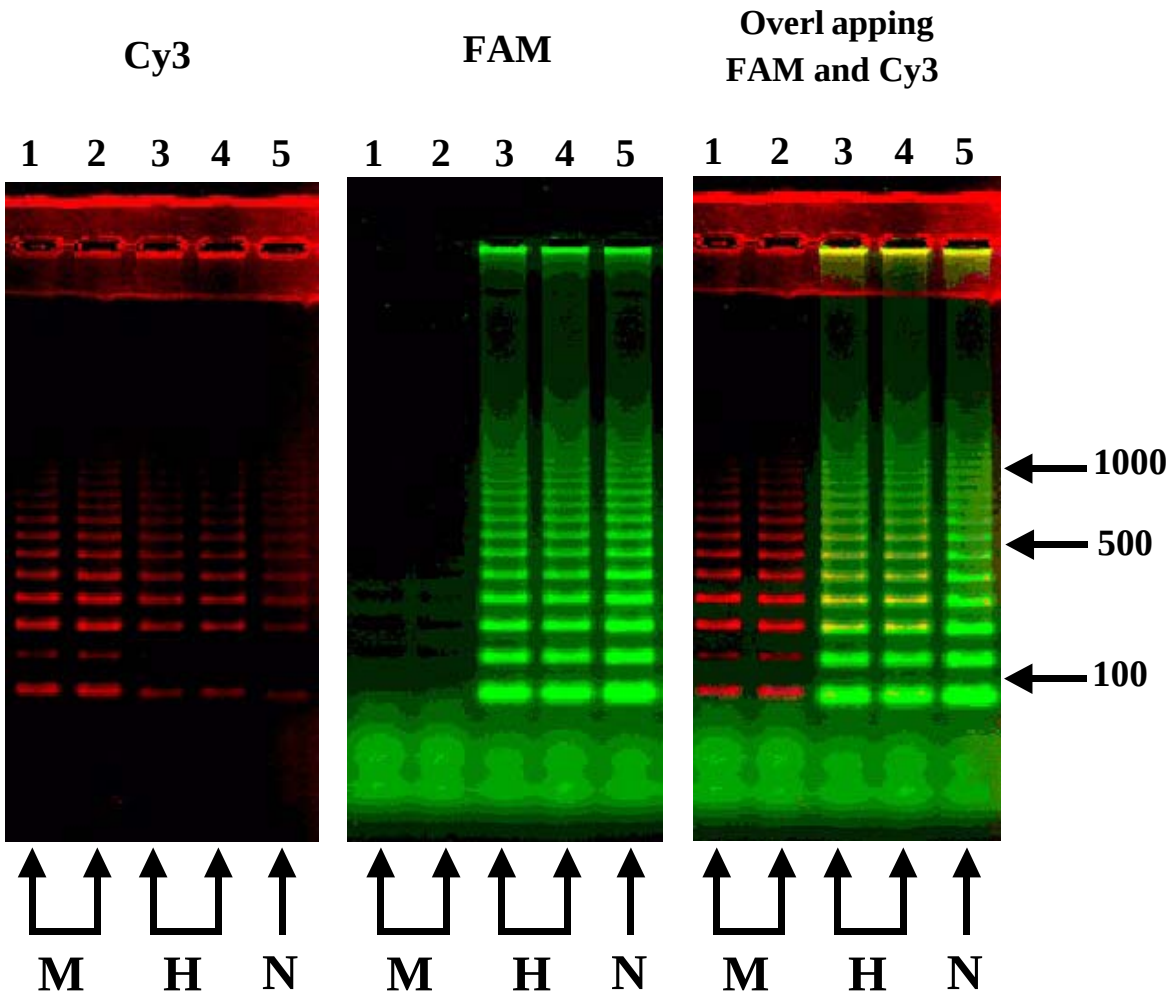

Figure 3
Agarose gel analysis of ERCA reaction products. Assay conditions were as described in Methods for specific Factor V Leiden Mutation ligation/ERCA genotyping assay. Following ERCA, reaction products were separated on a 1.5% agarose gel in 1X TBE at 200 Volts for 2 hours. FAM and Cy3 fluorescence were visualized using a Molecular Dynamics FluorImager. FAM signal was produced during ERCA in response to the normal allele, and Cy3 signal in response to the mutant. Sample in lanes 1 and 2 contained genomic DNA target homozygous for the Factor V Leiden mutation (M), lanes 3 and 4 were heterozygous for the mutation (H), and lane 5 was homozygous normal (N). Three different exposures of the same gel, lanes 1–5, are shown. On the left, the Cy3 fluorescent product is shown in red. In the center, the FAM product is shown in green. On the right, both exposures are shown overlapped. Overlapping FAM and Cy3 produces a yellow signal in this exposure. Migration of 100, 500, and 1000 base pair DNA size markers (not shown) is indicated by arrows at right.

rectly (Figure 8). In both cases, the non-hairpin genotyping reactions showed a dramatic increase in the incidence of late appearing, non-specific signal. When the hairpin containing OCPs were used, all 96 samples were genotyped correctly (Figures 7 and 9).

To increase the accuracy of genotyping, probes for both alleles were combined in a single tube format. For heterozygotes, the two probes are concurrently amplified in the ERCA reaction. Exponential amplification of circularized OCP requires primers P1 and P2 (see Figure 2). The concurrent amplification of mutant and normal probes is

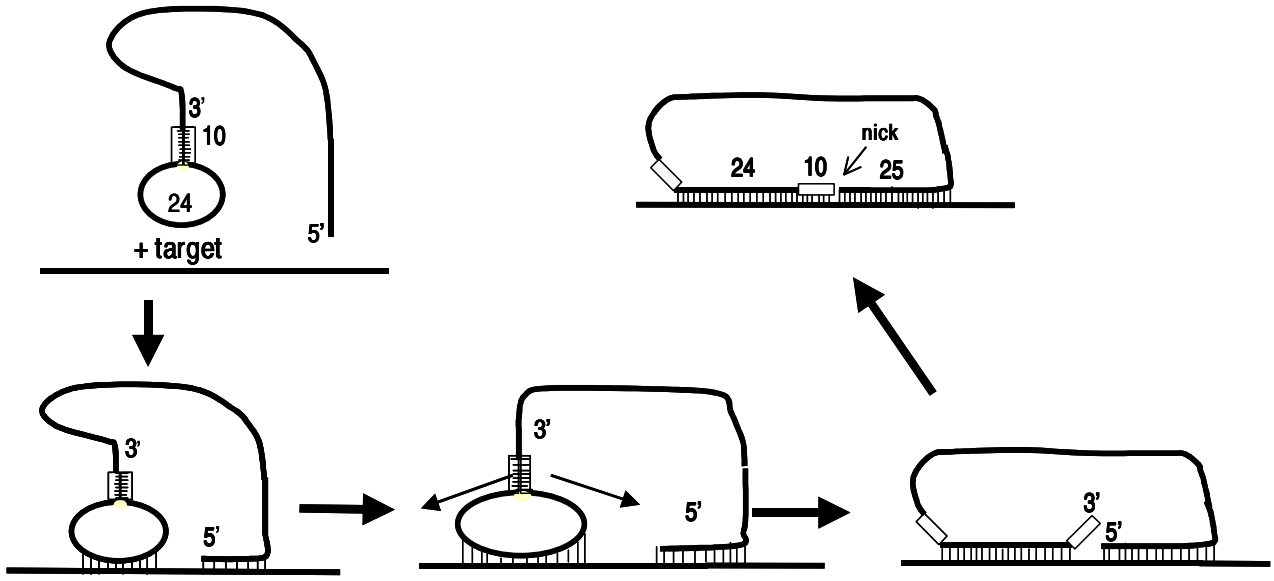

Figure 4
Specific annealing of 3' hairpin OCP to its genomic target followed by ligation. The hairpin structure can increase OCP discrimination during ligation by providing a two state equilibrium for the OCP, where OCP is either in the hairpin form or annealed to target. A mismatch between target and OCP favors the intra-molecular hairpin formation. A match between the OCP and the target provides an extended region of complementarity that is sufficient to maintain stable binding of the 3' end of the OCP to the target, allowing ligase to circularize the OCP. This balance contributes to accurate and sensitive allele discrimination.

kept in balance by using the same P2 for each probe. P1 primers were routinely titrated to determine optimum concentration [4]. This approach greatly reduces the difficulty of balancing the kinetics of biallelic ERCA amplifications. If the P2 primer is common to both OCPs, only P1 priming efficiencies need to be similar, greatly simplifying the process of obtaining balanced reaction kinetics in a biallelic assay.

