Volume 16 Number 21 1988 Nucleic Acids Research

Errors in the Polymerase Chain Reaction

Alison M.Dunning, Philippa Talmud and Steve E.Humphries

Charing Cross Sunley Research Centre, Lurgan Avenue, London, W6 8LW, UK
Submitted August 24, 1988

We have used the Polymerase Chain Reaction (PCR) to amplify a 798bp fragment of the gene for human Apolipoprotein B (Apo B), that contains sequences coding for the putative LDl-receptor binding domain.

5µg genomic DNA from 10 individuals was amplified[1] using 30mer oligonucleotides spanning bases 9599-10397 (inclusive) of the Apo B gene. 30 rounds of amplification were carried out using 5U of Taq Polymerase (Anglian Biotech.) per sample, in a buffer containing : 67mM Tris-HCl (pH 8.8), 6.7mM $MgCl_2$, 16.7mM $(NH_4)_2SO_4$, 10mM β mercaptoethanol, 6.7µM disodium EDTA, 4mg/ml BSA, 10% Dimethylsulphoxide and 330µM (each) dATP, dCTP, dGTP, dTTP, under the regime: 2mins @ 95^0C, 1min @ 55^0C, 5mins @ 70^0C. Amplified DNA was digested with EcoRI and ScaI (Anglian Biotech.) and force-cloned into EcoRI/SmaI cut M13 mp10 (Amersham) using standard techniques. At least 10 clones from each subject were purified. Clones were sequenced using the Sequenase kit (USB Inc.) and analysed on 8% denaturing polyacryamide gels.

Initially the sequence of one clone per individual was determined.Out of the total of 8000 bases sequenced (10 individuals), 22 differences were detected (Table). No clone was identical to the published sequences[2]. Since any genuine base change should be present in approximately half the clones analysed (assuming the individual to be heterozygous), we subsequently analysed all 10 clones from each individual. None of the initial differences found were present in any of the other clones, although all of them were reproduced upon resequencing of the original clones. This implies that all the base differences seen were artefacts generated by the PCR.

The most common changes found were from A to G and from T to C. 17/22 (77%) of the changes noted were associated with a run of bases of the same sequence (Table). This may be an indication of the mechanism by which the errors are inserted.

These observations indicate that the interpretation of sequence changes from cloned, amplified DNA must be made with caution. Direct sequencing of the PCR material would probably overcome these artefacts[3].

Base Change	Sequence
A > C	ATTTA**A**GT
A > G	TCACA**A**AT
	CATCA**A**AT
	TACAA**A**GC
	GATTG**A**AG
	GCCAC**A**GC

C > T	CTCCC**C**AG
	AGGCA**C**AG
G > A	GAGCT**G**CC
	TGCCA**G**TC
T > A	AGTTG**T**CA
	CAGCA**T**GC
	TGCAC**T**GC

T > C	GTTTA**T**CA
	GGACC**T**TT
	ATGAT**T**TC
	ATTGT**T**GC
	CTTCA**T**TG
	TGAAG**T**TA
Lost A	GAAAA**A**GG
Lost G	TTCCA**G**TT
Lost T	AAGTT**T**GA

Acknowledgements: This work was funded by the British Heart Foundation & the Charing Cross Sunley Research Trust.

References

1. Saiki RK, Gelfand DH, et al. (1987) Science 239: 487-491.
2. Olofsson S-O, Bjursell G, et al.(1987) Atherosclerosis 68: 1-17.
3. Paabo S. and Wilson A.C. (1988) Nature 334:387.