

Accessing genomic information: alternatives to PCR

Anders Isaksson and Ulf Landegren

The growing abundance of genomic sequence data invites increasingly large-scale genetic analyses. Studies of genetic variation in large sets of genes can illuminate important disease mechanisms and serve to identify novel drug targets or predict therapeutic responses. At present mostly a concern in extensive research projects, large-scale genetic analyses will gradually also find their way into clinical practice as an aid to the physician. It is timely, therefore, to take stock of methods that are becoming available for analyses of large sets of gene sequences. Clearly PCR remains the workhorse for molecular genetic analysis, and several modifications such as homogenous amplification assays and parallel detection on DNA microarrays further increase throughput. Recent developments, however, also offer hope that other methods will become available for genomic investigations, providing substantially increased analytical capacity.

Address

The Beijer Laboratory, Department of Genetics and Pathology, Box 589, Uppsala Biomedical Center, Se-751 23 Uppsala, Sweden

Current Opinion in Biotechnology 1999, 10:11–15

<http://biomednet.com/elecref/0958166901000011>

© Elsevier Science Ltd ISSN 0958-1669

Abbreviation

RCA rolling-circle amplification

Introduction

Extensive research efforts both within academia and at companies such as The Institute for Genomic Research (TIGR), Incyte, Genset and Celera are geared to rapidly expand knowledge about the human genome, and of human genetic variation. In addition to compiling DNA sequences for all human genes and their regulatory regions, hundreds of thousands of DNA sequence variants will be identified and used to establish high-density maps of the human genome [1]. In particular, the scope of normal variation in all human coding sequences will be recorded. The genetic variation will also be characterized within separate human populations and can be used to trace their histories [2]. A lot of excitement has been created due to the potential for these genetic variants to be used in population-based association studies as a means to locate genes involved in common disorders, such as hypertension and Alzheimer's disease [3,4].

A whole-genome perspective, where numerous potentially interacting genetic factors are assessed in parallel, assumes increasing importance also in studies of model organisms and of animals and plants of agricultural importance. Clinical application of genetic variation found to be relevant for diagnosing and prognosticating disease again necessitates a capacity to monitor large sets of gene sequences.

Here we discuss some recent papers presenting approaches that promise increased capacity in genome analyses.

Developments in genotyping

Sequences to be analyzed are usually first amplified from genomic DNA, most commonly by PCR, in order to increase copy numbers and effectively reduce the complexity of the sample. The two alternative amplification methods, strand-displacement amplification (SDA) and nucleic-acid sequence-based amplification (NASBA), may in some cases offer advantages over PCR, for example, isothermal reaction conditions that simplify instrumentation [5]. PCR, however, remains the predominant amplification technique in genotyping.

So-called homogenous amplification assay formats simplify detection of amplification products and thereby increase throughput [6]. Extensive post-amplification processing and carry-over contamination between amplification reactions are both largely avoided by monitoring the amplification products via fluorescence measurements during or immediately at the conclusion of the reaction, with no need for isolation of reaction products on solid phases and washes. The assays all hinge on target-dependent changes in the characteristics of fluorophore-labelled probes, measured via fluorescence energy transfer, or through fluorescence polarization spectroscopy. Examples of techniques building on these principles are the 5' exonuclease (Taqman) assay [7,8], molecular beacons [9–11], fluorescence-based minisequencing or oligonucleotide ligation [12,13*], and strand displacement amplification with fluorescence polarization [14,15].

Homogenous PCR-based assays can also be performed in silicon microstructures [16*,17,18]. The miniaturized amplification format allows shorter cycling times, reduces reagent use, and offers a possibility to monitor more reactions in parallel. Another way forward would be to detect an increased number of sequences using multiple fluorophores. To this end, molecular beacon probes with distinct emission spectra have been developed, that allow simultaneous detection of four probes [19].

DNA microarray-based tests

With the aid of DNA microarrays, numerous sequences can be detected in parallel by confining individual analyses to specified positions on a flat surface interacting with the sample. In array-based minisequencing, oligonucleotide primers are immobilized in patches on a support. The target-dependent, polymerase-assisted incorporation of a specific detectable nucleotide at the 3' end of the primers is taken as a measure of the presence of alternate sequence variants of the amplified target sequences. The method has been shown to clearly distinguish between closely similar amplified DNA

sequence variants [20•], and it has been applied to study the association of polymorphisms in candidate genes with myocardial infarction in Finnish patients [21].