The use of several P1 and P2 primers, which initiate DNA synthesis from multiple positions on the probe and product, increases ERCA rate, sensitivity and specificity ([15], data not shown). Therefore, secondary amplification primers were designed for each probe. In genotyping reactions where two probes were present in a single reaction, the secondary primers were designed to anneal to target specific sequences shared by both OCPs, in the 5' and/or 3' arms (Table 1).

Genotyping Accuracy

The improvements to the OCP ligation/ERCA genotyping reaction outlined above were incorporated into four genotyping assays, and a study was performed to test the reaction design for genotyping accuracy. Probes for the Factor V Leiden (FVL) mutant and normal alleles (see Table 1) were designed using the parameters outlined in Methods. These probes were used to screen a 216 patient sample set that had been previously genotyped using RFLP. The sample groups were enriched for the mutant genotypes that normally occur at very low frequency in the population. For example, the FVL sample group contained 105 normal, 98 heterozygous, and 13 homozygous mutant patient samples for the FVL allele. Of 352 genotyping reactions performed in triplicate (1056 total FVL reactions), only 6 samples gave inconsistent genotyping results and had to be repeated. A repeat was indicated if any of the three independent experiments did not agree. By performing reactions in triplicate, the assay yielded results that were 100% accurate, with a 1.7% repeat rate. Reactions performed only once had an accuracy of 99.4 %. Results were visualized as both real time and end-point, and a representative FVL genotyping experiment is shown in Figure 10. The genotype for each sample was determined

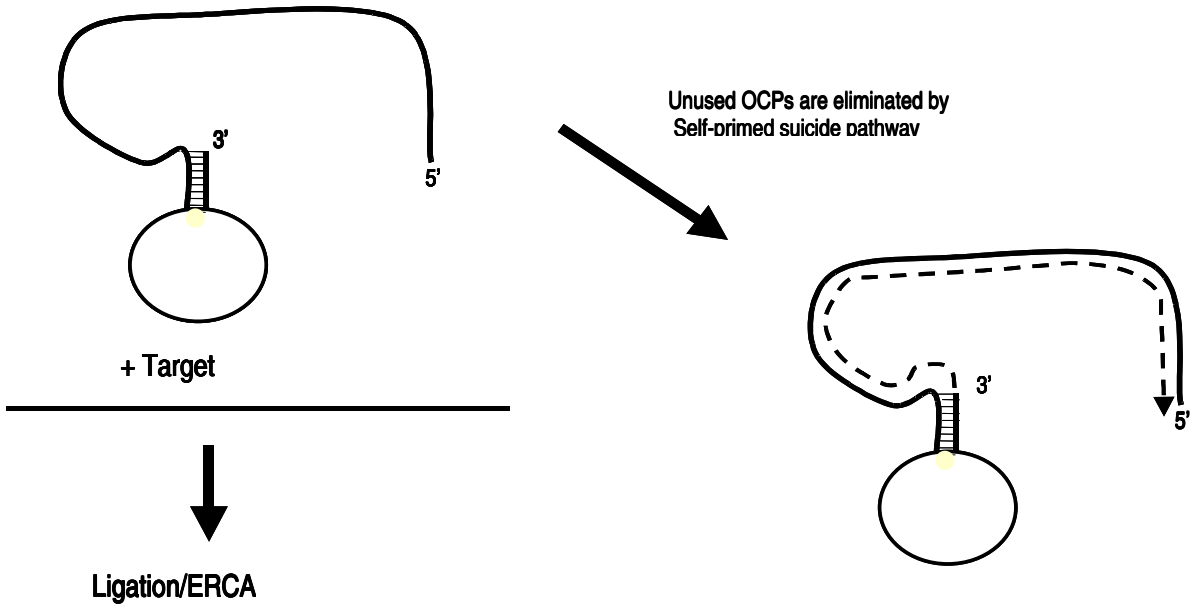

Figure 5
3' Hairpin OCPs design allows conversion of unused OCPs to inert double-stranded form. If the OCP is not circularized, the single stranded region of the unligated linear probe can bind P1 primer and be copied, a process that can lead to misamplification of unligated probe. To reduce the potential for misamplification, a 3' hairpin has been designed into the OCP. During ERCA reaction, 3' hairpin is extended by polymerase, resulting in an extended hairpin with minimal single strandedness, which is less likely to participate in faulty priming reactions that result in misamplification.

as outlined in Methods, and compared to the known genotype. The number of samples genotyped and the accuracy of genotyping are summarized in Table 2.

Probes for the Factor II prothrombin (FII) mutant and normal alleles (see Table 1) were designed as discussed in Methods. These probes were used to screen a 298 patient sample set that was previously genotyped using RFLP. There were two mutant genotypes, 12 heterozygotes, and 284 homozygous normal genotypes. Because mutant patient samples were limiting, a separate sample group that consisted of multiple repeats of the same Factor II mutant and heterozygous samples was screened. As with the genotyping assays described above, Factor II prothrombin (FII) mutant genotyping results were also visualized either in real time or as an end-point assay (Figure 11). Summary of results is shown in Table 2. Samples probed for the Factor II mutation were genotyped with 99.5% accuracy within a single assay, 100% accuracy if performed in triplicate.

Probes for the two Hemochromatosis mutations H63D and C282Y were designed using the parameters outlined in Methods and used to screen patient sample sets of known genotype (probe sequences in Table 1). There were 138 samples screened for H63D and 190 samples screened for C282Y. This sample group was also enriched for the mutant genotype, containing 89 homozygous normal, 48 heterozygous and 1 homozygous mutant (the single mutant was repeated 29 times) for the H63 allele and 84 normal, 93 heterozygous and 13 homozygous mutant patient samples for the C282 allele. Examples of typical genotyping results for H63D are shown in Figure 12 and results for C282Y are shown in Figure 13. Genotyping results were visualized as both real time reaction traces and as an end-point assay. Samples screened for both hemochromatosis mutations were genotyped with 99.5% accuracy within a single assay, 100% accuracy if performed in triplicate (Table 2).

Discussion

OCP ligation/ERCA can be used to directly probe genomic DNA for SNPs or mutations, and the entire process can be

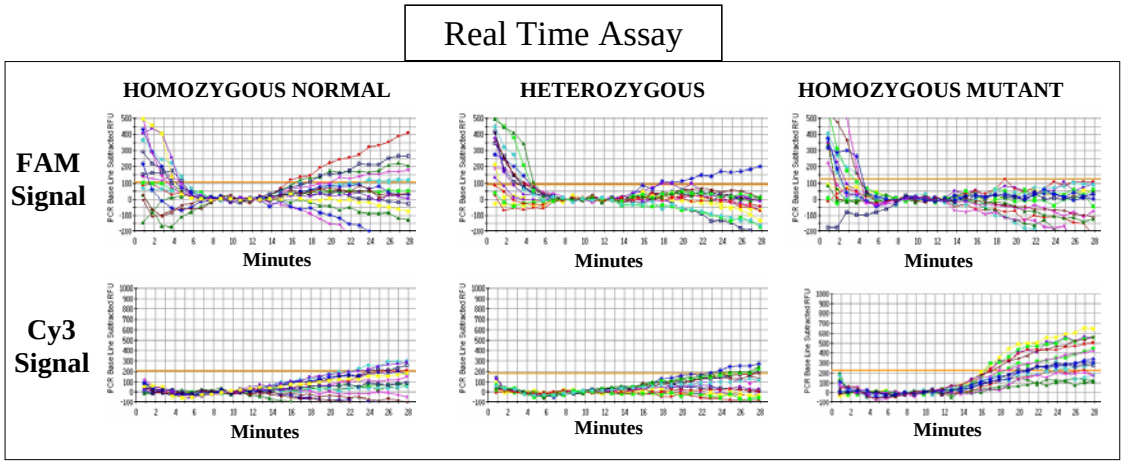

Figure 6
OCPs without hairpins, Factor V Leiden Mutation Ligation/ERCA Assay. Optimal assay conditions were as described in Methods. 1000 ng of MDA genomic DNA was incubated with probes for both alleles in the presence of ligase. ERCA mix was added and the reaction was followed in real time at 60°C in a Bio-Rad I-Cycler. The PCR baseline subtracted relative fluorescence units are shown as a function of time for both FAM (top row) and Cy3 (bottom row). Each trace (colored lines) corresponds to a different sample. Real time traces from reactions containing the same genotypes (either normal, heterozygous, or homozygous mutant) were grouped together according to genotype in each of the three columns. Genotypes were assigned as described in Methods.

performed in a single tube. The ERCA amplification reaction is rapid, generating signal in as little as 10 minutes, and is incubated at a single temperature. These characteristics make OCP ligation/ERCA easily adaptable to high throughput automated genotyping platforms. Accurate genotyping was obtained in screens designed to detect four clinically relevant mutations, Factor V Leiden, Factor II prothrombin, Hemochromatosis C282Y and Hemochromatosis H63D.

