



UPPSALA
UNIVERSITET

*Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Medicine 65*

Selector Technology

For Multiplex DNA Analysis

FREDRIK DAHL



ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2005

ISSN 1651-6206
ISBN 91-554-6330-4
urn:nbn:se:uu:diva-5921

Dissertation presented at Uppsala University to be publicly examined in Rudbecksalen, Rudbecklaboratoriet, Uppsala, Friday, October 7, 2005 at 09:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract

Dahl, F. 2005. Selector Technology. For Multiplex DNA Analysis. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 65. 44 pp. Uppsala. ISBN 91-554-6330-4.

A majority of methods for identifying sequences in the human genome involve target sequence amplification through PCR. This work presents novel methods for amplifying circularized DNA and presents solutions for some major limitations of PCR.

We have developed a novel method to amplify circularized DNA molecules based on a serial rolling-circle replication reaction, called circle to circle amplification (C2CA). Amplified DNA circles can be detected in array-based analyses or in real-time using molecular beacons. The amplification mechanism allows higher precision in quantification than in exponential amplification methods like PCR, and more products can be generated than in PCR.

A major limitation of PCR is that amplification artifacts arise when large numbers of specific primer pairs are simultaneously added to a reaction. We have developed a solution to this problem that enables multiplex PCR amplification of specific target sequences without producing amplification artifacts. The procedure is based on oligonucleotide constructs, called selectors. The selectors identify defined target nucleic acid sequences, and they act as ligation templates to direct circularization of these targets. The selectors contain a general primer-pair motif that allows the circularized targets to be amplified in multiplex using a universal PCR primer pair. We also developed a computer program, PieceMaker, that finds an optimal design of selector probes for a given selector application. We demonstrate the method by performing a 96-plex PCR of specific DNA sequences with high success-rate and reproducibility.

Keywords: DNA technology

Fredrik Dahl, Department of Genetics and Pathology, Rudbecklaboratoriet, Uppsala University, SE-75185 Uppsala, Sweden

© Fredrik Dahl 2005

ISSN 1651-6206

ISBN 91-554-6330-4

urn:nbn:se:uu:diva-5921 (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-5921>)

To my family

List of Papers:

This thesis is based on the following publications, which are referred to in the text by their roman numerals.

- I. Nilsson, M., Gullberg, M., **Dahl, F.**, Szuhai, K. and Raap, A.K. (2002) Real-time monitoring of rolling-circle amplification using a modified molecular beacon design. *Nucleic Acids Res*, **30**, e66.
- II. **Dahl, F.**, Baner, J., Gullberg, M., Mendel-Hartvig, M., Landegren, U. and Nilsson, M. (2004) Circle-to-circle amplification for precise and sensitive DNA analysis. *Proc Natl Acad Sci U S A*, **101**, 4548-4553.
- III. **Dahl, F.**, Gullberg, M., Stenberg, J., Landegren, U. and Nilsson, M. (2005) Multiplex amplification enabled by selective circularization of large sets of genomic DNA fragments. *Nucleic Acids Res*, **33**, e71.
- IV. Stenberg, J., **Dahl, F.**, Landegren, U. and Nilsson, M. (2005) PieceMaker: selection of DNA fragments for selector-guided multiplex amplification. *Nucleic Acids Res*, **33**, e72.

The first two authors of paper III contributed equally to the work.

Contents

Introduction.....	9
Multiplex DNA Analysis	10
Polymerase chain reaction.....	10
The multiplex PCR problem.....	11
Rolling circle amplification.....	14
Whole genome amplification	15
Multiplex probe-based DNA analysis methods.....	16
Present Investigations	19
Paper I. Real-time monitoring of rolling-circle amplification using a modified molecular beacon design.....	19
Perspectives on Paper I.....	20
Paper II. Circle-to-circle amplification for precise and sensitive DNA analysis.....	21
Perspectives on Paper II.....	23
Paper III. Multiplex amplification enabled by selective circularization of large sets of genomic DNA fragments.	25
Perspectives on Paper III.	26
Recent application developments	27
Paper IV. PieceMaker: selection of DNA fragments for selector-guided multiplex amplification.	31
Perspectives on Paper IV	32
Future Perspectives	33
Acknowledgements.....	35
References.....	37

Abbreviations

C2CA	Circle-to-circle amplification
cDNA	Complementary DNA
DNA	Deoxyribonucleic acid
LDR	Ligase detection reaction
MB	Molecular beacon
MDA	Multiple displacement amplification
MIP	Molecular inversion probe
MLGA	Multiplex ligation genome amplification
MLPA	Multiplex ligation probe amplification
mtDNA	Mitochondrial DNA
NT	Nucleotide
OLA	Oligonucleotide ligation assay
PCR	Polymerase chain reaction
RE	Restriction enzyme
RO	Replication oligonucleotide
SNP	Single nucleotide polymorphism
WGA	Whole genome amplification
WGS	Whole genome sequencing

Introduction

DNA sequencing (1) and polymerase chain reaction (2), PCR, are the two molecular techniques that have made the largest contribution to the understanding of the human genome. These two techniques along with the strategy for gene cloning (3) made it possible to sequence genomes from a wide variety of organism, and recently also the human genome (4,5). The sequence of the human genome consists of 3×10^9 base pairs and offers an invaluable source of information that describes the molecular function of our species. Our next major challenge is to understand what the sequence signifies, what proteins the sequence encodes and how the sequence varies between individuals. Because of the very complex arrangement of the human genome we need new improved techniques to sort and compare the enormous information that the DNA sequence contains.

There are several methods available to identify sequences in the human genome, ranging from techniques to investigate single nucleotide variation to methods for resequencing whole genomes. Currently, the vast majority of these methods involve a target sequence amplification step, because of the limitation with current instrument of detecting single molecules in a high throughput manner. To be able to study many specific positions in the human genome from several individuals, a multitude of DNA targets have to be amplified. Inexpensive and rapid methods to copy DNA in a high throughput manner are thus a central goal of molecular technology development.

The aim of this work has been to develop methods for DNA amplification that allows many DNA targets to be copied in multiplex. The following section will list some of the methods that are relevant for this work, and their respective advantages and disadvantages.

Multiplex DNA Analysis

Polymerase chain reaction

The most widely used method for DNA amplification in genetic research and diagnostic is the polymerase chain reaction, PCR (2). PCR was developed in the 1980s and has revolutionized the understanding of genomes and is today involved in almost every type of applications coupled to DNA analysis. PCR has proven very specific and sensitive and is very easy to handle. The method utilizes a thermostable DNA polymerase that copies DNA by extending primers that hybridizes to the targets of interest. The reaction is then carried out in cycles of denaturation, annealing, and extension using the replicated products as templates in later cycles, resulting in an exponential growth of sequences.

During amplification, re-hybridization of denatured targets competes with the annealing of primers to the amplification product. At a certain concentration of products, the reaction reaches a plateau phase where the amplification more or less stops. High concentrations of targets have also been reported to have inhibitory effects on the PCR polymerase used (6). This phenomenon is herein referred as “product inhibition” (Figure 1) and is one major limitation of the PCR method.

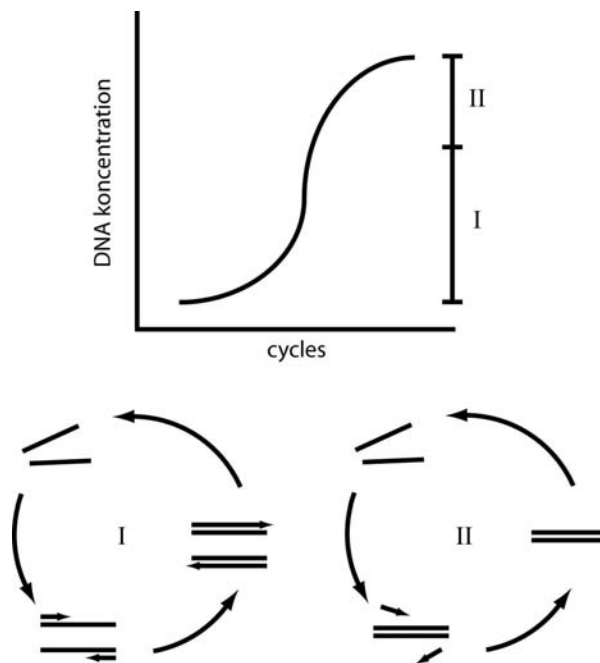


Figure 1. Product inhibition. Under optimal conditions, PCR copies DNA exponentially (I) to a certain concentration where the reaction reaches a plateau (II). The reason for the plateau is that at high concentrations of amplification products, annealing of primers competes with re-hybridization of denatured targets. At this point the reaction essentially stops and few or no more targets can be copied regardless the concentration of input reagents.

The multiplex PCR problem

To increase assay throughput and allow more efficient use of DNA samples, simultaneous amplification of many targets can be carried out by combining many specific primer pairs in individual reactions (7,8). However, it is one of the crucial problems with PCR that when large numbers of specific primer pairs are added to the same reaction, undesired amplification products arise. The amplification artifacts in multiplex PCR tends to increase roughly as the square of the number of added primer-pairs (9). Even with careful attention paid to the design of the primers, PCR is usually limited to 10-20 simultaneous amplification reactions before amplification yield is compromised by the accumulation of irrelevant amplification products (10,11). Therefore, in research projects that comprise identification and analysis of many nucleic acids sequences, a large number of separate PCRs must be performed. To-

day, a PCR generally takes about two hours and requires a defined amount of target material. In investigations where many PCRs must be performed, the projects often prove time consuming, expensive, and require a large amount of target DNA. Because of this problem there is a requirement for methods that permit amplification of multiple specific DNA sequences in the same reaction without producing amplification artifacts. Different methods have been developed to work out the problem associated with multiplex PCR, but so far none that satisfactorily solves the problem.

A number of PCR-based DNA amplification methods using so called adaptor-ligation have been developed. An adaptor is a short piece of double stranded DNA that contains a general PCR primer motif and is suitable for ligation to digested genomic DNA. The adaptor is not target specific but complementary to the ends of every digested fragment. The concept of adaptor-ligation was first introduced in cloning strategies (12) to reduce background and increase efficiency of constructing cloning libraries. Adaptor-ligation was later combined with PCR to amplify unknown regions by using a universal primer pair complementary to the target ligated adaptor (13,14). The concept has subsequently been used for PCR-based DNA fingerprinting (15), and for genotyping using genomic sub fractions prepared by size selection on gels followed by adaptor-ligation PCR (16). Kennedy, *et al.* (17) presented a similar method, termed whole genome sampling analysis (WGSa), for genotyping ~15000 SNPs without locus specific primers or gel purification. The method utilizes fragment selection and complexity reduction through adaptor ligation on digested whole genome samples (Figure 2).



