

Elsevier Editorial System(tm) for Journal of Chromatography A
Manuscript Draft

Manuscript Number:

Title: Ultrafast haplotyping of putative microRNA-binding sites in the WFS1 gene by multiplex PCR and capillary gel electrophoresis

Article Type: MSB 2012

Keywords: WFS1 gene; miRNA-binding site; haplotype; SNP; double-tube allele-specific PCR; capillary gel electrophoresis

Corresponding Author: Professor Andras Guttman,

Corresponding Author's Institution: University of Debrecen

First Author: Marta Kerekyarto

Order of Authors: Marta Kerekyarto; Nora Nemeth; Tamas Kerekes; Zsolt Ronai; Andras Guttman

Suggested Reviewers: Julia Khandurina PhD
Senior scientist, Genomatica
jkhandurina@genomatica.com
Genomic expertises

Huba Kalasz PhD
Professor , United Arab Emirates University
huba.huba.kalasz@gmail.com
Separation science

Eva Szoko PhD
Professor , Semmelweis University
eva.szoko@net.sote.hu
Capillary electrophoresis expertises

Opposed Reviewers:

Highlights

- Simultaneous haplotyping of two miRNA-binding sites in the WSF1 is presented.
- A combination allele-specific amplification and capillary electrophoresis was used.
- Ultra-fast size determination of the generated PCR fragments was done by CGE.
- Excellent detection limit of 2 ng/ml was demonstrated.

1 Ultrafast haplotyping of putative microRNA-binding sites in the WFS1 gene 2 by multiplex PCR and capillary gel electrophoresis

3
4 Márta Kerékgyártó^a, Nóra Németh^b, Tamás Kerekes^c, Zsolt Rónai^b and András Guttman^{a*}
5

6 ^a*Horváth Laboratory of Bioseparation Sciences, University of Debrecen, H-4032 Debrecen,*
7 *Nagyerdei krt. 98, Hungary;* ^b*Department of Medical Chemistry, Semmelweis University, H-*
8 *1094 Budapest, Tuzolto u. 37-47, Hungary;* ^c*Centre for Clinical Genomics and Personalized*
9 *Medicine, Medical and Health Science Center, University of Debrecen, H-4032 Debrecen,*
10 *Nagyerdei krt. 98, Hungary*

11
12 * To whom correspondence should be addressed: Phone: +36(52) 414-717/64182; Fax: +36 (52)
13 414-717/55539; E-mail: guttman.andras@hlbs.org
14
15

16 **ABSTRACT**

17 The transmembrane protein wolframin (WSF1) plays a crucial role in cell integrity in pancreatic
18 beta cells and maintaining ER homeostasis. Genetic variations in the WFS1 gene have been
19 described to be associated with Wolfram syndrome or type 2 diabetes mellitus. In this paper we
20 report on an efficient double-tube allele-specific amplification method in conjunction with
21 ultrafast capillary gel electrophoresis for direct haplotyping analysis of the SNPs in two
22 important miRNA-binding sites (*rs1046322* and *rs9457*) in the WFS1 gene. An automated
23 single-channel capillary gel electrophoresis system was utilized in the method that provided
24 dsDNA fragment analysis in less than 240 sec. The light-emitting diode induced fluorescence
25 (LEDIF) detection system enabled excellent sensitivity for automated haplotyping of a large
26 number of clinical samples. The detection limit was 0.002 ng/μL using field amplified injection
27 from water diluted samples. The dynamic quantitation range was 0.08 - 10.00 ng/μL ($R^2=0.9997$)
28 in buffer diluted samples.
29
30

31 **Keywords:** WFS1 gene; miRNA-binding site; haplotype; SNP; double-tube allele-specific PCR;
32 capillary gel electrophoresis
33

34 **1. Introduction**

35

36 Wolframin (WFS1) is a transmembrane protein in the endoplasmic reticulum (ER), which is
37 produced at higher levels in pancreatic beta cells and specific neurons in the central nervous
38 system [1]. It plays an important role in the ER calcium homeostasis [1-3] and in the ER stress
39 response [4]. As an ER stress signaling suppressor it affects the negative regulatory feedback
40 loop of the ER stress signaling network [5], which is strictly controlled in pancreatic beta cells to
41 produce adequate amounts of insulin in case of blood glucose levels fluctuation [6,7]. Moreover

42 it plays an essential role in the cell integrity of pancreatic beta cells and maintains ER
43 homeostasis [8]. When the wolframin gene (WFS1) is inactivated in beta cells of rodents it
44 causes ER stress and death of the beta cells by accelerated apoptosis [9]. Mutations in the WFS1
45 gene are causing the so called Wolfram syndrome [10], which includes young onset non-
46 autoimmune insulin dependent diabetes mellitus, diabetes insipidus, optic atrophy, deafness or
47 other neurological and endocrine abnormalities [11,12]. An increased prevalence of diabetes
48 mellitus in reference to Wolfram syndrome was reported in first-degree relatives of patients [13],
49 suggesting a probable effect of WFS1 mutation heterozygosis. The occurrence of single
50 nucleotide polymorphisms (SNPs) in WFS1 has recently been demonstrated to be associated
51 with type 2 diabetes mellitus in populations of European descents [14-16].
52

53 MicroRNAs are non-coding short ribonucleic acids, which are responsible for the translation
54 regulation of gene expression. The homeostatic protein level is modified due to the interaction
55 between miRNAs and its targets, resulting in possible phenotype changes, such as disease. This
56 modified interaction can be caused by SNPs either in the gene of the miRNA or its target. SNPs
57 are rare in miRNA-coding genes [17] and referred to as miRSNPs. They were shown to be
58 associated with different illnesses, such as various types of cancers [18], autoimmune diseases
59 [19], or neurological disorders [20].
60

61 Multiplex PCR techniques gained recent popularity in assessing genetic variation by
62 simultaneous analysis of two or more DNA regions or genetic variations of interest [21,22].
63 Development of a multiplex PCR reaction involves the design of the relevant primer sets and
64 examination of their various combinations, different reaction components and/or thermal cycling
65 conditions. Multiplexing in this way increases the throughput of the amplification steps
66 especially when capillary gel electrophoresis is utilized with rapid separation and quantitation
67 capability for the analysis of the resulting fragments [23]. Multiplexing on the other hand may
68 lead to unequal amplification, particularly at the larger DNA fragment range, so the above
69 mentioned reaction design is of high importance [24].
70

71 Simultaneous study of multiple polymorphisms, such as haplotyping, is getting more and more
72 attention to analyze the genetic background of complex diseases [25]. Haplotype, the relative
73 chromosomal localization of the alleles of the polymorphic loci, can serve as very effective
74 genetic markers [26]. Haplotype identification can be accomplished by several ways. One of the
75 oldest methods is based on the theory of Mendelian inheritance of families or larger pedigrees;
76 however, this approach has several drawbacks [27,28]. Other methods, such as computer-based
77 haplotype prediction can also be suitable for haplotypes determination, but haplotypes of
78 individual samples cannot be obtained by this approach [28,29]. Direct haplotype determination
79 by allele-specific amplification (ASA), also referred to as molecular haplotyping, is one of the
80 most efficient and reliable methods that is based on appropriate amplification providing the
81 required haplotype information without the need of biological parents' genotype information
82 [30]. Moreover this technique provides fast and reliable genotyping data of any SNP in a single

83 tube polymerase chain reaction (mPCR) followed by electrophoresis analysis [31,32]. This
84 amplification method is based on the use of an allele-specific primer as its 3'-end hybridizes to
85 the SNP site. This is followed by amplification using a DNA-polymerase enzyme, which is
86 lacking 3'-exonuclease activity, thus, amplification can only be carried out in the case the primer
87 completely matches with the template. The technique when two allele-specific primers are used
88 for convenient allelic variant determination in two separate reactions is referred to as double-tube
89 specific-allele amplification [33]. This novel haplotyping technique was introduced earlier to
90 investigate the -616CG and -521CT SNPs in the Dopamine D4 Receptor gene by Szantai et al
91 [24]. The resulting DNA fragments after the amplification process are regularly analyzed by
92 conventional agarose/polyacrylamide slab gel electrophoresis for genotype or haplotype
93 determination. However, these methods are labor intensive and time consuming, also requiring
94 improvements in terms of resolving power and analysis throughput. Recent developments in the
95 field of capillary gel electrophoresis resulted in rapid electrophoresis-based fragment analysis
96 techniques that can readily speed up this process. In addition to its speed, capillary gel
97 electrophoresis offers further advantages over traditional slab gel electrophoresis, such as low
98 reagent consumption, small sample volume requirement and the option of multiplexing [30].
99 CGE combined with light-emitting diode induced fluorescence (LEDIF) detection enables
100 sensitive detection of dsDNA fragments and can be readily applied for automated large scale
101 analyses in clinical settings [34].