Several improvements to the ligation/ERCA method [8] have increased accuracy to levels acceptable for diagnostic applications and reduced reaction time. Other reports using ERCA based genotyping require PCR amplification of the locus of interest prior to genotyping [3,16], which is prohibitively expensive, time consuming, and more difficult to automate. We have designed ERCA primers that are optimized for minimal misamplification and artifact formation. With any exponential signal amplification method, nonspecific amplification due to exponential artifacts presents a potential problem. In ERCA as in PCR, amplification of primer-primer artifacts can mimic specific signal if primers are not selected carefully. To avoid this problem, commercially available primer design software was used to design P1 and P2 primers that are optimal for amplification. This step does not necessarily have to be repeated for each target, however, which is an advantage over PCR based methodologies. The primary amplification primer sequences are present in the OCP backbone, not in the portion of the OCP that anneals to target, allowing optimized P1 and P2 pairs, to be used for many different targets. In principle, a small collection of optimized P1/P2 pairs should contain primers that can be used for any given target.

The Q-PNA detection system [9] was used as a fluorescent reporter during ERCA. During ERCA, Q-PNA is rapidly displaced from the ERCA product and is physically separated from the fluor in a bimolecular reaction, resulting in detectable signal in as little as 10 minutes. By comparison, Amplifluors typically required 60–120 minutes to generate signal[8]. The substantial difference in time to generate signal may be due to the nature of Amplifluor design. In an Ampliflour, complementary DNA sequences, fluorescent reporter, and quencher are all covalently linked on the same DNA strand. The comparative

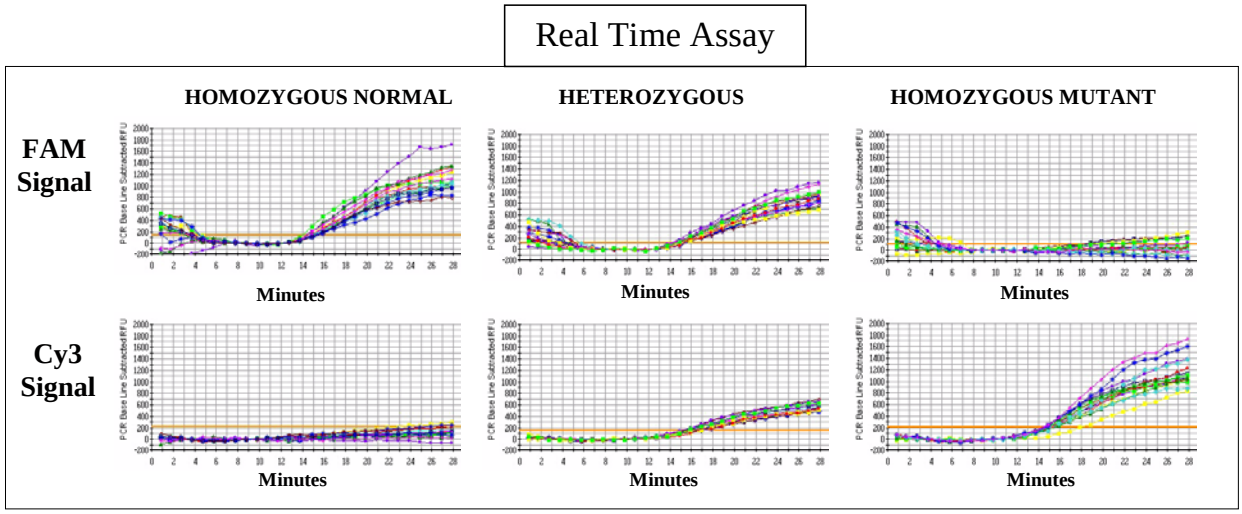

Figure 7
Hairpin OCPs, Factor V Leiden Mutation Ligation/ERCA Assay-. Assay was performed as described in Figure 6.

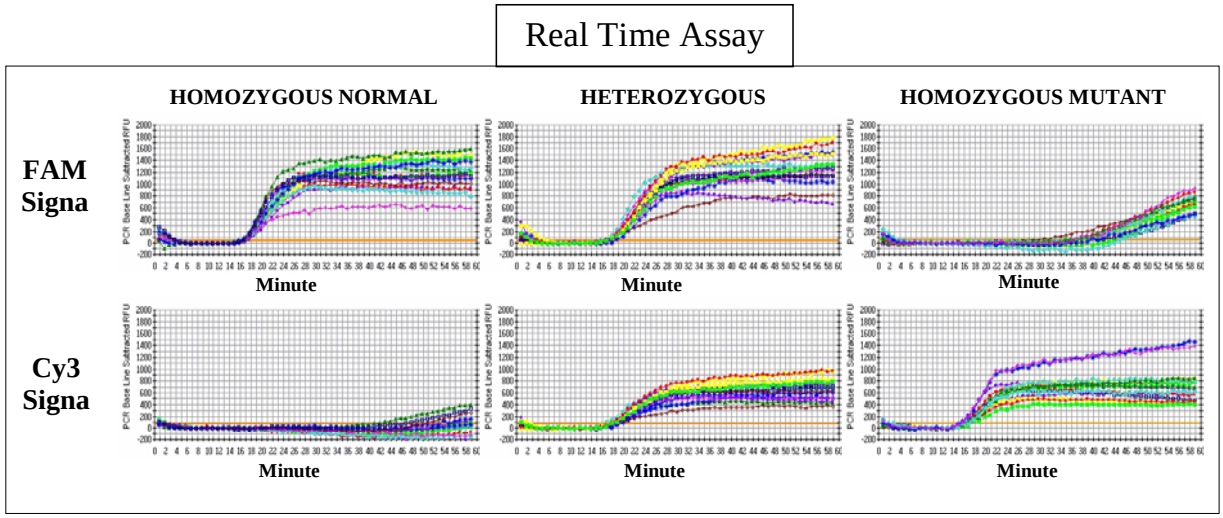

Figure 8
OCPs without hairpin, Hemochromatosis H63D Mutation Ligation/ERCA Assay-s. Assay was performed as described in Figure 6.

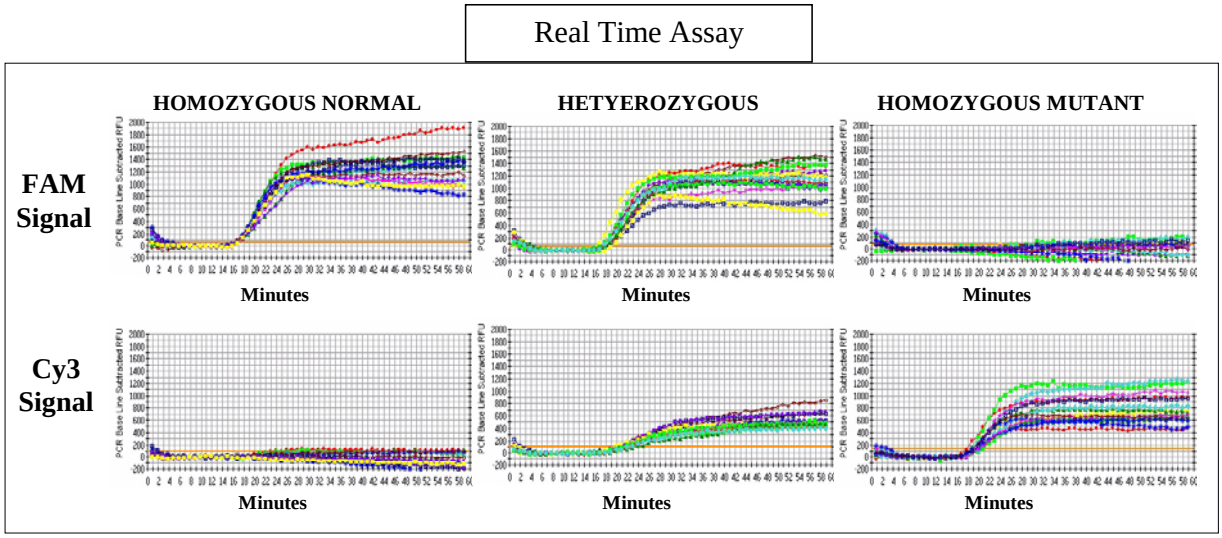

Figure 9
Hemochromatosis H63D Mutation Ligation/ERCA Assay-Hairpin OCPs. Assay was performed as described in Figure 6.

stability of the unimolecular Amplifluor hairpin is likely to hinder displacement synthesis when compared to the bimolecular Q-PNA system. In addition, the high local concentration of fluor and quencher may serve to supress signal from the Amplifluor. The Q-PNA based reporter system consistently produces signal far more rapidly than the Amplifluor based system, allowing time to result of less than an hour for the entire assay.