Wang *et al.* [22••] combined multiplex DNA amplification with DNA microarray-based hybridization analysis. In so doing, up to 500 single nucleotide polymorphisms (SNPs) could be screened in a single hybridization of PCR products from 12 pooled amplification reactions to a high-density oligonucleotide microarray.

As more amplification primer pairs are combined in a reaction, there is an increasing risk of generating irrelevant PCR products from primer pairs. Any such false products compete with the intended products for access to the polymerase, and they may also interfere at the hybridization-based detection step. Furthermore, some primer pairs function less well than others and may thus have a competitive disadvantage in the amplification reaction. This frequently leads to a failure to detect certain amplicons [23].

Wang *et al.* [22••] took several measures to minimize problems connected with multiplex PCR. First, primers were designed to hybridize very closely to the location of the single nucleotide polymorphisms, yielding short, uniform PCR products. Second, the primers all had similar melting temperatures to ensure more uniform primer hybridization. Third, general sequences were added at the 5′ ends of each primer pair, a measure shown to improve multiplex PCR [24]. The relevant fragments were first amplified with the specific primers, followed by labeling through amplification with biotinylated general primers that represent the sequences at the 5′ ends of the primers. In order to improve amplifications in multiplex reactions, PCRs performed separately were classified as strong, medium or weak, depending on the yield of PCR products, and primer pairs from each of those groups were pooled for multiplex amplification. With 46 loci amplified in parallel, 90% of the expected products were possible to score. If even higher levels of multiplexing were attempted the success ratio dropped significantly. Nonetheless, this shows that a considerable number of sequences can be amplified in parallel.

Genotyping directly on unamplified genomic DNA

A radical means to avoid problems with multiplex PCR would be to analyze sequences in genomic DNA without prior target amplification. Most simply, genomic DNA could be labeled and allowed to hybridize to allele-specific oligonucleotides immobilized in a DNA microarray, in order to monitor genetic variation in the genomic DNA. In a recent paper by Winzler *et al.* [25••] two different yeast strains were compared in just this way. An array of oligonucleotides was prepared representing a total 21.8% of the non-repetitive genomic sequence of the sequenced strain of *Saccharomyces cerevisiae*, which is very similar to one used in the study. Genomic DNA from the other, less well-characterized strain was also hybridized to the array, but at 3714

loci a reduced hybridization signal was observed, probably reflecting sequence differences between the two strains. At one locus, the difference was shown to represent variation in a single nucleotide position. Loci where the two strains differed were used as genetic markers. The two strains were mated and segregation of the markers and an antibiotic resistance trait was studied. In this manner, the antibiotic resistance gene could be mapped to within a 57 kb region allowing the gene to be identified.

