Nobody is perfect: Comparison of the accuracy of PCR-RFLP and probe-based method for genotyping. *ADH1B* and *FTO* polymorphisms as examples.

Short communication

Key words: genotyping, PCR-RFLP, KASP™ genotyping assay, accuracy

J.A. HUBÁČEK[1], H. PIKHART[2], A. PEASEY[2], R. KUBÍNOVÁ[3], M. BOBÁK[2]

[1] Center for Experimental Medicine, Institute for Clinical and Experimental Medicine, Prague, Czech Republic.
[2] Department of Epidemiology and Public Health, University College London, London, United Kingdom
[3] National Institute of Public Health, Prague, Czech Republic

This work was supported by the project (Ministry of Health, Czech Republic) for the development of research organisation 00023001 (IKEM, Prague, Czech Republic) – Institutional support.

Corresponding author:
Jaroslav A. Hubacek; IKEM-DEM-LMG
Videnska 1958/9 140 21 Prague 4 Czech Republic
Phone: +420 261 363 379 Fax: +420 241 721 666 E-mail: jahb@ikem.cz

Abbreviations: ADH1B - alcohol dehydrogenase, DNA – deoxyribonucleic acid, EDTA - ethylenediaminetetraacetic acid, FTO – fat mass and obesity related protein, RFLP – restriction fragment length polymorphism, SNP - single nucleotide polymorphism.

Abstract.

DNA genotyping is among the most common analyses currently performed in scientific research. Two high throughput genotyping techniques are widely used – the "classic" PCR-RFLP and probe based methods such as TaqMan(®) PCR assay or KASP™ genotyping. The probe based techniques are claimed to be more accurate then PCR-RFLP; however, the evidence for this claim is sparse. We have directly compared results of genotyping of two SNPs (rs1229984 and rs17817449) obtained by the PCR-RFLP and KASP™ on 1,502 adult Caucasians. The results were identical in 97.3% and 95.9% cases. Discrepancies (either different results or result obtained with one but not with the second method) were addressed by confirmatory analysis using by direct sequencing. The sequencing revealed that both methods can give incorrect results but the frequency of incorrect genotyping of rs1229984 and rs17817449 was very low for both methods – 0.1% and 0.5%, respectively, for PCR-RFLP and 0.1% and 0.3%, respectively, for KASP™. These results confirm that the probe-based technique is slightly more accurate but it achieves slightly lower call rates than PCR-RFLP. When carefully set up, both PCR-RFLP and KASP™ could have accuracy of 99.5% or higher.

Introduction

The analysis of the single nucleotide polymorphisms (SNPs) is among the most common laboratory procedures currently used in molecular genetic research. The methods for genotyping developed distinctly over the last few decades. The radioactively labelled probes (southern blotting) (Southern 1975) were very expensive and the whole procedure was time- and material-consuming. When the polymerase chain reaction (PCR) method was described (Saiki et al., 1985), it became the gold standard for all subsequent genotyping approaches. First, and probably still the most commonly used method, is the restriction of the PCR product with exact bacterial endonucleases, the polymerase chain reaction – restriction fragment length polymorphism method (PCR-RFLP) (Shi et al. 1999). More recently, „modern“ real time PCR methods, using different modification of probes hybridising to the PCR product, are rapidly expanding (Shi et al. 1999, Jenkins and Gibson, 2002). The commonly cited is a TaqMan(®) PCR assay (Maubaret et al. 2013, Dušátková et al. 2013). Other methods, such as high resolution melting (Obeidová et al. 2012, Safaříková et al. 2013), NanoChip electronic microarray (Schrijver et al. 2003) or direct sequencing (Ohmoto et al., 2014 Tomašov et al. 2014) are also based on the analyses of PCR products. These methods are sometimes cheaper (although analysers are usually more expensive), may be quicker (depending on the number of samples and equipment) and are often claimed to be much more accurate than the „classic” PCR-RFLP (Ali et al. 2010). However, exact assessments of the methods’ accuracy are sparse and inconclusive (Osaki et al. 2011, Johnson et al. 2004, Bianchi et al. 2010). PCR-RFLP method remains widely accepted in impact journals (for example Hubacek et al. 2013, Bloudickova et al. 2014, Holmes et al. 2014, Ergen et al. 2014, Zheng et al. 2014, Drogari et al. 2014, Yenmis et al. 2015), but most

recent meta-analyses or consortia seem to prefer probe based methods of genotyping (see e.g. Patel et al. 2014).

More recently, the KASP method for genotyping becomes popular. The method is based on competitive allele-specific PCR amplification with one universal primer and two allele specific, differentially labelled primers (FAM™ and HEX™).