Secondary allele specific, nonfluorescent, primers can also be used to increase the speed and specificity of ERCA. These primers significantly advanced the rate of ERCA amplification, cutting the reaction time in half for the Factor V Leiden probe set (data not shown), resulting in a 20-minute assay. Although maximum fluorescent signal decreased slightly as each additional nonfluorescent primer was added, the addition of one or two more primers to the reaction did not adversely influence signal strength to the extent that genotyping was compromised.

Each of the four assays developed for this report contained two probes, one for each allele. Using both probes in a single reaction has been demonstrated to reduce the levels of nonspecific amplification due to primer dimers and misamplification of unligated probe [3]. At least one of the probes will always be amplified, suppressing low levels of nonspecific signal.

The introduction of a hairpin into the design of the open circle probes, similar to approaches taken in molecular beacon design [12], provides a means to regulate the degree of ligation discrimination. A 3' and/or 5' stem-loop structure may be designed to increase specific binding to difficult sequences. The ratio of the hairpin stability to target annealing stability directly affects the overall target annealing specificity. By modulating the stability of the hairpin any OCP can be fine tuned for ligation specificity.

The OCP 3' hairpin also decreases the level of nonspecific DNA synthesis caused by misamplification of OCP sequences. Analysis of misamplified products has shown that OCP sequences are usually found in the misamplified DNA products. This means that uncircularized OCP is able to provide a template for primer-based misamplification. Removing the unligated OCP helps to reduce background amplification [13], but digestion of the unused OCP or purification of the ligated circles introduces time consuming and expensive steps. Designing a 3' hairpin into the OCP creates a self-priming sequence in unligated OCP. During ERCA, the polymerase will extend unligated OCP starting at the 3' end in a suicide pathway that renders the OCP double stranded and inert. This approach eliminates nonspecific amplification and improves genotyping accuracy without the introduction of separate isolation or purification steps. As demonstrated above, removal of the hairpin sequence either by

Table 1: DNA sequences

Open Circle Probes	Length	Nucleotide Sequence
FV Normal	78 bases	5'-GCCTGTCCAGGGATCTGCTTCTTCGGTGCCCATCGCGCAGACACGATA GAGGAATACAACAAAATACCTGTATTCCTC-3'
FV WT-no hairpin	68 bases	5'-GCCTGTCCAGGGATCTGCTTCTTCGGTGCCCATCGCGCAGACACGATA ACAAAATACCTGTATTCCTC-3'
FV Mutant	78 bases	5'-GCCTGTCC AGGGATCTGCTCTGTTATCGGCCGTCGCGCAGACACGATA AAGGAATA-CAACAAAATACCTGTATTCCTT-3'
FV Mut-no hairpin	68 bases	5'-GCCTGTCCAGGGATCTGCTCTGTTATCGGCCGTCGCGCAGACACGATA ACAAAATACCTGTATTCCTT-3'
FII Normal	79 bases	5'-agcctcaatgctcccagtgcACAAGACCGAAAGGGTAGTCGCGGATTGTTG CGCTGAGAaataaaagtgactctcagcg-3'
FII Mutant	79 bases	5'-agcctcaatgctcccagtgcACTCAATCCCAGGCGAGTCGCGGATTGTTG TGCTGAGAGaataaaagtgactctcagca-3'
H63 Normal	99 bases	5'-ATCATAGAACACGAACAGCTGGTAGTATTGATCGGCCACTGGTATCTCGTCCG TGCTAGAAGGAAACATGAGAGTAGGCTCCACACGGCGACTCTCATG-3'
H63 WT-no hairpin	99 bases	5'-ATCATAGAACACGAACAGCTGGTAGTATTGATCGGCCACTGGTATCTCGTCCG TGCTAGAAGGAAATCAAGAGATGGGCTCCACACGGCGACTCTCATG-3'
H63 Mutant	99 bases	5'-ATCATAGAACACGAACAGCTGGTAATGTTGACTGGTCACACGTCGTATCGTCCG TGCTAGAAGGAAAGATGAGAGTCAGGCTCCACACGGCGACTCTCATC-3'
H63 Mut-no hairpin	99 bases	5'-ATCATAGAACACGAACAGCTGGTAGTATTGATCGGCCACTGGTATCTCGTCCG TGCTAGAAGGAAATCAAGAGATGGGCTCCACACGGCGACTCTCATC
C282 Normal	89 bases	5'-CCAGGTGGAGCACCCAGGTCAGTTATCGGGCCGAGGCGCGAGTGTAAA CACGTATATCTCTCCCCTGGGGAAGAGCAGAGATATACGTG-3'
C282 Mutant	90 bases	5'-CCAGGTGGAGCACCCAGGTTCTACCTCAGCCCGAAGGCGCGAGTGTAAA CACGTATATCTCTCCCCTGGGGAAGAGCAGAGATATACGTA-3'

Primers	Length	Nucleotide Sequence
FV Normal P1	31 bases	5'-/56-FAM/ TCATTCGCAATCAATGGGCACCGAAGAAGCA-3'
FV Mutant P1	31 bases	5'-/5Cy3/TCATTCGCAATCAACGGCCGATAACAGAGCA-3'
FV P2	15 bases	5'-CGCGCA GACACGATA-3'
FV 5' allele specific primer	18 bases	5'-GAGGAATACAACAAAATA-3'
FV 3' allele specific primer	15 bases	5'-GATCCCTGGACAGGC-3'
FII Normal P1	28 bases	5'-/56-FAM/ TCATTCGCAATCACCCTTTCGGTCTTGT-3'
FII Mutant P1	28 bases	5'-/5Cy3/ TCATTCGCAATCACGCCTGGGATTGAGT-3'
FII P2	15 bases	5'-AGTCGCGGATTGTTG-3'
FII 5' allele specific primer	15 bases	5'-CACTGGGAGCATTGA-3'
H63 Normal P1	29 bases	5'-/56-FAM/ TCATTCGCAATCAGATACCAGTGGCCGAT-3'
H63D Mutant P1	27 bases	5'-/5Cy3/ TCATTCGCAATCAACGACGTGTGACCA-3'
H63 P2	18 bases	5'-GTCCGTGCTAGAAGGAAA-3'
H63 5' allele specific primer	17 bases	5'-TGTTCGTGTTCTATGAT-3'
H63 3' allele specific primer	15 bases	5'-GGCTCCACACGGCGA-3'
C282 Normal P1	28 bases	5'-/56-FAM/ TCATTCGCAATCACGGCCCGATAACTGA-3'
C282Y Mutant P1	28 bases	5'-/5Cy3/ TCATTCGCAATCA CGGGCTGAGGTAGAA-3'
C282 P2	15 bases	5'-AGGCGCGAGTGTAAA-3'
C282 5' allele specific primer	15 bases	5'-GGGTGCTCCACCTGG-3'
C282 3' allele specific primer	14 bases	5'-GCAGAGATATACGT-3'

deletion or point mutation of the base paired region resulted in greatly decreased genotyping accuracy for both Factor V Leiden and Hemochromatosis H63D genotyping assays.

