HAL
open science

Efficiency of heterologous capture for targeted resequencing of the Major Histocompatibility Complex in Suidae and Tayassuidae

Jaime Gongora, Núria Mach, Jordi Estellé, Per Wahlberg, Sylvain Marthey, Jerôme Lecardonnel, Marie-Thérèse Bihoreau, Claire Gaillard, Marco Moroldo

To cite this version:

Jaime Gongora, Núria Mach, Jordi Estellé, Per Wahlberg, Sylvain Marthey, et al.. Efficiency of heterologous capture for targeted resequencing of the Major Histocompatibility Complex in Suidae and Tayassuidae. 20. Plant and Animal Genome, Labo/service de l'auteur, Ville service, Pays service., Jan 2012, San Diego, CA, United States. pp.P0623. hal-01019000

HAL Id: hal-01019000

https://hal.science/hal-01019000v1

Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

P0623

Efficiency of heterologous capture for targeted resequencing of the Major Histocompatibility Complex in Suidae and Tayassuidae

[1]Gongora, J., [2,3,4]Mach, N., [2,3,4]Estellé, J., [2,3,4]Wahlberg, P., [2,3,4]Marthey, S., [2,3,4]Lecardonnel, J., [5]Bihoreau, M.T., [2,3,4]Rogel-Gaillard, C. and [2,3,4]Moroldo, M.

[1] Faculty of Veterinary Science, University of Sydney, NSW 2006, Australia
[2] INRA, UMR GABI, CRB-GADIE, Domaine de Vilvert, 78350, Jouy-en-Josas, France
[3] CEA/DSV/iRCM/SREIT/GABI, Domaine de Vilvert, 78350, Jouy-en-Josas
[4] AgroParisTech, Domaine de Vilvert, 78350, Jouy-en-Josas
[5] CEA/IG/CNG, 2 rue Gaston Crémieux, 91057 Evry, France

(1) INTRODUCTION

The Major Histocompatibility Complex (MHC) is one of the most gene dense and polymorphic regions in mammalian genomes. It plays an important role in the development and regulation of immune response to pathogens. To date, studies on the MHC have mostly focused on model species. Within Suidae, this region has been extensively sequenced and annotated only in the domestic pig (*Sus scrofa*)[1].

The aim of this study was to gather information about the MHC locus from other Suidae and its sister family Tayassuidae. We approached this by using what we defined as the 'heterologous sequence capture' which is a solid phase array designed based on *S. scrofa* to target MHC orthologous regions across species from those two families. Here we present the preliminary analyses of the efficiency of this method.

(2) METHODS

SAMPLING

A total number of 96 samples representing 86 specimens from 9 species of Suidae (n=69) from Eurasia and Africa and 2 species of Tayassuidae (n=19) from the Americas were sampled.

SEQUENCING AND DATA ANALYSIS

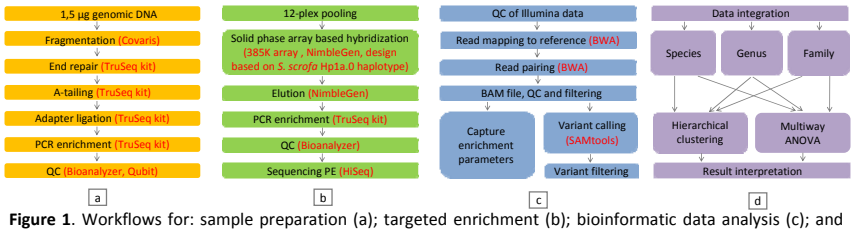
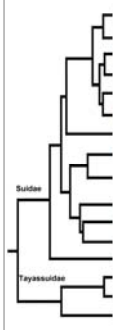


Figure 1. Workflows for: sample preparation (a); targeted enrichment (b); bioinformatic data analysis (c); and statistical analysis (d).

(3) RESULTS



	Samples	Total reads (x10⁶)	Mapped (%)	Specificity (%)	Evenness	Coverage (X)	Duplicate (%)
Sus cebifrons	4	38	77	33	68	157	37
Sus philippensis	N/A						
Sus verrucosus	N/A						
Sus barbatus	10	266	86	31	65	94	45
Sus celebensis	10	271	81	32	66	107	42
Sus scrofa	18	31	83	28	62	54	62
Porcula salvania	N/A						
Potamochoerus porcus	1	12	76	33	62	29	51
Potamochoerus larvatus	3	26	76	33	63	100	42
Hylochoerus meinertzhageni	3	23	61	35	47	55	39
Phacochoerus africanus	10	38	76	34	64	103	48
Phacochoerus aethiopicus	N/A						
Babyrousa babyrussa	13	31	73	34	64	83	48
Tayassu pecari	4	23	12	22	29	6	6
Catagonus wagneri	N/A						
Pecari tajacu	19	18	12	19	34	4	7