In this investigation, we compared the accuracy of these two types of genotyping in a large group if adult individuals, using two SNPs: rs1229984 (G>A; Arg47>His exchange with a significant effect on the activity of the enzyme and associated with alcohol consumption) within the alcohol dehydrogenase gene (*ADH1B*, OMIM acc N. 103720) and rs17817449 (G>T substitution within the 1st intron of the gene, with a strong impact on body weight; each T allele is associated with a mean of about 1.2 kg of body weight growth) within the fat mass and obesity associated gene (*FTO*, OMIM acc N. 610966).

Material and Methods

Two polymorphisms (rs1229984 and rs17817449) were genotyped in 1,502 individuals, a random subsample of the Czech branch of the HAPIEE study (Peasey et al. 2006), independently by two methods: PCR-RFLP (for exact details see Hubacek et al., 2008 and Hubacek et al., 2012) and KASP™ genotyping assay (LGS Genomics, Germany) were used.

Briefly, primers 5′ ACA ATC TTT TCT GAA TCT GAA CAG CTT CTC and 5′ TTG CCA CTA ACC ACG TGG TCA TCT GCG) were used to amplify a 97 bp fragment of the ADH1B gene containing the rs1229984 polymorphisms. PCR product was cut with

the restriction enzyme Hin6I; restriction fragments of 65bp and 27bp refer to the common G allele, while uncut PCR product is characteristic for the allele A.

For PCR-RFLP analysis of the *FTO* rs17817449 variant, primers 5′ GGT GAA GAG GAG GAG ATT GTG TAA CTG G and 5′ GAA GCC CTG AGA AGT TTA GAG TAA ATT GGG were used. This fragment (198 bp) was cut with the restriction enzyme *Alw*NI (uncut PCR product 198 bp represent allele G, restriction fragments of 99 bp and 99 bp allele T).

In both cases, PCR products have not been purified before the restriction and restriction fragments have been separated on 10% polyacrylamide gele using the MADGE electrophoresis (Day and Humphries, 1994).

All used chemicals were produced by Fermentas, Burlington, Canada. Restrictions have been performed according the conditions as recommended by the manufacturer on untreated PCR product.

For the KASP™ genotyping assay, the universal KASP Master mix was added to the DNA samples (http://www.lgcgroup.com/products/kasp-genotyping-chemistry/#.VT-sCmdO7Z4). For the allelic specific amplification of the rs1229984 SNP common primer 5’ GKT TGC CAC TAA CCA CGT GGT CAT was used with either 5’ ATG GTG GCT GTA GGA ATC TGT CA (allele A specific) or 5’ GGT GGC TGT AGG AAT CTG TCG (allele G specific).

For the allelic specific amplification of the rs17817449 SNP common primer 5’ CTT TGT GTT TCA GCT TGG CAC ACA GAA was used with either 5’ AAG GAG CTG GAC TGT TAA ATT AAA ACC (allele G specific) or 5’ AAA GGA GCT GGA CTG TTA AAT TAA AAC A (allele T specific).

Results obtained by different genotyping methods were compared and in the case of the discrepancy, or in the case that only one method was successful, PCR product was reamplified, treated by MinElute PCR purification Kit (Qiagen, Hilden, Germany) and custom sequenced by GATC Biotech, Germany. Oligonucleotides used for sequencing have been as follows 5′ ACA ATC TTT TCT GAA TCT GAA CAG CTT CTC for ADH1B polymorphism and 5′ GGT GAA GAG GAG GAG ATT GTG TAA CTG G for the *FTO* polymorphism. Sequencing results were than compared with the original results.

Results and Discussion

For both polymorphisms, both methods achieved call rates between 97.5% and 99.3%. Surprisingly, slightly higher call rates were observed for the PCR-RFLP method for both SNPs.

In the case of the *ADH1B* gene (rs1229984), 97.3% of the samples were genotyped with identical results (Table 1). For 2 samples results could not be obtained with any method. In the five cases where discrepancies were observed, in four cases direct sequencing confirmed the KASP™ genotyping results and in one case the PCR-RFLP was correct. Eight 8 samples were genotyped successfully with KASP™ but not with PCR-RFLP; of these, 4 were correct and in 4 cases sequencing failed. Among 17 cases where KASP™ genotyping assay failed but which were successfully genotyped by PCR-RFLP, 13 genotyped correctly, no sample was misgenotyped and in 4 cases the sequencing failed.

