AnnotQTL: a new tool to gather functional and comparative information on a genomic region

Frédéric Lecerf, Anthony Bretaudeau, Olivier Sallou, Colette Désert, Yuna Blum, Sandrine Lagarrigue, Olivier O. Demeure

To cite this version:

Frédéric Lecerf, Anthony Bretaudeau, Olivier Sallou, Colette Désert, Yuna Blum, et al.. AnnotQTL: a new tool to gather functional and comparative information on a genomic region. 12. Journées Ouvertes en Biologie, Informatique et Mathématiques, Jun 2011, Paris, France. 2011, JOBIM 2011. ⟨hal-02746728⟩

HAL Id: hal-02746728

https://hal.inrae.fr/hal-02746728v1

Submitted on 3 Jun 2020

AnnotQTL: a New Tool to Gather Functional and Comparative Information on a Genomic Region

Frédéric LECERF[1,2], Anthony BRETAUDEAU[3], Olivier SALLOU[3], Colette DESERT[1,2], Yuna BLUM[1,2], Sandrine LAGARRIGUE[1,2] and Olivier DEMEURE[1,2]

[1] INRA, UMR598 Génétique Animale, F-35000 Rennes, France
`{blum, demeure}@rennes.inra.fr`
[2] Agrocampus OUEST, UMR598 Génétique Animale, F-35000 Rennes, France
`{lecerf, desert, lagarrigue }@agrocampus-ouest.fr`
[3] GenOuest Platform, INRIA/Irisa – Campus de Beaulieu, F-35042 Rennes Cedex, France
`{anthony.bretaudeau, olivier.sallou}@irisa.fr`

Keywords Genomic regions, Annotation, Gene ontology, QTL.

The final steps of genetic mapping research programs require close analysis of several QTL regions to select candidate genes for further studies. Despite several websites (NCBI genome browser, Ensembl Browser, UCSC Genome Browser) or web tools (Biomart, Galaxy) developed to achieve this task, the selection of candidate genes remains a laborious process. The information made available on the more prominent websites differs slightly in terms of gene prediction and functional annotation, while other websites provide extra information that researchers may want to use (HGNC approved gene symbols, Gene Ontology Annotation or functional data, conservation of synteny with other species, etc.). It is possible to manually merge and compare this information for one QTL containing few genes, but not for many different QTL regions containing dozens of genes.

Here, we propose a web tool that, for a given region of interest, merges the list of genes available in NCBI and Ensembl, removes redundancy, adds functional annotations from different prominent web sites, and highlights the genes for which functional annotation fits the biological function or diseases of interest. The tool is dedicated to sequenced species of livestock including cattle, pig, chicken, and horse as well as dog, i.e. species that have been extensively studied (with over 8000 QTLs detected; see http://www.animalgenome.org/cgi-bin/QTLdb/index). Nevertheless, the family designs and the low number of animals used in these species, most of the studies use linkage analysis, and the QTL regions identified remain large (containing dozens of genes). Conversely, in human and model species, most analyses now draw heavily on association studies involving large cohorts, thus providing more power and accuracy, and the web tools already available focus on these species through functional annotation of SNPs in association with the trait [1-3]. As most of these tools focus on the SNP annotation itself, describing whether the SNP is located in a gene, then a coding sequence could have a functional effect, etc. While these web tools are highly efficient in providing a good annotation for specific SNPs, they clearly cannot be used to collect information on the large regions obtained in livestock species.

AnnotQTL is a web tool designed to gather the functional annotation of different prominent websites while minimizing redundant information. Using all known information substantially accelerates the gene analysis of QTL regions for livestock species traits and improves the selection of candidate genes. The AnnotQTL web tool is available at http://annotqtl.genouest.org.

[1] S. Goodswen, C. Gondro, N. Watson-Haigh, and H. Kadarmideen, FunctSNP: an R package to link SNPs to functional knowledge and dbAutoMaker: a suite of Perl scripts to build SNP databases. *BMC Bioinformatics*, 11(1): 311-311, 2010.

[2] J. Reumers, S. Maurer-Stroh, J. Schymkowitz, and F. Rousseau, SNPeffect v2.0: a new step in investigating the molecular phenotypic effects of human non-synonymous SNPs. *Bioinformatics (Oxford, England)*, 22(17): 2183-2185, 2006.

[3] M. Ryan, M. Diekhans, S. Lien, Y. Liu, and R. Karchin, LS-SNP/PDB: annotated non-synonymous SNPs mapped to Protein Data Bank structures. *Bioinformatics (Oxford, England)*, 25(11): 1431-1432, 2009.