Figure 2. Adaptor ligation PCR. The DNA adaptor (black) contains a general PCR primer motif and is connected by ligation to the ends of every piece of the digested DNA sample (grey). A general PCR primer-pair, complementary to the DNA adaptor, is added to the following PCR to amplify the DNA sample. The complexity of the DNA sample is reduced during the PCR step; short fragments form hairpins and fail to be amplified, and amplification of long fragments can be avoided by using a short extension step. The resulting amplified product is a reduced representation of the digested sample and can now be analyzed using e.g. high-density oligonucleotide arrays.

The ligation of adaptors to the digested sample is then followed by a PCR designed to amplify fragments of a certain size using only one primer pair. The size is determined by the PCR protocol; very long fragments are not amplified due to a short extension step and very short fragments fail to be

amplified due to formation of hairpins for these targets. The complexity of the human genome is reduced by a factor of 50, using one restriction enzyme followed by a PCR designed to amplify fragments of 250-1000 bp in size (18). The reduced complexity of the DNA sample allows the amplified product to be analyzed using high-density oligonucleotide arrays (19,20). By recent advances in array technology, this approach has enabled parallel genotyping of over 100,000 SNPs on a pair of arrays (21).

The main limitation of adaptor-ligation approaches is that the panel of targets to be amplified in parallel is fixed for a given restriction enzyme(s) and PCR protocol. In principle – you analyze what you get. However, targets that are being processed through complexity reduction seem to be quite reproducibly amplified with a given set of restriction enzyme(s) and PCR protocol (17,18,21).

A number of techniques have been developed for multiplex PCR of specific targets. Callow *et al.* (22) present a method with two rounds of amplification using semi-specific adaptors. This method provides sufficient specificity to amplify single fragments from the human genome but it scales very poorly due to problems with cross-reactivity when combining different adaptors. Another solution of the multiplex PCR problem, based on two rounds of amplification, are described by both Shuber *et al.* and Brownie *et al.* (23,24). In the first round of amplification, specific primers containing a general sequence motif at their 5'-ends are used in relatively low concentration using long annealing times. In the second round, high concentration of a general primer is used having a sequence corresponding to the general sequence motif in the specific primers. These methods are an improvement compared to standard multiplex PCR, but none of them have demonstrated reactions over ~100-plex.

Broude, *et al.* (11) presented an approach to amplify 10-20 specific targets in parallel where each fragment is amplified using one specific primer and one general primer complementary to an adaptor attached to the target. A similar approach was recently described by Shapero, *et al.* (25) that present impressive data with parallel amplification of 4000 specific fragments. The authors claim that the content of the amplified product is sufficiently defined to be used for genotyping using high-density oligonucleotide arrays (19). But, even though the complex content of the amplified product suits this particular application, the amplification method still copies a large collection of undesired DNA fragments that will generate problems in other types of DNA analyses. Therefore, large numbers of separate PCRs are typically performed whenever many genomic sequences need to be analyzed. A solution of the multiplex PCR problem is presented in paper III.

Rolling circle amplification

Several viruses and plasmids use a rolling circle mechanism to replicate their circular genomes (26). These biological systems have stimulated development of a group of techniques that is based on copying circular DNA molecules through rolling-circle amplification (RCA) procedures. The amplification reaction requires a DNA polymerase, nucleotides, and one primer, and it generates a single stranded product with repeated complementary copies of the circle (27,28).

Circularized padlock probes are very efficiently copied by RCA (29,30). A padlock probe contains two target specific ends of 20 nucleotides (nt) each and a 50 nt long linking sequence in between. When hybridizing to a target molecule the ends of the probe are brought in close proximity forming a nick that is sealed by DNA ligase. The DNA ligase is able to discriminate between single mismatches (31,32) ensuring in that the linear padlock probe becomes circular if and only if a correct hybridization has occurred. Using the very processive phi 29 DNA polymerase (33) in the RCA reaction, 90 nt long circularized padlock probes are copied about 1000 times during one hour of incubation. RCA is a linear amplification process and has proven insufficient to report reacted padlock probes for genotyping human genomes in solution, and monitored with standard DNA detection systems like gel-electrophoresis or oligonucleotide arrays. A method that increases the sensitivity of the RCA procedure is presented in paper II, and a system to monitor the RCA in real-time is presented in paper I.

Another solution to the sensitivity problem is to use a more sensitive analysis instrument such as epifluorescence microscopy. Using this instrument, padlock probes can be combined with RCA, permitting detection of DNA molecules *in situ* (34). Also, circular DNA molecules created in proximity ligation reactions (35,36) *in situ*, are suitable for amplification by RCA (Söderberg *et al.* in preparation).

By including a second primer in the reaction, so called hyperbranched rolling circle amplification, HRCA, can be obtained (37). The second primer, complementary to the rolling circle product, primes an extension that subsequently displaces other downstream extension products, resulting in a displaced single stranded product complementary to the first primer (Figure 3). This cascade-process amplifies the circles much faster than traditional RCA and is sensitive enough to detect single-copy genes in the human genome (37) and to be used for genotyping (38-40). However, amplification of a multitude of reacted padlock probes using HRCA has never been reported, which is probably due to high cross-reactivity between reacted and unreacted padlock probes.

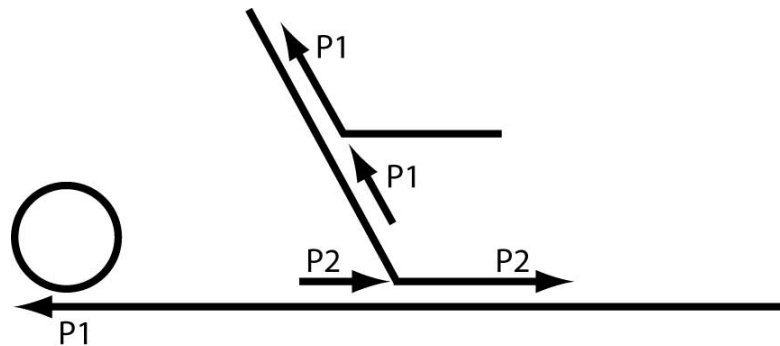


Figure 3. HRCA. A DNA circle and two primers, P1 and P2, are added together. Primer 1 (P1) is complementary to the circle and primes the RCA reaction. A single stranded product of repeated copies of the circle is formed and templates extension of primer 2 (P2). This extension product is displaced by another P2, resulting in a single stranded product suitable for hybridization and extension of P1.

Whole genome amplification

The amount of DNA accessible from patients may become a limiting source in large scale genetic studies. Therefore, much effort has been invested in developing methods for whole genome amplification (WGA) that nonspecifically amplifies whole genomes. Most of these methods relate to one or the other of the two groups of amplification techniques mentioned above (PCR- or RCA techniques).

PCR-based whole genome amplification methods that are based on addition of random (41) or partially degenerate primers (42) have been developed. However, these techniques suffer from biased representation of the amplified source DNA (43). The HRCA method has recently led to a more popular whole genome amplification technique known as Multiple Displacement Amplification, MDA (43). The isothermal reaction uses random hexamers as primers and also amplifies linear DNA, not just DNA circles. The method yields about 20-30 μg of DNA from 10 copies of intact human genomic DNA obtained from cells or fresh tissues and provides a uniformed representation across the genome. However, it has been reported that the MDA method amplifies degraded DNA inefficiently (44). Therefore, a twist of the MDA method has been developed, called RCA-RCA (45). The principle of RCA-RCA is to digest the DNA sample and circularize DNA fragments to double stranded circles. The circles are then denatured to enable HRCA using random hexamers and phi 29 DNA polymerase. This technique

seems to achieve almost complete genome coverage even when templated by degraded DNA.

Multiplex probe-based DNA analysis methods

Many DNA analyses methods are based on oligonucleotide probes that recognize DNA targets by hybridization. To achieve higher specificity, many of these techniques also rely on one or more enzyme reactions to exploit target recognition by the probes. By amplifying the reacted probes instead of the target molecules, problems associated with multiplex PCR can be avoided.

One of the pioneering methods that utilize a probe-target recognizing step along with a guiding enzyme reaction is the oligonucleotide ligation assay, OLA (31). This method uses two hybridizing probes and a DNA ligase to investigate single nucleotide polymorphism directly on genomic DNA. A multiplexed variant of this method has been developed for profiling splice-variants (46), expression profiling (47), and genotyping (48,49) using the GoldenGate assay (Illumina). The genotyping reactions generate templates that can be amplified by PCR using a universal primer-pair that corresponds to a universal sequence in every GoldenGate probe. Another variant of OLA, the multiplex ligation-dependent probe amplification method, MLPA (MRC Holland), has been used for relative quantification of specific genomic sequences in parallel (50).

A further development of OLA is the padlock probe technique (29,51-53)(previously described), which also has been used for analyzing a multitude of SNPs in parallel in the human genome (54). This concept has recently also been applied for expression analysis (55), microbial detection (56), and SNP analysis in plant research (57). A twist of the padlock probe technique that utilizes a gap-fill polymerization step together with the ligation reaction is the molecular inversion probe (MIP) assay (58) (Parallel Biosciences, recently acquired by Affymetrix). This assay has recently allowed multiplexing levels exceeding 10,000 SNPs (59).

To overcome the problems associated with traditional multiplex PCR, all reacted probes (e.g. padlock probes, MIPs, MLPA probes, and GoldenGate-probes) are amplified using a universal primer pair. However, an obstacle with these multiplex PCR methods is that the total amount of products formed after amplification is approximately the same as the amount of product formed from each single template when amplified individually one by one. This is probably due to product inhibition, caused by high concentration of total amount of product. The general primer-pair motif in every target hybridizes to each other and competes with primer annealing (Figure 4). As a consequence, all individual reactions stop being amplified at approximately the same cycle, reducing the dynamic range, and thus preventing simultaneous analysis of high-and low-abundant targets. Yeakley *et al* (46)

addressed this problem by separating the pool of probes in one high-abundant and one low-abundant reaction, prior to PCR amplification.

