LARGE-SCALE GENOTYPING BY NEXT GENERATION SEQUENCING: HOW TO OVERCOME THE CHALLENGES TO RELIABLY GENOTYPE INDIVIDUALS?

M. Ferrandiz-Rovira[1,2], T. Bigot[1], D. Allainé[1], M-P. Callait-Cardinal[1,2], J. Radwan[3] and A. Cohas[1]

[1]Université de Lyon, F-69000, Lyon; Université Lyon 1; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69622, Villeurbanne, France.
[2] Univ Lyon, VetAgro Sup Campus Vet, F-69280 Marcy L'étoile, France.
[3]Institute of Environmental Sciences, Jagiellonian University, Krakow, Poland.
E-mail: mariona.ferrandiz@univ-lyon1.fr

INTRODUCTION

Large-scale and reliable genotypes at individual level are required to understand the role of selective pressures on the evolution of highly polymorphic functional genes. Next-generation sequencing techniques allow for such large-scale sequencing, but post-processing treatment is still a challenge.

OBJECTIVES

→ We develop a methodology to obtain reliable genotypes after an optimized 454 sequencing and post-processing.
→ We apply this methodology to an extensive dataset.
→ We assess the quality of assigned genotypes through three different procedures.

DATA

DNA extracted from hairs of 1,096 Alpine marmots (*Marmota marmota*).

454 SEQUENCING

Four loci of the major histocompatibility complex [1].

Well-optimized primers [1] and individual tags.

Moderate number of PCR cycles (24).

1,152 amplicons at each loci (4,608 amplicons) on 8 parts of a Roche® 454 FLX sequencing instrument.

→ Low amplicon coverage to improve the cost-efficiency.

COMPARISON OF LARGE-SCALE 454 SEQUENCING

Species	cost/genotype	% assigned	Ref.
Ficedula albicollis	26.07 €	100%	[2]
Parus major	8.38 €	56%	[3]
Myodes glareolus	92.39 €	82%	[4]
11 Rodent Genera	5.19 €	90%	[5]
Microcebus murinus	5.63 €	91%	[6]
Marmota marmota	**2.37 €**	**67%**	**This study**

454 POST-PROCESSING PROPOSED METHODOLOGY

Processing description	UB reads (N amp)	UD reads (N amp)	DRB1 reads (N amp)	DRB2 reads (N amp)
Reads with individual tags and 3-bp loci markers	152,960 (1067)	113,862 (1069)	79,114 (966)	113,207 (1001)
Reads with > 95% and < 105% of locus sizes	72,214 (948)	59,642 (944)	49,212 (903)	64,110 (901)
Amplicons with > 11 reads for a given and locus	71,611 (835)	58,911 (793)	48,291 (720)	63,254 (726)
Reads with within a given amplicon with > 24% of allelic frequency with respect to the main read	70,839 (835)	58,481 (793)	47,667 (720)	62,682 (726)

→ High variability of coverage per amplicon decreased the number of retained reads and amplicons.

→ The optimization of 454 post-processing increased the number of retained reads and amplicons.

→ No artefacts, PCR chimeras or new alleles were found after 454 post-processing.

VALIDATION OF GENOTYPES

1,460 out of 3,074 obtained genotypes could be validated through intra-individual 454 sequencing repeatability (51 genotypes); 454 and Sanger sequencing repeatability (36 genotypes); and comparing the consistency of mother-father-offspring triads determined through paternity analysis (1,373 genotypes) [7].

RESULTS

High quality MHC genotypes were obtained at the four MHC sequenced loci for 895 genotyped individuals, with a mean of 3.43 ± 0.96 (± SD) genotypes per individual and an error rate estimated to 0.3%.

CONCLUSIONS

→ Methods applicable methods to any organism for which the number of loci is known.

→ Our results emphasize the potential of 454 sequencing and 454 post-processing to obtain large-scale highly reliable genotypes.

→ The use of the proposed methodology could improve research on the evolution of highly polymorphic functional genes.

REFERENCES

[1] Kuduk K et al. (2012) *J* Evol Biol*,* 25, 1686-1693.
[2] Zagalska-Neubauer M et al. *BMC* Evol Biol, 10, 395.
[3] Sepil I et al. Evol Biol, 12, 68.
[4] Babik W et al. *Mol Ecol Resour*, 9, 713-719.
[5] Galan M et al. *BMC Genomics*, 11, 296.
[6] Huchard E et al. *Immunogenetics*, 64, 895-913.
[7] Cohas A et al. (2008). *Anim Behav* 76, 87-95

ACKNOWLEDGEMENTS

We warmly thank all people who helped in marmots catching (specially the Earthwatch volunteers). We are grateful to the authorities of the Vanoise National Park for granting us permission to work in La Grande Sassière Nature Reserve.