



OPEN

# Author Correction: Comparative genomic analysis of eutherian connexin genes

Marko Premzl

Correction to: *Scientific Reports* <https://doi.org/10.1038/s41598-019-53458-x>, published online 15 November 2019

This Article contains errors.

In the Introduction,

“For example, the protocol was applicable in initial descriptions of human genes<sup>50,52</sup>. There was positive correlation between genomic sequence redundancies of 35 public eutherian reference genomic sequence data sets respectively and published complete coding sequence numbers<sup>50</sup>.”

should read:

“For example, the protocol was applicable in initial descriptions of human genes<sup>50,51</sup>. There was positive correlation between genomic sequence redundancies of 35 public eutherian reference genomic sequence data sets respectively and published complete coding sequence numbers<sup>51</sup>.”

In the Results,

“For example, the present analysis initially described human *CXNK1* gene as complete coding sequence that disagreed with Fishman *et al.*<sup>53</sup>.”

should read:

“For example, the present analysis initially described human *CXNK1* gene as complete coding sequence that disagreed with Fishman *et al.*<sup>52</sup>.”

“Second, among eutherian *CXN* major gene clusters including orthologues respectively, there were nucleotide sequence identity calculations typical in comparisons between eutherian orthologues ( $\approx 0,65-0,9$ )<sup>49,50,52</sup>.”

should read:

“Second, among eutherian *CXN* major gene clusters including orthologues respectively, there were nucleotide sequence identity calculations typical in comparisons between eutherian orthologues ( $\approx 0,65-0,9$ )<sup>49,50,51</sup>.”

“Specifically, the major gene clusters *CXNH* (*GJA4*, *CX37*) and *CXNK* (*GJA1*, *CX43*) respectively included close eutherian orthologues and paralogues ( $\approx 0,7-0,85$ )<sup>49,50,52</sup>, but major gene clusters *CXNJ* (*GJA3*, *CX46*) and *CXNP* (*GJC3*, *CX30.2*, *CX31.3*) respectively included typical eutherian orthologues and paralogues ( $\approx 0,45-0,7$ ). Fourth, in comparisons between eutherian *CXN* major gene clusters, there were nucleotide sequence identity patterns of very close ( $>0,5$ ), close ( $\approx 0,35-0,5$ ), typical ( $\approx 0,25-0,35$ ), distant ( $\approx 0,15-0,25$ ) and very distant ( $<0,15$ ) eutherian homologues<sup>49,50,52</sup>.”

should read:

“Specifically, the major gene clusters *CXNH* (*GJA4*, *CX37*) and *CXNK* (*GJA1*, *CX43*) respectively included close eutherian orthologues and paralogues ( $\approx 0,7-0,85$ )<sup>49,50,51</sup>, but major gene clusters *CXNJ* (*GJA3*, *CX46*) and *CXNP*

Published online: 26 June 2020

(GJC3, CX30.2, CX31.3) respectively included typical eutherian orthologues and paralogues ( $\approx 0.45\text{--}0.7$ ). Fourth, in comparisons between eutherian *CXN* major gene clusters, there were nucleotide sequence identity patterns of very close ( $>0.5$ ), close ( $\approx 0.35\text{--}0.5$ ), typical ( $\approx 0.25\text{--}0.35$ ), distant ( $\approx 0.15\text{--}0.25$ ) and very distant ( $<0.15$ ) eutherian homologues<sup>49,50,51</sup>.”

In the Conclusions,

“Using eutherian comparative genomic analysis protocol and 35 public eutherian reference genomic sequence assemblies<sup>49,50,52</sup>, the present analysis attempted to update and revise comprehensive eutherian *CXN* gene data sets, and address and resolve major discrepancies in their descriptions.”

should read:

“Using eutherian comparative genomic analysis protocol and 35 public eutherian reference genomic sequence assemblies<sup>49,50,51</sup>, the present analysis attempted to update and revise comprehensive eutherian *CXN* gene data sets, and address and resolve major discrepancies in their descriptions.”

In the Methods,

“The eutherian comparative genomic analysis protocol RRID:SCR\_014401 integrated gene annotations, phylogenetic analysis and protein molecular evolution analysis with new genomics and protein molecular evolution tests into one framework of eutherian gene descriptions<sup>49,50,52</sup>.”

should read:

“The eutherian comparative genomic analysis protocol RRID:SCR\_014401 integrated gene annotations, phylogenetic analysis and protein molecular evolution analysis with new genomics and protein molecular evolution tests into one framework of eutherian gene descriptions<sup>49,50,51</sup>.”

