# In silico inference of multi-locus genotypes from SSCP markers

Panzitta F.[1*,2], Gandini G.C.[2] , Mariani P. [1], Boettcher P.J.[3] and A. Stella[1]
*[1] PTP–CERSA, Lodi, Italy ; [2] VSA, University of Milan, Italy; [3] CNR-IBBA, Segrate, Italy.*

## INTRODUCTION

Single-Stranded Conformation Polymorphisms (SSCP) are a cost-efficient type of genetic marker. However, depending on the eventual information desired, sequencing of the resulting fragments may be necessary, and this can be a time-consuming procedure. Traditional approaches for estimation of genetic distances are often based on genotypes at specific sites of polymorphism. Within the SSCP fragments obtained by PCR amplification, several single nucleotide polymorphisms can be present.

The objective of the study was to develop an in silico method for inference of multi-locus genotypes from SSCP markers.

## MATERIALS AND METHODS

A software was developed to analyze DNA sequences data in order to predict migration of fragment during electrophoresis

### *STEPS*

- The lengths of the fragments expected to be visualized after electrophoresis were obtained based on the embedded restriction sites
  - Sequences were analyzed using the *Restrict* option of the Sequence Analysis Tools at http://bioweb.pasteur.fr/
- The three-dimensional structure of each potential fragment was predicted by applying an energy minimization algorithm (Barash J Comput Biol. 2004;11(6):1169-74), (Kamashev et al.Biochemistry. 2004 Jun 29;43(25):8160-8)
  - Multiple suboptimal candidate structures were predicted
- Fragment length and 3D structure were used to determine the migration distance
  - Different mass / volume ratios were used to estimate migration

### *DATA*

- Genomic DNA from 150 animals coming from 10 Italian Goat breeds
- Three pairs of primers (named GA, 2/3 and GB) were designed
  - These primers produced *amplicons* covering all the gene
- PCR products were digested with two restriction enzymes (MSpI and AvaII)
- Fragments were electrophoresed and visualized by silver staining (Fig. 1)

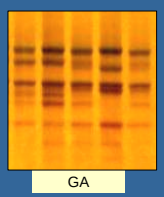


Figure. 1 : SSCP gels of PCR products obtained with the three sets of GH Primers : GA

- PCR products were sequenced and sequences were analyzed using Chromas and ClustalW
- Simulated migration and electrophoresis visualization were compared for 150 sequences of the GA region
  - Based on number of fragments obtained and their relative distances

## RESULTS AND DISCUSSION

- Fragments of less than 50 bases were excluded from the analyses
- The overall success rate in the assignment of the fragments was 81%
  - Ranging from 98% (Trentina) to 68% (Camosciata)
- Simulated distances were proportional to visualized distances among fragments
  - r = .93 (p< 0.0001)
  - Higher correlation when mass was given higher weight over volume
- Figure 2 shows fragments obtained by *in silico* analysis simulating restriction sites for a specific individual of the Ciavenasca breed
- Figures 3 and 4 compare distances obtained by simulation to distances among fragments shown on the electrophoretic gel
  - The number of fragments resulting for this animal corresponds for the two methods
  - Two distinct sets of fragments are visible based on the distance run

```
>GA_Cia_06
AACATTTGGCCAAGTTTGAAATGTTCYCAGTCCCTGGAGGGAAGGGCAGGCGGGGGCTGGCAGGAGATCAGGCRTCCAGC
TCTCTGGGCCCCTCCGTCGCGGCCCTCCTGGTCTCTCCCTAGGCCCC
CG
BGACCTCCYTGCTCCTGGCTTTCACCCTGCTCTGCCTGCCCTGGACTCAGGTGGTGGGCGCCTTCCCAGCCATGTCCTTGTC
CCGGGCTCAGCACCTGCATCAACTGGCTGCTGACACCTTCAAAGAGTTTGTAAGCTCCCCAGAGATGTGTCCTAGAGGTGG
GGAGGCAGGAAGGGGTGAATCCGCACCCCCTCCACACAATGGGAGGGAACTGAG
DGACCTCAGTGGTATTTTATCCAAGTAAGGATGTGGTCAGGGGAGTAGAAATGGGGGTGTGTGGGGTGGGGAGGGTTCYG
AATAAGGCAGTGAGGGGAACCACACACCAGCTTAGACC
ECGGAGGGACAGAGATACTCCRTCCARAACACCCAGGTTGCCTTCTGCTTCTCYGAAACCATCCCAGGCCCCCACGAGCAA
GAAYGAGGCCCAGCAGAAWTCAGTGAG
```

Figure 2. Example of fragments obtained by in silico analysis simulating restriction sites from a Ciavenascaa goat.

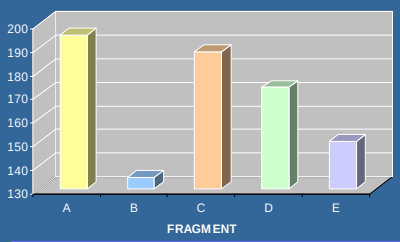


Figure 3. Distances of fragments obtained by simulation.

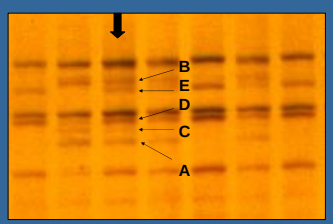


Figure 4. Actual banding pattern obtained.

## CONCLUSIONS

The Energy Minimization algorithm can be used to predict folding of DNA and movement of fragments on electrophoresis gel. The method may be used to infer genotypes from SSCP data

ABSTRACT NUMBER: 477 - Session G6.21, 56th EAAP Annual Meeting, Uppsala, Sweden