

Effect of climate change on the spatial distribution of genomic variants involved in the resistance to East Coast Fever in Ugandan cattle

Estelle Rochat^{1*}, Elia Vajana^{2*}, Licia Colli², Charles Masembe³, Riccardo Negrini², Paolo Ajmone-Marsan², Stéphane Joost¹ and the NEXTGEN Consortium

¹ Laboratory of Geographic Information Systems (LASIG), School of Architecture, Civil and environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

² Institute of Zootechnics and BioDNA Research Centre, Faculty of Agricultural, Food and Environmental Sciences, Università Cattolica del S. Cuore, Piacenza, Italy

³ Institute of Environment & Natural Resources, Makerere University, Kampala, Uganda

* These authors contributed equally to this work

East Coast Fever (ECF) is a major livestock disease caused by *Theileria parva* Theiler, 1904, an emo-parasite protozoan transmitted by the tick *Rhipicephalus appendiculatus* Neumann, 1901. This disease provokes high mortality in cattle populations of East and Central Africa, especially in exotic breeds and crossbreds (Olwoch et al., 2008). Here, we use landscape genomics (Joost et al., 2007) to highlight genomic regions likely involved into tolerance/resistance mechanisms against ECF, and we introduce *Spatial Areas of Genotype Probability* (SPAG) to delimit territories where favourable allelic variants are predicted to be present.

Between 2010 and 2012, the NEXTGEN project (nextgen.epfl.ch) carried out the geo-referencing and genotyping (54K SNPs) of 803 Ugandan cattle, among which 496 were tested for *T. parva* presence. Moreover, 532 additional *R. appendiculatus* occurrences were obtained from a published database (Cumming, 1998). Current and future values of 19 bioclimatic variables were also retrieved from the WorldClim database (www.worldclim.org/).

In order to evaluate the selective pressure of the parasite, we used MAXENT (Phillips et al., 2006; Muscarella et al., 2014) and a mixed logistic regression (Bates et al., 2014) to model and map the ecological niches of both *T. parva* and *R. appendiculatus*. Then, we used a correlative approach (Stucki et al., 2014) to detect molecular markers positively associated with the resulting probabilities of presence and built the corresponding SPAG. Finally, we considered bioclimatic predictors representing two different climate change scenarios for 2070 - one moderate and one severe - to forecast the simultaneous shift of both SPAG and vector/pathogen niches.

While suitable ecological conditions for *T. parva* are predicted to remain constant, the best environment for the vector is predicted around Lake Victoria. However, when considering future conditions, parasite occurrence is expected to decrease because of the contraction of suitable environments for the tick in both scenarios.

Landscape genomics' analyses revealed several markers significantly associated with a high probability of presence of the tick and of the parasite. Among them, we found the marker ARS-BFGL-NGS-113888, whose heterozygous genotype AG showed a positive association.

Interestingly, this marker is located close to the gene IRAK-M, an essential component of the Toll-like receptors involved in the immune response against pathogens (Kobayashi et al., 2002). If the implication of this gene into resistance mechanisms against ECF is confirmed, the corresponding SPAG (Figure 1) represents either areas where the variant of interest shows a high probability to exist now, or areas where ecological characteristics are the most favorable to induce its presence under future climatic conditions.

Beyond the results presented here, the combined use of SPAG and niche maps could help identifying critical geographical regions that do not present the favourable genetic variant in the present, but where a parasite is likely to expand its range in the future. This may represent a valuable tool to support the identification of current resistant populations and to direct future targeted crossbreeding schemes.

