

# Characterization of Y-chromosome SNP duplications

M.R. Silva <sup>a</sup>, S. Serra <sup>b</sup>, T. Ribeiro <sup>b</sup>, H. Geada <sup>a,\*</sup>

<sup>a</sup> *Legal Medicine and Forensic Sciences, Faculty of Medicine, University of Lisbon, Lisbon, Portugal*

<sup>b</sup> *Forensic Genetics Laboratory, National Institute of Legal Medicine, Lisbon Delegation, Lisbon, Portugal*

---

**Abstract.** In Forensic Genetics, it is important that SNP markers have only one polymorphic site for result interpretation, but P25 and 92R7, two of the most widely used Y-SNPs for population studies and forensic purposes, can present two signals. These two loci are Paralogous Sequence Variants. Different polymorphisms detected at P25 and 92R7 loci when performing population studies are reported. © 2005 Elsevier B.V. All rights reserved.

*Keywords:* Y-chromosome; Single nucleotide polymorphism; Duplication; P25; 92R7

---

## 1. Introduction

Single nucleotide polymorphisms (SNPs) have become widely useful in population genetics as they are the most frequent sequence variation encountered in DNA.

In Forensic Genetics, it is important that SNP markers have only one polymorphic site for result interpretation, as simple transitions or transversions. However, P25 and 92R7, two of the most widely used Y-SNPs for population studies and forensic purposes, can present two signals [1] which are alternatively presented in almost all samples. These duplication segments occur for both loci in the same multiplex reaction and at a first glance could interfere in result interpretation as those two SNP markers are implicated in the definition of P, Q or R haplogroups in the phylogenetic tree of the Y-chromosome consortium [2]. P25 and 92R7 are Paralogous Sequence Variants (PSVs) as two different but almost identical fragments were amplified during PCR reaction and one of these sequence variants is polymorphic [3].

One should keep in mind that some SNPs are not true binary markers and that could influence result interpretation specially when performing forensic casework analysis. Here

---

\* Corresponding author. Tel.: +351 21 8811800; fax: +351 21 8864493.

*E-mail address:* hgeada@dlinml.mj.pt (H. Geada).

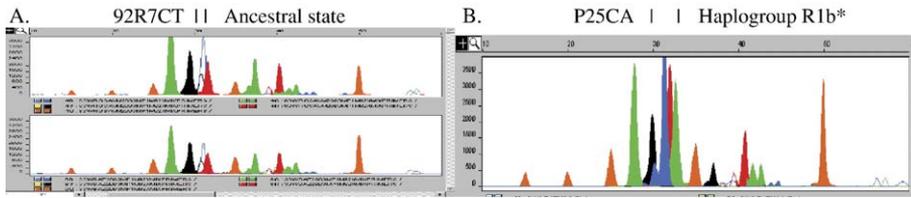


Fig. 1. A. 92R7 duplication CT in ancestral state; B. P25 duplication CA in haplogroup R1b\*.

we report the different polymorphisms detected at P25 and 92R7 loci when performing population studies.

**2. Materials and methods**

A total of 102 bloodstains from unrelated South Portuguese males belonging to four different districts were analysed, 28 being from Faro District(Algarve) and 19 from the nearby Beja District(Alentejo). One hundred and four samples (104) from an African population were also analysed. DNA was extracted by the Chelex method.

For characterization of Y-chromosome haplogroups, 20 Y-SNPs have been studied in three multiplex(MP) reactions [4]: MP1—M22, P25, SRY1532, 92R7, M173, M70, Tat, M213, M9; MP2—12f2, M170, M62, M172, M26, M201; MP3—M34, M81, M78, M35, M96, M123.

SNP typing method was based on multiplex PCRs performed in 25 ul final volume and carried out on an GeneAmpR PCR System 9600 (Applied Biosystems) followed by addition of Exo-SAP for removing unincorporated primers and dNTPs. Multiplex single base extension reactions were performed using SNaPshot Multiplex Kit from Applied Biosystems and specific extension primers with poly C tails, followed by addition of SAP for inactivating excess ddNTPs.

Some extension reaction primers are reverse primers, as is the case for 92R7 locus, showing a CT-T polymorphism in the electropherogrames, corresponding to a GA-A structure in direct sense. For typing P25 null samples redesigned P25 primers with Primer3 Program were also used—P25For CCATACGTGGAACAGACACG; P25Rev AATGCTTCCTGTTGGCAGAT.