102
103 In this paper we report on haplotyping (i.e., simultaneous multiple genotyping) of two adjacent
104 putative miRNA-binding SNPs in the WFS1 gene by combining double-tube allele-specific
105 amplification and rapid capillary gel electrophoresis with LED-induced fluorescent detection to
106 analyze the resulting DNA fragments. The detection limit of the method was as low as 0.002
107 ng/ μ L using field amplified injection method.

108 109 **2. Materials and Methods**

110 **2.1. Chemicals**

111 The HotStar *Taq* DNA polymerase including the 10 \times reaction buffer and the Q-solution was
112 used from Qiagen (Valencia, CA, USA) for the allele-specific PCR reaction. The oligonucleotide
113 primers were obtained from Sigma Genosys (Woodlands, TX, USA). For agarose slab gel
114 electrophoresis, the 100 base pair DNA ladder (GeneRuler, Thermo Fisher Scientific, FL, USA)
115 was diluted to a final concentration of 0.5 μ g/ μ L and stored at -20 $^{\circ}$ C. In CGE separations, the
116 Qsep100 DNA-CE high-resolution gel buffer and Qsep100 DNA-CE running buffer were used
117 (BiOptic, New Taipei City, Taiwan). The DNA alignment marker (20 base pair, 1.442 ng/ μ L and
118 5000 base pair, 1.852 ng/ μ L) and the DNA size marker (50–3000 bases, 10.5 ng/ μ L) were from
119 BiOptic and stored at -20 $^{\circ}$ C. The WFS1 PCR samples (576 bp, 253.19 ng/ μ L) were diluted to
120 the appropriate concentrations with MilliQ-grade water (Millipore, Billerica, MA, USA) or
121 dilution buffer (BiOptic) for the detection limit and linearity studies and stored at -20 $^{\circ}$ C.

122

123 **2.2. Non-invasive DNA sampling and DNA extraction**

124 DNA samples were obtained using non-invasive DNA sampling (buccal swabs) from healthy
125 Hungarian volunteers. The study protocol was approved by the Scientific and Research Ethics
126 Committee of the Medical Research Council of Hungary (ETT TUKEB). DNA samples were
127 purified by standard procedure as described earlier [35,36].
128

129 **2.3. Molecular haplotype analysis**

130 Direct haplotype determination of the *rs1046322* and *rs9457* SNPs was carried out by allele-
131 specific amplification. The HotStarTaq polymerase kit (Qiagen) was used for the PCR
132 amplification and each DNA sample was analyzed in two separate reactions. Both reaction
133 mixtures contained approximately 4 ng gDNA template, 200 μ M deoxyadenosine triphosphate
134 (dATP), deoxycytidine triphosphate (dCTP), deoxyguanosine triphosphate (dGTP) and
135 deoxythymidine triphosphate (dTTP), 0.5 U HotStar *Taq* DNA polymerase with 1 \times reaction
136 buffer and 1 \times Q-solution, as well as 1 μ M of each outer primer (sense: 5' TCT GTC CAC TCT
137 GAA TAC 3' and antisense: 5' CAG GCT CTT CTA AAC ACT 3'). Reaction mixture-I was
138 used to analyze the presence of the *rs1046322A* and *rs9457C* alleles, as well as their haplotype,
139 thus it contained the *rs1046322A* specific sense (5' GAG CCT GAC CTT TCT GAA 3') and the
140 *rs9457C* specific antisense (5' CCA CTA CCT GCT GGA G 3') primers. Reaction mixture-II
141 was employed to investigate the other possible variants (*rs1046322G*-specific sense primer: 5'
142 GAG CCT GAC CTT TCT GAG 3', *rs9457G*-specific antisense primer: 5' CCA CTA CCT
143 GCT GGA C 3'). PCR amplification reactions were carried out in a total volume of 10 μ L. The
144 primers were tested by the Oligo 5.0 software (Molecular Biology Insides, Cascade, CO, USA).
145 Thermocycling was initiated at 95 $^{\circ}$ C for 15 min, this step also served for the activation of the
146 hot-start DNA polymerase. It was followed by 40 cycles of denaturation at 94 $^{\circ}$ C for 30 s,
147 annealing at 55 $^{\circ}$ C for 30 s and then extension at 72 $^{\circ}$ C for 1 min. The last step of the
148 amplification was a final extension at 72 $^{\circ}$ C for 10 min after that the PCR products were kept at
149 8 $^{\circ}$ C.
150

151 For the detection limit and linearity studies, the PCR reaction mixture contained approximately 4
152 ng gDNA template, 200 μ M of each deoxyribonucleotide triphosphate (dATP, dCTP, dGTP and
153 dTTP), 0.05 U/ μ L HotStar *Taq* DNA polymerase with 1 \times reaction buffer and 1 \times Q-solution, as
154 well as 1 μ M of each primer (sense: 5' GCC CTT CTC GAG TCT TGC AGC GCC GGA ATA
155 GGC 3' and antisense: 5' GCA GAA GCT TAA GTT GTT CGG GAG CAG CTG AAC G 3').
156 The amplification reaction was carried out in a total volume of 100 μ L. The first step was the
157 initial denaturation of the gDNA at 95 $^{\circ}$ C for 15 min; it was followed by 40 cycles of
158 denaturation (94 $^{\circ}$ C, 30 sec), annealing (65 $^{\circ}$ C, 30 sec) and then extension (72 $^{\circ}$ C, 1 min). The
159 last step of the PCR was a final extension at 72 $^{\circ}$ C for 10 min after that the sample was kept at 8
160 $^{\circ}$ C.
161

162

163 **2.4. PCR- fragment analysis by agarose slab gel electrophoresis**

164 The PCR products were first analyzed by agarose slab gel electrophoresis. Agarose powder (final
165 concentration: 2% w/v) was mixed with electrophoresis buffer (1× TAE buffer; 40 mM Tris, 20
166 mM acetic acid, and 1 mM EDTA, pH 8.0) and heated until the agarose completely dissolved.
167 Ethidium bromide was added to the melted gel in a final concentration of 0.5 µg/mL. After
168 solidification at room temperature, 20 ng of PCR products and the 100 bp DNA sizing ladder
169 (100–1000 bp, 0.5 µg/uL) containing DNA loading Dye (6× loading Dye: 10 mM Tris-HCl (pH
170 7.6), 0.03% bromophenol blue, 0.03% xylene cyanol FF, 60% glycerol and 60 mM EDTA) were
171 loaded into the sample wells followed by electrophoresis (100 V for 45 min, BioRad PowerPac
172 300; Hercules, CA, USA). The separated DNA bands were visualized in a UV light box (Bio
173 Rad Gel-Doc XR System).

