**ExoSeq: Protocols – ExoSeq Primer Design**

Exons and up to 1 kb of flanking sequence for all known human protein-coding genes, novel coding sequences and transcripts are being extracted and submitted for primer design. Initially primer design targeted the 14 finished human chromosomes for which manual annotation is available in the Vega database (http://vega.sanger.ac.uk/index.html). The Vega database contains high-quality, frequently updated, manually curated annotation for vertebrate finished genome sequence. Primer design is now focussed on the remaining exons in the human genome, using the best quality annotation data available in addition to Ensembl.

Repeats in the extracted sequence are masked using RepeatMasker (http://wwwrepeatmasker.org/) prior to submission for design. Primers are designed automatically to the extracted sequence using Primer3 (http://frodo.wi.mit.edu/primer3/primer3_code.html) to amplify the exon and at least 125 base pairs either side of the exon to capture the splice sites and regions which may contain other signals for splicing or be of other functional importance (see illustration below). Primers with exact or 1 bp mismatches (or SNPs) in their sequence are excluded and re-designed. Larger exons may need to have a series of overlapping primer pairs designed to them to obtain complete coverage.


**Primers are designed to amplify the exon plus 125 bp of flanking sequence.**

The optimum amplimer size for design is between 450 - 550 bp. Any exons failing automatic primer design at this size range are submitted for re-design with an increased permitted product size of up to 700 bp. Primers that fail this second round are submitted for manual primer design. The success rate for automatic primer design is approximately 80 - 90%. Manual primer design adds an additional 10 - 15% of successful designs leaving only a small percentage of exons that require alternative methodology (e.g. nesting) for successful amplification.

Ipcress (http://www.ebi.ac.uk/~guy/exonerate/ipcress.man.html) is used to check the primer pairs are unique in the genome prior to ordering in 96-well format. Primers are pre-screened (see pre-screening protocol) to determine the optimum conditions for amplification with the majority of exons being successfully amplified at 60°C.

Over 280,000 primer designs covering over 80% of the exons annotated in Vega as well as 15,000 exons currently only annotated in Ensembl are available for downloading.