In the case of the *FTO* gene (rs17817449), 95.9% of samples were genotyped with identical results (Table 2). For 2 samples results were not obtained with either

method. Results were discrepant in 12 cases; among these, sequencing confirmed that KASP™ genotyping assay was correct in 9 cases and PCR-RFLP was correct in 1 case (however, see paragraph below). One sample was incorrect both if genotyped by KASP™ or PCR-RFLP and in one case the sequencing failed. From 12 samples genotyped successfully with KASP™ genotyping assay but not with PCR-RFLP, were 4 correct, two were misgenotyped and in 6 cases sequencing was not successful. Among 37 cases where KASP™ genotyping assay failed but which were successfully genotyped with PCR-RFLP, 27 were genotyped correctly, 8 were misgenotyped and in two cases sequencing did not provide clear results.

To detect the possible mistakes caused by the “human factor”, we search the archives for results where PCR-RFLP results and KASP™ genotyping assay/sequencing were not identical. We detected one mistake in *ALD1B* and 5 mistakes in the *FTO* procedures where DNA were correctly genotyped by PCR-RFLP but incorrectly entered in the database (although the database has been checked three times by three different members of staff).

The final frequency of incorrect genotyping was very low for both methods – 0.1% and 0.5%, respectively, for PCR-RFLP and 0.1% and 0.3%, respectively, for KASP™. The total numbers of correct, incorrect and not verified/unclear results are summarized in Table 3 and Figure 1.

In all cases of discordance between results, subsequent electrophoretic analysis (0.7% agarose in tris-EDTA buffer) revealed that all these DNA samples were partially degraded. This suggests that a careful processing, storage and/or selection of DNA prior analysis can further minimize the risk of false results for any genotyping method. In the case of the PCR-RFLP method, a careful quality control through manual allelic detection

is particularly important. We hypothesize that some of the discordant results reported in the literature (Osaki et al. 2011, Johnson et al. 2004, Bianchi et al. 2010) could be the results from suboptimal PCR or restriction analysis conditions.

Based on the results of this investigation, we conclude that the relatively “old fashioned” PCR-RFLP method is well suited to genotype DNA samples with very high accuracy, comparable with “modern” KASP™ genotyping assay.

References

Ali, M.M., Hasan, F., Ahmad, S., Al-Nakib, W. (2010) Comparative evaluation of INNO-LiPA HBV assay, direct DNA sequencing and subtractive PCR-RFLP for genotyping of clinical HBV isolates. *Virol. J.* **7,** 111.

Bianchi, M., Emanuele, E., Davin,A., Gagliardi, S., Cova, E., Meli, V., Trotti, R., Cereda, C. (2010) Comparison of three methods for genotyping of prothrombotic polymorphisms. *Clin. Exp. Med.* **10,** 269-272.

Bloudíčková, S., Kuthanová, L., Hubáček, J.A. (2014) MIA group. MTHFR and HFE, but not preproghrelin and LBP, polymorphisms as risk factors for all-cause end-stage renal disease development. *Folia Biol. (Praha).* **60,** 83-88.

Day, I.N., Humphries, S.E. (1994) Electrophoresis for genotyping: microtiter array diagonal gel electrophoresis on horizontal polyacrylamide gels, hydrolink, or agarose. *Anal. Biochem.* 222, 389-395.

Dušátková, L., Zamrazilová, H., Sedláčková, B., Včelák, J., Hlavatý, P., Aldhoon Hainerová, I., Korenková, V., Bradnová, O., Bendlová, B., Kunešová, M., Hainer, V. (2013) Association of obesity susceptibility gene variants with metabolic syndrome and related traits in 1,443 czech adolescents. *Folia Biol. (Praha).* **59,** 123-133.

Drogari, E., Ragia, G., Mollaki, V., Elens, L., Van Schaik, R.H., Manolopoulos, V.G. (2014) POR*28 SNP is associated with lipid response to atorvastatin in children and adolescents with familial hypercholesterolemia. *Pharmacogenomics.* **15,** 1963-1972.

Ergen, A., Karagedik, H., Karaali, Z.E., Isbir, T. (2014) An association between MPO -463 G/A polymorphism and type 2 diabetes. *Folia Biol. (Praha).* **60,** 108-112.

Holmes, M.V., Frikke-Schmidt, R., Melis, D., Luben, R., Asselbergs, F.W., Boer, J.M., Cooper, J., Palmen, J., Horvat, P., Engmann, J., Li, K.W., Onland-Moret, N.C., Hofker, M.H., Kumari, M., Keating, B.J., Hubacek, J.A., Adamkova, V., Kubinova, R., Bobak, M., Khaw, K.T., Nordestgaard, B.G., Wareham, N., Humphries, S.E., Langenberg, C., Tybjaerg-Hansen, A., Talmud, P.J. (2014) A systematic review and meta-analysis of 130,000 individuals shows smoking does not modify the association of APOE genotype on risk of coronary heart disease. *Atherosclerosis.* **237,** 5-12.

