View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Ghent University Academic Bibliography

FACULTEIT FARMACEUTISCHE WETENSCHAPPEN

Whole exome sequencing the dog

Bart J.G. Broeckx[1], Christophe Hitte[2], Frank Coopman[3], Geert Verhoeven[4], Sarah De Keulenaer[1], Ellen De Meester[1], Thomas Derrien[2], Jessica Alfoldi[5], Kerstin Lindblad-Toh[5,6], Tim Bosmans[7], Valérie Bavegems[7], Ingrid Gielen[4], Henri Van Bree[4], Bernadette Van Ryssen[4], Jimmy H. Saunders[4], Filip Van Nieuwerburgh[1‡], Dieter Deforce[1‡]

[1] Laboratory for Pharmaceutical Biotechnology, Faculty of Pharmaceutical Sciences, Ghent University, Ghent, Belgium.
[2] Institut de Génétique et Développement de Rennes, CNRS-URM6290, Université Rennes1, Rennes, France.
[3] Faculty of Bioscience Engineering, University College Ghent, Ghent, Belgium.
[4] Department of Medical Imaging and Small Animal Orthopaedics, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium.
[5] Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA.
[6] Science for Life Laboratory, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden.
[7] Department of Medicine and Clinical Biology of Small Animals, Faculty of Veterinary Medicine, Ghent University, Merelbeke, Belgium.
[‡] These authors contributed equally. Corresponding authors: dieter.deforce@ugent.be, bart.broeckx@ugent.be

Definition:
Targeted resequencing of those regions transcribed to mRNA.

History:
- 2014: publication of the report detailing the design and performance of a whole exome sequencing (WES) enrichment assay for the dog: the exome -1.0[1]
- 2015: release of updated canine annotation[2]
  => development of 2 novel WES desigsn:
  - Exome-CDS
  - Exome-plus

Probe design: Roche Nimblegen SeqCap EZ Developer

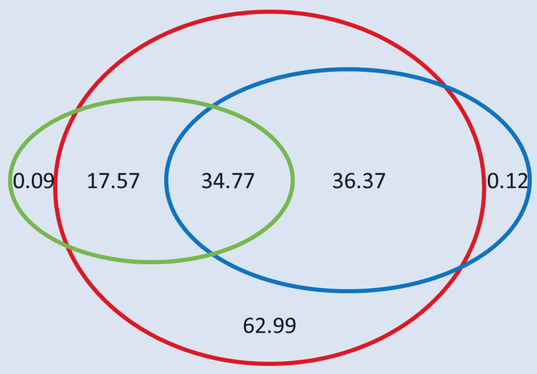
Sequencing:
- Exome-1.0: 8 dogs, 2 captures (HiSeq 2500 PE 100 bp)
- Exome-plus: 16 dogs, 4 captures (NextSeq 500 PE 75 bp)
- Capture = 4 pre-capture pooled, barcoded Illumina libraries
- Performance of the exome-CDS: estimated from the exome-plus

Average output:

	Reads (in million)				Sequencing depth (x)
	Total	Mapped	Duplicate	Remaining (%)	
1.0	90	86	6	80 (88.2%)	101.6
plus	243	226	14	212 (87.2%)	68.3

Design:

