

# Intra-genic Exon Duplications in the Human Transcriptome

Anke Busch & Peter F. Stadler

studla@bioinf.uni-leipzig.de

## Supplemental File 6

Examples of problems in RefSeq gene annotation. a) identical gene location but different gene names; b) different gene locations of the same gene (e.g. in case of gene duplications); c) identical gene names referring to several distinct but overlapping exons (alternative splicing). Differentiating numbers are given in bold.

name	chrom	strand	exon Count	exonStarts	exonEnds	name2
a)						
<b>NR_026818</b>	chr1	-	3	24474,25139,25583,	25037,25344,25944,	<b>FAM138A</b>
<b>NR_026820</b>	chr1	-	3	24474,25139,25583,	25037,25344,25944,	<b>FAM138F</b>
<b>NR_026822</b>	chr1	-	3	24474,25139,25583,	25037,25344,25944,	<b>FAM138C</b>
b)						
NM_001005224	chr1	+	1	<b>357521,</b>	<b>358460,</b>	OR4F3
NM_001005224	chr1	-	1	<b>610958,</b>	<b>611897,</b>	OR4F3
or						
NM_000451	<b>chrX</b>	+	6	505078,511200,515352,521555, 521733,525125,	505337,511909,515561,521613, 521822,527558,	SHOX
NM_000451	<b>chrY</b>	+	6	505078,511200,515352,521555, 521733,525125,	505337,511909,515561,521613, 521822,527558,	SHOX
or						
NM_173571	chrX	-	3	<b>119895375,119896653,119897696,</b>	<b>11989515,119896809,119898693,</b>	CT47A11
NM_173571	chrX	-	3	<b>119900236,119901514,119902557,</b>	<b>119900376,119901670,119903554,</b>	CT47A11
NM_173571	chrX	-	3	<b>119905095,119906374,119907417,</b>	<b>119905236,119906530,119908414,</b>	CT47A11
NM_173571	chrX	-	3	<b>119909957,119911235,119912278,</b>	<b>119910097,119911391,119913275,</b>	CT47A11
NM_173571	chrX	-	3	<b>119924561,119925839,119926882,</b>	<b>119924701,119925995,119927879,</b>	CT47A11
c)						
NR_026988	chr1	+	3	87595447,87599296, <b>87599647,</b>	87595677,87599401, <b>87602350,</b>	LOC339524
NR_026989	chr1	+	3	87595447,87599296, <b>87633818,</b>	87595677,87599401, <b>87634884,</b>	LOC339524