174

175 **2.5 PCR- fragment analysis by capillary gel electrophoresis**

176 Rapid capillary gel electrophoresis analysis of the PCR-products were accomplished in a single-
177 channel capillary cartridge Qsep100 DNA-CE unit (BiOptic) with an 11 cm effective length
178 (total length: 15 cm) fused silica capillary (internal diameter: 75 µm). The capillary was washed
179 with 5 mL of 70 °C MilliQ-grade water (Millipore) for 500 s before the first use. Then 5 mL of
180 Qsep100 DNA-CE high-resolution gel buffer was transferred into the gel reservoir and the
181 capillary was purged twice for 1000 s. The gel-buffer system contained ethidium-bromide to
182 accommodate fluorescent detection. Prior to each injection the sieving matrix was replaced in the
183 capillary by means of a 10 s purge step, followed by injection of the DNA alignment marker (4
184 kV for 10 s). After that the separation capillary was immersed into MilliQ-grade water (0 kV for
185 1 sec) as a washing step to avoid any sample cross-contamination. The samples (as well as the
186 DNA size marker and/or PCR-products) were introduced electrokinetically from a 96-well plate
187 (4 kV for 10 s). Separations were carried out at ambient temperature by applying 6 kV electric
188 potential. Data analysis was performed using the Q-Expert software package (BiOptic). All
189 buffers and reagents were filtered through 0.22 µm pore size Acrodisc syringe filters (Millipore,
190 Billerica, MA, USA) and degassed prior to use. Other reagents and chemicals for sample
191 preparation were purchased from Sigma-Aldrich (St. Louis, MO, USA).

192

193 **3. Results and Discussion**

194 **3.1. Direct haplotype determination by allele-specific PCR**

195 Haplotype determination of adjacent polymorphic loci is of high importance, especially in case
196 of SNPs with biological significance. The *rs9457* and *rs1046322* SNPs, located in the WFS1
197 gene 3' UTR, are assumable miR-SNPs and their *in silico* data analysis suggested that they may
198 alter the binding of miR-185 and miR-668, respectively. Consequently in case of double
199 heterozygote samples (*rs1046322AG* and *rs9457CG*) haplotype determination is essential, since
200 otherwise it is uncertain if the two allelic variants possibly affecting miRNA-binding are located
201 on the same mRNA (“cis”) or can be found on two different chromosomes (“trans”) as
202 delineated in Figure 1.

203 An allele-specific PCR based approach was elaborated for the haplotype determination of the
204 two SNPs of interest. The principle of the technique was the simultaneous application of two
205 outer and two allele-specific primers in a multiplex PCR reaction as shown in Figure 2. The
206 allele-specific primers were designed to anneal to the SNP by their 3' end. Based on
207 chromosomal localization, a sense *rs1046322*- and antisense *rs9457*-specific primer were applied
208 in the reaction. One reaction mixture tested the presence of one allele at each loci as well as one
209 haplotype combination, consequently two reaction mixtures were required for genotype and
210 haplotype determination, whereas further two can be applied for conformation (Figure 2). Panel
211 A in Figure 2 depicts the analysis using reaction mixture-I containing the sense *rs1046322A*- and
212 the antisense *rs9457C*-specific primers. In case of the presence of an A allele at the *rs1046322*
213 site, a 488-bp-long fragment was generated by the *rs1046322A*-specific and the antisense outer
214 primers. Similarly if the sample possessed the C allele at the *rs9457* locus, the primer specific for
215 this variant together with the sense outer primer could amplify a 437-bp-long fragment. More
216 importantly, if the *rs1046322A* and *rs9457C* alleles are located on the same chromosome, a 384-
217 bp-long product could also be observed as this product is generated by the two allele-specific
218 primers and suggested the presence of the A–C haplotype. The longest, 541-bp outer fragment is
219 a control product synthesized independently of the genotype and haplotype of the sample of
220 interest.

221
222 Reaction mixture-II worked similarly; however, it contained the *rs1046322G*- and *rs9457G*-
223 specific primers in combination with the outer oligos as shown Figure 2 Panel B. Thus, a 488-bp-
224 long product could be observed in case of the *rs1046322G* allele, a 437-bp-long product
225 produced if the *rs9457G* allele was present, whereas the 384-bp-long product suggested the G–G
226 haplotype. Genotype and haplotype information could be unambiguously determined by these
227 two reactions. For additional validation, two redundant combinations were also applied in a
228 subset of 24 samples (i.e. *rs1046322G* allele-, *rs9457C* allele- and thus G–C haplotype specific
229 reaction and *rs1046322A* allele-, *rs9457G* allele- and consequently A–G haplotype specific
230 mixture). Results of these analyses confirmed the data obtained by the original setup. Then 95
231 healthy Hungarian individuals were analyzed by the described method and the obtained results
232 were in 98.9% concordance with the genotype data determined earlier by an independent
233 approach employing sequence specific TaqMan probes (data not shown). The single discordant
234 result could be resolved by a repeated genotype and haplotype determination.

235
236 Figure 2 Panel C shows the conventional agarose slab gel electrophoresis based genotype and
237 haplotype determination of the *rs9457* and *rs1046322* SNPs in case of double heterozygote
238 samples. The 100 bp DNA sizing marker (M) was used with the PCR samples (1 and 2) to assess
239 the size of the double allele-specific amplicons in the case of both haplotypes verification. One
240 of the haplotype (*rs1046322A-rs9457C*) was labeled with A and the other (*rs1046322G-rs9457G*)
241 was indicated with B in Figure 2 Panel C.

242

243 **3.2. Haplotype determination by capillary gel electrophoresis**

244 The final step of the haplotyping protocol was capillary gel electrophoresis based size
245 determination of the dsDNA fragments from the multiplex amplification reaction. Figure 3
246 depicts the capillary gel electrophoresis traces of the PCR fragments generated during haplotype
247 determination. A DNA sizing ladder in the range of 50–3000 bp was used for fragment size
248 assessment in a final concentration of 10.5 ng/ μ L (upper trace). The analysis of the mPCR
249 samples is shown in the middle and lower traces. The samples were coinjected with the lower
250 and upper alignment markers (M_1 : 20 bp dsDNA and M_2 : 5000 bp dsDNA) to attain high
251 fragment sizing accuracy. The middle trace in Figure 3 shows the separation of three dsDNA
252 fragments from the multiplex amplification reaction mixture-1 with calculated sizes of 454, 500
253 and 583 bp fragments, corresponding to 437, 488 and 541 bp of the actual PCR reactions (see
254 variance data in Table 1). The lower trace in Figure 3 depicts the separation of four dsDNA
255 fragments from amplification reaction mixture-2 with calculated sizes of 399, 457, 504 and 591
256 bp fragments (corresponding to the actual fragment sizes of 384, 437, 488 and 541 bp with better
257 than 95% average accuracy) by the rapid CE-LEDIF based method (see variance data in Table
258 1). In Table 1 the size (bp) of each multiplex PCR sample was calculated by Q-Expert software
259 package (BiOptic) with the accuracy range of 2.4–9.2%. Furthermore the concentration of each
260 DNA fragment was calculated based on their peak areas as listed in Table 1.

261

262 **3.3. Limit of detection (LOD) and detector linearity**

263 Figure 4, Panel A compares the resulting signal from the electropherograms after the injection of
264 different concentration samples from 0.01 ng/ μ L to as low as 0.002 ng/ μ L, this latter being the
265 detection limit. In this instance the dilution of the 576 bp DNA fragment was done in water.
266 When the detector linearity experiments were conducted with the same water diluted samples,
267 the linear detection range was quite narrow (1.5 orders of magnitude) due to the effect of field
268 amplification. Detection linearity was therefore determined by using a dilution series in sample
269 buffer (BiOptic) in which case a linear detector response was obtained in a large interval of 0.08
270 ng/ μ L to 10.0 ng/ μ L with an $R^2 = 0.9997$, as shown in Figure 4, Panel B and in Table 2. Again,
271 we would like to emphasize that injection from water diluted samples results in much larger
272 sample intake as the buffer co-ions do not compete with the sample molecules, resulting in
273 excellent LOD. Sample concentration measurement on the other hand was more precise from
274 buffer diluted samples as shown in Table 1.