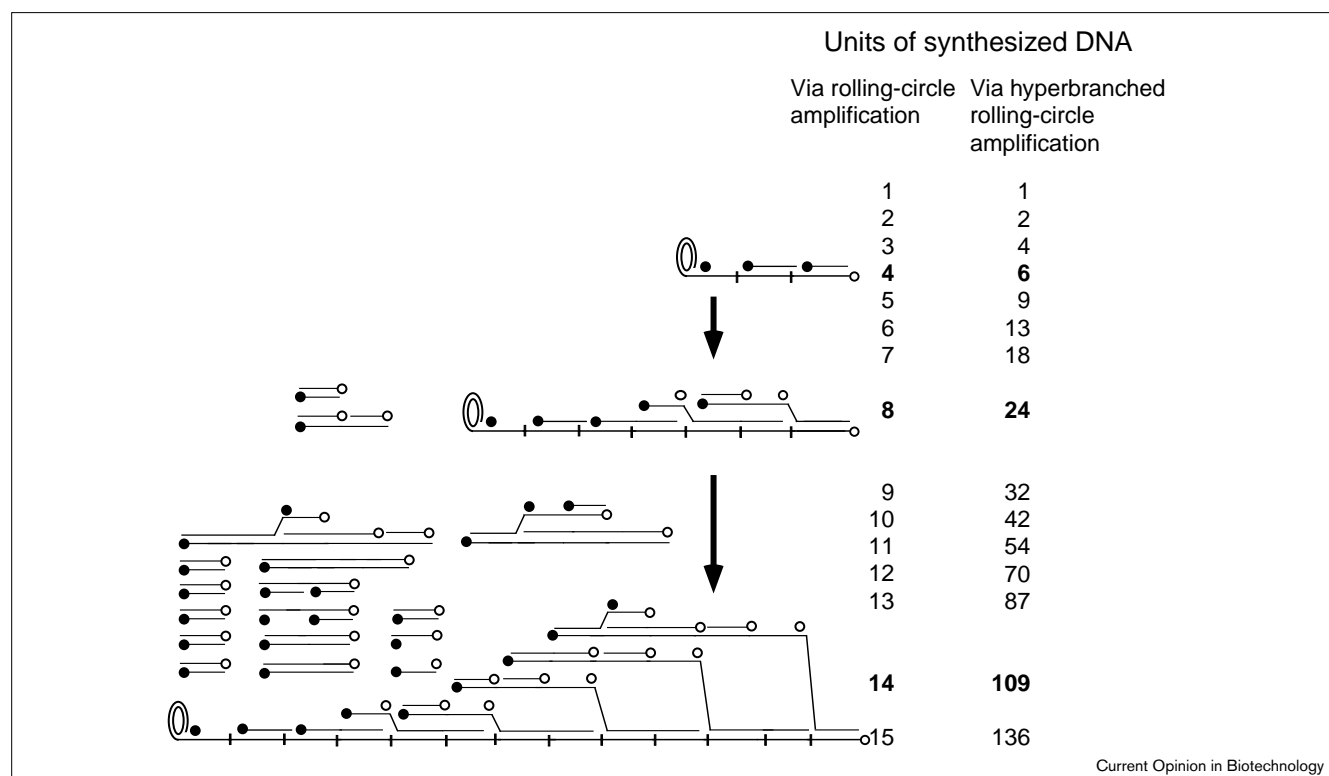
Before contemplating similar array-hybridization experiments with total human genomic DNA some fundamental differences between the yeast and human genomes should be considered. Firstly, the haploid human genome is about 250 times larger than that of yeast, thus representing a far greater complexity, and including many more cases of closely similar sequences. Secondly, the difference in genome size also implies that more human DNA will be needed to reach the same number of hybridizing molecules. Finally, whereas in the yeast study haploid samples were used, the human genome is diploid and this will complicate the analysis by requiring that also heterozygous sequences can be identified. It is pertinent to ask whether 15–25-mer oligonucleotide hybridization probes are sufficiently specific to search through the human genome.

Almost ten years ago Wu *et al.* [26] reported differential hybridization of short allele-specific oligonucleotide probes to total human genomic DNA spotted on a membrane. In the presence of an excess of unlabelled probes, specific for the opposite allele, they were able to type individuals as heterozygous or homozygous for either of two sequence variants that differed in a single nucleotide position. Despite this success, it appears doubtful if arrayed oligonucleotide probes can be used to screen for multiple sequence variants in total human genomic DNA due to problems with both specificity and the signal to background ratio.

A probe design that may offer some interesting advantages in analyses of unamplified genomic DNA are circularizable oligonucleotides or padlock probes. These are linear oligonucleotides, the two ends of which can hybridize next to each other on a target DNA strand, allowing the ends to be joined by enzymatic ligation [27]. Only when both end segments correctly recognize nearby target segments can the probes be converted to circular molecules that are wound around the target strands. The probes are promising for multiplex analyses, as they are expected to be less susceptible to the problem with crossreactions seen in multiplex PCR, as only intramolecular ligation reactions are scored.

Padlock probes have been used to investigate *in situ* single nucleotide variation in a repeated sequence in the human genome by taking advantage of the requirement for correct hybridization for the ligation to take place, but in this analysis detection sensitivity was insufficient to detect single-copy sequences [28•].