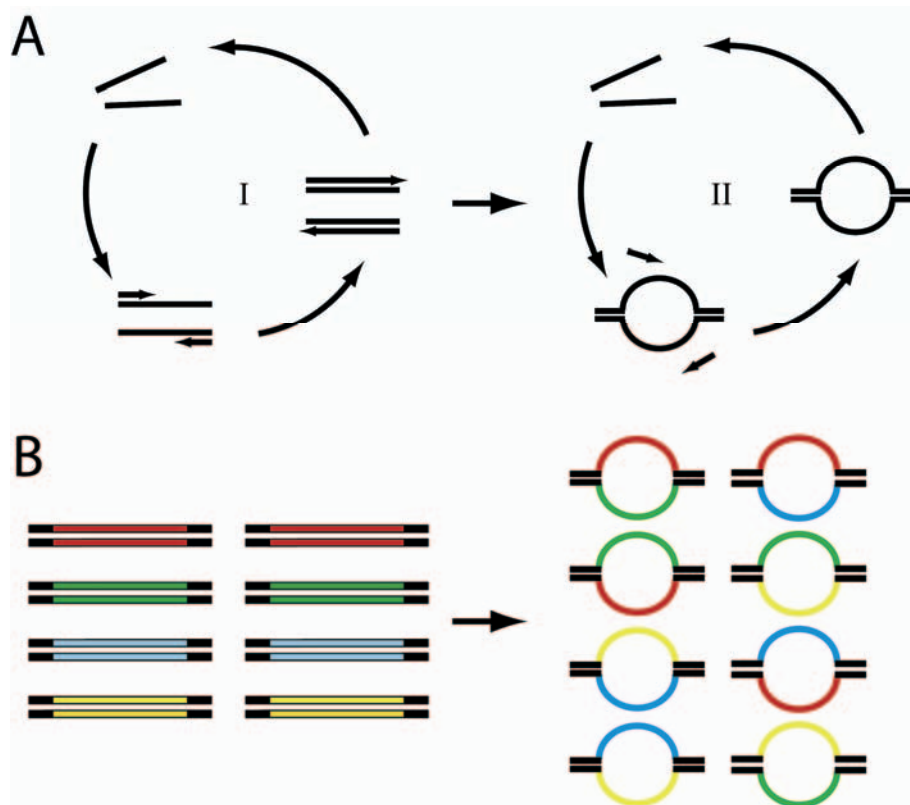


Figure 4. Product inhibition in multiplex PCR. (A) With optimal conditions, the PCR copies targets exponentially (I) to a certain concentration. At high concentration of targets, annealing of primers is competing with re-hybridization of the primer-pair motif in every denaturated target, and the reaction reaches a plateau (II). (B) illustrates the products formed in the beginning (left) and in the end (right) of a multiplex PCR.

A common feature of padlock probe, MIP, and GoldenGate assays, are that all probes in these assays contain unique tag sequences that are used to capture the corresponding amplification products by hybridization to complementary tag-microarrays. The use of generic microarrays with tag sequences was first applied for PCR-based analysis of bacterial genomes and expressed yeast sequences (60,61). The concept was later combined with the ligase detection reaction, LDR, (62) for genotyping (63) and for detection of small

insertions and deletions (64). For the padlock probe assay developed by Banér *et al.*(54), a microarray of immobilized, pre-synthesized, tag sequences (65) was used to capture the PCR product. For the MIP assay, a microarray of *in situ* synthesized tag sequences (19,20)(GeneChip, Affymetrix) is used, and for the GoldenGate assay the reaction products are captured on tag sequences that are chemically coupled on beads attached to arrays (BeadArrays, Illumina). The beads are randomly ordered on the arrays and are decoded using a combinatorial hybridization scheme (66). In paper II, a similar microarray as the one used in Baner *et al.* was employed to analyze the C2C-amplified product. The use of microarrays along with standard tag sequences allows a flexible design for any of these assays and simplifies manufacturing of DNA arrays.

For the described probe-based genotyping procedures, the concept of including a general amplification sequence in each target, overcomes the problems associated with traditional multiplex PCR. This idea together with enzyme directed probe processing has further been established in the selector technology described in papers III and IV, bringing the probe-based concept beyond analyzing SNPs to analyzing multiple selected segments of DNA in parallel.

Present Investigations

Paper I. Real-time monitoring of rolling-circle amplification using a modified molecular beacon design.

In this paper we describe a procedure to monitor rolling circle amplification (RCA) in real time using so called molecular beacons (67). We demonstrate that the procedure can be used for precise measurements of amplified circles, in this case circularized padlock probes, and we describe how the molecular beacon design can be adopted for RCA detection.

Molecular beacons are hairpin-shaped hybridization probes that contain a fluorophore at one end and a quencher at the other. The fluorophore and the quencher remain in close proximity as long as the molecular beacon adopts a stem-loop structure. In this configuration, the fluorescence is transformed to heat by the nearby quencher (by fluorescence resonance energy transfer, FRET) and no emitted light can be detected. However, upon hybridization to a target molecule the stem is denatured, removing the fluorophore from the quencher, allowing fluorescence emission. When molecular beacons, designed for real-time PCR, were added together with the rolling circle reaction mix, we found a template-independent signal. This signal was caused by the 3'-5' exonucleolytic activity of the added phi29 DNA polymerase (68) that degraded the molecular beacons and therefore separated the fluorophore and the quencher. To overcome this problem we protected the molecular beacons from degradation by using nuclease resistance 2'-O-Me RNA nucleotides at the 3'-ends. After solving that problem, we next found that the protected beacons gave rise to unexpectedly low signals. This was shown to be caused by inter-molecular hybridization between neighboring molecular beacons; one molecular beacon's fluorophore was quenched by another molecular beacon's quencher. To solve this problem we designed the padlock probe such that one of the molecular beacon's stem-sequences hybridized to the RCA product (Figure 5).

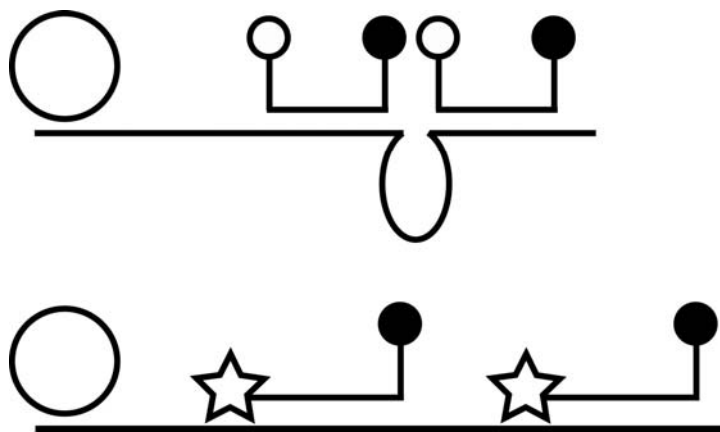


Figure 5. Modified molecular beacon design. The upper picture represents the inter-molecular hybridization between neighboring molecular beacons and results in lack of fluorescence signal. The lower illustration presents the modified design where one of the stem-sequences of the molecular beacon hybridizes to the RCA product, resulting in emitted fluorescence.

After we modified the molecular beacon design, we used the probes to show how to quantify amplification of circularized padlock probes in real-time. By adding the modified molecular beacons in the RCA reaction we could very precisely calculate the starting amount of circles. The precision of the assay showed an average CV of 2.7% calculated on number of start circles.

Perspectives on Paper I.

The paper presents a convenient means to monitor the quality and amount of padlock probes. However, we have seen that some padlock probes do not give rise to a signal upon molecular beacon hybridization even though the RCA generates a lot of amplified material (measured by gel electrophoresis). The reason for this can be secondary structures in the RCA product that hinder the molecular beacon hybridization. A possible solution to this problem is to use so called peptide nucleic acid beacons that have a unique DNA duplex-invading ability (69). Compared to standard real-time PCR (70) the presented assay has a very high precision, but the poor sensitivity and dynamic range limits the method to very few applications for DNA analysis. In order to increase the sensitivity of the assay in solution to levels comparable to PCR, we developed the circle-to-circle amplification method that is presented in paper II.

Paper II. Circle-to-circle amplification for precise and sensitive DNA analysis.

In paper II we present a novel method to amplify circularized padlock probes, called circle-to-circle amplification, C2CA. We describe the method in detail and demonstrate its utility for multiplex genotyping of human genomic DNA and for quantitative measurements of DNA molecules. The procedure is not product inhibited, and can yield around 100-fold higher concentrations of amplification product compared to PCR.

The process starts with a rolling-circle-amplification step that generates a single-stranded product composed of repeated copies of the amplified circle. The next step is to digest the single stranded product by adding a short replication oligonucleotide (RO) complementary to a region in each monomer of the RCA-product. This partially double stranded product contains the restriction sequence for a restriction enzyme, added to the reaction. After heat inactivation of the restriction enzyme and dissociation of the digested RO fragments from the digested RCA-product, a remaining intact RO hybridizes to one end of each monomer. Upon addition of DNA ligase, the RO guides circularization of the monomers to DNA circles. The newly formed circles are now suitable for new rounds of rolling-circle-amplification, primed by the RO. The RO is the key player in the C2CA process; it is involved in every step of the reaction and must be added in concentration that assures saturation of complementary monomer copies in the RCA-product as well as any previously added RO of opposite polarity.

To investigate the circle-to-circle conversion rate in the procedure we radiolabeled the 1st and 2nd generation of the C2CA products. Both products were monomerized and circularized and the conversion rate was analyzed by gel electrophoresis. For both generations the conversion from RCA product to monomer circles was close to 100%. Compared to the hyperbranched rolling-circle-amplification (HRCA) that produces a partly double stranded rolling-circle-product, C2CA produces a single stranded product with a polarity that alternates between generations. To prove that this is the case, we demonstrated that no signal could be obtained when a molecular beacon of the wrong polarity was added to the RCA step in each generation of the C2CA. Furthermore, using a double strand intercalating dye, the amount of double stranded products from a true HRCA (100fmol start circles) and a three generation C2CA (100zmol start circles) was measured (Figure 6). The temperature gradient analysis show that the HRCA produces a high molecular weight double stranded DNA product after one hour of amplification, that re-associates at temperatures below around 85°C. In contrast, the third generation C2CA product showed just a slight increase in signal as temperature decreased, indicating a minor formation of a secondary structured RCA product.