275

276 **4. Conclusions**

277 Capillary gel electrophoresis is an automated, high-throughput DNA fragment analysis method
278 that can be readily applied for the investigation of a large number of samples. In this paper we
279 introduced a rapid CE-LEDIF based method in conjunction with multiplex PCR amplification for
280 genotyping and haplotyping of two important, adjacent miRNA-binding sites (*rs1046322* and
281 *rs9457*) in the WSF1 gene. The separation performance of the system was demonstrated by
282 ultrafast (<240 sec) and accurate (2.4–9.2%) sizing analysis of multiplex PCR samples, also
283 exhibiting excellent detector linearity ($R^2=0.9997$) from 0.08–10.0 ng/ μ L concentration. The

284 LOD of the system was 0.08 ng/ μ L for samples in dilution buffer and 0.002 ng/ μ L for samples in
285 water. In summary, this CGE-LEDIF system is a sensitive and easy to use bio-analytical tool for
286 automated haplotyping of a large number of clinical samples.

287

288 **5. Acknowledgements**

289 This project was supported by the Hungarian Grant OTKA grants of K81839 and K83766 as
290 well as the János Bolyai Research Scholarship (BO/00089/10/5) of the Hungarian Academy of
291 Sciences. Provision of the capillary gel electrophoresis system by BiOptic, Inc. is also greatly
292 appreciated. The authors have declared no conflicts of interest.

293

294 **6. References**

- 295 [1] K. Takeda, H. Inoue, Y. Tanizawa, Y. Matsuzaki, J. Oba, Y. Watanabe, K. Shinoda, Y.
296 Oka, *Hum Mol Genet* 10 (2001) 477.
- 297 [2] A.A. Osman, M. Saito, C. Makepeace, M.A. Permutt, P. Schlesinger, M. Mueckler, *Journal*
298 *of Biological Chemistry* 278 (2003) 52755.
- 299 [3] D. Takei, H. Ishihara, S. Yamaguchi, T. Yamada, A. Tamura, H. Katagiri, Y. Maruyama,
300 Y. Oka, *Febs Letters* 580 (2006) 5635.
- 301 [4] S.G. Fonseca, M. Fukuma, K.L. Lipson, L.X. Nguyen, J.R. Allen, Y. Oka, F. Urano,
302 *Journal of Biological Chemistry* 280 (2005) 39609.
- 303 [5] S.G. Fonseca, S. Ishigaki, C.M. Osowski, S. Lu, K.L. Lipson, R. Ghosh, E. Hayashi, H.
304 Ishihara, Y. Oka, M.A. Permutt, F. Urano, *J Clin Invest* 120 (2010) 744.
- 305 [6] S.G. Fonseca, K.L. Lipson, F. Urano, *Antioxid Redox Signal* 9 (2007) 2335.
- 306 [7] K.L. Lipson, S.G. Fonseca, S. Ishigaki, L.X. Nguyen, E. Foss, R. Bortell, A.A. Rossini, F.
307 Urano, *Cell Metab* 4 (2006) 245.
- 308 [8] N. Cheurfa, G.M. Brenner, A.F. Reis, D. Dubois-Laforgue, R. Roussel, J. Tichet, O.
309 Lantieri, B. Balkau, F. Fumeron, J. Timsit, M. Marre, G. Velho, *Diabetologia* 54 (2011)
310 554.
- 311 [9] A.C. Riggs, E. Bernal-Mizrachi, M. Ohsugi, J. Wasson, S. Fatrai, C. Welling, J. Murray,
312 R.E. Schmidt, P.L. Herrera, M.A. Permutt, *Diabetologia* 48 (2005) 2313.
- 313 [10] D.J. Wolfram, H.P. Wagener, *Mayo Clinic Proceedings* (1938) 715.
- 314 [11] T.G. Barrett, S.E. Bunday, A.F. Macleod, *Lancet* 346 (1995) 1458.
- 315 [12] L. Rigoli, F. Lombardo, C. Di Bella, *Clin Genet* 79 (2011) 103.
- 316 [13] F.C. Fraser, T. Gunn, *J Med Genet* 14 (1977) 190.
- 317 [14] P.W. Franks, O. Rolandsson, S.L. Debenham, K.A. Fawcett, F. Payne, C. Dina, P. Froguel,
318 K.L. Mohlke, C. Willer, T. Olsson, N.J. Wareham, G. Hallmans, I. Barroso, M.S. Sandhu,
319 *Diabetologia* 51 (2008) 458.
- 320 [15] V. Lyssenko, A. Jonsson, P. Almgren, N. Pulizzi, B. Isomaa, T. Tuomi, G. Berglund, D.
321 Altshuler, P. Nilsson, L. Groop, *New England Journal of Medicine* 359 (2008) 2220.
- 322 [16] M.S. Sandhu, M.N. Weedon, K.A. Fawcett, J. Wasson, S.L. Debenham, A. Daly, H. Lango,
323 T.M. Frayling, R.J. Neumann, R. Sherva, I. Blech, P.D. Pharoah, C.N. Palmer, C. Kimber,
324 R. Tavendale, A.D. Morris, M.I. McCarthy, M. Walker, G. Hitman, B. Glaser, M.A.
325 Permutt, A.T. Hattersley, N.J. Wareham, I. Barroso, *Nature Genetics* 39 (2007) 951.
- 326 [17] M.A. Saunders, H. Liang, W.H. Li, *Proceedings of the National Academy of Sciences of*
327 *the United States of America* 104 (2007) 3300.

- 328 [18] L.J. Chin, E. Ratner, S.G. Leng, R.H. Zhai, S. Nallur, I. Babar, R.U. Muller, E. Straka, L.
329 Su, E.A. Burki, R.E. Crowell, R. Patel, T. Kulkarni, R. Homer, D. Zelterman, K.K. Kidd,
330 Y. Zhu, D.C. Christiani, S.A. Belinsky, F.J. Slack, J.B. Weidhaas, *Cancer Research* 68
331 (2008) 8535.
- 332 [19] S. Tan, J.M. Guo, Q.L. Huang, X.P. Chen, J. Li-Ling, Q.W. Li, F. Ma, *Febs Letters* 581
333 (2007) 1081.
- 334 [20] K.P. Jensen, J. Covault, T.S. Conner, H. Tennen, H.R. Kranzler, H.M. Furneaux, *Molecular*
335 *Psychiatry* 14 (2009) 381.
- 336 [21] M.C. Edwards, R.A. Gibbs, *PCR Methods Appl* 3 (1994) S65.
- 337 [22] C.P. Kimpton, P. Gill, A. Walton, A. Urquhart, E.S. Millican, M. Adams, *PCR Methods*
338 *Appl* 3 (1993) 13.
- 339 [23] J.M. Butler, C.M. Ruitberg, P.M. Vallone, *Fresenius Journal of Analytical Chemistry* 369
340 (2001) 200.
- 341 [24] E. Szantai, A. Szilagyi, A. Guttman, M. Sasvari-Szekely, Z. Ronai, *J Chromatogr A* 1053
342 (2004) 241.
- 343 [25] B. Lewin, *Genes VI*, Oxford University Press, Oxford ; New York, 1997.
- 344 [26] E. Szantai, Z. Ronai, M. Sasvari-Szekely, A. Guttman, *Anal Biochem* 352 (2006) 148.
- 345 [27] C.L. Barr, Y. Feng, K.G. Wigg, R. Schachar, R. Tannock, W. Roberts, M. Malone, J.L.
346 Kennedy, *Am J Med Genet* 105 (2001) 84.
- 347 [28] S.E. Hodge, M. Boehnke, M.A. Spence, *Nature Genetics* 21 (1999) 360.
- 348 [29] L. Excoffier, M. Slatkin, *Molecular Biology and Evolution* 12 (1995) 921.
- 349 [30] E. Szantai, Z. Ronai, A. Szilagyi, M. Sasvari-Szekely, A. Guttman, *J Chromatogr A* 1079
350 (2005) 41.
- 351 [31] Z. Ronai, C. Barta, A. Guttman, K. Lakatos, J. Gervai, M. Staub, M. Sasvari-Szekely,
352 *Electrophoresis* 22 (2001) 1102.
- 353 [32] Z. Ronai, E. Szantai, R. Szmola, Z. Nemoda, A. Szekely, J. Gervai, A. Guttman, M.
354 Sasvari-Szekely, *American Journal of Medical Genetics Part B-Neuropsychiatric Genetics*
355 126B (2004) 74.
- 356 [33] E. Szantai, A. Guttman, *Electrophoresis* 27 (2006) 4896.
- 357 [34] M. Kerekyarto, T. Kerekes, E. Tsai, V.D. Amirkhanian, A. Guttman, *Electrophoresis* 33
358 (2012) 2752.
- 359 [35] Z. Ronai, A. Guttman, Z. Nemoda, M. Staub, H. Kalasz, M. Sasvari-Szekely,
360 *Electrophoresis* 21 (2000) 2058.
- 361 [36] M. Sasvari-Szekely, A. Gerstner, Z. Ronai, M. Staub, A. Guttman, *Electrophoresis* 21
362 (2000) 816.