Figure 2. Overall results of the sequence capture per species. Left, a phylogenetic tree showing the relationships among species of Suidae and Tayassuidae based on nDNA and mtDNA[2,3]. The symbol alongside the scientific name is the key for the dendogram in Fig. 3. Right, a table showing the effect of species on capture parameters. The model included DNA-sequencing libraries as covariates, species as fixed effect, and animal as a random effect. Species illustrations by Toni Llobet were sourced with permission from the *Handbook of the Mammals of the World*[4]. N/A means not applicable as these species were not included in the analysis.

- Heterologous capture was significantly influenced by species, genera and families (Fig. 2)
- The average values of coverage and specificity were 86X and 32% within Suidae and 4X and 20% within Tayassuidae
- Hierarchical clustering allowed samples to be classified according to family and even genus when using % of mapped reads (Fig. 3) and coverage. In contrast, specificity (% reads mapping on target versus all mapped reads) did not allow samples to be clustered in the same way. This may indicate that stronger effects (i.e. capture protocol) may affect this parameter.

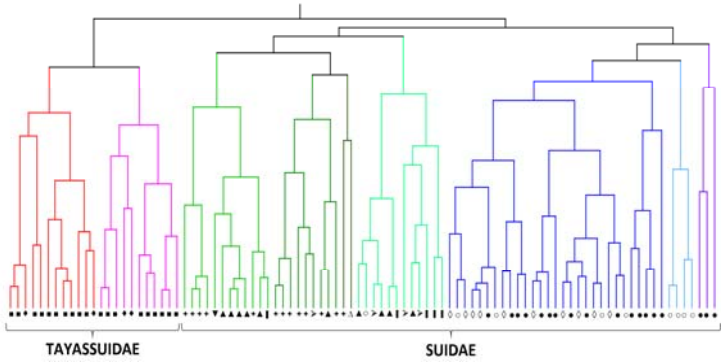


Figure 3. Hierarchical clustering analyses based on % of mapped MHC reads for the different samples sequenced. Symbols correspond to the species as described in Fig. 2. Possible clustering patterns are represented by branch colours. The families Tayassuidae and Suidae clustered separately. Within Suidae, reads from individuals belonging to the same species tended to cluster together, although there was some heterogeneity. This result indicates that heterologous capture performs well within Suidae, particularly for individuals of the genus *Sus*. However, phylogenetic reconstruction analyses will be required in order to make conclusions on the evolutionary relationships of MHC genes across species.

(4) DISCUSSION

- Heterologous capture appears to be efficient for the family Suidae
- Further optimization will be required before application in Tayassuidae
- The difference in capture efficiency is likely to reflect the divergence of the Suidae and Tayassuidae MHC which may have occurred after they diverged from the common ancestor ~35-39 Ma[2]

(5) FUTURE DEVELOPMENTS

- Reduction of duplicate reads is needed to increase the levels of coverage
- Higher levels of multiplexing paired with increased coverage generated per run would allow the heterologous capture to be an even more cost-effective technique
- The incorporation of species-specific probes into the capture array could improve the capture efficiency in Tayassuidae

References

1. Renard *et al*. (2006). The genomic sequence and analysis of the swine major histocompatibility complex. *Genomics*. 88: 96-110.
2. Gongora *et al*. (2011). Rethinking the evolution of extant sub - Saharan African suids (Suidae, Artiodactyla). *Zoologica Scripta*. 40: 327-335.
3. Gongora *et al*. (2006) Phylogenetic divisions among Collared peccaries (*Pecari tajacu*) detected using mitochondrial and nuclear sequences. *Molecular Phylogenetics and Evolution*. 41: 1-11.
4. Wilson, D.E. & Mittermeier, R.A. eds. (2011). *Handbook of the Mammals of the World*. Vol. 2. Hoofed Mammals. Lynx Edicions, Barcelona.

Acknowledgements

- INRA for making funding available which was provided by The French National Research Agency (PSC-08-GENO-CapSeqAn) for this research.
- The University of Sydney for providing research funding for sampling and financial support allowing Dr Gongora to undertake preliminary experiments on the cross-species approach in Sydney and to do this project at INRA for 6 months.
- All samples were provided by Dr Jaime Gongora, some of which he collected himself. Other samples were accessed through Dr Stewart Lowden, Dr Joeke Nijboer or through collaboration with various institutions from Eurasia, Africa and the Americas.