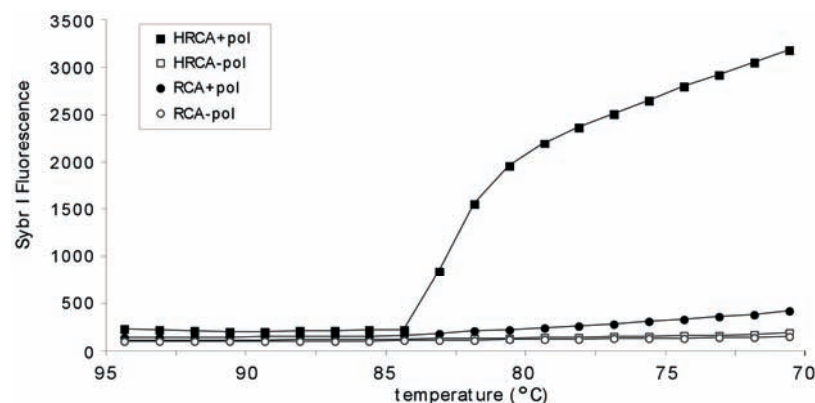


Figure 6. Thermal profile of amplification products from a HRCA and a third generation C2CA in the presence of a double stranded intercalating dye (Sybr 1).

The amplification rate was calculated by comparing a first generation C2CA to a million-fold diluted third-generation C2CA, analyzed in real-time using molecular beacons. The analysis showed a similar start amount of circles for both reactions indicated that the phi29 DNA polymerase copies DNA circles at a rate of 1.6×10^3 nt/min, resulting in a billion-fold amplification of a 93nt circle after three generations of C2CA with one out of RCA in each generation. Because only one polarity of DNA is produced in each generation, the product inhibition problem seen in PCR is avoided. We compared the amount of product after amplifying 500 zmol of DNA circles either by three generations of C2CA or a corresponding 27 cycles of PCR. At least $15\mu\text{M}$ of monomer RCA product was formed in the C2CA compared to 150nM of product formed by the 27 cycled PCR.

To investigate whether C2CA could be used for genotyping, the same set of 26 padlock probes used in Banér *et al.* (54) were reacted with nine human DNA samples. The results from the nine samples were in good concordance with the results obtained from PCR amplified padlock probes, demonstrating that the C2CA can be used for multiplex genotyping.

To determine the precision and sensitivity of the method; we used padlock probes complementary to mitochondrial DNA (mtDNA). Circularized padlock probes were subjected to three generations of C2CA together with a reference circle, and monitored in real-time by using two product specific molecular beacons, as described in Paper I. The sensitivity of the method using pre-ligated circles was around 10 zmol and the precision of the method using a dilution series of total cellular DNA as template for the padlock probe was calculated to an average CV of 10%.

Perspectives on Paper II.

The presented method introduces an alternative approach for amplifying circular molecules. The advantages compared to PCR are that C2CA is more precise and can produce more DNA. Molecular inversion probes can be applied for genotyping up to 10.000 SNPs per reaction (59). How far the multiplex level can be increased is unknown but will depend on factors like the amount of material to be analyzed and the sensitivity of the analysis instrument. Nevertheless, the product inhibition of the PCR will set the limit for how far you can press the levels of analyzing many DNA molecules simultaneously. Because C2CA can produce more DNA than PCR, the signal in highly multiplex assays can be enhanced and the dynamic range of the assay can be extended.

It is important to remove excess unreacted linear padlock probes regardless of whether PCR or C2CA is used. To obtain fast and efficient ligation detection, relatively high concentrations of padlock probes are needed. The problem is that high concentrations of remaining linear probes give rise to a high background and false positive signals when C2CA or PCR is used. Using C2CA, remaining linear probes can hybridize to the rolling-circle product and be incorporated in the reaction, resulting in an uncontrolled HRCA-like reaction (Figure 7). The most effective procedure so far of removing the linear probes is through exonuclease treatment, but the method still suffer from problems with remaining probes. Therefore, lower concentrations of probes than optimum must be used in the ligation reaction, resulting in lower ligation efficiency.

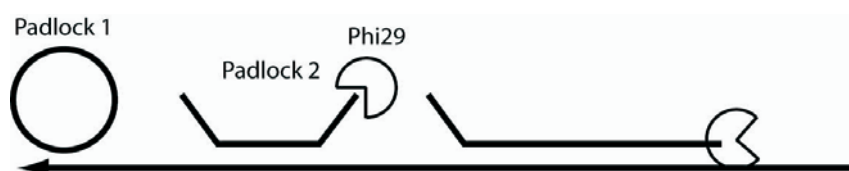


Figure 7. Padlock probes used in multiplex contain a universal region for amplification and detection. This region can hybridize to complementary regions of RCA products generated from other probes. The added Phi 29 DNA polymerase contains a 3'-5' single stranded exonucleolytic activity and will start to degrade the free 3'-end of the partly hybridized padlock. When the polymerase reaches a double stranded region it switches to polymerase activity resulting in amplification also of remaining linear unreacted probes.

The disadvantage of the C2CA method compared to PCR is the large numbers of steps needed to obtain the same amplification level. A total number of nine additions of reagents are needed for three generations of C2CA. The

C2CA furthermore consumes more reagents and is therefore potentially more expensive than PCR. To be able to reduce the number of generations needed for analysis of a C2CA product, a novel and more sensitive method to monitor single rolling circle products in a homogeneous detection system has been reported by Blab *et al.*(71), and is currently further miniaturized in a microfluidic system (Jarvius and Melin *et al.* in preparation).

Paper III. Multiplex amplification enabled by selective circularization of large sets of genomic DNA fragments.

In this paper we present a technique that enables multiplex PCR of specific DNA targets without producing amplification artefacts. The method introduces a novel tool, the so called selector that guides circularization of selected targets and introduces a universal primer-pair in the selected and circularized sequences. We demonstrate the procedure by performing a specific and robust 96-plex PCR without producing any amplification artefacts associated with traditional multiplex PCR.

The selector contains two oligonucleotides, one selector probe and one general vector oligonucleotide. The selector probe has two target-complementary end-sequences that are linked by a general sequence motif, which is base-paired to the general vector sequence. In the selection process (Figure 8), a pool of selectors is combined with denatured restriction digested DNA. Each selector probe hybridizes to its respective target together with the vector oligonucleotide, forming individual circular complexes that are covalently closed by DNA ligase. This process can be performed in two different ways, indicated in the figure as A and B. The general sequence that is introduced into the circularized fragments then allows multiplex PCR using a universal primer-pair.

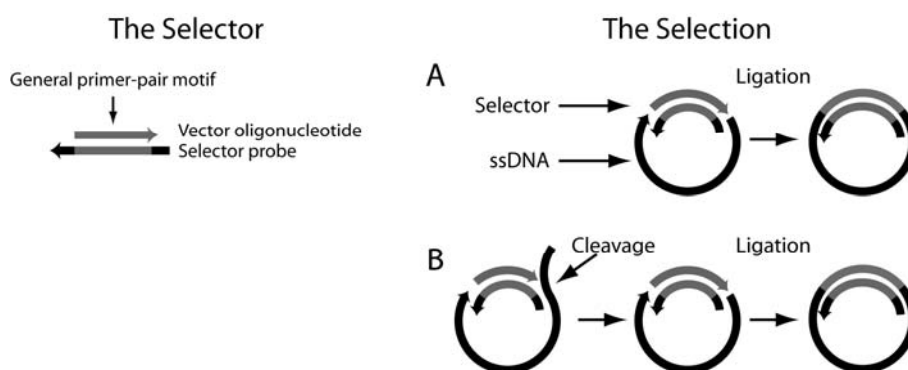


Figure 8. The selector and the selection process. The black parts of the selector probe represent target specific ends and the grey part represent the general primer-pair motif. The selection process can be carried on using two different approaches. (A) Both ends of the selected fragment connect to the vector oligonucleotide, or (B) the vector oligonucleotide forms a branched structure in an optional position at the 5'-end of the fragment. This structure is recognized by the added endonucleolytic enzyme (72-74), forming ends suitable for ligation as in (A).

To demonstrate the performance of the method we selected an unbiased set of 96 genomic fragments for multiplex PCR amplification, and then analyzed the PCR product by gel electrophoresis and by hybridization to a cDNA array. The first step of the experiment was to design the 96 selector probes needed for the selection process. We choose targets in the genome that corresponded to cDNA sequences at 96 positions on the array in the pattern of UU (as in Uppsala University). We developed a computer program, described in paper IV, which identify combinations of restriction enzymes that generate suitable fragments for the targets of interest. In this case we digested the human genome using in total four different restriction enzymes that generated fragments for selection that contained the targets of interest. We reproduced the 96-plex selection and PCR using five different DNA samples and hybridized the Cy-3 labeled PCR product on five separate cDNA arrays. The results showed reproducible signals from 89% of the selected sequences, and the gel analysis provide no evidence of amplification artifacts.

Perspectives on Paper III.

As discussed, one important problem with traditional multiplex PCR, when many specific primer-pairs are added in one reaction, is that amplification artifacts arise. This problem has been known for many years and different approaches have been developed to get around the problem, but so far none that are entirely satisfactory. The selector technique that is presented in paper III and IV demonstrates a solution to this problem.

The selector technique should be of great value for a wide range of PCR based applications, particular in combination with highly parallel DNA analysis platforms. One class of parallel DNA analysis is large-scale sequencing and resequencing platforms (75), like sequencing by hybridization (76,77), sequencing by ligation (78), or sequencing by synthesis (79-81) systems. Currently, much effort is being invested in developing low cost sequencing methods, with the goal of sequencing an entire human genome for US \$1,000. A more realistic goal in the shorter term is to sequence only the interesting parts of the genome, for example all exons. We believe that the selectors can be combined with several of the sequencing techniques to carry out the "US \$1,000 interesting genome".

The selector technology also highlights promising properties to be combined with PCR-intense genotyping methods (10,82), like minisequencing (65,83) and primer extension-based methods in concert with mass spectrometry analysis (84). Furthermore, the selectors could be combined with mutation detection techniques that require many single PCRs, (85,86) to increase throughput.

To select and amplify the 300.000 exons in the human genome around 0.5 million selectors would be needed to be synthesized, depending on the se-

quencing platform used. This production cost can be reduced by synthesizing the selector probes in parallel using array based synthesis (87-89). The quality of the oligonucleotides and the amount of material produced from the array synthesis are not fully evaluated. However, a potential way to enrich for the correctly synthesized probes are to perform enzymatic synthesis of the array synthesized pool of probes (F. Dahl unpublished results) (Figure 9) before use in the selection reaction.

