Parco Tecnologico Padano
La ricerca si fa impresa
Entrepreneurial research in ag-biotech

Insights in the genome of *Mycobacterium avium subsp. paratuberculosis* by Next Generation Sequencing approaches

Isabella Della Noce[1], Karen Stevenson[2], Julian Parkhill[3], Matteo Ricchi[4], Barbara Lazzari[1], Marcello Del Corvo[1], Enrico Zanetti[1], Valeria Messina[1], Norma Arrigoni[4], John L. Williams[1], Giulietta Minozzi[1,5]

[1] Parco Tecnologico Padano (PTP), Lodi, Italy - [2]Moredun Research Institute, Pentlands Science Park, Penicuik, UK - [3] Wellcome Trust Sanger Institute, Genome Campus, Cambridge, UK - [4] Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna (IZSLER), National Reference Centre for Paratuberculosis, Gariga di Podenzano (PC), Italy - [5] Department of Veterinary Science and Public Health (DIVET), University of Milan, Italy.

Corresponding author: Isabella.DellaNoce@tecnoparco.org

Introduction

Mycobacterium avium subsp. *paratuberculosis* (MAP) is the causative agent of paratuberculosis - or Johne's disease – that affects farmed and wild animals worldwide, causing negative economic consequences particularly relevant in the livestock sector of dairy cattle and beef (1). Recent estimates say that more than 50 % of the herds in Europe and North America are infected (3). In Italy, a study conducted in the Lombardy and Veneto regions reveals that about 70 % of dairy herds are infected (2).

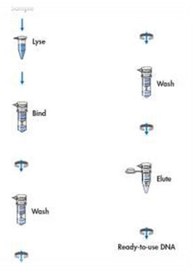
The disease shows high variability in the progression and symptoms that may be due to the genetic variability of the host, the pathogen, or a combination of the two. Understanding the mechanism responsible of this variability could be of paramount importance for the control of the disease (1, 3).

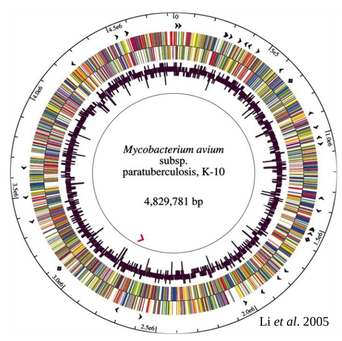
Aim of this work was to study the genomic variability of MAP isolated from dairy cattle from different farms distributed in several Italian regions through the use of Next Generation Sequencing (NGS) techniques.

The preliminary results on 15 strains are presented.

Methods

All the MAP strains were isolated from single animals from cattle herds located in 6 provinces in the north of Italy, in collaboration with the National Reference Centre for Paratuberculosis (Piacenza, Italy). Nine samples were paired end sequenced at 150bp end and six samples at 75bp per end on the Illumina Miseq. The K-10 (NC_002944.2) strain was used as the reference, and bioinformatic analyses were performed using the BWA-MEM, FreeBayes, GATK, and SnpEff software.





Results

The reads were mapped to the K-10 reference sequence with 99.96% reads finding a match, with a mean coverage of 138.26 X. In total 844 variants were identified, of which 698 SNP (Single Nucleotide Polymorphisms), 23 MNP (Multi-Nucleotide Polymorphisms), 45 INS (Insertions) , and 78 DEL (Deletions). Each strain has 25.90% private SNP on average. About 43% of variants were located in coding regions, showing the following effects on gene products: 1.40% missense, 1.14% nonsense, and 37.46% silent. Another 25% of variants were in genes -60 upstream regions.

Variants identified

SNP	698
MNP	23
INS	45
DEL	78
Total	**844**

Location on the genome

Coding regions	43%
Upstream	25%

Effect of the variants

Missense	1.40%
Nonsense	1.14%
Silent	37.46%

Conclusions

Even if the number of identified variants was quite low, the fact that almost 92% of the MAP genome is covered by genes must be taken into consideration. The need to maintain unaltered gene functions plays a significant role in reducing the number of missense and nonsense variants. In the future, phenotypic information linked to MAP strains and their hosts coupled with strain-specific genomic information may help to disentangle the genetic variability linked to virulence and MAP population substructure.

References: (1) Collins et al., Vet Clin Food Anim 27, Issue 3, Pages 525-664 (November 2011) Johne's Disease; (2) Pozzato *et al.*, Prev Vet Med. 2011, Oct 1; 102(1):83-6; (3) Kirkpatrick and Shook, Vet Clin North Am Food Anim Pract 2011, 27(3):559–571.

Acknowledgments

this work was developed in the framework of the research projects PON01_02589 – MicroMap and PON01_01841 – EpiSud
funded by the Italian Ministry of Education, University and Research (MIUR). New Generation Sequencing of 6 of the 15 strains was funded by the Scottish Government Rural and Environment Science and Analytical Services Division and The Sanger Research Institute.