STUDIES ON THE HUMFABP AND HUMTHO1 POLYMORPHISMS

M G SAVILL

Environmental Science and Research
Christchurch Science Centre
Christchurch
New Zealand

G J SENSABAUGH

School of Public Health
Department of Biomedical and
Environmental Health Sciences
University of California
Berkeley

1. INTRODUCTION

Genetic variation at loci containing short tandem repeats (STR) have been demonstrated to be highly informative for purposes of human identification and for the study of population structure [1, 2]. We describe here studies on two STR loci and their expression in three New Zealand populations.

HUMFABP is the gene for human intestinal fatty acid binding protein [3]; it is located on chromosome 4 at q31. STR variation is found in an ATT trinucleotide repeat at the 3' end of intron 2. Eight alleles, defined by sequences containing 8-15 repeats, were identified in the survey by Edwards et al of four US populations (Euro-americans, Afro-americans, Mexican-americans, and Asian-americans).

HUMTHO1 is the gene for human tyrosine hydroxylase [4]; it is located on chromosome 11 at p15. STR variation is found in a CATT tetranucleotide repeat in intron 1. The Edwards survey identified eight alleles defined by sequences contain 5-12 repeats. One of these alleles is anomolous: the THO1*10 allele contains a single base deletion in one of the repeat units and is referred to as the 10-1 allele.

FABP and/or THO1 typing has been performed on over 250 individuals representing US and New Zealand populations. The three New Zealand populations studied are:

Maori The original peoples of new Zealand; samples were collected from unrelated individuals from North and South island.

Samoan The original peoples of the Samoan islands (Polynesia) located approximately 3000 km northeast of New Zealand. Samples were collected from Samoans residing in New Zealand.

Caucasian Migrants to New Zealand from Europe, mostly of British extraction. Samples were collected from unrelated individuals in the Christchurch area of the South Island.

2. METHODS

The amplification and typing conditions used were those defined by Edwards et al [1,2] with the exception of one primer being end labelled and amplification products run on 5% M13 sequencing gels. Allelic Bands were revealed with autoradiography.

3. STR TYPING RESULTS

The use of the M13 sequence as a sizing ladder allows allelic band sizes to be assigned with single base resolution; any deviations in allele size are readily noted. Under the electrophoretic conditions used, no differences in allele band mobility were noted between PCR products end-labelled on either primer.

FABP typing patterns typically show a major and minor band for each allele; the two bands differ in length by a single base. The absolute band sizes indicate that the larger band is one base longer than the expected sequence; we demonstrated this experimentally. The double banding does not introduce ambiguity into typing although it can confound the typing of mixtures as might be encountered in case samples.

We have encountered one individual, an Hispanic , who appears to possess three FABP alleles; the nominal type would be 10, 13, 14. The sequences of each of the allelic PCR products have been determined. The sequences of the flanking regions in each are identical and the STR regions differ only in repeat number. This is consistent with a gene duplication on one chromosome.

The THO1 patterns show single bands for each allele. Each allelic band sizes one base longer than the nominal allele size based on sequence; the addition of an extra base at the end of each PCR product appears to be virtually complete. All "THO1*10" alleles sized as *10-1; no true THO1*10s were observed.

4. POPULATION STUDIES

Table 1 **FABP ALLELE FREQUENCIES (%)**

Allele	NZCau	Am-Ind	Maori	Samoa
8	-	-	-	-
9	-	2	1	4
10	50	55	47	67
11	12	8	3	3
12	5	-	2	-
13	27	29	28	27
14	6	8	19	-
15	-	-	-	-
N	47	49	57	51
Het (%)	62	59	60	47
DP (ca%)	78	76	77	67

Table 2 **THO1 ALLELE FREQUENCIES (%)**

Allele	NZCau	Maori	Samoa
5	-	-	-
6	25	25	6
7	23	27	47
8	10	7	18
9	9	13	15
10-1	34	28	15
11	-	-	-
12	-	-	-
N	46	69	34
Het (%)	83	71	71
DP (ca%)	94	85	85

There was no evidence of interloci allelic association.

5. SEQUENCE ANALYSIS AND STABILITY OF THE VARIANTS

Allelic sequences at FABP and THO1 loci were analysed to determine if there was sequence microheterogeneity within the length defined alleles.

FABP

Representative samples of FABP*10 (n=7) and FABP* 13 (n=6) were sequenced.
In every case no variation is seen in the repeat units nor in the flanking sequences. As noted elsewhere the individual exhibiting each of the three alleles was sequenced. The sequences there also identical, variation only observed in number of repeats.

THO1

Sequence analysis of the *10-1 allele in caucasians (n=4) shows a deletion of a T in a core repeat unit. The sequences of the *9 and *10-1 *11 alleles can be represented thusly:

*9 *****$[CATT]_9$*****
*10-1 *****$[CATT]_3$CAT $[CATT]_6$*****
*11 *****$[CATT]_{11}$*****

Sequence analysis of the 10-1 allele in Maoris (n=4) shows the same base deletion at the same site. Thus the *10-1 allele in caucasians and Maoris is identical at the sequence level.

6. IMPLICATIONS FOR ALLELIC STABILITY AT THO1

The common sequence of the *10-1 allele in caucasians and Maoris indicates that a mutation giving rise to its existence predates the splitting of these two populations. Moreover these sequence data provide evidence of allelic stability at the THO1 locus. Genetic variation at STR loci is believed to be generated by sequence slippage during strand replication in meiosis and or unequal crossing over. Both processes would be expected to yield x-1 variants of the other alleles, *10-1 alleles with the deletion site in different repeat units, and a true *10 allele.

That none of these products were observed in this diverse population sampling indicates that neither process occurs at significant frequency. This observation is consistent with the mutation rate estimated by Edward *et al* of 3.5 x 10^{-5}[4]. It is to be expected that as the number of individuals tested at the THO1 locus increases, that some of these variants may appear at low frequency. When observed these rare variants can be used to estimate the age of the THO1 polymophisim.

7. CONCLUSIONS

1. All three New Zealand populations are in Hardy-Weinberg Equilibrium at both loci.
2. The heterozygosity levels found are comparable with other reported populations.
3. The sequence deletion of the *10-1 repeat common to caucasian and Maori indicates an ancient origin of the allele.
4. Non templated single base addition can be removed by Klenow at certain loci.

8. REFERENCES

1. Edwards A, Civitello A, Hammond HA, and Caskey CT. *DNA typing and genetic mapping with trimeric and tetrameric repeats.* American Journal of Human Genetics 49: 746-756. 1991
2. Edwards A, Hammond HA, Jin L, Caskey CT, and Charkraborty R. *Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups.* Genomics 12: 241-253. 1992.
3. Sweetser DA, Birkenmeier EH, Klisak IJ, Zollman S, Sparkes RS, Mohandas T, Lusis AJ, Gordon JI. *The Human and Rodent Intestinal Fatty Acid Binding Protein Genes.* Journal of Biology chemistry 262, No33, 16060-16071, 1987.
4. Kobayashi K, Kaneda N, Ichinose H, KishiF, Nakazawa A, Kurosawa Y, Fujita K, Nagatsu T, *Structure of Human Tyrosine Hydroxylase Gene: Alternative Splicing from a Single Gene accounts for Generation of four MRNA types.* J Biochem 103, 907-912, 1988.
5. Clark JM, *Novel non-templated nucleotide addition reactions catalysed by procarvotic and eucaryotic DNA polymerases.* Nucleic Acids Research 16: 9677-9686, 1988.
6. Haggi TM, *Direct ligation of PCR products for cloning and sequencing, Nucleic Acids Research 20*: 6427-.1992.