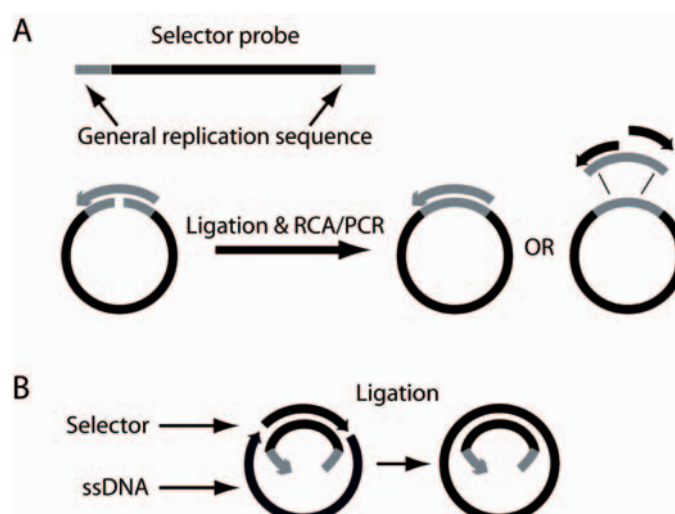


Figure 9. A) A general replication sequence is added to every selector probe (gray). By adding this sequence to the probes, a general amplification step can be performed using either PCR or RCA. Only the probes that contain the general replication sequence on both ends can be circularized and amplified, resulting in an enrichment of full-length probes. B) The general end-sequences in the enzymatic synthesized probes do not participate in the selection procedure.

Recent application developments

Another interesting application of the selector technology is to apply it for relative copy number measurements of specific DNA targets. Schouten *et al.* (50) described the multiplexed ligation-dependent probe amplification method (MLPA), for detecting copy number changes in specific chromosomal sequences, as for example chromosome deletions and duplications. The procedure is based on hybridizing target specific probe-pairs to interesting regions of genomic DNA. Each probe-pair includes a universal primer-pair motif, and one of the two probes has a variable length. If the target sequence

is present, the two probes hybridize adjacent to each other forming a substrate for the added DNA ligase. The quantitative analysis is based on separating a set of PCR-amplified size-tagged MLPA probes by gel electrophoresis. The amount of each ligated probe-pair will be proportional to the corresponding target copy number, and after PCR amplification the relative peak areas obtained from the gel electrophoresis analysis can be quantified. The method has recently been used for gene dosage determination in genes such as BRCA1 (90,91), MLH1 and MSH2 (92), prenatal detection of aneuploidy such as Down's syndrome (93), and for detection of CpG methylation (94). A problem with this method is that long oligonucleotides are needed to prepare a standard 40-plex assay. Because chemically synthesized probes longer than 100 nt are difficult to produce, a time-consuming cloning process is used to produce the long probes (95).

Instead of using MLPA probes, differently sized PCR products can be generated from selector reactions and quantified using the same type of gel electrophoresis instrument. This process is called multiplex ligation-dependent genome amplification, MLGA (Isaksson *et al.* in preparation), and is described in Figure 10.

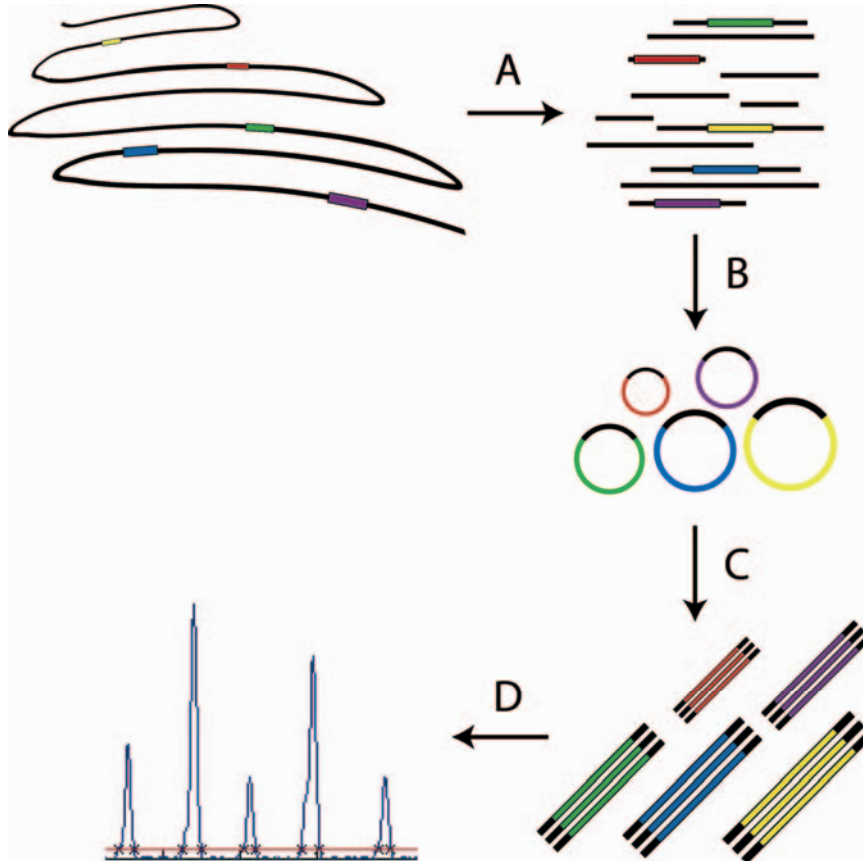


Figure 10. (A) Genomic DNA is restriction digested to fragments with defined ends. (B) Selectors are added to select the interesting targets and circularize each target into unique sizes. (C) A multiplex PCR are carried out followed by separation and quantification of the PCR product using gel electrophoresis (D).

The quantification can be performed using gel electrophoresis based instruments such as ABI's 3700 capillary electrophoresis system (as exemplified in Figure 11, Isaksson *et al.* in preparation) or Agilent's 2100 Bioanalyzer. Each set of selectors are chemically synthesized and are of uniform length, but they are designed to generate PCR-products of different sizes through the selection reaction described in figure 10. This quality circumvents the problem of obtaining long oligonucleotides as in the MLPA approach, and results in an inexpensive and convenient process to design custom MLGA assays.

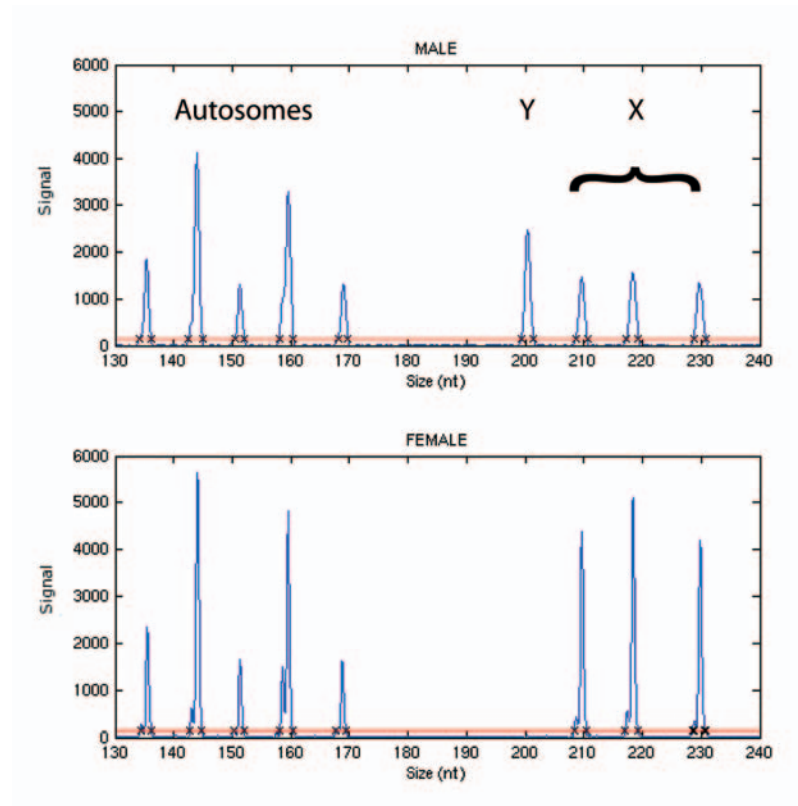


Figure 11. The figure presents data for 9-plex MLGA assays of male and female DNA samples. Five selectors analyze autosomal sequences (five peaks between 130 and 170 nt in size), one the Y-chromosome (200 nt) and three corresponds to the X-chromosome (210-230 nt). Relative quantification is carried out by measuring each peak area.

Paper IV. PieceMaker: selection of DNA fragments for selector-guided multiplex amplification.

A laborious task with the selector technology, described in paper III, is to design sets of selector probes for multiplex DNA analysis applications. This paper presents PieceMaker, a computer program that finds an optimal combination of restriction enzymes that generates suitable fragments for a selection reaction, and presents the selected sequences required for probe design.

Different selector applications usually have different design criteria. For example, selectors used for minisequencing must select only fragments that contain the SNP and a defined adjacent region, while selectors used for MLGA (see paper III) allows selection of all fragments that contain any sequence of interest. Therefore, functions to regulate different design parameters are included in the PieceMaker program to match the application. Target sequences are restriction digested *in silico*, and the resulted restriction fragments are evaluated based on the design parameter settings; selected fragment length, minimum content of sequence of interest (target sequence), maximum flap length (length of any sub-fragment that is removed), and nucleotide at cleavage position. These parameters are visualized in figure 12.