Figure 1



Current Opinion in Biotechnology

Hyperbranched rolling-circle amplification. Rolling-circle replication is initiated from a primer (open circle) that hybridizes to a circular single-stranded DNA molecule. The growing extension product represents a concatamer of complements to this circular DNA molecule. A second primer (filled circle) hybridizes to each unit of the rolling-circle replication product as it is being displaced. Upon extension, this primer can displace other copies of the primer

hybridizing downstream on the extension product. Displaced extension products in turn can template replication initiated from the primers. Double-stranded molecules can not be further replicated. In this manner, various-size linear concatemers of the circular DNA molecule and its complement accumulate. The number of unit size linear molecules, identical in sequence or complementary to the circular DNA molecule are indicated.

The circular nature of the reacted padlock probes offers a valuable opportunity for signal amplification by using the probes to template a rolling-circle amplification (RCA) [29,30]. In this reaction, a primer hybridized to the circularized probe is extended by a DNA polymerase, generating a strand that represents multiple head-to-tail copies of the probe sequence. As this mechanism is critically dependent on circularization, only reacted probes can template synthesis of long single-stranded DNA. Both Zhang *et al.* [31[•]] and Lizardi *et al.* [32^{••}] have extended the rolling-circle mechanism further by using both a primer suitable to initiate copying of the circularized probe and another one, capable of priming replication of the RCA product. In the strategy taken by Zhang *et al.* [31[•]] an amplified signal is obtained by cyclically altering the reaction temperature, as in conventional PCR. By contrast, Lizardi *et al.* [32^{••}] relied solely on a strand-displacement mechanism to allow a so-called hyperbranched RCA at a fixed temperature (Figure 1). This reaction progresses faster than a linear amplification and produces double stranded DNA molecules that represent multiples of the probe sequence. Lizardi *et al.* [32^{••}] designed their probes so that an intermediary fragment must be in place before

the ends of the padlock probe can be joined and the probe circularized. With this approach the reaction was shown to be both sufficiently specific and sensitive to detect single-copy sequences in human genomic DNA.

Banér *et al.* [33] demonstrated that while the efficient RCA from a single circularized probe can generate half a megabase of DNA in an overnight incubation, topological problems can prevent RCA. Probes bound to target molecules without nearby free ends remain catenated to the targets, and this essentially abolishes replication. By contrast, if the target is interrupted near the probe-complementary sequence, then the probes can slip off the end, whereupon efficient RCA follows. This effect has to be taken into account in assays using RCA to enhance detection.

The RCA mechanism has also been used for signal-amplification in the absence of probe circularization in a method referred to as RCA-CACHET, which avoids the above topological problem [32^{••}]. In this technique, the 5' end of an oligonucleotide bound to a solid support can be ligated to another oligonucleotide added in solution if the probes are properly aligned by hybridizing next to each other on a

target sequence. The second oligonucleotide was designed to have two 3' ends, one capable of ligating to the immobilized probe, and another one at the opposite end of the molecule, which was used to prime RCA by adding preformed circular DNA molecules, complementary to the free 3' end of the ligated probe. While the use of two separate ligation probes may complicate multiplex analyses, the assay offers detection specificity that should be adequate to analyze total human genomic DNA, and it simplifies detection by resulting in an amplified DNA strand that is contiguous with the immobilized probe.

Conclusion

The need for extremely high-throughput genetic analyses is being met by further developments of PCR-based techniques, but also by the introduction of new molecular strategies that could circumvent some of the limitations of current methods. Any such assays must combine adequate specificity and sensitivity with selectivity for sequence variants, and they should be applicable in parallel to large numbers of target sequences in a sample. On the basis of recent development there is excellent hope that assays will indeed become available where a small amount of genomic DNA can be added to an extensive set of probes, revealing quantitative or qualitative variation in many or all human genes.

