

# LNCipedia 2.0 update

Pieter-Jan Volders

January 15, 2013

## 1 What's new

In LNCipedia version 2.0 we introduce over 10,000 new lncRNA sequences, thereby expanding our lncRNA set to 32,198. This corresponds to 6,000 new lncRNA genes/loci. All the integrated tools have been run for the new sequences apart from MirTarget2 and the PRIDE reprocessing pipeline. Output for these two tools will be integrated in the near future.

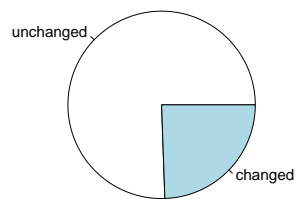
	Genes	Transcripts
Version 1.3	11439	21504
Version 2.0	17513	32198

## 2 LncRNA name changes

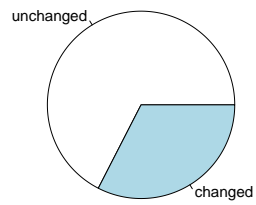
As discussed in Volders et al., 2013[1], LNCipedia employs a unique naming schema for lncRNAs. LncRNA transcripts are clustered into genes and named after the closest protein-coding gene on the same strand. Therefore this numbering can change when new transcripts are introduced. A portion of the transcripts and genes present in LNCipedia version 1.3 has a different name in version 2.0 as displayed in figure 1. Due to the major 2.0 update, gene names were changed. For future releases, gene names will not change and can hence be considered stable identifiers.

## References

- [1] Pieter-Jan Volders, Kenny Helsens, Xiaowei Wang, Björn Menten, Lennart Martens, Kris Gevaert, Jo Vandesompele and Pieter Mestdagh, *LNCipedia: a database for annotated human lncRNA transcript sequences and structures*. Nucleic Acids Research, 2013.



(a) Genes



(b) Transcripts

Figure 1: Overview of the fractions of genes and transcripts with changed and unchanged names the in LNCipedia 2.0.