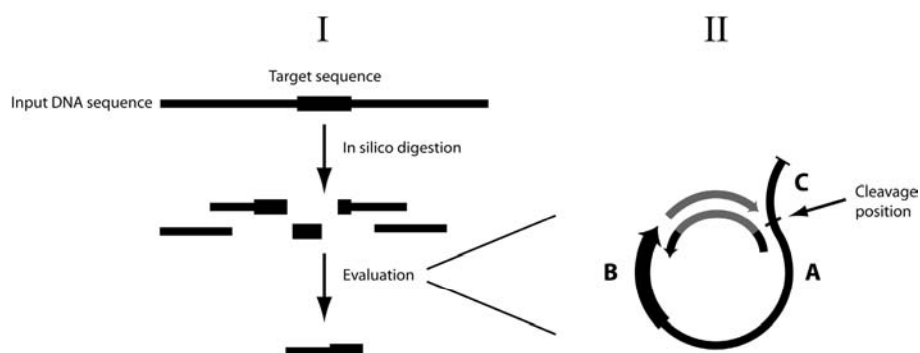


Figure 12. Illustration of selector probe design. Panel I describes the selection process from input DNA sequence to selected fragment. Panel II describes the design parameters used in the evaluation of the digested fragments. (A) Represents the length of the selected fragment from the cleavage position at the 5'-end to the 3'-end, (B) represents the interested target sequence, and (C) represents the flap, which will be removed before circularization.

As a result of the evaluation of each fragment, a set of accepted fragments is generated for each combination of input DNA sequence and restriction enzymes. The next step in the design process is to select the best combination of restriction enzymes, i.e. the combination that maximizes the number of

input sequences for which there are accepted fragments. The sequences of the selected fragments can then be used for designing the selector probe sequences.

Each application has its own design requirements. To investigate the impact of parameter choice on success rate, designs with different criteria were carried out for two sets of target sequences. As expected, the design success rate increases by allowing longer flaps, a wider range of fragment length, and when more digestion reactions are used in parallel. We have also tried to establish some design rules experimentally, where we have successfully circularized fragments up to 1000 nucleotides in length and used flap lengths up to 1000 nucleotides.

Perspectives on Paper IV

Designing selector probes for a particular application has three general limitations. First, the actual sequence of the input targets will set the limit for finding suitable fragments for a given set of restriction enzymes and parameter settings. For example, in paper IV, design was unsuccessful for one target because it was located in a region with highly repeated A/T-sequences, resulted in that no suitable fragment for this target could be found. Secondly, the reaction mechanism of the selector technology has limitations that have to be considered when designing the probes. Very short fragments (less than 100 nt) will be difficult to circularize because the rigidity of the double stranded part of the selector probe-DNA target hybridization, and very long fragments will have low circularization efficiency because of long distance between the ends of the target. Finally, the particular application has limitations that must be considered when designing the selector probes. For example, the MLGA approach must select fragments within a certain size-range because limitations of resolution of the gel electrophoresis-based instrument used.

So far, only a few sets of selector probes have been designed and experimentally tested. The design rules are thus slowly expanding, and with more sets being tested the knowledge of how to design the probes will hopefully further improve, and result in a higher success rate of the selector method than the current rate of ~90%.

Future Perspectives

Following the final publication of the complete sequence of the human genome, large numbers of sequences are available for analysis, serving to improve the understanding of genetic disease mechanisms. As a consequence, there are increasingly demands for methods that can perform simultaneous analysis of large numbers of sequences directly in complex genomes.

DNA arrays have had a major impact on throughput in genetic research and are today the method of choice to investigate many different DNA sequences simultaneously. However, for most application involving DNA arrays, processing of the genome is essential prior to analysis. In the field of large scale SNP-analysis, a various number of genome processing procedures has successfully been developed based on DNA complexity reduction schemes (10,21). Also, a number of probe-based methods employ microarrays to encode the identity of recognized target sequences, *e.g.* padlock probes (54), molecular inversion probes (59), and GoldenGate assays (49). A genotyping technique that does not require processing of DNA samples or a probe-based target recognition event before array hybridization has recently been developed by Illumina (96). They demonstrate a highly flexible genome-wide genotyping approach directly on unprocessed whole-genome samples using microarrays combined with a powerful signal amplification procedure carried out on the array. However, as the method needs a large number of DNA copies (10^6), a whole genome amplification procedure is usually needed before hybridization on the microarray. On the other hand, the scalability of the method is limited only by the density of the microarray.

Despite the recent success of simultaneously genotype unprocessed DNA samples, multiplex analysis of longer stretches of sequences has proven difficult, partially because of ineffective DNA processing procedures. One problem is that no method has been available that performs specific amplification of many targets in parallel, and the bottleneck has therefore been that large numbers of separate reactions must be performed whenever many genomic sequences needs to be analyzed. Novel techniques for DNA processing, such as the selector technique, open up for new solutions to also investigate a multitude of specific DNA targets simultaneously. One such solution has recently been published by the company 454 Life Sciences (79). They present a novel platform for parallel large scale DNA sequencing that allows *de novo* sequencing of bacterial genomes in a four hour run with high accuracy. The platform performs an emulsion-based procedure (97) to clone and

amplify DNA fragments *in vitro* on beads, followed by pyrosequencing (98,99) of the fragments on solid support in picolitre-sized wells (100). So far the platform is limited to less-complex genomes, such those of bacteria, but if combined with the selector technology, selected parts of the human genome can simultaneously be sequenced (Dahl *et al.* in preparation). For investigation of many DNA targets in parallel, this platform can be one serious competitor to microarray-based analysis methods.

The selector technology and some of the mentioned multiplexed probe-based methods are limited by the production cost of oligonucleotides, since they require rather long probes. Therefore, the next potentially bottleneck, and the last link in the multiplex analysis chain, is to obtain inexpensive and high quality reagents needed for the parallel amplification techniques. One potential solution to this problem is to produce the oligonucleotides using parallel *in situ*-synthesis methods. This could reduce the cost per oligonucleotide, although the quality and amount of produced probes still needs to be evaluated.

Acknowledgements

This work was performed at the department of Genetics and Pathology, Rudbeck laboratory. I would like to express my sincere gratitude to all those who have contributed in one way or another to this thesis.

In particular I would like to thank:

Mats Nilsson, my supervisor, for your amazing enthusiasm for our research and for all encouraging pep-talks that has kept me motivated during these years. Thanks for great support and for giving me free hands. I have very much enjoyed our discussions on various topics over our own-brewed coffee.

Ulf Landegren, my co-supervisor, for always keeping the door open for encouraging discussions, and for creating, together with Mats, such an inspiring and open environment to work in. I also very much enjoyed your endless source of exciting anecdotes.

All members of the Landegren/Nilsson group, I will never forget how fun it was. Johan S, for a productive and pleasant collaboration on the selectors, Olle, for sharing stimulating thoughts about science, politics and Weissbier, and Henrik, for introducing me to every korpen-team in the city. Sigrun, for teaching me how to feed horses with Sangria in Montpellier, Jonas and Malin J for nice barbecues, and Mats G, for good collaboration on all possible projects. Lena, thanks for keeping it all together in the lab, and for all help with my projects, Magnus, for fruitful collaboration on the MLGA stuff, and Masood, Edith, Jenny, Chatarina, Jonas M, Carolina R, Carolina W, Ola, Sara, Mathias, Kalle, Carlos, Mia, Anna, Anders, and Tim for making the group very pleasant. Thanks also to the former members, Simon, Johan, Maritha, Marek and DOA. Simon, for being an inspiring “business-mentor”, and for helping me getting a postdoc position, and Johan B, for great golf and funny parties, and for teaching me how to behave like Carl XVI Gustaf. Thanks guys for great fun during these years.

All other people at Rudbeck, for creating such a friendly environment to work in.

The administrators at Rudbeck for always being kind and helpful.

My collaborators during the years, Snaevar Sigurdsson at Akademiska Sjukhuset, the people at 454, Sequenom, and Agilent. Thanks also to The Genome Center at Rudbeck for all help with sequencing.

The people at Olink, especially Björn Ekström, for an inspiring and fruitful collaboration around selectors.

The people at Forskarpatent, for teaching me a lot of patent-strategies.

Thanks to all my friends in Uppsala, especially Fredrik Hjelm my former lab-companion, and golf-partner. I would also like to thank my friends from Sundsvall (med omnejd), for unforgettable valborgs-firanden, and other interesting parties, during my years in Uppsala.

My family, Mum, Dad, and Markus, for your constant support and encouragement

And finally, thank you Linda, for your support and patience through all these years. You're simply the best!

References

1. Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, C.A., Hutchison, C.A., Slocombe, P.M. and Smith, M. (1977) Nucliotide sequence of bacteriophage phi X174 DNA. *Nature*, **265**, 687-695.
2. Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, **230**, 1350-1354.
3. Cohen, S.N., Chang, A.C., Boyer, H.W. and Helling, R.B. (1973) Construction of biologically functional bacterial plasmids in vitro. *Proc Natl Acad Sci U S A*, **70**, 3240-3244.
4. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.
5. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A. *et al.* (2001) The sequence of the human genome. *Science*, **291**, 1304-1351.
6. Kainz, P. (2000) The PCR plateau phase - towards an understanding of its limitations. *Biochim Biophys Acta*, **1494**, 23-27.
7. Chamberlain, J.S., Gibbs, R.A., Ranier, J.E., Nguyen, P.N. and Caskey, C.T. (1988) Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res*, **16**, 11141-11156.
8. Shigemori, Y., Mikawa, T., Shibata, T. and Oishi, M. (2005) Multiplex PCR: use of heat-stable *Thermus thermophilus* RecA protein to minimize non-specific PCR products. *Nucl. Acids Res.*, **33**, e126-.
9. Landegren, U. and Nilsson, M. (1997) Locked on target: strategies for future gene diagnostics. *Ann Med*, **29**, 585-590.
10. Syvanen, A.C. (2005) Toward genome-wide SNP genotyping. *Nat Genet*, **37 Suppl**, S5-10.
11. Broude, N.E., Zhang, L., Woodward, K., Englert, D. and Cantor, C.R. (2001) Multiplex allele-specific target amplification based on PCR suppression. *Proc Natl Acad Sci U S A*, **98**, 206-211.