Acknowledgements

We are grateful to A-C Syvänen and M Nilsson for critical reading of the manuscript. Our work is supported by the Beijer Foundation, the Swedish Research Councils for the Engineering Sciences and Medicine, and by the Swedish Cancer Fund. Anders Isaksson is a recipient of an EMBO long-term fellowship.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Collins FS, Guyer MS, Charkravarti A: **Variations on a theme: cataloging human DNA sequence variation.** *Science* 1997, **278**:1580-1581.
 2. Cavalli-Sforza LL: **The DNA revolution in population genetics.** *Trends Genet* 1998, **14**:60-65.
 3. Strittmatter WJ, Roses AD: **Apolipoprotein E and Alzheimer's disease.** *Annu Rev Neurosci* 1996, **19**:53-77.
 4. Risch N, Merikangas K: **The future of genetic studies of complex human diseases.** *Science* 1996, **273**:1516-1517.
 5. Landegren U: **The challengers to PCR: a proliferation of chain reactions.** *Curr Opin Biotechnol* 1996, **7**:95-97.
 6. Landegren U, Nilsson M, Kwok P-Y: **Reading bits of genetic information: methods for single-nucleotide polymorphism analysis.** *Genomes Res* 1988, **8**:769-776.
 7. Lee LG, Connell CR, Bloch W: **Allelic discrimination by nick-translation PCR with fluorogenic probes.** *Nucleic Acids Res* 1993, **21**:3761-3766.
 8. Livak K, Marmaro J, Todd JA: **Towards fully automated genome-wide polymorphism screening.** *Nat Genet* 1995, **9**:341-342.
 9. Tyagi S, Kramer FR: **Molecular beacons: probes that fluoresce upon hybridization.** *Nat Biotechnol* 1996, **14**:303-308.
 10. Leone G, van Schijndel H, van Gemen B, Kramer FR, Schoen CD: **Molecular beacon probes combined with amplification by NASBA enable homogeneous, real-time detection of RNA.** *Nucleic Acids Res* 1998, **26**:2150-2155.
 11. Kostrikis LG, Tyagi S, Mhlanga MM, Ho DD, Kramer FR: **Spectral genotyping of human alleles.** *Science* 1998, **279**:1228-1229.
 12. Chen X, Zehnbauser B, Gnirke A, Kwok P-Y: **Fluorescence energy transfer detection as a homogeneous DNA diagnostic method.** *Proc Natl Acad Sci USA* 1997, **94**:10756-10761.
 13. Chen X, Livak KJ, Kwok P-Y: **A homogenous, ligase-mediated DNA diagnostic test for genome analysis.** *Genome Res* 1998, **8**:549-556.
- By slightly adjusting the temperature conditions for the last cycles, fluorophore-labelled ligation probes added to a PCR become enzymatically joined in a target-dependent reaction. The assay allows allelic sequence variants to be resolved, and potentially the results can be recorded by transferring reaction plates to an inexpensive plate reader.
14. Walker GT, Nadeau JG, Linn CP, Devlin RF, Dandliker WB: **Strand displacement amplification (SDA) and transient-state fluorescence polarization detection of *Mycobacterium tuberculosis* DNA.** *Clin Chem* 1996, **42**:9-13.
 15. Spears PA, Linn CP, Woodard DL, Walker GT: **Simultaneous strand displacement amplification and fluorescence polarization detection of *Chlamydia trachomatis* DNA.** *Anal Biochem* 1997, **247**:130-137.
 16. Taylor TB, Winn-Deen ES, Picozza E, Woudenberg TM, Albin M: **Optimization of the performance of the polymerase chain reaction in silicon-based microstructures.** *Nucleic Acids Res* 1997, **25**:3164-3168.
- This paper demonstrates a miniaturized PCR format with on-board detection, yielding high-throughput amplification of numerous target sequences.
17. Ibrahim MS, Lofts RS, Jahrling PB, Henchal EA, Weedn VW, Northrup MA, Belgrader P: **Real-time microchip PCR for detecting single-base differences in viral and human DNA.** *Anal Chem* 1998, **70**:2013-2017.
 18. Northrup MA, Benett B, Hadley D, Landre P, Lehew S, Richards J, Stratton P: **A miniature analytical instrument for nucleic acids based on micromachined silicon reaction chambers.** *Anal Chem* 1998, **70**:918-922.
 19. Tyagi S, Bratu DP, Kramer FR: **Multicolor molecular beacons for allele discrimination.** *Nat Biotechnol* 1998, **16**:49-53.
 20. Pastinen T, Kurg A, Metspalu A, Peltonen L, Syvänen AC: **Minisequencing: a specific tool for DNA analysis and diagnostics on oligonucleotide arrays.** *Genome Res* 1997, **7**:606-614.
- This paper demonstrates that excellent allele distinction can be obtained between amplification products added to an oligonucleotide array through a minisequencing reaction.
21. Pastinen T, Perola M, Niini P, Terwilliger J, Salomaa V, Vartiainen E, Peltonen L, Syvänen A: **Array-based multiplex analysis of candidate genes reveals two independent and additive genetic risk factors for myocardial infarction in the finnish population.** *Hum Mol Genet* 1998, **7**:1453-1462.
 22. Wang DG, Fan J-B, Siao C-J, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J *et al.*: **Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome.** *Science* 1998, **280**:1077-1082.
- A landmark paper, testing the limits of multiplex PCR, combined with high density oligonucleotide hybridization. The paper also describes a set of more than 2000 polymorphic genetic markers.
23. Walsh PS, Erlich HA, Higuchi R: **Preferential PCR amplification of alleles: mechanisms and solutions.** *PCR Methods Appl* 1992, **1**:241-250.
 24. Shuber AP, Grondin VJ, Klinger KW: **A simplified procedure for developing multiplex PCRs.** *Genome Res* 1995, **5**:488-493.
 25. Winzler E, Richards DR, Conway AR, Goldstein AL, Kalman S, McCullough MJ, McCusker JH, Stevens DA, Wodicka L, Lockhart DJ, Davis RW: **Direct allelic variation scanning of the yeast genome.** *Science* 1998, **281**:1194-1197.
- In this paper, sequence differences between two yeast strains were analyzed by whole genome hybridization to dense oligonucleotide arrays, and used to map an antibiotic resistance trait.
26. Wu DY, Nozari G, Schold M, Conner BJ, Wallace RB: **Direct analysis of single nucleotide variation in human DNA and RNA using *in situ* dot hybridization.** *DNA* 1989, **8**:135-142.