363
364
365
366
367
368
369
370
371
372

373 **Figure Captions**

374

375 **Figure 1. Schematic representation of the putative effect of the two SNPs on miRNA**
376 **binding.** Genotypes and haplotypes can be determined by allele-specific amplification using
377 sense *rs1046322*- and antisense *rs9457*-specific primers in different combination in case of
378 double heterozygote samples (Sample 1 and Sample 2). The four thick gray lines indicate the
379 four haplotypes on the same chromosomes: A–C, G–G, A–G and G–C.

380

381 **Figure 2. Allele-specific multiplex PCR-based direct haplotype determination of the**
382 ***rs1046322* and *rs9457* SNPs in the *WSF1* gene. (A)** Fragments expected in the presence of the
383 sense *rs1046322A*- and the antisense *rs9457C*-specific primers in combination with the outer
384 oligos. The 384-bp-long fragment was generated by *rs1046322A* and *rs9457C* primers and
385 demonstrated the presence of the A–C haplotype. **(B)** PCR products obtained in the presence of
386 the *rs1046322G*- and *rs9457G*-specific primers in combination with the outer oligos. The 384-
387 bp-long product demonstrated the presence of the G–G haplotype. **(C)** Genotype and haplotype
388 readings by agarose slab gel electrophoresis. M: 100 base pair DNA sizing marker; Lanes 1-2:
389 PCR samples: A: 437, 488, 541 bp dsDNA fragments; B: 384, 437, 488, 541 bp dsDNA
390 fragments from the multiplex amplification reaction. Separation conditions: 2% agarose gel in 1
391 × TAE containing 0.5 µg/mL ethidium bromide; U=100V; t=45 min; room temperature.

392

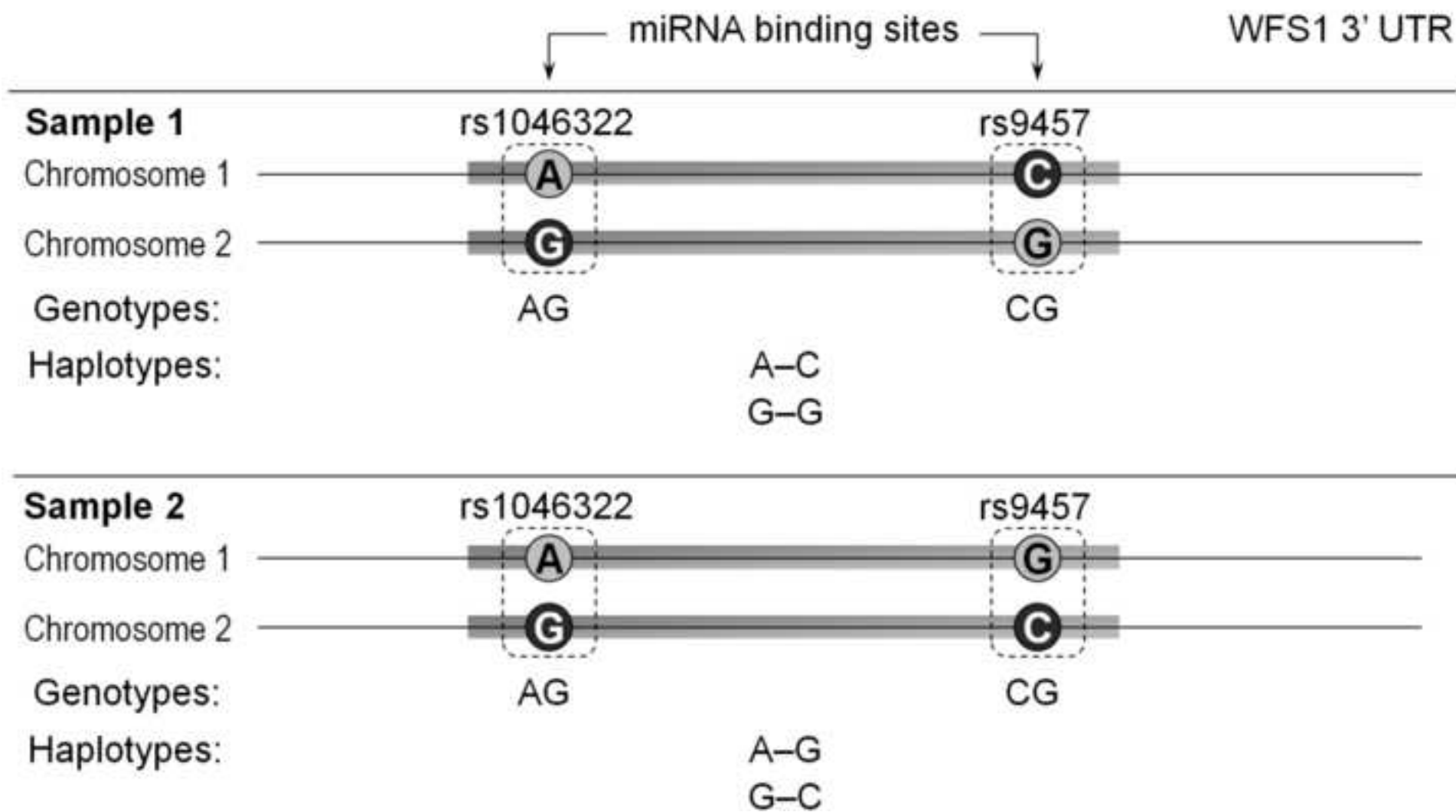
393 **Figure 3. Capillary gel electrophoresis based fragment analysis of representative multiplex**
394 **PCR amplicons.** Upper trace: DNA sizing ladder (M: 50 to 3000 bp) co-injected with the lower
395 (M1 = 20 bp) and upper (M2 = 5000 bp) alignment markers; Middle and lower traces:
396 representative PCR fragments of 1 and 2 were the same as in Figure 2, respectively, with the
397 respective alignment markers. Separation conditions: marker and sample injection: 4kV/10sec;
398 separation voltage 6 kV; capillary: 75-µm i.d., total length of 15 cm length (effective separation
399 length: 11 cm); ambient temperature.

