Table 1. Recently originated de novo genes discovered in diverse eukaryotic lineages.

organisms	number of de novo genes	genes found in previous studies	notable examples and comments	references
Drosophila				
D. melanogaster	5	—	four are X-linked; all five have testis expression bias	[16]
D. yakuba and *D. erecta*	7 + 3	—		[17]
mainly *D. Yakuba*	11	—	seven are X-linked	[18]
D. melanogaster subgroup	1	—	*hydra*; testis expression	[19]
D. melanogaster subgroup	14	5	—	[20]
D. melanogaster group and *D. willistoni*	16	—		[21]
D. melanogaster	248 (106 fixed) proto-genes	—	discovered based on testis expression. Male-biased and underrepresented on X chromosome	[22]
mammals				
primates (*H. sapiens, P. troglodytes, M. mulatta)*	15	—	*PART1;* prostate carcinogenesis	[23]
hominoids	24	2	regulated RNA expression predates protein-coding potential. Transcription in cerebellum	[24]
hominids	1	—	*NCYM;* neuroblastoma pathogenesis	[25]
H. sapiens	3	—	*CLLU1;* upregulated in chronic lymphocytic leukaemia	[26]
H. sapiens	1	—	*FLJ33706* (*C20orf203*); expressed in brain; protein found in neurons.	[27]
H. sapiens	60	1		[28]
H. sapiens	1	—	*PBOV1;* mitigates cancer outcomes	[29]
H. sapiens	1		*ESRG;* essential for maintenance of pluripotency	[30]
M. musculus	1	—	*Poldi*; testis expression	[31]
M. musculus and *R. norvegicus*	69 + 6	—		[32]
plants				
Oryza	1	—	*OsDR10*; defence gene	[33]
A. thaliana	1	—	*QQS*; starch biosynthesis pathway	[34]
A. thaliana and Brassicaceae	25			[35]
Plasmodium				
P. vivax	13	—	5/13 have introns within the coding sequence	[36]
Yeast				
S. cerevisiae	1	—	*BSC4*; DNA repair, synthetic lethal	[37]
S. cerevisiae	1	—	*MDF1*; functional role in promoting vegetative growth	[38]
S. cerevisiae	1	—	*RDT1*; ORF is absent in some strains of *S. cerevisiae*	[39]
S. cerevisiae	~1900 proto-genes	—		[40]

16. Levine MT, Jones CD, Kern AD, Lindfors HA, Begun DJ. 2006 Novel genes derived from noncoding DNA in *Drosophila melanogaster* are frequently X-linked and exhibit testis-biased expression. *Proc. Natl Acad. Sci. USA* **103**, 9935–9939. (doi:10.1073/pnas.0509809103)