12. Bahl, C.P., Wu, R., Brousseau, R., Sood, A.K., Hsiung, H.M. and Narang, S.A. (1978) Chemical synthesis of versatile adaptors for molecular cloning. *Biochem Biophys Res Commun*, **81**, 695-703.
13. Kinzler, K.W. and Vogelstein, B. (1989) Whole genome PCR: application to the identification of sequences bound by gene regulatory proteins. *Nucleic Acids Res*, **17**, 3645-3653.
14. Ko, M.S., Ko, S.B., Takahashi, N., Nishiguchi, K. and Abe, K. (1990) Unbiased amplification of a highly complex mixture of DNA fragments by 'lone linker'-tagged PCR. *Nucleic Acids Res*, **18**, 4293-4294.
15. Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*, **23**, 4407-4414.
16. Dong, S., Wang, E., Hsie, L., Cao, Y., Chen, X. and Gingeras, T.R. (2001) Flexible use of high-density oligonucleotide arrays for single-nucleotide polymorphism discovery and validation. *Genome Res*, **11**, 1418-1424.
17. Kennedy, G.C., Matsuzaki, H., Dong, S., Liu, W.M., Huang, J., Liu, G., Su, X., Cao, M., Chen, W., Zhang, J. *et al.* (2003) Large-scale genotyping of complex DNA. *Nat Biotechnol*.
18. Matsuzaki, H., Loi, H., Dong, S., Tsai, Y.Y., Fang, J., Law, J., Di, X., Liu, W.M., Yang, G., Liu, G. *et al.* (2004) Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. *Genome Res*, **14**, 414-425.
19. Fodor, S.P., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T. and Solas, D. (1991) Light-directed, spatially addressable parallel chemical synthesis. *Science*, **251**, 767-773.
20. Lipshutz, R.J., Fodor, S.P., Gingeras, T.R. and Lockhart, D.J. (1999) High density synthetic oligonucleotide arrays. *Nat Genet*, **21**, 20-24.
21. Matsuzaki, H., Dong, S., Loi, H., Di, X., Liu, G., Hubbell, E., Law, J., Berntsen, T., Chadha, M., Hui, H. *et al.* (2004) Genotyping over 100,000 SNPs on a pair of oligonucleotide arrays. *Nat Methods*, **1**, 109-111.
22. Callow, M.J., Drmanac, S. and Drmanac, R. (2004) Selective DNA amplification from complex genomes using universal double-sided adaptors. *Nucleic Acids Res*, **32**, E21.
23. Shuber, A.P., Grondin, V.J. and Klinger, K.W. (1995) A simplified procedure for developing multiplex PCRs. *Genome Res*, **5**, 488-493.
24. Brownie, J., Shawcross, S., Theaker, J., Whitcombe, D., Ferrie, R., Newton, C. and Little, S. (1997) The elimination of primer-dimer accumulation in PCR. *Nucleic Acids Res*, **25**, 3235-3241.
25. Shapero, M.H., Zhang, J., Loraine, A., Liu, W., Di, X., Liu, G. and Jones, K.W. (2004) MARA: a novel approach for highly multiplexed locus-specific SNP genotyping using high-density DNA oligonucleotide arrays. *Nucleic Acids Res*, **32**, e181.

26. Gilbert, W. and Dressler, D. (1968) DNA replication: the rolling circle model. *Cold Spring Harb Symp Quant Biol*, **33**, 473-484.
27. Fire, A. and Xu, S.Q. (1995) Rolling replication of short DNA circles. *Proc Natl Acad Sci U S A*, **92**, 4641-4645.
28. Kool, E.T. (1996) Circular oligonucleotides: new concepts in oligonucleotide design. *Annu Rev Biophys Biomol Struct*, **25**, 1-28.
29. Nilsson, M., Malmgren, H., Samiotaki, M., Kwiatkowski, M., Chowdhary, B.P. and Landegren, U. (1994) Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science*, **265**, 2085-2088.
30. Baner, J., Nilsson, M., Mendel-Hartvig, M. and Landegren, U. (1998) Signal amplification of padlock probes by rolling circle replication. *Nucleic Acids Res*, **26**, 5073-5078.
31. Landegren, U., Kaiser, R., Sanders, J. and Hood, L. (1988) A ligase-mediated gene detection technique. *Science*, **241**, 1077-1080.
32. Luo, J., Bergstrom, D.E. and Barany, F. (1996) Improving the fidelity of *Thermus thermophilus* DNA ligase. *Nucleic Acids Res*, **24**, 3071-3078.
33. Blanco, L., Bernad, A., Lazaro, J.M., Martin, G., Garmendia, C. and Salas, M. (1989) Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. *J Biol Chem*, **264**, 8935-8940.
34. Larsson, C., Koch, J., Nygren, A., Janssen, G., Raap, A.K., Landegren, U. and Nilsson, M. (2004) In situ genotyping individual DNA molecules by target-primed rolling-circle amplification of padlock probes. *Nat Methods*, **1**, 227-232.
35. Fredriksson, S., Gullberg, M., Jarvius, J., Olsson, C., Pietras, K., Gustafsdottir, S.M., Ostman, A. and Landegren, U. (2002) Protein detection using proximity-dependent DNA ligation assays. *Nat Biotechnol*, **20**, 473-477.
36. Gullberg, M., Gustafsdottir, S.M., Schallmeiner, E., Jarvius, J., Bjarnegard, M., Betsholtz, C., Landegren, U. and Fredriksson, S. (2004) Cytokine detection by antibody-based proximity ligation. *Proc Natl Acad Sci U S A*, **101**, 8420-8424.
37. Lizardi, P.M., Huang, X., Zhu, Z., Bray-Ward, P., Thomas, D.C. and Ward, D.C. (1998) Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nat Genet*, **19**, 225-232.
38. Faruqi, A.F., Hosono, S., Driscoll, M.D., Dean, F.B., Alsmadi, O., Bandaru, R., Kumar, G., Grimwade, B., Zong, Q., Sun, Z. *et al.* (2001) High-throughput genotyping of single nucleotide polymorphisms with rolling circle amplification. *BMC Genomics*, **2**, 4.
39. Pickering, J., Bamford, A., Godbole, V., Briggs, J., Scozzafava, G., Roe, P., Wheeler, C., Ghouze, F. and Cuss, S. (2002) Integration of DNA ligation and rolling circle amplification for the homogeneous, end-point detection of single nucleotide polymorphisms. *Nucleic Acids Res*, **30**, e60.

40. Alsmadi, O.A., Bornarth, C.J., Song, W., Wisniewski, M., Du, J., Brockman, J.P., Faruqi, A.F., Hosono, S., Sun, Z., Du, Y. *et al.* (2003) High accuracy genotyping directly from genomic DNA using a rolling circle amplification based assay. *BMC Genomics*, **4**, 21.
41. Zhang, L., Cui, X., Schmitt, K., Hubert, R., Navidi, W. and Arnheim, N. (1992) Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A*, **89**, 5847-5851.
42. Telenius, H., Carter, N.P., Bebb, C.E., Nordenskjold, M., Ponder, B.A. and Tunnacliffe, A. (1992) Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics*, **13**, 718-725.
43. Dean, F.B., Hosono, S., Fang, L., Wu, X., Faruqi, A.F., Bray-Ward, P., Sun, Z., Zong, Q., Du, Y., Du, J. *et al.* (2002) Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci U S A*, **99**, 5261-5266.
44. Lage, J.M., Leamon, J.H., Pejovic, T., Hamann, S., Lacey, M., Dillon, D., Segraves, R., Vossbrinck, B., Gonzalez, A., Pinkel, D. *et al.* (2003) Whole genome analysis of genetic alterations in small DNA samples using hyperbranched strand displacement amplification and array-CGH. *Genome Res*, **13**, 294-307.
45. Wang, G., Maher, E., Brennan, C., Chin, L., Leo, C., Kaur, M., Zhu, P., Rook, M., Wolfe, J.L. and Makrigiorgos, G.M. (2004) DNA amplification method tolerant to sample degradation. *Genome Res*, **14**, 2357-2366.
46. Yeakley, J.M., Fan, J.B., Doucet, D., Luo, L., Wickham, E., Ye, Z., Chee, M.S. and Fu, X.D. (2002) Profiling alternative splicing on fiber-optic arrays. *Nat Biotechnol*, **20**, 353-358.
47. Fan, J.B., Yeakley, J.M., Bibikova, M., Chudin, E., Wickham, E., Chen, J., Doucet, D., Rigault, P., Zhang, B., Shen, R. *et al.* (2004) A versatile assay for high-throughput gene expression profiling on universal array matrices. *Genome Res*, **14**, 878-885.
48. Oliphant, A., Barker, D.L., Stuelpnagel, J.R. and Chee, M.S. (2002) BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques*, **Suppl**, 56-58, 60-51.
49. Fan, J.B., Oliphant, A., Shen, R., Kermani, B.G., Garcia, F., Gunderson, K.L., Hansen, M., Steemers, F., Butler, S.L., Deloukas, P. *et al.* (2003) Highly parallel SNP genotyping. *Cold Spring Harb Symp Quant Biol*, **68**, 69-78.
50. Schouten, J.P., McElgunn, C.J., Waaijer, R., Zwijnenburg, D., Diepvens, F. and Pals, G. (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res*, **30**, e57.
51. Nilsson, M., Krejci, K., Koch, J., Kwiatkowski, M., Gustavsson, P. and Landegren, U. (1997) Padlock probes reveal single-nucleotide differences, parent of origin and in situ distribution of centromeric

- sequences in human chromosomes 13 and 21. *Nat Genet*, **16**, 252-255.
52. Nilsson, M., Baner, J., Mendel-Hartvig, M., Dahl, F., Antson, D.O., Gullberg, M. and Landegren, U. (2002) Making ends meet in genetic analysis using padlock probes. *Hum Mutat*, **19**, 410-415.
 53. Baner, J., Nilsson, M., Isaksson, A., Mendel-Hartvig, M., Antson, D.O. and Landegren, U. (2001) More keys to padlock probes: mechanisms for high-throughput nucleic acid analysis. *Curr Opin Biotechnol*, **12**, 11-15.
 54. Baner, J., Isaksson, A., Waldenstrom, E., Jarvius, J., Landegren, U. and Nilsson, M. (2003) Parallel gene analysis with allele-specific padlock probes and tag microarrays. *Nucleic Acids Res*, **31**, e103.
 55. Baner, J., Marits, P., Nilsson, M., Winqvist, O. and Landegren, U. (2005) Analysis of T-cell receptor V beta gene repertoires after immune stimulation and in malignancy by use of padlock probes and microarrays. *Clin Chem*, **51**, 768-775.
 56. Szemes, M., Bonants, P., de Weerd, M., Baner, J., Landegren, U. and Schoen, C.D. (2005) Diagnostic application of padlock probes--multiplex detection of plant pathogens using universal microarrays. *Nucleic Acids Res*, **33**, e70.
 57. van Eijk, M.J., Broekhof, J.L., van der Poel, H.J., Hogers, R.C., Schneiders, H., Kamerbeek, J., Verstege, E., van Aart, J.W., Geerlings, H., Buntjer, J.B. *et al.* (2004) SNPWave: a flexible multiplexed SNP genotyping technology. *Nucleic Acids Res*, **32**, e47.
 58. Hardenbol, P., Baner, J., Jain, M., Nilsson, M., Namsaraev, E.A., Karlin-Neumann, G.A., Fakhrai-Rad, H., Ronaghi, M., Willis, T.D., Landegren, U. *et al.* (2003) Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat Biotechnol*.
 59. Hardenbol, P., Yu, F., Belmont, J., Mackenzie, J., Bruckner, C., Brundage, T., Boudreau, A., Chow, S., Eberle, J., Erbilgin, A. *et al.* (2005) Highly multiplexed molecular inversion probe genotyping: Over 10,000 targeted SNPs genotyped in a single tube assay. *Genome Res*, **15**, 269-275.
 60. Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E. and Holden, D.W. (1995) Simultaneous identification of bacterial virulence genes by negative selection. *Science*, **269**, 400-403.
 61. Shoemaker, D.D., Lashkari, D.A., Morris, D., Mittmann, M. and Davis, R.W. (1996) Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat Genet*, **14**, 450-456.
 62. Barany, F. (1991) Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc Natl Acad Sci U S A*, **88**, 189-193.
 63. Gerry, N.P., Witowski, N.E., Day, J., Hammer, R.P., Barany, G. and Barany, F. (1999) Universal DNA microarray method for multiplex detection of low abundance point mutations. *J Mol Biol*, **292**, 251-262.