400

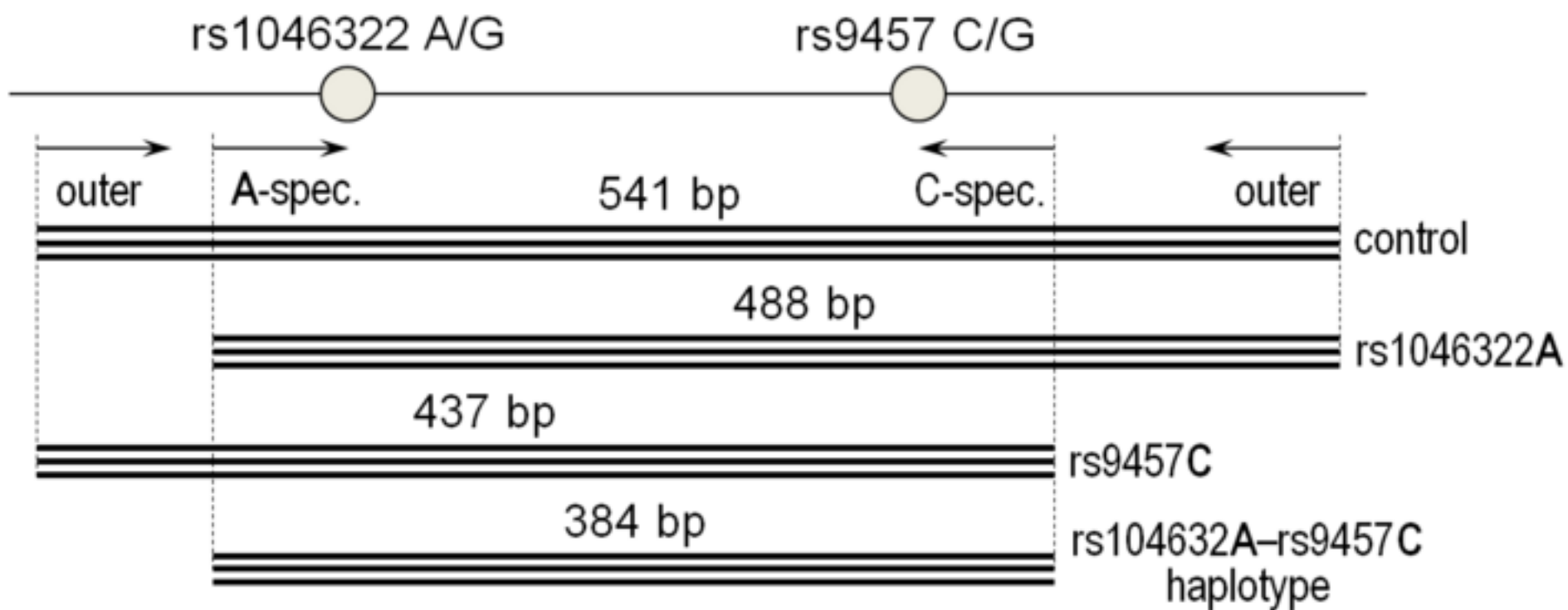
401 **Figure 4. LOD and detection linearity measurements. (A)** Determination of the limit of
402 detection (LOD) with a representative PCR fragment (576 bp) serially diluted in water compared
403 to the sizing ladder. **(B)** Detection linearity study using the of the 576 bp PCR fragment serially
404 diluted in the sample buffer. Separation conditions, sizing ladder and lower and upper alignment
405 markers were the same as in Figure 3.

406

Figure 1
[Click here to download high resolution image](#)



Reaction mixture I



Reaction mixture II

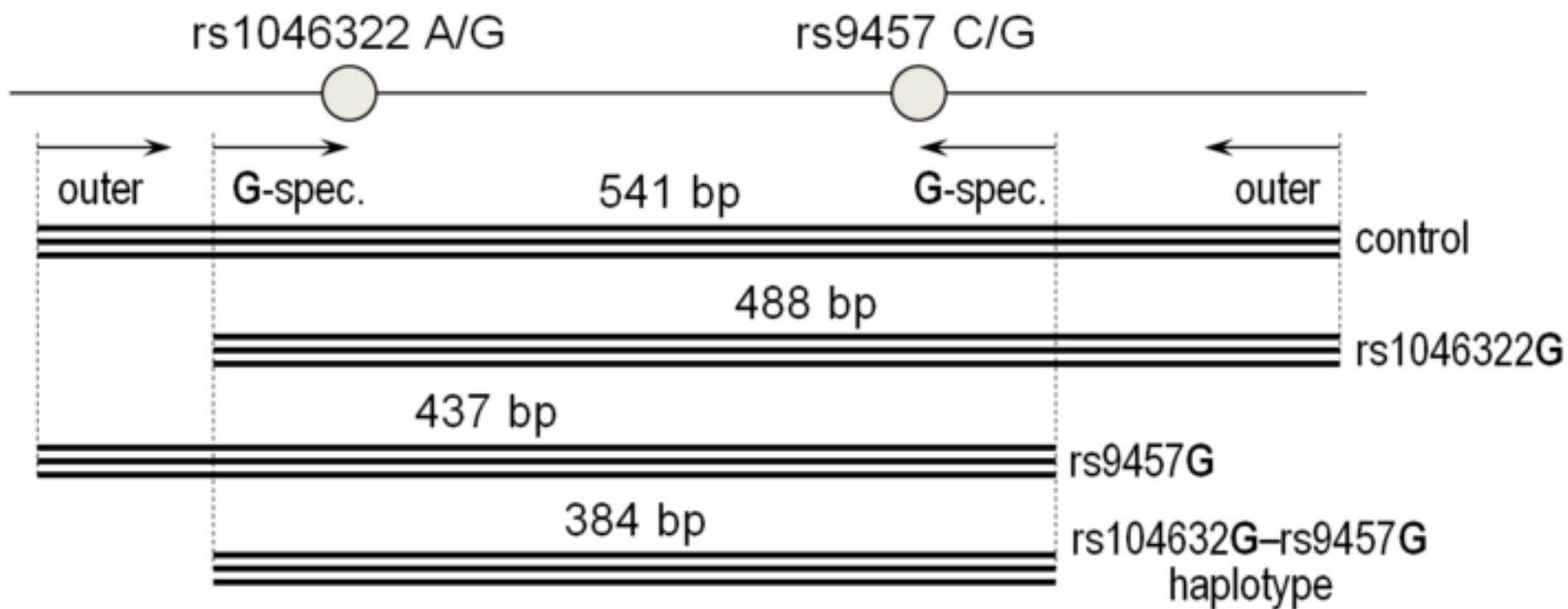


Figure 2C
[Click here to download high resolution image](#)

