"Evaluating the use of massively parallel sequencing technology for gene doping testing"

Anna Baoutina, K. Emslie (Commonwealth of Australia, Australia)

Project overview

We will develop a new approach to gene doping detection based on the leading-edge technology, targeted massively parallel sequencing (MPS). Similarly to PCR-based methodology, doping genes will be detected by identifying sequences that do not feature introns present in natural genes. However, the MPS approach will be more reliable since multiple splice sites will be analyzed simultaneously.

Its sensitivity is likely to be superior due to target enrichment during library preparation and to greater flexibility in choosing targeted regions. The MPS multiplexing capability will allow simultaneous analysis of many samples and genes, reducing test's cost and turnaround time. We will develop a reference material for the MPS detection of five genes, most likely candidates for doping, and validate the test using blood samples.

Results and Conclusions:

In this project we developed and tested the protocol for a targeted MPS-based method for simultaneous detection of five transgenes in solutions of doping genes and genomic DNA that we used to mimic DNA extracts from athletes' blood samples. Enrichment is achieved by PCR that amplifies a large portion of each transgene covering several exon-exon junctions, which should increase the likelihood of transgene detection. This also allows greater flexibility in selecting regions targeted by primers for enrichment PCR and provides a potential advantage of MPS over real-time PCR methods, where assay design is confined to a small area around the exon-exon border. Enrichment PCR for each transgene is performed separately in simplex, before the materials from five PCR, each for one transgene, are mixed prior to being subjected to library preparation and sequencing. With this experimental workflow, it will be relatively easy to add more genes to the detection panel; this will require designing primers and optimising enrichment PCR for new target genes.

We optimised all wet lab steps and designed, produced and tested a prototype reference material that is suitable for use in positive controls in the test.

We developed a tailored bioinformatics pipeline that reliably distinguishes doping genes from the corresponding endogenous genes and from the reference material. The pipeline enabled step-wise analysis of global distribution of alignments for sample and the RM, checking whether the sample is positive for the RM and doping genes, followed by analysis of distribution of alignments across the individual doping genes and their

corresponding fragments in the RM and, lastly, making a final positive/negative call for individual doping genes. The process allows optimum elimination of false positives due to accidental contamination of sample with the RM.

Using ‘mock’ samples, we showed that MPS-based approach can detect down to five copies of doping genes in a background of genomic DNA similar in quantity to that in one mL of blood. We demonstrated that the amount of genomic DNA in a sample affects target enrichment for some transgenes more than others and this can affect the sensitivity of their detection.

As part of the evaluation of the long-term potential of the MPS for transgene detection, we performed preliminary assessment of the method’s cost, sensitivity, reliability, turn-around time and required infrastructure and expertise, and compared it with the real-time PCR-based transgene detection.