64. Favis, R., Day, J.P., Gerry, N.P., Phelan, C., Narod, S. and Barany, F. (2000) Universal DNA array detection of small insertions and deletions in BRCA1 and BRCA2. *Nat Biotechnol*, **18**, 561-564.
65. Pastinen, T., Raitio, M., Lindroos, K., Tainola, P., Peltonen, L. and Syvanen, A.C. (2000) A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. *Genome Res*, **10**, 1031-1042.
66. Gunderson, K.L., Kruglyak, S., Graige, M.S., Garcia, F., Kermani, B.G., Zhao, C., Che, D., Dickinson, T., Wickham, E., Bierle, J. *et al.* (2004) Decoding randomly ordered DNA arrays. *Genome Res*, **14**, 870-877.
67. Tyagi, S. and Kramer, F.R. (1996) Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol*, **14**, 303-308.
68. de Vega, M., Lazaro, J.M., Salas, M. and Blanco, L. (1998) Mutational analysis of phi29 DNA polymerase residues acting as ssDNA ligands for 3'-5' exonucleolysis. *J Mol Biol*, **279**, 807-822.
69. Smolina, I.V., Demidov, V.V., Cantor, C.R. and Broude, N.E. (2004) Real-time monitoring of branched rolling-circle DNA amplification with peptide nucleic acid beacon. *Anal Biochem*, **335**, 326-329.
70. Heid, C.A., Stevens, J., Livak, K.J. and Williams, P.M. (1996) Real time quantitative PCR. *Genome Res*, **6**, 986-994.
71. Blab, G.A., Schmidt, T. and Nilsson, M. (2004) Homogeneous detection of single rolling circle replication products. *Anal Chem*, **76**, 495-498.
72. Holland, P.M., Abramson, R.D., Watson, R. and Gelfand, D.H. (1991) Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci U S A*, **88**, 7276-7280.
73. Lyamichev, V., Brow, M.A. and Dahlberg, J.E. (1993) Structure-specific endonucleolytic cleavage of nucleic acids by eubacterial DNA polymerases. *Science*, **260**, 778-783.
74. Lyamichev, V., Brow, M.A., Varvel, V.E. and Dahlberg, J.E. (1999) Comparison of the 5' nuclease activities of taq DNA polymerase and its isolated nuclease domain. *Proc Natl Acad Sci U S A*, **96**, 6143-6148.
75. Shendure, J., Mitra, R.D., Varma, C. and Church, G.M. (2004) Advanced sequencing technologies: methods and goals. *Nat Rev Genet*, **5**, 335-344.
76. Patil, N., Berno, A.J., Hinds, D.A., Barrett, W.A., Doshi, J.M., Hacker, C.R., Kautzer, C.R., Lee, D.H., Marjoribanks, C., McDonough, D.P. *et al.* (2001) Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science*, **294**, 1719-1723.
77. Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X.C., Stern, D., Winkler, J., Lockhart, D.J., Morris, M.S. and Fodor, S.P. (1996) Ac-

- cessing genetic information with high-density DNA arrays. *Science*, **274**, 610-614.
78. Shendure, J., Porreca, G.J., Reppas, N.B., Lin, X., McCutcheon, J.P., Rosenbaum, A.M., Wang, M.D., Zhang, K., Mitra, R.D. and Church, G.M. (2005) Accurate Multiplex Polony Sequencing of an Evolved Bacterial Genome. *Science*, 1117389.
 79. Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z. *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, **advanced online publication**.
 80. Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M. *et al.* (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat Biotechnol*, **18**, 630-634.
 81. Mitra, R.D., Shendure, J., Olejnik, J., Edyta Krzymanska, O. and Church, G.M. (2003) Fluorescent in situ sequencing on polymerase colonies. *Anal Biochem*, **320**, 55-65.
 82. Syvanen, A.C. (2001) Accessing genetic variation: genotyping single nucleotide polymorphisms. *Nat Rev Genet*, **2**, 930-942.
 83. Syvanen, A.C., Aalto-Setälä, K., Harju, L., Kontula, K. and Soderlund, H. (1990) A primer-guided nucleotide incorporation assay in the genotyping of apolipoprotein E. *Genomics*, **8**, 684-692.
 84. Tost, J. and Gut, I.G. (2002) Genotyping single nucleotide polymorphisms by mass spectrometry. *Mass Spectrom Rev*, **21**, 388-418.
 85. Faham, M., Baharloo, S., Tomitaka, S., DeYoung, J. and Freimer, N.B. (2001) Mismatch repair detection (MRD): high-throughput scanning for DNA variations. *Hum Mol Genet*, **10**, 1657-1664.
 86. Fakhrai-Rad, H., Zheng, J., Willis, T.D., Wong, K., Suyenaga, K., Moorhead, M., Eberle, J., Thorstenson, Y.R., Jones, T., Davis, R.W. *et al.* (2004) SNP discovery in pooled samples with mismatch repair detection. *Genome Res*, **14**, 1404-1412.
 87. Tian, J., Gong, H., Sheng, N., Zhou, X., Gulari, E., Gao, X. and Church, G. (2004) Accurate multiplex gene synthesis from programmable DNA microchips. *Nature*, **432**, 1050-1054.
 88. Zhou, X., Cai, S., Hong, A., You, Q., Yu, P., Sheng, N., Srivannavit, O., Muranjan, S., Rouillard, J.M., Xia, Y. *et al.* (2004) Microfluidic PicoArray synthesis of oligodeoxynucleotides and simultaneous assembling of multiple DNA sequences. *Nucleic Acids Res*, **32**, 5409-5417.
 89. Cleary, M.A., Kilian, K., Wang, Y., Bradshaw, J., Cavet, G., Wei, G., Kulkarni, A., Paddison, P.J., Chang, K., Sheth, N., Leprost, E., Coffey, E.M., Burchard, J., McCombie, W.R., Linsley, P. and Hannon, G.J. (2004) Production of complex nucleic acid libraries using highly parallel in situ oligonucleotide synthesis. *Nat Methods*, **1**, 241-248.
 90. Hogervorst, F.B., Nederlof, P.M., Gille, J.J., McElgunn, C.J., Grippe, M., Pruntel, R., Regnerus, R., van Welsem, T., van Spaen-

- donk, R., Menko, F.H. *et al.* (2003) Large genomic deletions and duplications in the BRCA1 gene identified by a novel quantitative method. *Cancer Res*, **63**, 1449-1453.
91. Bunyan, D.J., Eccles, D.M., Sillibourne, J., Wilkins, E., Thomas, N.S., Shea-Simonds, J., Duncan, P.J., Curtis, C.E., Robinson, D.O., Harvey, J.F. *et al.* (2004) Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification. *Br J Cancer*, **91**, 1155-1159.
 92. Gille, J.J., Hogervorst, F.B., Pals, G., Wijnen, J.T., van Schooten, R.J., Dommering, C.J., Meijer, G.A., Craanen, M.E., Nederlof, P.M., de Jong, D. *et al.* (2002) Genomic deletions of MSH2 and MLH1 in colorectal cancer families detected by a novel mutation detection approach. *Br J Cancer*, **87**, 892-897.
 93. Slater, H.R., Bruno, D.L., Ren, H., Pertile, M., Schouten, J.P. and Choo, K.H. (2003) Rapid, high throughput prenatal detection of aneuploidy using a novel quantitative method (MLPA). *J Med Genet*, **40**, 907-912.
 94. Nygren, A.O.H., Ameziane, N., Duarte, H.M.B., Vijzelaar, R.N.C.P., Waisfisz, Q., Hess, C.J., Schouten, J.P. and Errami, A. (2005) Methylation-Specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. *Nucl. Acids Res.*, **33**, e128-.
 95. Sellner, L.N. and Taylor, G.R. (2004) MLPA and MAPH: new techniques for detection of gene deletions. *Hum Mutat*, **23**, 413-419.
 96. Gunderson, K.L., Steemers, F.J., Lee, G., Mendoza, L.G. and Chee, M.S. (2005) A genome-wide scalable SNP genotyping assay using microarray technology. *Nat Genet*, **37**, 549-554.
 97. Dressman, D., Yan, H., Traverso, G., Kinzler, K.W. and Vogelstein, B. (2003) Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc Natl Acad Sci U S A*, **100**, 8817-8822.
 98. Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. and Nyren, P. (1996) Real-time DNA sequencing using detection of pyrophosphate release. *Anal Biochem*, **242**, 84-89.
 99. Ronaghi, M., Uhlen, M. and Nyren, P. (1998) A sequencing method based on real-time pyrophosphate. *Science*, **281**, 363, 365.
 100. Leamon, J.H., Lee, W.L., Tartaro, K.R., Lanza, J.R., Sarkis, G.J., deWinter, A.D., Berka, J. and Lohman, K.L. (2003) A massively parallel PicoTiterPlate based platform for discrete picoliter-scale polymerase chain reactions. *Electrophoresis*, **24**, 3769-3777.

Acta Universitatis Upsaliensis

*Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Medicine 65*

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-5921



ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2005