Capillary electrophoresis detection of single base extension products was performed on an ABI Prism 3100 Genetic Analyser using filter set E5. Separation time was approximately 25 min on a 36 cm array using POP<sup>TM</sup>4 (Applied Biosystems). GeneScan

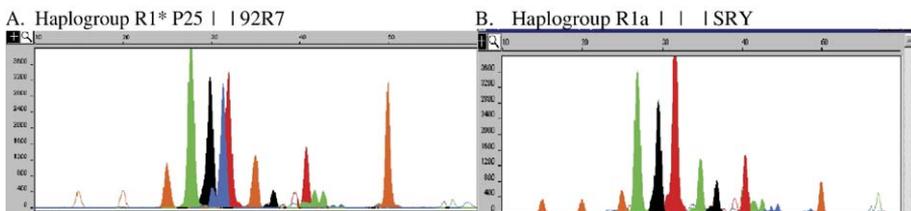


Fig. 2. A. Dashes indicate P25C and 92R7T polymorphisms; B. The third dash indicates SRY1532A polymorphism.

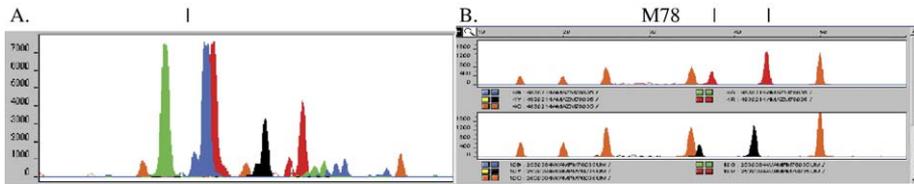


Fig. 3. A. Dash indicates the missing P25; B. Dashes indicate M78 duplication TT in singleplex with the first extension primer batch used.

3.7 software was used for analysing electropherograms, determining the size of the fragments with GeneScan-120 LIZ size standard.

### 3. Results and discussion

P25 polymorphism was initially considered to be a C–A transversion, while 92R7 polymorphism was a G–A transition, but almost all samples studied presented duplication in one of these SNP markers with the following polymorphism 92R7 GA–A and P25 C–CA. The most common Y-chromosome SNP haplogroups defined by these markers are R1\*, R1a and R1b\*.

The ancestral state is characterized by 92R7GA(CT reverse) and P25 C when studying Multiplex1. The most common haplogroup in the Portuguese population R1b\* (frequency 0.4901) is characterized by the derived state 92R7A (T reverse) and P25 CA (Fig. 1).

Very few samples belonging to haplogroup R1\* or haplogroup R1a, this last one having a recurrent mutation SRY1532 G–A, present no duplication (Fig. 2). No P25A or 92R7G has been detected in the Portuguese population. In an African population two samples without P25 peak were detected—in the first one (Fig. 3A) we have to rule out a mutation at the extension primer site but in the second sample occurred a microdeletion in the AZF region [5] as other Y-SNPs and Y-STRs are also missing. These samples were reanalysed by a Reference Laboratory (Santiago de Compostela). We have also redesigned new P25 PCR primers leading to the same result.

In our first study, a possible duplication was detected in M78 while analysing Multiplex3 (Fig. 3). This duplication was specifically noted in E3b1 haplogroup—detected in Faro and Beja samples and presenting a double M78 T. However, the use of a new extension primer batch defines M78 polymorphism as a C–T transition.

These duplication processes should be capped in mind when performing SNP typing as duplication can occur, which could lead to a misleading interpretation of a SNP electropherogram. Rather than an obstacle, these problems should be a challenge when using this useful new methodology—base primer extension, for SNP typing.

### References

- [1] J.J. Sanchez, et al., *Forensic Sci. Int.* 137 (2003) 74–84.
- [2] M.A. Jobling, C. Tyler-Smith, The Y chromosome: an evolutionary marker comes of age, *Nat. Genet. Rev.* 4 (2003) 598–612.
- [3] J.J. Sanchez, et al., Duplications of the Y-chromosome specific loci P25 and 92R7 and forensic implications, *Forensic Sci. Int.* 140 (2004) 241–250.
- [4] M. Brión, et al., Hierarchical analysis of 30 Y-chromosome SNPs in European populations, *Int. J. Legal Med.* 119 (2004) 10–15.
- [5] T.E. King, et al., Inadvertent diagnosis of male infertility through genealogical DNA testing, *J. Med. Genet.* 42 (2005) 366–368.