Design and optimisation of a validated primer set for automated mutation scanning of the BRCA1 and BRCA2 genes.

Mattocks C, Ward D, Harvey J, Cross N

National Genetics Reference Laboratory (Wessex), Salisbury District Hospital, Salisbury, Wiltshire, SP2 8BJ

Introduction

The increasing demand for fast throughput mutation scanning has necessitated the introduction of automation in many diagnostic laboratories. A key requirement to facilitate effective automation is standardisation of PCR amplifications. To address this issue NGRL (Wessex) have developed a Standardised Primer Optimisation and Design Specification, which includes design, optimisation and validation parameters. We have used this system to set up a fully automated screen, by conformation sensitive capillary electrophoresis (CSCE) and sequencing, that covers the entire coding regions of BRCA1 & 2 in 79 fragments.

Development

The primary aim was to develop a standard protocol for design, optimisation and validation of primers. The primers should be suitable for as wide a range of methodologies as possible; in particular they should be suitable for plate work and automation. This required that, as far as possible, all primer pairs should work at the same temperature. To afford flexibility with fluorescent labelling of fragments and allow universal sequencing primers to be used, a tagged gene specific primer system was chosen.

System 1 (see figure 1)

PCR carried out with four primers:

- Two gene specific primers with M13 tails (GS1 and GS2) at limiting concentrations so that they are depleated during PCR
- Two M13 primers at higher concentration which can be modified to suit the application (e.g. flourescent label)

Key findings:

- M13 tags could not be used to amplify NGRL plasmid controls for BRCA and HNPCC genes as the M13 sequences in the plasmids competed with the target sequences for the GS primers (figure 2).

System 2 (see figure 1)

Four primers as system 1 but M13 tags replaced with widely used Uniseq tags (US1 and US2)

Key findings:

- Standard concentrations of the primers did not give robust amplification accros a range of different fragments
- Excessive primer dimer was observed in all fragments amplified with this system. This appears to be due to a two base pair complementarity at the 3' end of the Uniseq tags (figure 3).

System 3 (see figure 4)

PCR carried out with three primers

- GS1 with US1 tag at limiting concentration
- GS2 at much higher concentration (usually 20x to 80x GS1)
- US1 with appropriate modification (flourescent label US1-F) at a standard concentration (15 fmol/ul rxn)

Key findings:

- This system was found to perform well for a wide range of fragments
- Optimisation by cross titration of GS primer concentrations allows eliminaton of dimers, modulation of peak shape and intensity, and gives an indication of the robustness of the amplification.

This system (3) was carried forward for developing diagnostic primer sets; initially BRCA1 and BRCA2.

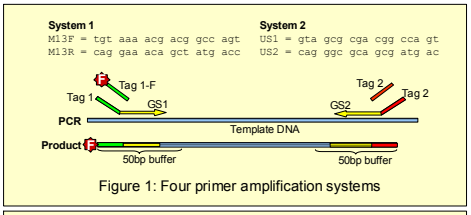

Figure 1: Four primer amplification systems

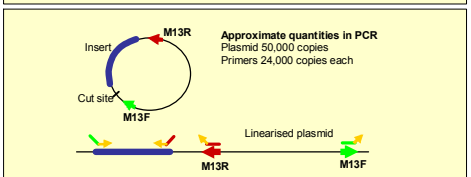

Figure 2: Competition for M13 tagged primers with plasmid controls

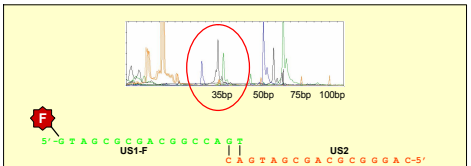

Figure 3: Uniseq primer dimer

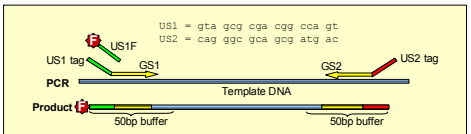

Figure 4: Three primer amplification system

Primer design (Abstacted)

Amplicons/primers	
•Amplicon length	250 - 510bp
•Buffer zone including primer (can be ignored for analysis purposes)	Minimum 50bp for overlapping fragments
•Primer location	3' end minimum of 20 bp from intron / exon boundary
•Universal tails	Each pair to have one US1tail and one US2 tail

Primer Design	
•Programme for primer design	Parameters giver for Oligo 5
○Minimum acceptable stringency	Moderate
○Most stable 3' dimer	ΔG = -6.0 kcal/mol
○Most stable overall dimer	ΔG = -10.0 kcal/mol
○Most stable hairpin	Tm = 40°C
○Lowest acceptable maximum T_a	68°C
•General	'Stuffer' sequences placed between the Gsportion of the pimer and the universal tail can be used to increase the ROI in a fragment (i.e. the analysable sequence between buffers) Very AT rich sequences near the end of amplicons tend to give poor peak shapes for flourescent analysis (stutter peaks and shoulders) – consider splitting fragment.