Hubacek, J.A., Bohuslavova, R., Kuthanova, L., Kubinova, R., Peasey, A., Pikhart, H., Marmot, M.G., Bobak, M. (2008) The FTO gene and obesity in a large Eastern European population sample: the HAPIEE study. *Obesity (Silver Spring).* **16,** 2764-2766.

Hubacek, J.A., Pikhart, H., Peasey, A., Kubinova, R., Bobak, M. (2012) ADH1B polymorphism, alcohol consumption, and binge drinking in Slavic Caucasians: results from the Czech HAPIEE study. *Alcohol Clin. Exp. Res.* **36,** 900-905.

Hubáček, J.A., Neškudla, T., Klementová, M., Adámková, V. (2013) Tagging rs10811661 variant at CDKN2A/2B locus is not associated with type 2 diabetes mellitus in Czech population. *Folia Biol. (Praha)* **59,** 168-171.

Jenkins, S., Gibson, N. (2002) High-throughput SNP genotyping. *Comp. Funct. Genomics.* **3,** 57-66.

Johnson, V.J., Yucesoy, B., Luster, M.I. (2004) Genotyping of single nucleotide polymorphisms in cytokine genes using real-time PCR allelic discrimination technology. *Cytokine* **27,** 135-141.

Maubaret, C.G., Salpea, K.D., Romanoski, C.E., Folkersen, L., Cooper, J.A., Stephanou, C., Li, K.W., Palmen, J., Hamsten, A., Neil, A., Stephens, J.W., Lusis, A.J., Eriksson, P., Talmud, P.J., Humphries, S.E., Simon Broome Research Group; EARSII consortium. (2013) Association of TERC and OBFC1 haplotypes with mean leukocyte telomere length and risk for coronary heart disease. *PLoS One.* **8,** e83122.

Obeidová, L., Reiterová, J., Lněnička, P., Štekrová, J., Šafránková, H., Kohoutová, M., Tesař, V. (2012) TRPC6 gene variants in Czech adult patients with focal segmental glomerulosclerosis and minimal change disease. *Folia Biol. (Praha).* **58,** 173-176.

Ohmoto, M., Takahashi, T., Kubota, Y., Kobayashi, S., Mitsumoto, Y. (2014) Genetic influence of dopamine receptor, dopamine transporter, and nicotine metabolism on smoking cessation and nicotine dependence in a Japanese population. *BMC Genet.* **15,** 151.

Osaki, R., Imaeda, H., Ban, H., Aomatsu, T., Bamba, S., Tsujikawa, T., Sasaki, M., Fujiyama, Y., Andoh, A. (2011) Accuracy of genotyping using the TaqMan PCR assay for single nucleotide polymorphisms responsible for thiopurine sensitivity in Japanese patients with inflammatory bowel disease. *Exp. Ther. Med.* **2,** 783-786.

Patel, R.S., Asselbergs, F.W., Quyyumi, A.A., Palmer, T.M., Finan, C.I., Tragante, V., Deanfield, J., Hemingwaz, H., Hingorani, A.D., Holmes, M.V. (2014) Genetic variants at chromosome 9p21 and risk of first versus subsequent coronary heart

disease events: a systematic review and meta-analysis. *J. Am. Coll. Cardiol.* **63,** 2234-2245

Peasey, A., Bobak, M., Kubinova, R., Malyutina, S., Pajak, A., Tamosiunas, A., Pikhart, H., Nicholson, A., Marmot, M. (2006) Determinants of cardiovascular disease and other non-communicable diseases in Central and Eastern Europe: rationale and design of the HAPIEE study. *BMC Public Health.* **6,** 255.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., Arnheim, N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*. **230,** 1350-1354.

Schrijver, I., Lay, M.J., Zehnder, J.L. (2003) Diagnostic single nucleotide polymorphism analysis of factor V Leiden and prothrombin 20210G > A. A comparison of the Nanogen Eelectronic Microarray with restriction enzyme digestion and the Roche LightCycler. *Am. J. Clin. Pathol.* **119,** 490-496.

Shi, M.M., Bleavins, M.R., de la Iglesia, F.A. (1999) Technologies for detecting genetic polymorphisms in pharmacogenomics. *Mol. Diagn.* **4,** 343-351.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98,** 503–517.

