



# The Open Bioinformatics Journal

## Supplementary Material

Content list available at: [www.benthamopen.com/TOBIOIJ/](http://www.benthamopen.com/TOBIOIJ/)

DOI: 10.2174/1875036201811010182, 2018, 11, i-iii



## RESEARCH ARTICLE

## Long Non-Coding RNAs in the Human Genome Acquired by Horizontal Gene Transfer

Rian Pierneef, Frederick J. Clasen, Oliver K. I. Bezuidt and Oleg N. Reva\*

*Centre for Bioinformatics and Computational Biology (CBCB), Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa*

Received: May 25, 2018

Revised: July 31, 2018

Accepted: August 10, 2018

## SUPPORTIVE/SUPPLEMENTARY MATERIAL

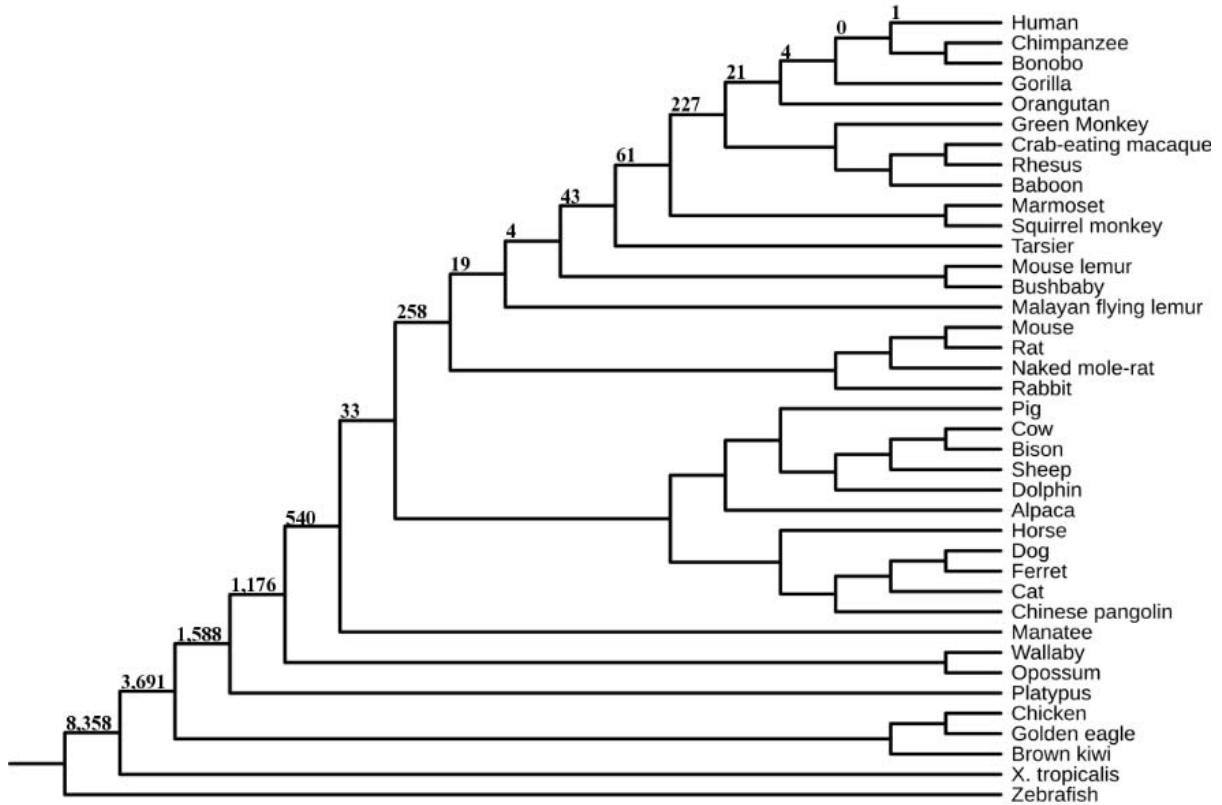
tRNAs located in GIs in the human genome.

Table S1. General statistics regarding GIs identified in human chromosomes.