Additional methods to improve ERCA specificity were integrated into the genotyping protocol. Careful design allowed the same P2 sequence to be used in the ERCA reaction for both alleles. For each set of OCPs, the P2 sequence was generic. The generic P2 was not part of the target specific portion of the probe, which allowed it to be used in both the normal and mutant OCP. Because both probes in each assay contained the same P2, the reaction kinetics could be more easily balanced, promoting uniform amplification for both alleles. Uniform amplification is especially important when the genotype is heterozygous so that both alleles are accurately represented.

Table 2: Table of patient sample numbers and genotyping accuracy.

Assay	Number of Individual Samples Mutant:Hetero:Normal	Single Assay Accuracy	Observed Assay Accuracy in Triplicate
Factor V Leiden	13:98:105	99.5%	100%
Hemochromatosis C282Y	13:93:84	99.5%	100%
Hemochromatosis H63D	1:48:89	99.5%	100%
Factor II Prothrombin	2:12:284	99.4%	100%

For each of the assays (column 1) the number of different samples screened are grouped according to genotype (column 2). Mutant (homozygous mutant), heterozygous, normal (homozygous normal). The accuracy obtained for each single genotyping assay is shown in column three. The observed genotyping accuracy for the samples as tested in triplicate is shown in column four.

Genotyping accuracy depends on using sufficient quantities of genomic DNA target. As expected, OCP ligation/ ERCA results were influenced by the amount of genomic DNA used in the assay. Typically, genotyping was possible with as little as 50 ng target DNA. However, specific signal strength increased with more DNA target, up to 1000 ng, after which no further benefit was seen (data not shown). WGA DNA can be produced in milligram quantities from nanograms [10] of original sample allowing use of high DNA concentration for optimal signal. As a result, 1000 ng WGA DNA was used to obtain optimal genotyping reaction conditions. The results obtained for the WGA DNA agreed in every instance with the known genotypes determined by RFLP. WGA product should also be compatible with other genotyping assays where better results can be obtained using more target DNA.

Many aspects of probe design are open to manipulation, and in practice probes can be successfully designed for most target sequences of interest. Pickering *et al.* demonstrated a 95% success rate in designing OCPs for 99 different targets on the first attempt [3]. Probes can be designed against either strand of genomic DNA, and the backbone portion can also be varied. Optimization of T_m parameters, and sequence allows rapid design of OCPs with a high degree of success.

The degree to which a genotyping assay needs to be optimized depends in part on the purpose of the assay. A genotyping system intended for diagnostic purposes needs to meet the highest standards of accuracy, robustness, and throughput, as compared to a system that is intended solely for research purposes. In recognition of the expectation that 100% accuracy is a requirement, even under suboptimal conditions, the assays developed for the targets in this report have undergone extensive optimization to maximize accuracy even under less than optimal conditions. As a result, the concentrations of primers, probes and polymerase vary between assays. After determining the best conditions for each individual genotyping assay, the expectation is that these conditions would be the best possible for genotyping large numbers of samples with high accuracy. The critical nature of the information derived from the output of the assay justifies the extra optimization effort. To minimize the potential for error, the accuracy for a single assay was compared to the accuray for the same assay performed in triplicate. Depending on the assay requirements, the level of accuracy can be selected by increasing or decreasing the number of repeats. The demonstration that high levels of input genomic DNA improve robustness of genotyping results is paralleled by PCR results. WGA can be used to generate sufficient DNA for genotyping using large quantities of template. The cost of performing WGA is offset by the fact that sample prep is unnecessary; WGA product can be performed directly on crude blood sample and yields enough DNA for dozens or hundreds of genotyping reactions[11]. Under different circumstances, it may be necessary to rapidly generate assays for a large number of SNPs. For this, the emphasis could be placed on optimization of assay design. Amounts of input DNA, primers and polymerase could be standardized, rapidly yielding accurate genotyping assays but without the highest accuracy under suboptimal conditions. The process of OCP ligation followed by ERCA has been designed such that genomic DNA can be genotyped directly in less than an hour. Hundreds of individual samples were genotyped with over 99% accuracy. Accurate genotyping was demonstrated for both genomic DNA and whole genome amplified DNA. Isothermal exponential amplification produces a fluorescent signal from ligated OCP using nanomolar amounts of probe. Because OCP ligation/ ERCA does not rely on thermal cycling for amplification, amplification kinetics are not limited by cycling rate. The exponential nature of the reaction produces millions of fold amplification in as little as 20 minutes. Output can be read on a fluorescent plate reader, real time PCR instrument, fluorescent imager, or other device equipped with the ability to measure fluorescent signals. OCP ligation/ ERCA probes can be multiplexed. OCP ligation and ERCA reactions for each SNP target are performed in a single

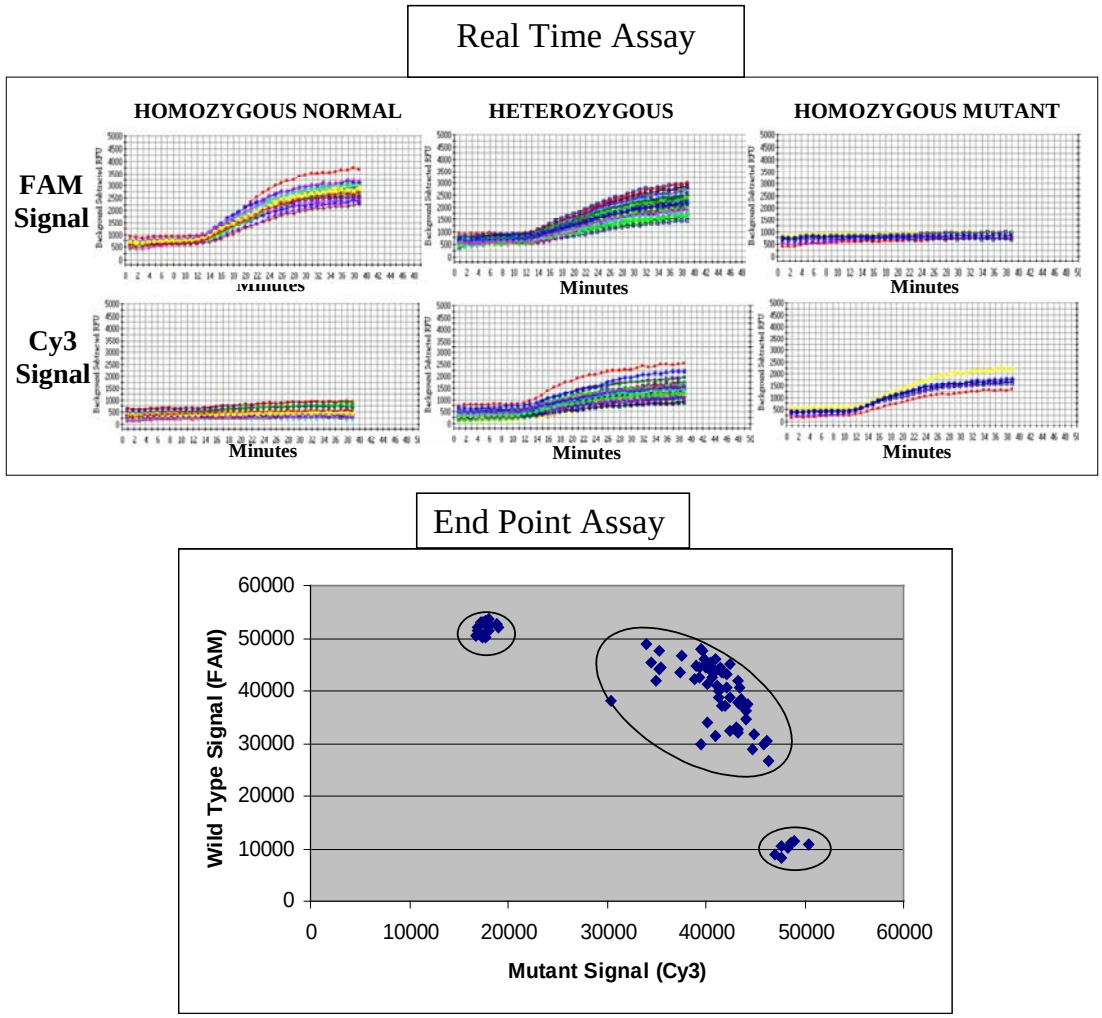

Figure 10
Factor V Leiden Mutation: 87 different patient samples, 3 different genotypes. Ligation/ERCA genotyping results visualized as both a real-time assay and as an end-point. Assay conditions were as described in Methods for specific genotyping assay targets. 1000 ng of WGA genomic DNA was incubated with OCPs for both alleles in the presence of ligase. ERCA reactions were monitored in real time in the Bio-Rad I-cycler. Each colored trace in the real time charts corresponds to a different sample. At the end of the assay, approximately one-hour after addition of DNA polymerase, the intensity of both FAM and Cy3 fluorescence were measured for each sample in a fluorescent plate reader. End-point genotyping results were analyzed by the fuzzy cluster program and grouped automatically into genotypes as described in Methods.

tube, and are easily scalable from 96 to 384 well formats, making OCP ligation/ ERCA ideal for high throughput screening.