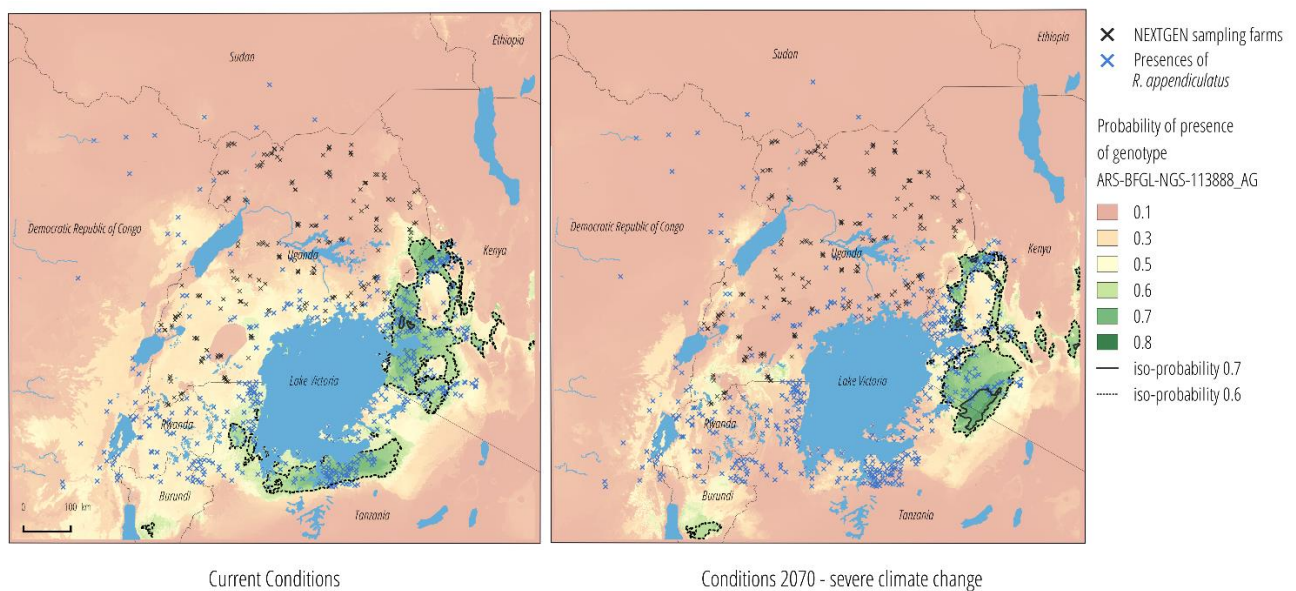


Figure 1 - Geographic Areas of Allele Frequency (SPAG) for the genotype AG of the SNP "ARS-BFGL-NGS-113888" (ARS-11), highlighting areas where this genotype shows a high probability to be present (Current Conditions), and where it may be distributed in the future (Conditions 2070). As the presence of ARS-11_AG is positively correlated with the presence of the tick *R. appendiculatus* ($\alpha = 0.01$; Efron pseudo $R^2 = 0.074$), we can estimate the probability of presence of this genotype also in regions without sampling points and thus without genetic data. At present, the areas of high probability of presence of ARS-11_AG are mainly observed in the North-East and the South of Lake Victoria. However, when considering environmental conditions in 2070 (assuming severe climate change), these areas are expected to be mainly restricted to the North-East of Lake Victoria, where favorable conditions for the presence of *R. appendiculatus* are supposed to be maintained.

References

- Bates D., Mächler M., Bolker B., & Walker S. (2014) Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*, .
- Cumming G.S. (1998) Host preference in African ticks (Acari: Ixodida): a quantitative data set. *Bulletin of Entomological Research*, **88**, 379.
- Gachohi J., Skilton R., Hansen F., Ngumi P., Kitala P., & others (2012) Epidemiology of East Coast fever (*Theileria parva* infection) in Kenya: past, present and the future. *Parasit Vectors*, **5**, 194.
- Joost S., Bonin A., Bruford M.W., Despré S. L., Conord C., Erhardt G., & Taberlet P. (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, **16**, 3955–3969.
- Kobayashi K., Hernandez L.D., Galán J.E., Janeway C.A., Medzhitov R., & Flavell R.A. (2002) IRAK-M is a negative regulator of Toll-like receptor signaling. *Cell*, **110**, 191–202.
- Muscarella R., Galante P.J., Soley-Guardia M., Boria R.A., Kass J.M., Uriarte M., & Anderson R.P. (2014) ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for MAXENT ecological niche models. *Methods in Ecology and Evolution*, **5**, 1198–1205.
- Olwoch J.M., Reyers B., Engelbrecht F.A., & Erasmus B.F.N. (2008) Climate change and the tick-borne disease, Theileriosis (East Coast fever) in sub-Saharan Africa. *Journal of Arid Environments*, **72**, 108–120.
- Phillips S.J., Anderson R.P., & Schapire R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Stucki S, Orozco-terWengel P, Bruford MW et al. (2014) High performance computation of landscape genomic models integrating local indices of spatial association. arXiv:1405.7658 [q-bio].