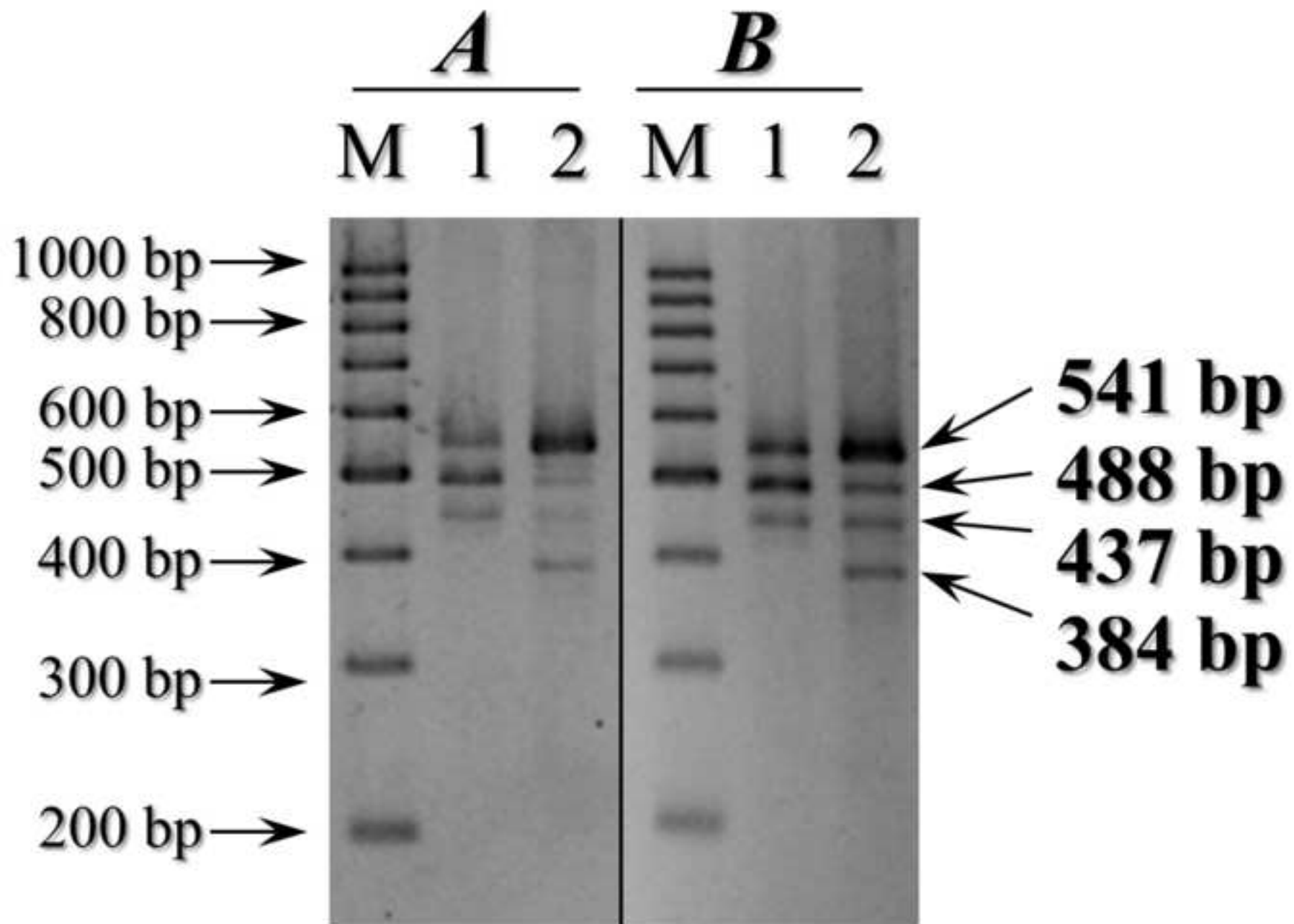


Figure 3
[Click here to download high resolution image](#)

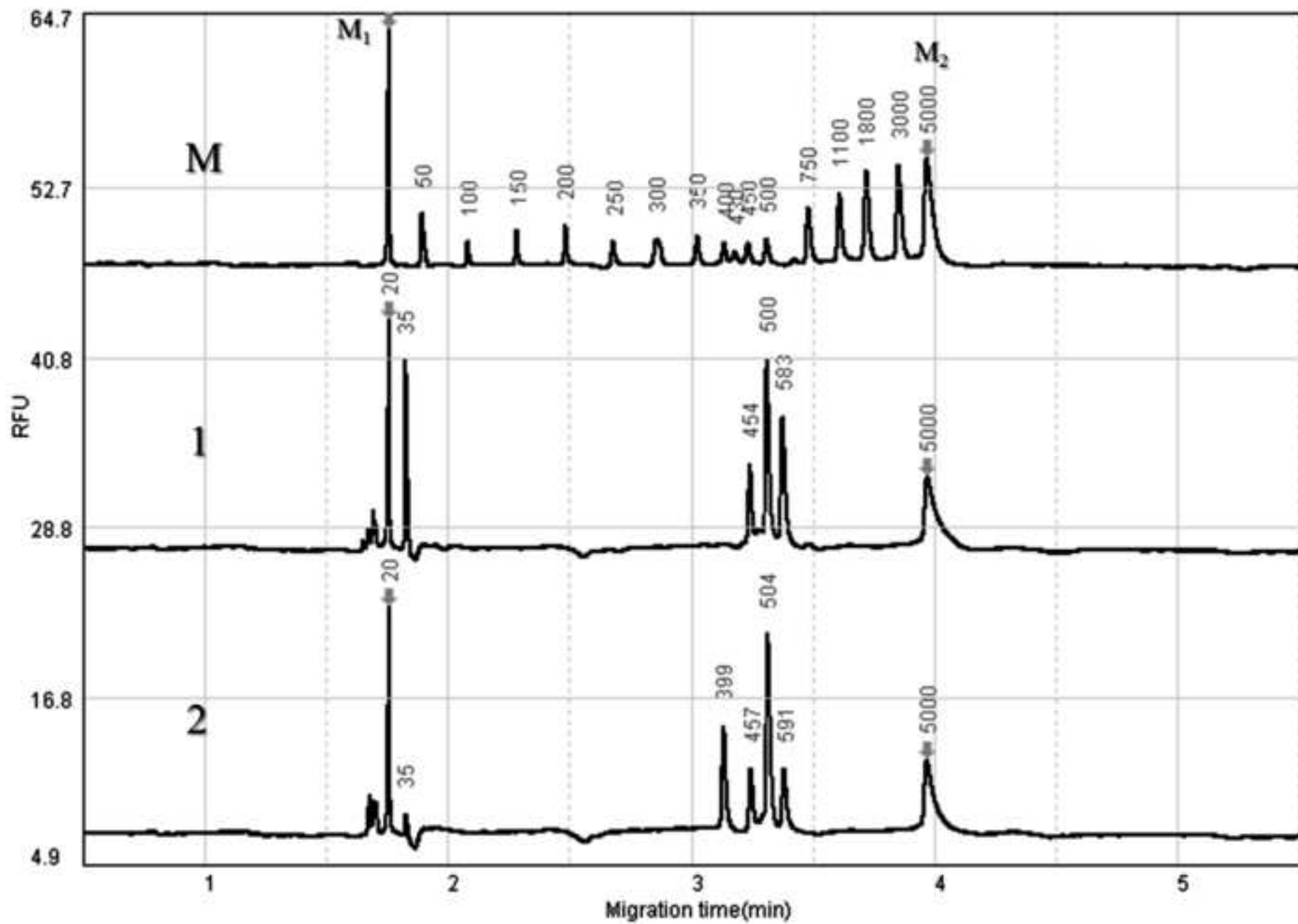


Figure 4A
[Click here to download high resolution image](#)