Methods

Hairpin OCP Design

All oligomers were ordered from Integrated DNA Technologies, Inc.(Coralville, IA). Sequences of oligomers are shown in Table 1. All OCPs were 5' phosphorylated and gel purified. OCPs were designed such that the 3' terminal nucleotide annealed opposite the mutation of interest (Figure 1). The T_m for the 3' arm of the probe annealed to target was designed to be between 60–65°C. The 5' arm of all OCPs annealed to target upstream of the mutation and was designed to have a T_m of approximately 70°C. Oligo

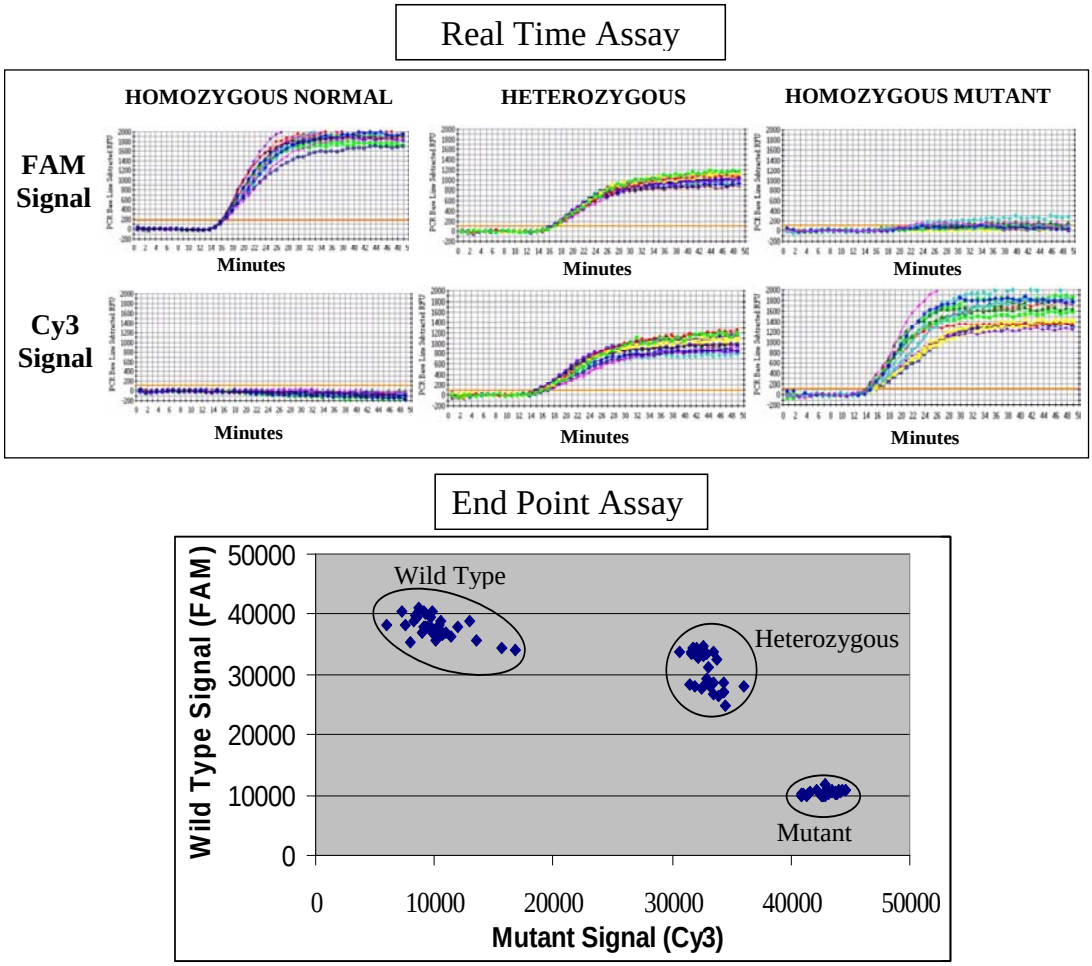

Figure 11
Factor II Prothrombin Mutation: 32 repeats of 3 patient samples, 3 different genotypes. Assay was performed as described in Figure 10.

6 (Molecular Biology Insights, Inc., Cascade, CO) was used to calculate T_m values.

To reduce or eliminate the artifactual amplification of unligated OCPs, a hairpin was designed into the 3' end of the OCP, with approximately 10 bases in the stem and 10 in the loop. The hairpin was calculated to be stable at 60°C in the context of the entire OCP using the DNA folding server at Michael Zucker's website http://bioinfo.math.rpi.edu/~zukerm/ OCP sequence was chosen such that the hairpin was the only stable structure in the OCP, with a ΔG of -0.5 to -1.5 kcal/mol.

Generic P1 and P2 design

ERCA primers P1 and P2 were chosen using Oligo 6. These were generic primer sequences, and could be used to amplify any circle sequence. First, 10–20 kb of randomly generated DNA sequence was scanned for primers using settings limiting 3' dimer ΔG to -1.0 kcal/mol, 3' dimer length to 3, 3' terminal stability range -5.5 to -8.0 kcal/mol, GC clamp stability to -10 kcal/mol, no acceptable loop, T_m range 52.1 to 56.2°C. Compatible primers were selected using the multiplex function of Oligo 6. A subset of these compatible primers was utilized as generic P1 and P2 primers (see Table 1). ERCA P1 primers were