Validation

Database information	
•Databases used to exclude polymorphic sites – Where this is problematic degeneracy has been introduced into the primer to cope with the polymorphism.	Various dependant on gene http://www.ncbi.nlm.nih.gov/projects/SNP/
•Exclude known mutations – no attempt has been made to exclude mutations from primers in an overlapping sequence of amplicons. Where there is doubt about the significance of a mutation it has been treated as if it is a mutation. It is therefore theoretically possible to miss a mutation because a rare polymorphism is located at a primer binding site in overlapping amplicons although the risk is very small.	Various dependant on gene http://archive.uwcm.ac.uk/uwcm/mg/hgmd/search.html BRCA http://research.nhgri.nih.gov/projects/bic/ http://www.humgen.nl/lab-devilee/UncVar/b12nlpol.htm http://snp500cancer.nci.nih.gov/home_1.cfm UKGTN sources WRGL database
•BLAST search for non specific binding	Using Blastn on nt database with default parameters

Experimental	
•Successful amplification using two primer system	PCR products run on 2% agarose
•Successful amplification using three primer system	Peak threshold on 3730>1000FU
•Clean signal	PCR products run on 2% agarose PCR products run on capillary (3730)
•No sign of non-specific amplification	No signals above 10% of main peak above 100bp. No primer dimer
•Size of product matches theoretical amplicon	PCR products sized on 3730
•Successful sequencing	Details to be specified

Optimisation

A fixed US1-F primer concentration is used (15 fmol/ul rxn). Optimisation of the PCRs are carried out by cross titration of the concentrations of the two GS primers. GS1 3, 6, 9, 27 fmol/ul rxn and GS2 40, 80, 160, 320 fmol/ul rxn are good ranges for most amplicon. Figure 5 shows the effect of primer concentration on dimers.

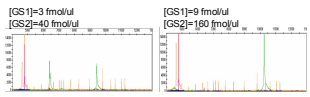

Figure 5: Effect of GS primer concentration on dimer

Disease	Gene	Fragments	Design 1 (%)	Design 2 (%)	Design 3 (%)
Breast cancer	BRCA1	33	64	29	7
Breast cancer	BRCA2	46	84	15	1
Marfans	FBN1	61	92	-	-

Table 1: Optimisation success rate for current screens

Experience has enabled refinement of the design and optimisation proceedure - A pass rate of > 90% is now achieved for new designs (table 1). The optimisation protocol has been fully automated giving the capacity to optimise 60 or more amplicons in a matter of days.

PCR setup

Reaction components	2 primer PCR	3 primer PCR
DNA	20-50ng	20-50ng
GS primers	500 fmol/ul rxn	Frag. specific
US1F primer	none	15 fmol/ul rxn
Amplitaq Gold master mix x2 (AB)	To x1	To x1
Thermal cycling		
Taq activation	95°C for 10mins	95°C for 10mins
Denature	95°C for 5 sec	95°C for 5 sec
GS primer annealing	61°C for 30 sec	61°C for 30 sec
GS cycles	40	30
Final extend	72°C for 5 mins	72°C for 5 mins

Application

This system has been used to design primers for BRCA mutation scanning in the SCOBEC High Throughput Screening facility. The BRCA primer set comprises 79 fragments ranging form 258bp to 509bp covering the entire coding regions of the two genes. To date the primers have been used to PCR amplify 474 samples for various combinations of fragments with analysis by both CSCE and sequencing. In total approximately 30,000 reactions have been performed (A pictorial representation of the first 282 samples is given in figure 6 together with summay statistics).

The estimated failure rate is approximately 4% which includes initial rates of >10%. The current failure rate is estimated to be <2% of PCRs performed.

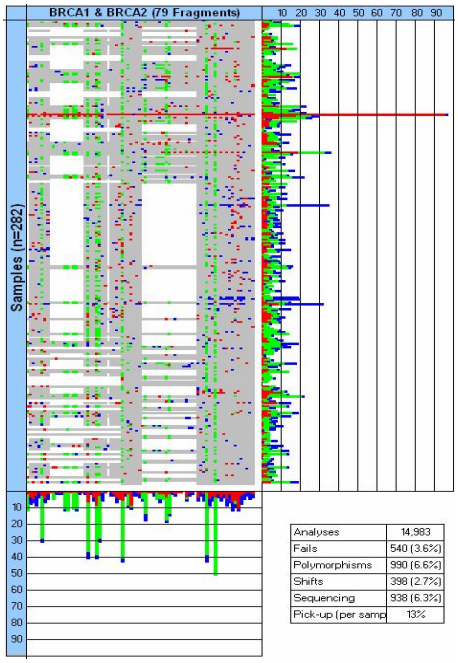

Figure 6: Application in the SCOBEC High Throughput Screening Facility. Grey represents analyses performed, Red=Fail, Blue=shift, and Green=polymorphism. The histograms to the right and below summarise the data for each sample and fragment respectively

Conclusions and future work

We have developed a prototype specification for validated primer design and optimisation and have demonstrated its utility in the context of a fully automated sample processing system. We have also implemented an automated optimisation pipeline that will expedite the development of new primer sets. Assuming continued NGRL funding, we propose to coordinate a network best practice consultation and initiate a centrally administered database of validated primer sets. This will have numerous benefits including compatibility with automation and reference materials, avoiding duplication of effort, Equity of testing between labs, transportability and mechanism for sharing of information between labs.

Futher information can be found at http://www.ngrl.org.uk/Wessex/downloads.htm