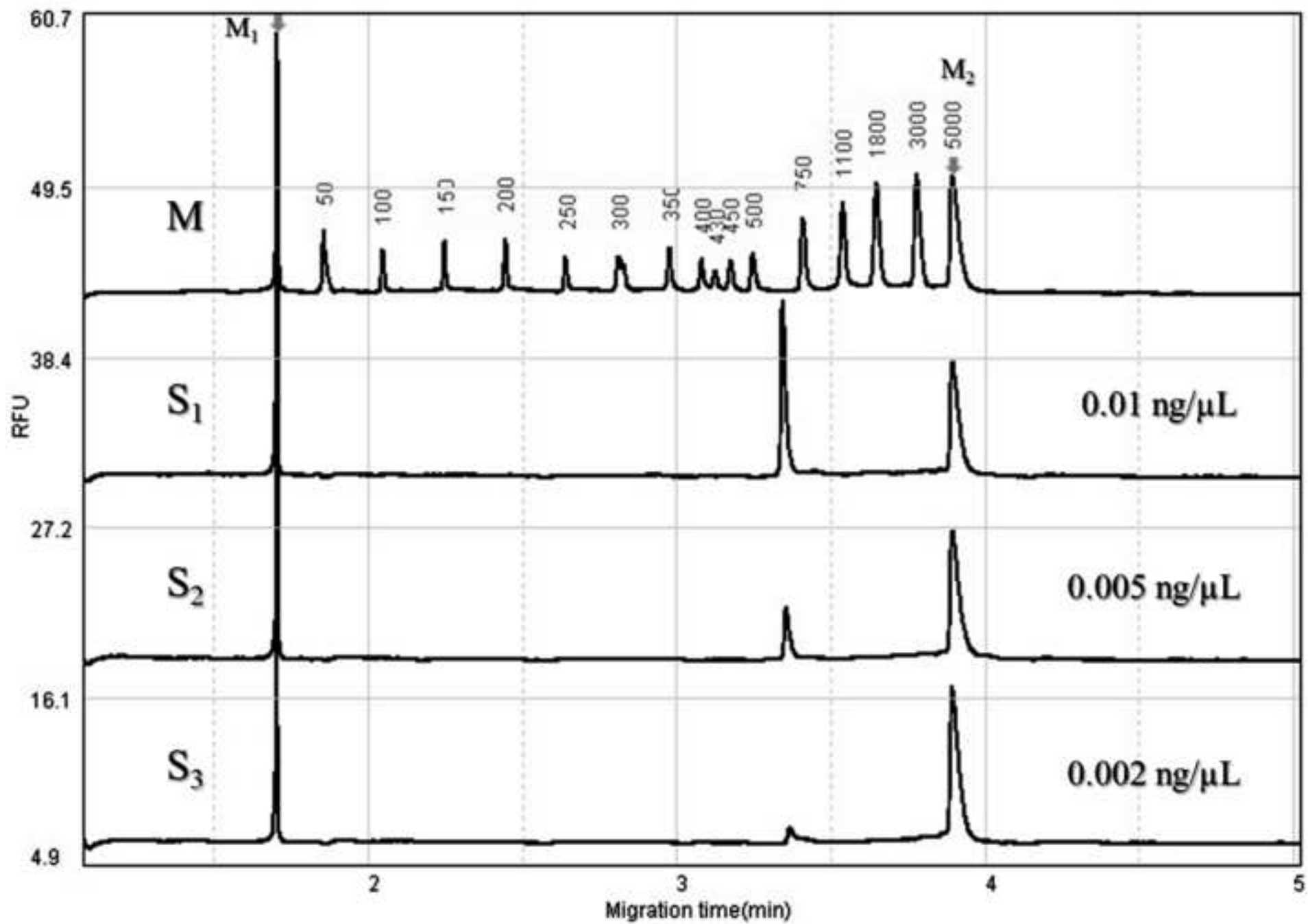
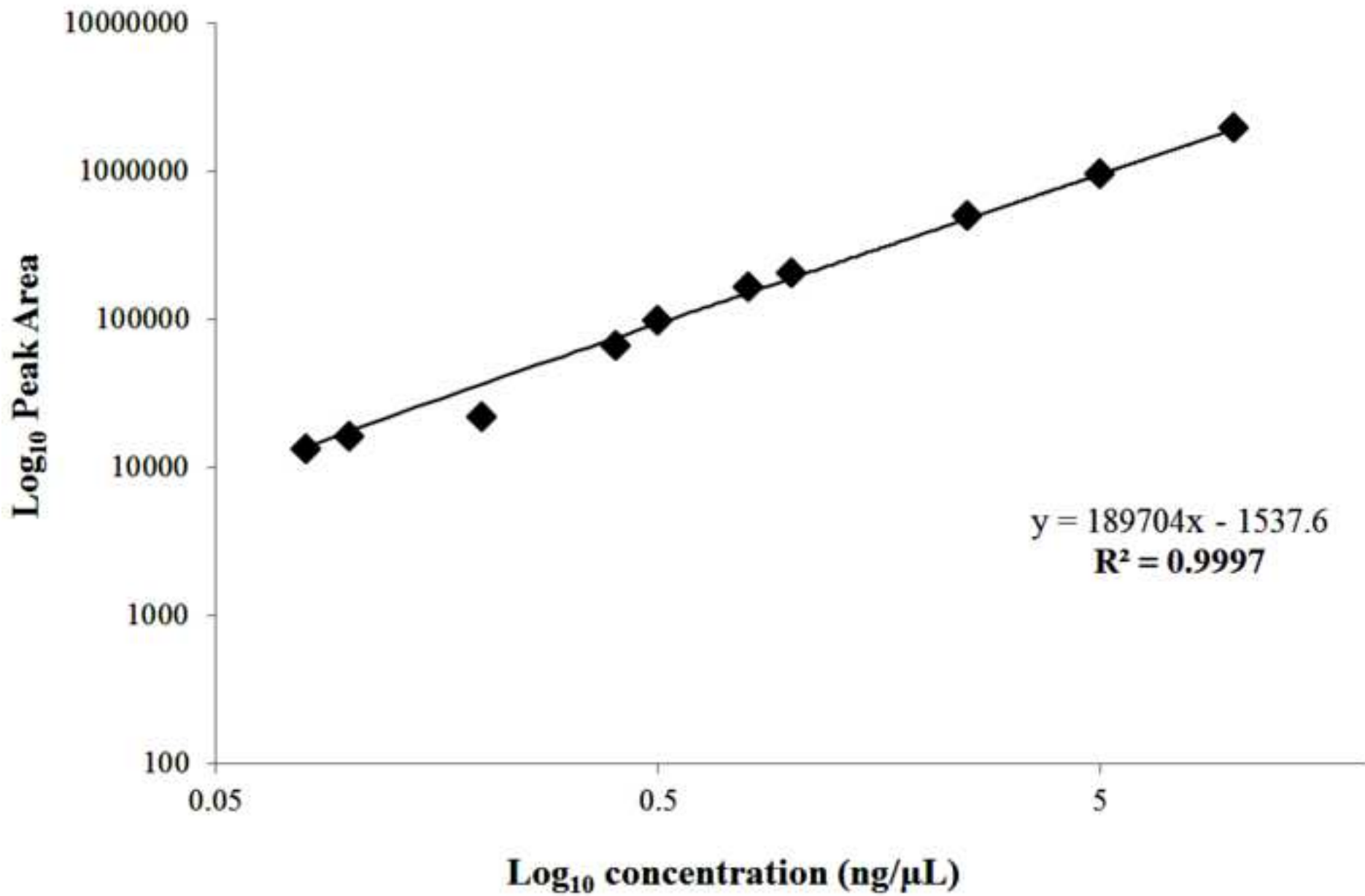


Figure 4B
[Click here to download high resolution image](#)



TABLES

Table 1. Base pair accuracy determination and calculated concentrations of the multiplex PCR samples using CGE.

1. PCR sample (Figure 3 middle trace)					2. PCR sample (Figure 3 lower trace)				
Fragment (bp)	Measured (bp)	Variance (bp)	Accuracy (%)	Concentration ($\mu\text{g}/\mu\text{L}$)	Fragment (bp)	Measured (bp)	Variance (bp)	Accuracy (%)	Concentration ($\mu\text{g}/\mu\text{L}$)
-	-	-	-	-	384	399	15	3.9	1.65
437	454	17	3.8	1.34	437	457	20	4.5	0.90
488	500	12	2.4	4.16	488	504	16	3.2	4.04
541	583	42	7.7	2.38	541	591	50	9.2	1.09

Table 2. CGE Detector linearity of measured by injecting the 576 bp PCR sample in the 10.00-0.08 ng/ μL concentration range.

Detector linearity of the representative PCR sample										
	No 1.	No 2.	No 3.	No 4.	No 5.	No 6.	No 7.	No 8.	No 9.	No 10.
Concentration (ng/μL)¹	10.0	5.00	2.50	1.00	0.80	0.50	0.40	0.20	0.10	0.08
Average peak area²	1898510	930498	487797	199921	160618	96405	65104	21230	15755	12894
SD	21552	7487	23358	8014	748	2960	11572	2038	1589	1007.
RSD%	1.13	0.80	4.78	4.00	0.46	3.07	17.77	9.60	10.08	7.81

¹WFS1 PCR samples were diluted with dilution buffer. ²Average peak area was determined from triplicate measurements for each concentration.