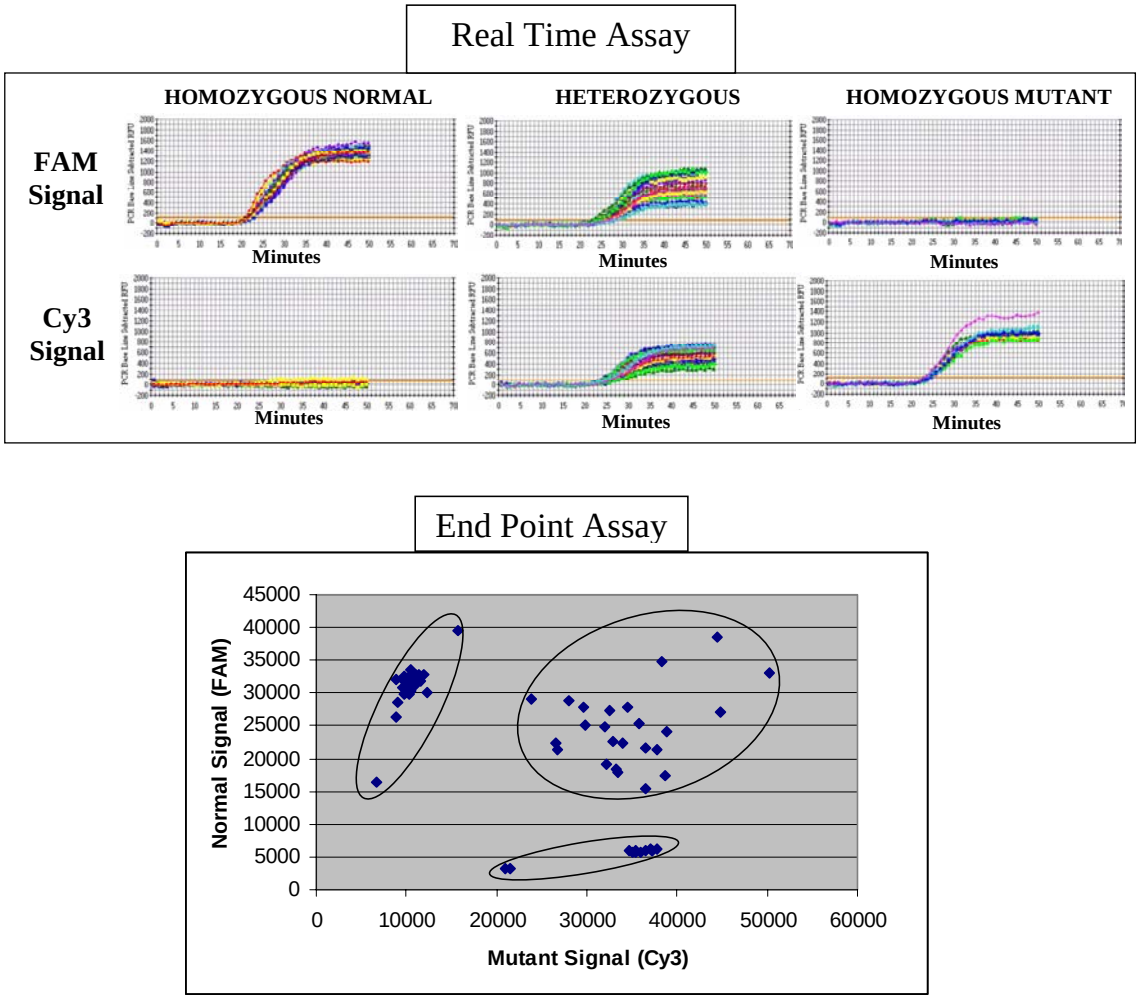

Figure 12
Hemochromatosis H63D Mutation: 71 different patient samples, 3 different genotypes. Assay was performed as described in Figure 10.

labeled with FAM (if paired with the normal OCP) or 5-Cy3 (if paired with the mutant OCP) and were RP-HPLC purified. P2 primers were purified by desalting.

Allele Specific ERCA primers

The sequence of these primers was chosen from the target specific 5' and 3'-arms of the OCP, either complementary to the OCP (5'-arm) or the exact sequence of the OCP (3'-arm). These primers annealed to both the normal and mutant OCPs. The sequences of these allele specific primers are shown in Table 1. Allele specific primers were desalted after synthesis.

Q-PNA

Exponential amplification of circularized OCP (see Figure 2) was accomplished using generic primers P1 and P2, along with allele specific primer(s). Each P1 primer contained a universal 13 base 5'-tail (see Table 1) with a fluorescent dye at the 5'-end. Fluorescence was quenched in unincorporated primers by the annealing of a 13-residue PNA molecule, Q-PNA-13 (Applied Biosystems/Boston Probes, Bedford, MA). The sequence of Q-PNA-13 is Ac-X-OO-TGA-TTG-CGA-ATG-A-Lys (Dabcyl). The C-terminal dabcyl was positioned so that it was proximal to the 5'-fluorescent moiety on the DNA reporter, quenching fluores-

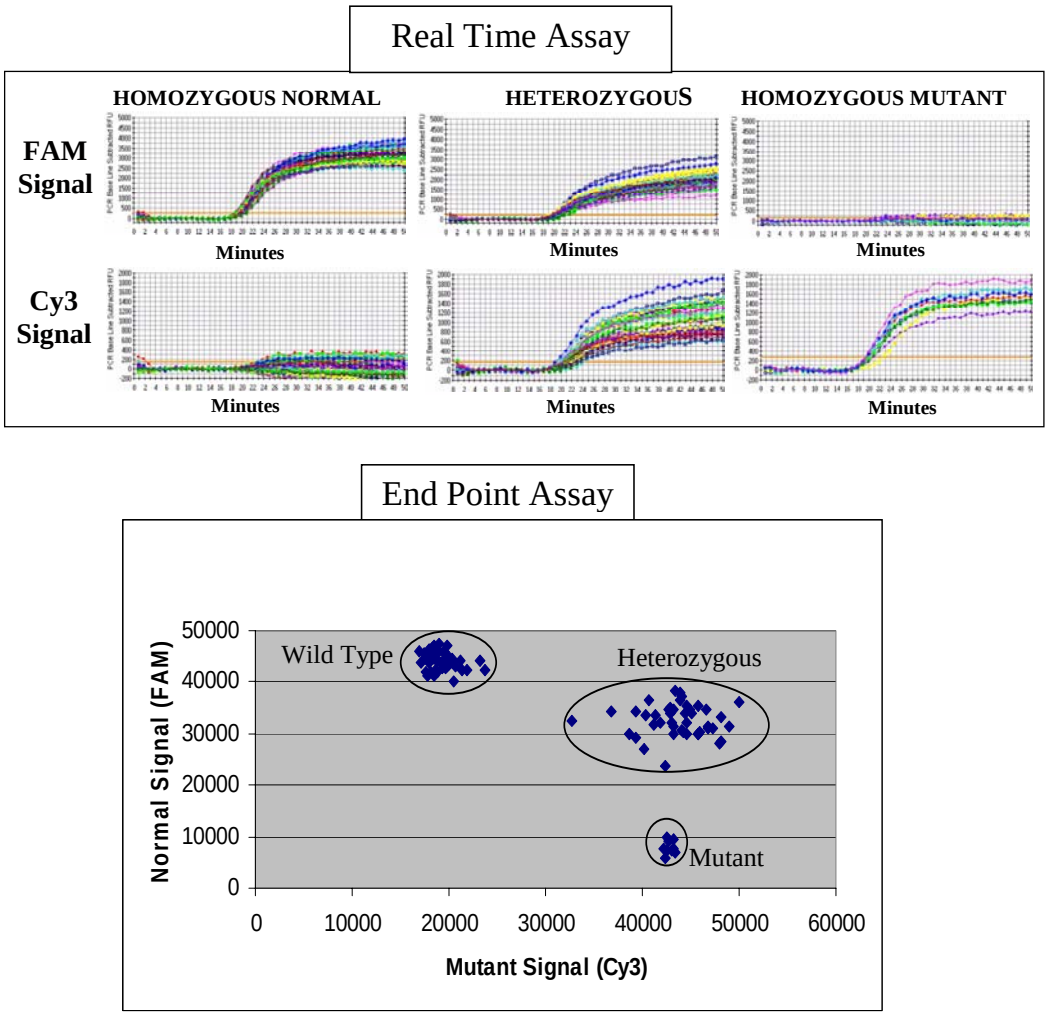

Figure 13
Hemochromatosis C282Y Mutation: 96 different patient samples, 3 different genotypes. Assay was performed as described in Figure 10.

cence when the reporter and Q-PNA-13 were annealed, as shown in Figure 2. After P1 primer is incorporated into ERCA product, Q-PNA is displaced, producing a fluorescent signal.

Genomic DNA

Genomic DNA samples were purchased from Coriell Cell Repositories, Camden, NJ. Human Variation Panel, Caucasian panel of 100, Reference # HD100CAU, was used as non-mutant controls and for general screening purposes. Human samples containing specific mutations were also purchased from Coriell: Factor V Leiden mutation (heterozygous NA14642 and homozygous NA14899); Hemochromatosis C282Y (heterozygous NA14642, homozygous NA14620); and Hemochromatosis H63D (heterozygous NA14641 and homozygous NA13591); Factor II Prothrombin mutation (heterozygous NA16028 and homozygous NA16000). Other patient DNA samples were genotyped using RFLP, and an aliquot of the purified DNA was subsequently used either directly in genotyping reactions, or used after Whole Genome Amplification (WGA).

Whole Genome Amplification (WGA) of genomic DNA samples

For some rare genotypes, it was only possible to obtain a few nanograms of sample, insufficient for more than one genotyping reaction. As a result, the rare samples were subjected to whole genome amplification (WGA), which yielded enough DNA for 200 or more reactions. WGA was performed as described previously [10]. Briefly, 10 ng of each individual genomic sample was amplified into 40 μg of product by adding 99 μl of amplification mix and incubating at 30°C for 6 hours. Quality control assays demonstrating that whole genome amplification was successful were performed for locus 979 and 1004 as described [10], and amplified DNA was used directly in genotyping reactions at the same concentration as the original genomic DNA, typically 200 ng/reaction.