Chromosome	Number of tRNAs	Number of tRNAs located in GIs	Percentage of GIs containing tRNAs	Odds Ratio	Adjusted p-value
1	154	59	38.31%	4.016715	3.62E-09
2	28	9	32.14%	2.055368	> 0.99
3	12	1	8.33%	0.3757768	> 0.99
4	5	0	0%	0	> 0.99
5	24	0	0%	0	1.82E-01
6	183	21	11.48%	0.450461	3.12E-02
7	26	0	0%	0	1.13E-01
8	14	2	14.29%	0.6927681	> 0.99
9	8	2	25%	1.40028	> 0.99
10	6	1	16.67%	0.8357041	> 0.99
11	19	1	5.26%	0.226578	> 0.99
12	16	4	25%	1.406895	> 0.99
13	7	1	14.29%	0.6952451	> 0.99
14	23	0	0%	0	3.00E-01
15	11	2	18.18%	0.9287638	> 0.99
16	34	6	17.65%	0.8915104	> 0.99
17	42	7	16.67%	0.8273512	> 0.99
18	4	0	0%	0	> 0.99
19	14	1	7.14%	0.3167633	> 0.99
20	7	1	14.29%	0.6952451	> 0.99
21	2	0	0%	0	> 0.99
22	1	0	0%	0	> 0.99
X	7	6	85.71%	26.12126	6.60E-03
Y	1	1	100%	Inf	> 0.99

**Supplementary File 2**

lncRNA insertion events in concordance with human speciation. The numbers on the branches indicate the amount of proposed lncRNA insertions after an evolutionary split. This dendrogram was obtained from TimeTree [40] and edited with iTOL [46].



Supplement Fig. (2).

**Supplementary File 3**

**Table S3. Genes located in pairwise non-aligning regions, within or spanning GIs, to at least 10 other species.**

Chromosome	Gene ID	Gene Name	Gene Description	Gene Biotype	Number of Species with Non-Alignment
4	ENSG00000186146	DEFB131A	defensin beta 131A	Protein coding	10
4	ENSG00000280043	AC118282.4	-	To be Experimental Confirmed	26
9	ENSG00000279561	AL845472.1	-	To be Experimental Confirmed	21
15	ENSG00000270824	IGHD5OR15-5B	immunoglobulin heavy diversity 5/OR15-5B	Immunoglobulin variable chain	11
15	ENSG00000282089	IGHD3OR15-3B	immunoglobulin heavy diversity 3/OR15-3B	Immunoglobulin variable chain	10
15	ENSG00000282268	IGHD2OR15-2B	immunoglobulin heavy diversity 2/OR15-2B	Immunoglobulin variable chain	11
15	ENSG00000270185	IGHD1OR15-1B	immunoglobulin heavy diversity 1/OR15-1B	Immunoglobulin variable chain	13
16	ENSG00000205457	TP53TG3C	TP53 target 3C	Protein coding	11
16	ENSG00000261509	TP53TG3B	TP53 target 3B	Protein coding	10

Supplementary File 4

Table S4. lncRNAs located in GIs the human genome.

Chromosome	Number of lncRNAs	Number of lncRNAs in GIs	Percentage of GIs containing lncRNAs	Odds Ratio	Adjusted p-value
1	12,828	2,477	80.07%	1.073476	6.46E-02
2	12,296	2,203	78.78%	0.97038	> 0.99
3	8,444	1,459	81.13%	0.9269721	2.47E-01
4	6,334	914	75.00%	0.7427451	5.28E-15
5	7,778	1,489	83.61%	1.058581	> 0.99
6	8,188	1,071	73.38%	0.6568522	5.28E-15
7	7,148	1,412	78.44%	1.10268	3.69E-02
8	6,147	1,103	80.80%	0.9735549	> 0.99
9	5,611	1,226	78.80%	1.258117	2.23E-10
10	6,396	1,195	82.48%	1.025223	> 0.99
11	6,787	1,114	83.63%	0.8697575	6.10E-04
12	7,182	1,230	89.30%	0.9172356	1.71E-01
13	3,970	796	77.71%	1.121469	1.12E-01
14	4,996	952	86.91%	1.051096	> 0.99
15	5,406	1,071	83.84%	1.105457	1.05E-01
16	5,617	1,109	87.68%	1.100721	1.31E-01
17	6,939	1,402	86.55%	1.135787	9.77E-04
18	2,950	554	84.06%	1.031225	> 0.99
19	6,069	1,040	92.54%	0.9185428	3.51E-01
20	3,653	605	83.12%	0.882	1.25E-01
21	2,327	397	76.62%	0.9155771	> 0.99
22	3,116	656	82.20%	1.193317	2.21E-03
X	3,689	839	68.44%	1.322327	1.86E-10
Y	557	152	74.19%	1.676669	5.06E-06

© 2018 Pierneef *et al.*

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: (<https://creativecommons.org/licenses/by/4.0/legalcode>). This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.