27. Nilsson M, Malmgren H, Samiotaki M, Kwiatkowski M, Chowdhary BP, Landegren U: **Padlock probes: circularizing oligonucleotides for localized DNA detection.** *Science* 1994, **265**:2085-2088.
28. Nilsson M, Krejci K, Koch J, Kwiatkowski M, Gustavsson P, Landegren U: **Padlock probes reveal single-nucleotide differences, parent of origin and *in situ* distribution of centromeric sequences in human chromosomes 13 and 21.** *Nat Genet* 1997, **16**:252-255.
 Padlock probes were used to monitor single-nucleotide differences among repeated centromeric sequences *in situ*.
29. Fire A, Xu S-Q: **Rolling replication of short DNA circles.** *Proc Natl Acad Sci USA* 1995, **92**:4641-4645.
30. Liu D, Daubendiek SL, Zillman MA, Ryan K, Kool ET: **Rolling circle DNA synthesis: small circular oligonucleotides as efficient templates for DNA polymerases.** *J Am Chem Soc* 1996, **118**:1587-1594.
31. Zhang DY, Brandwein M, Hsuih TCH, Li H: **Amplification of target specific, ligation dependent circular probe.** *Gene* 1998, **211**:277-285.
 This paper shows the utility of padlock probes for analysis of RNA, coupled with exponential rolling-circle replication via temperature cycling.
32. Lizardi PM, Huang X, Zhu Z, Bray-Ward P, Thomas DC, Ward DC: **Mutation detection and single-molecule counting using isothermal rolling-circle amplification.** *Nat Genet* 1998, **19**:225-232.
 This interesting paper demonstrates several genetic assays, involving replication of circularized DNA strands. Single-nucleotide discrimination in a single-copy gene in total human genomic DNA is demonstrated, and a localized amplification product is elicited from immobilized oligonucleotide probes.
33. Banér J, Nilsson M, Mendel-Hartvig M, Landegren U: **Signal amplification of padlock probes by rolling circle replication.** *Nucleic Acids Res* 1998, **26**:5073-5078.