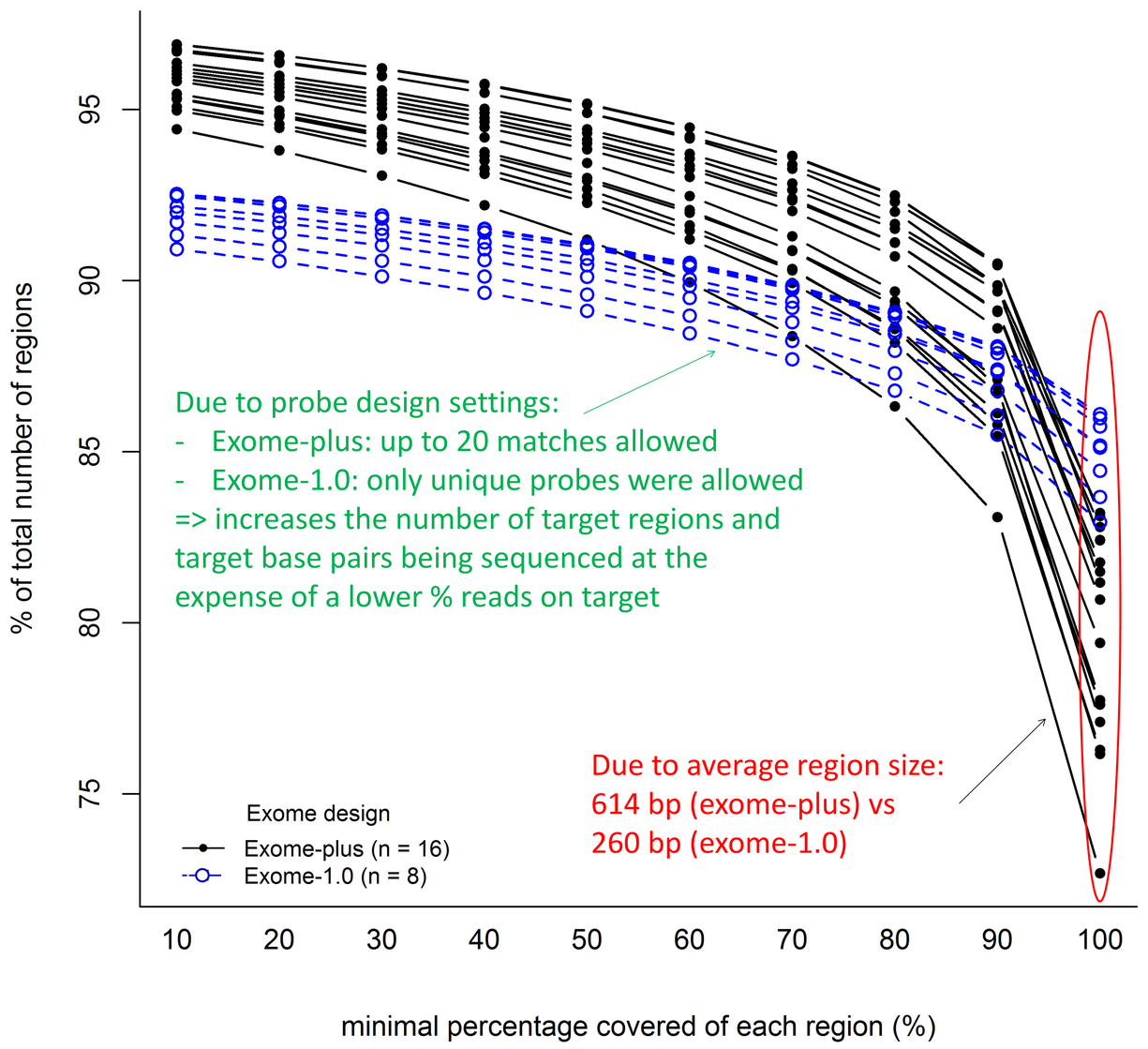
	Exome-1.0	Exome-CDS	Exome-plus
Ensembl Genes	✓	✓ (excl. UTR)	✓
RefSeq Genes	✓	✗	✗
Novel Broad genes[2]	✗	✓ (excl. UTR)	✓
mRNA	✓	✗	✗
microRNAs	✓	✓	✓
long non-coding RNAs[2]	✗	✓	✓
Antisense transcripts[2]	✗	✓	✓
Candidate CDS transdecoder	✗	✓	✗
Number of regions	203,059	244,543	242,914



Exome-plus (= 152 Mb)
Exome-1.0 (= 52 Mb)
Exome-CDS (= 71 Mb)

Comparison average performance:

	Exome-1.0	Exome-CDS	Exome-plus
fully covered regions (%)	84.9	85.4	79.7
base pairs covered (%)	90.2	93.4	95.1
% on target (%)	90.4	-	75.8
reproducibility regions (%)	79.9	71.4	63.5
reproducibility base pairs (%)	87.4	87.7	90.4
non-reference variants (n)	61,820	117,442	266,857



Conclusion:
- Overall: cost-efficient direct sequencing approach
- User-specific customization options
- Exome-1.0: the core set of protein coding genes
- Exome-plus, exome-CDS: contain many regulatory regions
⇒ added value for complex diseases where mutations influencing expression are more likely to be involved[2]
- Exome-plus: most comprehensive design
- Exome-CDS balances completeness and cost-efficiency

Reference list
1. Broeckx, B.J.G. et al. Development and performance of a targeted whole exome sequencing enrichment kit for the dog (Canis Familiaris Build 3.1). Sci. Rep. 4, 5597; doi:10.1038/srep05597.
2. Hoeppner MP, Lundquist A, Pirun M, Meadows JRS, Zamani N, et al. (2014) An Improved Canine Genome and a Comprehensive Catalogue of Coding Genes and Non-Coding Transcripts. PLoS ONE 9(3): e91172. doi:10.1371/journal.pone.0091172.

Acknowledgements
The authors would like to thank the Agency for Innovation by Science and Technology in Flanders (IWT) for providing a PhD grant. We would like to express our gratitude to all the veterinarians, assistance dogs, the Belgian rescue dogs federation, laboratories and owners that participated.