"33rd Annual Meeting EAAP 2012"
Bratislava, SLOVAKIA, August 27th-31st, 2012

SNP identification in Italian pig breeds using Next Generation Sequencing Technology

Guiatti, D.[1], Sgorlon, S.[1], Licastro, D.[2], Stefanon, B[1].

[1] Dipartimento di Scienze Agrarie ed Ambientali, Università di Udine, Italy [2] CBM S.c.r.l., Area Science Park, Basovizza - Trieste, Italy

Introduction

The dissection of complex traits of economic importance to the pig industry requires the availability of a significant number of genetic markers, such as single nucleotide polymorphisms (SNPs). Most of the Italian breeds were selected for heavy pig production to obtain animals with high aptitude for the PDO dry-cured ham production, such as Parma or S. Daniele ham. Identifying genetic variants should become increasingly feasible with improved sequencing methods on a genome-wide scale. Next-generation sequencing (NGS) tools are valuable for the discovery of genetic markers in animal genomes. An high number of SNPs and small indels (insertions/deletions) were found with these novel technologies. The Pig Genome Assembly 9.2 provides the research community with a high quality reference of Sus scrofa genome. In order to evaluate the global genetic variation, we have sequenced another pig genome at high coverage.

Materials & methods

Pig genomic DNA from a F1 (Italian Duroc x Italian Large White) crossbreed was extracted and normalized to 55ul at 20ng/ul and using a Covaris station we sheered the DNA to an average of 300 base pair. The purified obtained sample has been ends repaired, 3' adenilated and adaptor ligated using the the appropriate DNA Adapter Index tube. The obtained library has been PCR amplified to selectively enrich those DNA fragments that have adapter molecules on both ends. The final purified product has been quantitated using both qPCR and Agilent 2100 Bioanalyzer (Agilent) and normalized to 10nM and seeded on two lanes of a Illumina Flow cell to a final concentration of 12pM using a cBOT system with a TRuSeq PE Kit cluster Kit V3. After cluster generation the sequencing has been performer using a 200-Cycle Paired-End Run TruSeq SBS. The output from a sequencing run has been mapped against the reference genome Sscrofa9.2 using the open software BWA (Li *et al.*, 2009). The resulted mapped sequences have been quality filtered, duplicate removed and variants called using the software Samtools (Li *et al.*, 2009).

	Reads	Base pairs
Total Reads	~ 453298548	~ 90659709600
Total Mapped reads	~ 340467813	~ 68093562600
Mapped and Duplicated removed reads	~ 340413847	~ 68082769400
Covered genome	94%	94%

Tab.1: Resequencing results

Results and Discussion

- We produced 90.7 gigabases of sequence as reported in Table1. The consensus assembly based on these uniquely mapped reads represent a 94% coverage (Figure 1) of the whole pig genome and give a effective mean depth 26.8 fold for the covered regions. These short reads were mapped to Ensembl Sus scrofa9.2 genome.
- The resequencing of this target genome by using the Illumina® next generation sequencing approach provides us a global view of porcine genomic variance.
- Millions of SNPs (Figure 2), including many novel polymorphisms, were identified as a valuable resource to be used as high density genetic markers for genotyping or for investigating phenotypic variations.
- A resequencing of target regions containing QTLs for intramuscolar fat (IMF) – i.e. SSC 4 and 6 – for a larger number of pigs will be performed in order to validate SNP presence and their association with carcass traits under selection for Italian heavy pigs used in DOP production as San Daniele Ham.

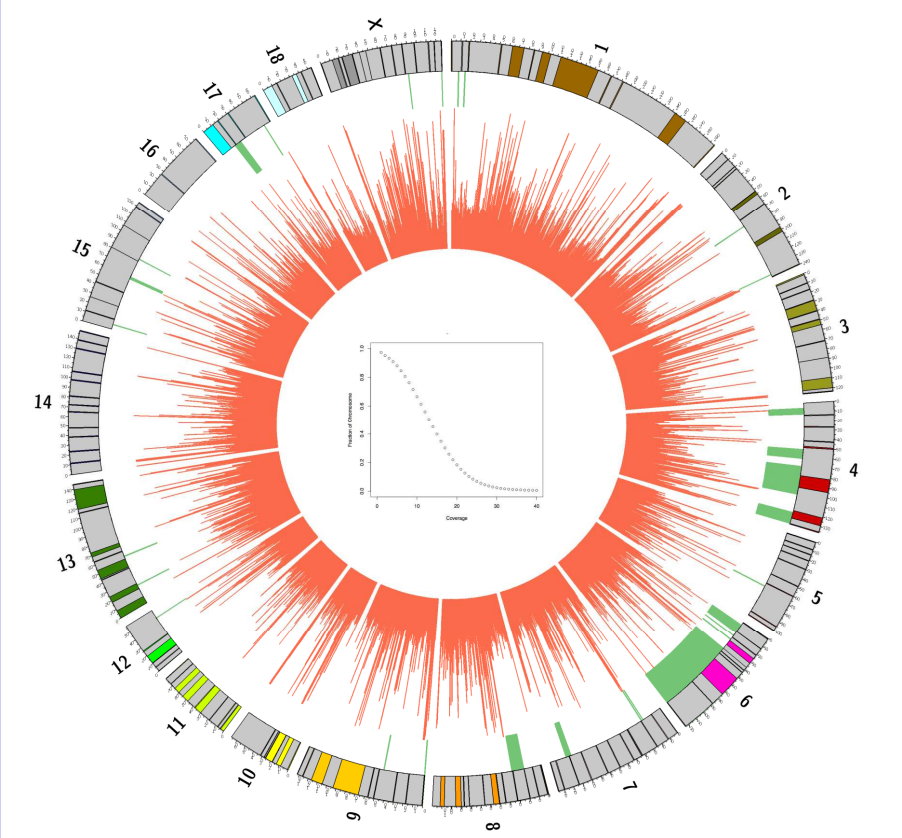

Fig 1: Sus scrofa Karyotype with resequencing coverage (red) and QTL regions for IMF (green). In the centre is plotted the fraction of chromosome as a function of coverage.

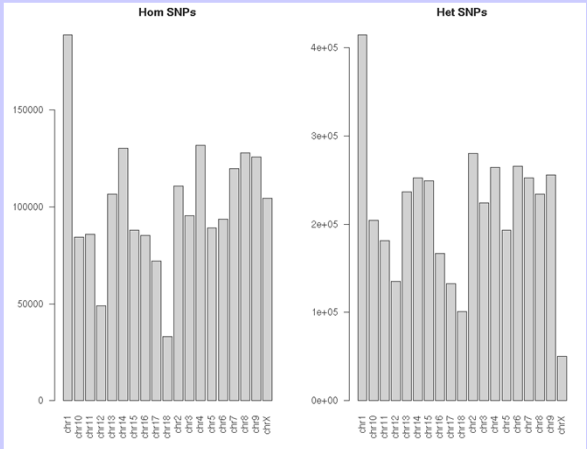

Fig 2: Barplots of number of homozygous and heterozygous SNPs identified for each chromosome.

References

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079