“In identifications of potential *CXN* coding sequences in 35 eutherian reference genomic sequence data sets, the protocol used National Center for Biotechnology Information’s (NCBI) BLAST Genomes<sup>35,36,54,55</sup> (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) or Ensembl genome browser’s BLAST or BLAT<sup>37</sup> (<https://www.ensembl.org>). Second, the potential *CXN* coding sequences were then used in tests of reliability of eutherian public genomic sequences. The first test steps analysed nucleotide sequence coverages of each potential *CXN* coding sequence, using BLASTN<sup>54,55</sup> and processed public Sanger DNA sequencing reads or traces deposited in NCBI’s Trace Archive<sup>35</sup> (<https://www.ncbi.nlm.nih.gov/Traces/trace.cgi>).”

should read:

“In identifications of potential *CXN* coding sequences in 35 eutherian reference genomic sequence data sets, the protocol used National Center for Biotechnology Information’s (NCBI) BLAST Genomes<sup>35,36,53,54</sup> (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) or Ensembl genome browser’s BLAST or BLAT<sup>37</sup> (<https://www.ensembl.org>). Second, the potential *CXN* coding sequences were then used in tests of reliability of eutherian public genomic sequences. The first test steps analysed nucleotide sequence coverages of each potential *CXN* coding sequence, using BLASTN<sup>53,54</sup> and processed public Sanger DNA sequencing reads or traces deposited in NCBI’s Trace Archive<sup>35</sup> (<https://www.ncbi.nlm.nih.gov/Traces/trace.cgi>).”

“The protocol then deposited complete *CXN* coding sequences in European Nucleotide Archive as one curated eutherian gene data set<sup>56,57,58</sup> (<https://www.ebi.ac.uk/ena/about/tpa-policy>). In updated human and eutherian *CXN* gene classification and nomenclature, the protocol used guidelines of human gene nomenclature<sup>59</sup> (<http://www.genenames.org/about/guidelines>) and guidelines of mouse gene nomenclature (<http://www.informatics.jax.org/mgihome/nomen/gene.shtml>). Specifically, the present eutherian *CXN* gene name assignments used both phylogenetic analysis (Fig. 1) and genomic sequence information (Supplementary Data File 1). Third, the protocol used mVISTA’s program AVID in multiple pairwise genomic sequence alignments using default settings<sup>51,60</sup> (<http://genome.lbl.gov/vista/index.shtml>). In pairwise alignments with base sequences (*Homo sapiens*), the cut-offs of detection of common genomic sequence regions were calculated *a posteriori* using analyses of 11 eutherian major gene data sets<sup>49,50,52</sup> including 95% along 100 bp (*Homo sapiens*, *Pan troglodytes*, *Gorilla gorilla*), 90% along 100 bp (*Pongo abelii*, *Nomascus leucogenys*), 85% along 100 bp (*Macaca mulatta*, *Papio hamadryas*), 80% along 100 bp (*Callithrix jacchus*), 75% along 100 bp (*Tarsius syrichta*, *Microcebus murinus*, *Otolemur garnettii*), 65% along 100 bp (Rodentia) or 70% along 100 bp in other pairwise alignments.”

should read:

“The protocol then deposited complete *CXN* coding sequences in European Nucleotide Archive as one curated eutherian gene data set<sup>55,56,57</sup> (<https://www.ebi.ac.uk/ena/about/tpa-policy>). In updated human and eutherian *CXN* gene classification and nomenclature, the protocol used guidelines of human gene nomenclature<sup>58</sup> (<http://www.genenames.org/about/guidelines>) and guidelines of mouse gene nomenclature (<http://www.informatics.jax.org/mgihome/nomen/gene.shtml>). Specifically, the present eutherian *CXN* gene name assignments used

both phylogenetic analysis (Fig. 1) and genomic sequence information (Supplementary Data File 1). Third, the protocol used mVISTA's program AVID in multiple pairwise genomic sequence alignments using default settings<sup>59,60</sup> (<http://genome.lbl.gov/vista/index.shtml>). In pairwise alignments with base sequences (*Homo sapiens*), the cut-offs of detection of common genomic sequence regions were calculated *a posteriori* using analyses of 11 eutherian major gene data sets<sup>49,50,51</sup> including 95% along 100 bp (*Homo sapiens*, *Pan troglodytes*, *Gorilla gorilla*), 90% along 100 bp (*Pongo abelii*, *Nomascus leucogenys*), 85% along 100 bp (*Macaca mulatta*, *Papio hamadryas*), 80% along 100 bp (*Callithrix jacchus*), 75% along 100 bp (*Tarsius syrichta*, *Microcebus murinus*, *Otolemur garnettii*), 65% along 100 bp (Rodentia) or 70% along 100 bp in other pairwise alignments.”



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020