Šafaříková, M., Reiterová, J., Šafránková, H., Štekrová, J., Zidková, A., Obeidová, L., Kohoutová, M., Tesař, V. (2013) Mutational analysis of ACTN4, encoding α-actinin 4, in patients with focal segmental glomerulosclerosis using HRM method. *Folia Biol. (Praha).* **59,** 110-115.

Tomašov, P., Minárik, M., Zemánek, D., Čadová, P., Homolová, S., Čurila, K., Pěnička, M., Benešová, L., Belšánová, B., Gregor, P., Veselka, J. (2014) Genetic testing in

the management of relatives of patients with hypertrophic cardiomyopathy. *Folia Biol. (Praha).* **60,** 28-34.

Yenmis, G., Oner, T., Cam, C., Koc, A., Kucuk, O.S., Yakicier, M.C., Dizman, D., Kanigur Sultuybek, G. (2015) Association of NFKB1 and NFKBIA polymorphisms in relation to susceptibility of Behçet's disease. *Scand. J. Immunol.* **81,** 81-86.

Zheng, H., Xie, N., Xu, H., Huang, J., Xie, X, Luo, M. (2013) Association of Gln223Arg polymorphism of the leptin receptor with hypertensive left ventricular hypertrophy. *Folia Biol. (Praha)* **59**, 246-252.

Figure legend

Summary of the results. Total numbers of correctly genotyped samples by both methods; confirmed results by at least one method; and unclear results are summarized under a) for *ADH1B* rs1229984 polymorphism and under b) for *FTO* rs17817449 polymorphism.

Figure 1a

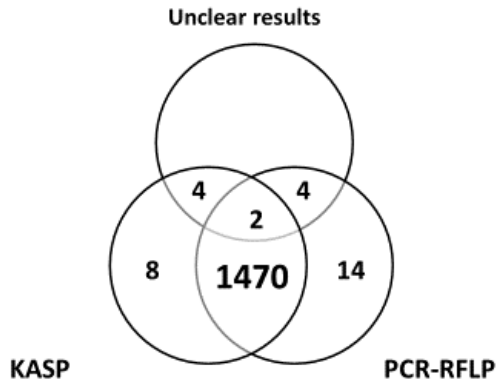

Figure 1b

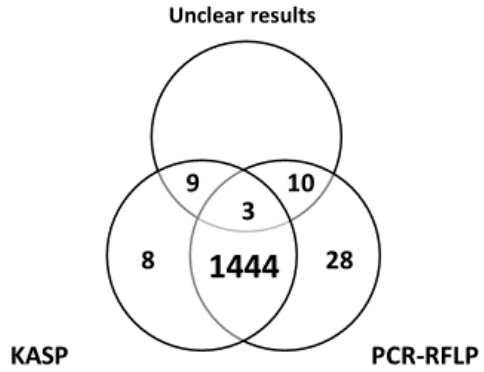

Table 1 Comparison of the results obtained by PCR-RFLP and KASP™ genotyping assay; rs1229984 polymorphism within the ADH1B gene.

KASP™	PCR-RFLP				
ADH GENOTYPE	AA	AG	GG	RESULT UNCLEAR	TOTAL
AA	**4**	0	0	0	4
AG	0	**153**	2	1	156
GG	0	3	**1.313**	7	1323
RESULT UNCLEAR	0	11	6	2	19
TOTAL	4	167	1.321	10	1.502

Table 2. Comparison of the results obtained by PCR-RFLP and KASP™ genotyping assay; rs17817449 polymorphism within the FTO gene.

KASP™	PCR-RFLP				
FTO GENOTYPE	GG	TG	TT	RESULT UNCLEAR	TOTAL
GG	**302**	5	0	3	310
TG	1	**693**	2	5	701
TT	1	3	**444**	4	452
RESULT UNCLEAR	1	34	2	2	39
TOTAL	305	735	448	14	1.502

Table 3. Final summary of correct results obtained by PCR-RFLP and KASP™ genotyping assay. Percentages are calculated for genotyped samples/total number of examined samples.

		Correct		Misgenotyped		Unclear/failed	
		N	%	N	%	N	%
PCR-RFLP	*ADH1B*	1,485	99.9/98.9	2	0.1/0.1	15	0.0/1.0
	FTO	1,472	99.5/98.0	13	0.5/0.9	17	0.0/1.1
KASP™	*ADH1B*	1,478	99.9/98.4	1	0.1/0.1	23	0.0/1.5
	FTO	1,452	99.7/96.7	4	0.3/0.3	46	0.0/3.1