17. Begun DJ, Lindfors HA, Thompson ME, Holloway AK. 2006 Recently evolved genes identified from *Drosophila yakuba* and *D. erecta* accessory gland expressed sequence tags. *Genetics* **172**, 1675–1681. (doi:10.1534/genetics.105.050336)
18. Begun DJ, Lindfors HA, Kern AD, Jones CD. 2007 Evidence for de novo evolution of testis-expressed genes in the *Drosophila yakuba/Drosophila erecta* clade. *Genetics* **176**, 1131–1137. (doi:10.1534/genetics.106.069245)
19. Chen S-T, Cheng H-C, Barbash DA, Yang H-P. 2007 Evolution of *hydra*, a recently evolved testis-expressed gene with nine alternative first exons in *Drosophila melanogaster*. *PLoS Genet.* **3**, e107. (doi:10.1371/journal.pgen.0030107)
20. Zhou Q *et al.* 2008 On the origin of new genes in *Drosophila*. *Genome Res.* **18**, 1446–1455. (doi:10.1101/gr.076588.108)
21. Chen S, Zhang YE, Long M. 2010 New genes in *Drosophila* quickly become essential. *Science* **330**, 1682–1685. (doi:10.1126/science.1196380)
22. Zhao L, Saelao P, Jones CD, Begun DJ. 2014 Origin and spread of de novo genes in *Drosophila melanogaster* populations. *Science* **343**, 769–772. (doi:10.1126/science.1248286)
23. Toll-Riera M, Bosch N, Bellora N, Castelo R, Armengol L, Estivill X, Albà MM. 2009 Origin of primate orphan genes: a comparative genomics approach. *Mol. Biol. Evol.* **26**, 603–612. (doi:10.1093/molbev/msn281)
24. Xie C *et al.* 2012 Hominoid-specific de novo protein-coding genes originating from long non-coding RNAs. *PLoS Genet.* **8**, e1002942. (doi:10.1371/journal.pgen.1002942)
25. Suenaga Y *et al.* 2014 *NCYM*, a cis-antisense gene of *MYCN*, encodes a de novo evolved protein that inhibits GSK3β resulting in the stabilization of *MYCN* in human neuroblastomas. *PLoS Genet.* **10**, e1003996. (doi:10.1371/journal.pgen.1003996)
26. Knowles DG, McLysaght A. 2009 Recent *de novo* origin of human protein-coding genes. *Genome Res.* **19**, 1752–1759. (doi:10.1101/gr.095026.109)
27. Li C-Y *et al.* 2010 A human-specific de novo protein-coding gene associated with human brain functions. *PLoS Comp. Biol.* **6**, e1000734. (doi:10.1371/journal.pcbi.1000734)
28. Wu D-D, Irwin DM, Zhang Y-P. 2011 *De novo* origin of human protein-coding genes. *PLoS Genet.* **7**, e1002379. (doi:10.1371/journal.pgen.1002379)
29. Samusik N, Krukovskaya L, Meln I, Shilov E, Kozlov AP. 2013 *PBOV1* Is a human de novo gene with tumor-specific expression that is associated with a positive clinical outcome of cancer. *PLoS ONE* **8**, e56162. (doi:10.1371/journal.pone.0056162)
30. Wang J *et al.* 2014 Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **516**, 405–409. (doi:10.1038/nature13804)
31. Heinen TJAJ, Staubach F, Häming D, Tautz D. 2009 Emergence of a new gene from an intergenic region. *Curr. Biol.* **19**, 1527–1531. (doi:10.1016/j.cub.2009.07.049)
32. Murphy DN, McLysaght A. 2012 De novo origin of protein-coding genes in murine rodents. *PLoS ONE* **7**, e48650. (doi:10.1371/journal.pone.0048650)
33. Xiao W, Liu H, Li Y, Li X, Xu C, Long M, Wang S. 2009 A rice gene of de novo origin negatively regulates pathogen-induced defense response. *PLoS ONE* **4**, e4603. (doi:10.1371/journal.pone.0004603)
34. Li L, Foster CM, Gan Q, Nettleton D, James MG, Myers AM, Wurtele ES. 2009 Identification of the novel protein QQS as a component of the starch metabolic network in *Arabidopsis* leaves. *Plant J.* **58**, 485–498. (doi:10.1111/j.1365-313X.2009.03793.x)
35. Donoghue MT, Keshavaiah C, Swamidatta SH, Spillane C. 2011 Evolutionary origins of Brassicaceae specific genes in *Arabidopsis thaliana*. *BMC Evol. Biol.* **11**, 47. (doi:10.1186/1471-2148-11-47)
36. Yang Z, Huang J. 2011 *De novo* origin of new genes with introns in *Plasmodium vivax*. *FEBS Lett.* **585**, 641–644. (doi:10.1016/j.febslet.2011.01.017)
37. Cai J, Zhao R, Jiang H, Wang W. 2008 De novo origination of a new protein-coding gene in *Saccharomyces cerevisiae*. *Genetics* **179**, 487–496. (doi:10.1534/genetics.107.084491)
38. Li D, Dong Y, Jiang Y, Jiang H, Cai J, Wang W. 2010 A de novo originated gene depresses budding yeast mating pathway and is repressed by the protein encoded by its antisense strand. *Cell Res.* **20**, 408–420. (doi:10.1038/cr.2010.31)
39. Wilson BA, Masel J. 2011 Putatively noncoding transcripts show extensive association with ribosomes. *Genome Biol. Evol.* **3**, 1245–1252. (doi:10.1093/gbe/evr099)
40. Carvunis A-R *et al.* 2012 Proto-genes and de novo gene birth. *Nature* **487**, 370–374. (doi:10.1038/nature11184)