Genotyping reactions

50 ng to 1 μg of either genomic DNA or WGA genomic sample was mixed with OCP (typically 0.5 nM final concentration) and 0.5 unit of Ampligase (Epicentre Technologies, Madison, WI) in 1x Ampligase buffer (Epicentre), for a total volume of 10 μl. The reaction was heated to 95°C for 10 seconds, and cooled to 63–68°C for 5–20 minutes, during which time OCP annealed to genomic target and was circularized by ligase. The reaction was subsequently heated to 95°C for 10 minutes to release ligated circles from genomic DNA. The reaction was cooled to 4°C, and 20 μl ERCA reaction mix was added (typically 16 units BST polymerase (New England Biolabs, Beverly, MA), 6 mM dNTPs, 0.5 μM P1, 0.5 μM P2, 4 μM Q-PNA, 7.5 μM TMAO in 1x ThermoPol Buffer II, all concentrations final). Reactions were incubated at 60°C for 1 hour in an I-Cycler (BioRad, Hercules, CA) reading both FAM and Cy3 channels. Signals typically appeared after 10–20 minutes.

Specific reaction component concentrations

FV ligation reaction conditions

0.5 nM wt OCP and 0.5 nM mut OCP (final concentration); 0.5 U ampligase.

FV ERCA reaction conditions

0.5 μM WT P1; 0.5 μM mutant P1; 0.75 μM P2; 0.75 μM 5'-allele specific primer; 0.75 μM 3'-allele specific primer; 4 μM PNA (final concentrations); 24–32 U DNA polymerase.

FII ligation reaction conditions

0.5 nM wt OCP and 0.5 nM mut OCP (final concentration); 0.5 U ampligase.

FII ERCA reaction conditions

0.5 μM WT P1; 0.7 μM mutant P1; 0.9 μM P2; 0.4 μM 3'-allele specific primer; 4 μM PNA (final concentrations); 16 U DNA polymerase.

Hemochromatosis H63 ligation reaction conditions

0.1 nM wt OCP and 1.2 nM mut OCP (final concentration); 0.5 U ampligase.

Hemochromatosis H63 ERCA reaction conditions

0.4 μM WT P1; 0.5 μM mutant P1; 0.75 μM P2; 0.5 μM 5'-allele specific primer; 0.5 μM 3'-allele specific primer; 4 μM PNA (final concentrations); 16 U DNA polymerase.

Hemochromatosis C282 ligation reaction conditions

0.1 nM wt OCP and 1.2 nM mut OCP (final concentration); 0.5 U ampligase.

Hemochromatosis C282 ERCA reaction conditions

0.4 μM WT P1; 0.5 μM mutant P1; 0.75 μM P2; 0.5 μM 5'-allele specific primer; 0.5 μM 3'-allele specific primer; 4 μM PNA (final concentrations); 16 U DNA polymerase.

Determination of genotype

In this retrospective study, the genotype of each sample was known. Genotyping by RFLP analysis (data not shown) was performed prior to OCP ligation/ ERCA genotyping, and the results of OCP ligation/ ERCA analysis were compared to the RFLP analysis results.

Real Time OCP ligation/ ERCA

The genotype for each sample was determined by amplitude of amplification. The average amplification threshold time for all amplified reactions was determined using the I-Cycler software. Fluorescence traces were normalized using early cycles as a baseline, and a threshold value was determined, typically at 10-fold above the average standard deviation of the baseline values. Threshold cycle for each trace was measured at the point where the trace crossed the threshold value. Threshold cycle values fell into three distinct clusters, one each for homozygous normal, heterozygous, and homozygous mutant. Reactions with times more than 2 standard deviations beyond the mean cluster value were considered failures and repeated. In general, amplification reactions with an increase of less than 300 units were scored as negative. Any reaction with amplification of 300 fluorescence units above baseline or greater was scored as positive. Reactions with fluorescent units between 101–299 above baseline were repeated. Any reaction where signal was baseline for both alleles was repeated.

End Point OCP ligation/ ERCA

ERCA reactions were allowed to incubate at 60°C until the reaction was expected to be complete (30 to 40

minutes, depending on the assay). After the reaction was complete, the results could be read at any time. Some reactions were allowed to remain at room temperature overnight, some were stored at 4°C, some were frozen at -20°C. In each case, the reaction was protected from light to prevent photobleaching of the fluorescent reporters. The genotype of each sample was determined automatically using a modified fuzzy *c*-means clustering algorithm [3], which groups the data into three genotypes plus a negative control, and assigns a confidence level to each genotyping call from 0 (not in cluster) to 1 (100% certainty that point belongs to cluster).

Authors' Contributions

OA contributed the hairpin OCP idea and performed initial hairpin OCP design and hairpin folding experiments, and edited the manuscript. CB designed and tested FVL OCPs and participated in generating figures and writing the manuscript. WS performed Factor II OCP genotyping. MW designed H63D and C282 probes and performed H63D OCP genotyping. JD assisted with FVL genotyping. JB assisted with OCP model validation. AFF and SH assisted in OCP design and troubleshooting. ZS, YD, XW performed WGA amplifications of genomic DNA. ME provided PNA idea, edited manuscript. PA participated in editing the manuscript, design and coordination of the hairpin OCP work. RSL edited manuscript and participated in study design and coordination. MD led study design and coordination, participated in writing the manuscript, designed and validated primers, probes. All authors read and approved the manuscript.

References

1. Isaksson A and Landegren U: **Accessing genomic information: alternatives to PCR** *Curr Opin Biotechnol* 1999, **10:**11-15.
2. Shi MM: **Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies** *Clin Chem (Washington, DC, US)* 2001, **47:**164-172.
3. Pickering J, Bamford A, Godbole V, Briggs J, Scozzafava G, Roe P, Wheeler C, Ghouze F and Cuss S: **Integration of DNA ligation and rolling circle amplification for the homogeneous, end-point detection of single nucleotide polymorphisms** *Nucleic Acids Res* 2002, **30:**e60.
4. Lizardi PM, Huang X, Zhu Z, Bray-Ward P, Thomas DC and Ward DC: **Mutation detection and single-molecule counting using isothermal rolling-circle amplification** *Nat Genet* 1998, **19:**225-232.
5. Baner J, Nilsson M, Mendel-Hartvig M and Landegren U: **Signal amplification of padlock probes by rolling circle replication** *Nucleic Acids Res* 1998, **26:**5073-5078.
6. Nilsson M, Baner J, Mendel-Hartvig M, Dahl F, Antson DO, Gullberg M and Landegren U: **Making ends meet in genetic analysis using padlock probes** *Hum Mutat* 2002, **19:**410-415.
7. Nilsson M, Malmgren H, Samiotaki M, Kwiatkowski M, Chowdhary BP and Landegren U: **Padlock probes: circularizing oligonucleotides for localized DNA detection** *Science* 1994, **265:**2085-2088.
8. Faruqi AF, Hosono S, Driscoll MD, Dean FB, Alsmadi O, Bandaru R, Kumar G, Grimwade B, Zong Q, Sun Z, Du Y, Kingsmore S, Knott T and Lasken RS: **High-throughput genotyping of single nucleotide polymorphisms with rolling circle amplification** *BMC Genomics* 2001, **2:**4.
9. Fiandaca MJ, Hyldig-Nielsen JJ, Gildea BD and Coull JM: **Self-reporting PNA/DNA primers for PCR analysis** *Genome Res* 2001, **11:**609-613.
10. Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore SF, Egholm M and Lasken RS: **Comprehensive human genome amplification using multiple displacement amplification** *Proc Natl Acad Sci U S A* 2002, **99:**5261-5266.
11. Hosono S, Faruqi AF, Dean FB, Du Y, Sun Z, Wu X, Du J, Kingsmore SF, Egholm M and Lasken RS: **Unbiased Whole Genome Amplification Directly From Clinical Samples** *Genome Res* 2003, **13:**954-964.
12. Bonnet G, Tyagi S, Libchaber A and Kramer FR: **Thermodynamic basis of the enhanced specificity of structured DNA probes** *Proc Natl Acad Sci U S A* 1999, **96:**6171-6176.
13. Zhang DY, Brandwein M, Hsuih TC and Li H: **Amplification of target-specific, ligation-dependent circular probe** *Gene* 1998, **211:**277-285.
14. Hafner GJ, Yang IC, Wolter LC, Stafford MR and Giffard PM: **Isothermal amplification and multimerization of DNA by Bst DNA polymerase** *BioTechniques* 2001, **30:**852-854.
15. Zhang DY, Brandwein M, Hsuih T and Li HB: **Ramification amplification: a novel isothermal DNA amplification method** *Mol Diagn* 2001, **6:**141-150.
16. Thomas DC, Nardone GA and Randall SK: **Amplification of padlock probes for DNA diagnostics by cascade rolling circle amplification or the polymerase chain reaction** *Arch Pathol Lab Med* 1999, **123:**1170-1176.