

Table of contents

1 Publications.....	3
2 Abstract.....	4
2.1 En français.....	4
2.2 In English.....	4
3 Introduction.....	5
3.1 What is a gene?.....	5
3.2 cDNA libraries screening.....	5
3.3 Expressed sequenced tags.....	6
3.4 Comparative genomics.....	6
3.5 Automated methods.....	7
3.6 Manual annotation.....	8
3.7 Experimental validation.....	9
3.8 Visualization and databases.....	10
4 Gene annotation and diseases.....	12
4.1 Introduction.....	12
4.2 Article 1.....	14
4.3 Article 2.....	20
4.4 Conclusion.....	26
5 Experimental validation of the GENCODE annotation.....	28
5.1 Introduction.....	28
5.1.1 Large-scale genomic projects.....	28
5.1.2 The ENCODE project.....	29
5.1.3 The GENCODE project.....	31
5.2 The RT-PCR-seq method.....	34
5.2.1 Background.....	34
5.2.2 Primer design.....	35
5.2.3 RT-PCR.....	36
5.2.4 Solexa sequencing.....	36
5.2.5 Validation of the exon junctions.....	37
5.3 Results.....	38
5.3.1 Accuracy of the RT-PCR-seq method.....	38
5.3.2 Primer dimer events.....	38
5.3.3 GENCODE gene set validation.....	41
5.3.4 Split mapping strategy.....	43
5.4 Conclusion.....	46
6 Chimeric induced transcripts.....	49
6.1 Introduction.....	49
6.2 Article.....	52
6.3 Conclusion.....	69
7 Conclusion.....	71
8 Appendix.....	73
Appendix A Experimental validation. Primer3 parameters.....	73
Appendix B Experimental validation. PCR parameters (multi-exonic).....	73
Appendix C Experimental validation. Positive controls.....	74
9 References.....	75