



## Transposable elements and circular DNAs

Tobias Mourier 

Pathogen Genomics Laboratory, Biological and Environmental Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

### ABSTRACT

Circular DNAs are extra-chromosomal fragments that become circularized by genomic recombination events. We have recently shown that yeast LTR elements generate circular DNAs through recombination events between their flanking long terminal repeats (LTRs). Similarly, circular DNAs can be generated by recombination between LTRs residing at different genomic loci, in which case the circular DNA will contain the intervening sequence. In yeast, this can result in gene copy number variations when circles contain genes and origins of replication. Here, I speculate on the potential and implications of circular DNAs generated through recombination between human transposable elements.

### ARTICLE HISTORY

Received 19 September 2016  
Accepted 20 September 2016

### KEYWORDS

Alu elements; circular DNAs; eccDNAs; evolution; genome; human; yeast

Although the presence of circular DNAs in the shape of fragments excised from the genome has been known for decades,<sup>1,2</sup> their prevalence has only recently been highlighted.<sup>3,4</sup> Circular DNAs are likely formed by non-homologous recombination events.<sup>5,6</sup> In this respect, the multitude and sequence redundancy of transposable elements should render them highly efficient agents for the generation of circular DNAs. We recently assessed the involvement of transposable elements in the genesis of circular DNAs in baker's yeast, *Saccharomyces cerevisiae*.<sup>7</sup> Detection of structural variants—of which circular DNAs are a subset—requires customized analysis and is not directly elucidated using conventional mapping of short sequence reads onto reference genomes.<sup>3,7</sup> Analysis is further complicated by the sequence redundancy of transposable elements.<sup>8,9</sup>

The baker's yeast genome contains 5 families of LTR elements occupying roughly 3% of the genome.<sup>10,11</sup> The LTR elements are retrotransposable elements encoding the proteins necessary for their movement and flanked by long-terminal repeats (LTRs). Most copies exist as solo LTR sequences,<sup>12</sup> presumably generated through recombination events.<sup>11</sup> When sequence reads highly enriched for circular

sequences<sup>3,13</sup> were mapped onto the yeast genome, a relative uniform coverage was observed across full-length LTR elements and this level of coverage did not extend outside the LTR borders.<sup>7</sup> This suggested that the majority of LTR sequences in circular DNAs exist as full-length LTR elements, although circles generated from LTR sequences residing at different genomic location have also been reported.<sup>3,14,15</sup>

LTR elements are structurally related to retroviruses, for which generation of circular structures from extra-chromosomal linear DNA is readily observed.<sup>16–18</sup> The presence of LTR sequences as circular DNA could therefore potentially both be a result of circularization of extra-chromosomal DNAs and the circularization arising from genomic recombination events. Circularization of linear extra-chromosomal LTR DNAs can happen in several ways, including nonhomologous end-joining, recombination between flanking LTR sequences, and through so-called auto-integration in which the LTR element inserts into its own sequence.<sup>16,18,19</sup> However, due to the fact that these scenarios are all preceded by transcription of the genomic LTR element, any resulting circular LTR sequence will display an apparent breakpoint at the transcription start site (despite the fact that they may not be created

by recombination events).<sup>7,20</sup> If nucleotide differences exist between the 2 LTR sequences flanking an element the apparent breakpoint site can be inferred,<sup>5</sup> and any apparent breakpoints outside the transcription start site region are therefore inconsistent with circles generated through circularization of linear extra-chromosomal DNA. Although only a few LTR sequences contained informative nucleotide differences, all tested sequences were inconsistent with being generated through circularization of linear extra-chromosomal DNA.<sup>7</sup> It therefore seems that yeast LTR elements are a source for circular DNAs through genomic recombination between the flanking LTR sequences. Interestingly, this provides a potential novel way of LTR element movement, in which full-length LTR elements in circular DNAs may recombine back into the genome at other LTR loci.<sup>21</sup>

### **Potential for DNA circularization by human transposable elements**

The human genome contains relatively few intact LTR elements.<sup>22</sup> However, as circles can be generated through recombination between repetitive sequences residing at different genomic loci,<sup>3</sup> recombination between the highly abundant human transposable elements is not an unlikely scenario.<sup>23</sup> One prominent candidate for generating circular DNAs would for example be the Alu family that is present in more than 1 million copies in human genomes.<sup>24</sup> Although several subfamilies of Alu elements have evolved during primate evolution, consensus sequences from the different families are highly similar but for a few diagnostic sites.<sup>25</sup> Alu elements are further known to participate in unequal homologous recombination, which has been associated with a range of human disease states.<sup>26</sup>

Although the sizes of coding sequences are highly uniform across eukaryotic genomes,<sup>27</sup> the presence of introns means that the total size of human genes can exceed several hundreds of kilobasepairs.<sup>28</sup> Although the possibility of entire human genes ending up in relatively small circular DNAs would at first appear to be limited, large circular DNAs as well as relatively short human genes would render this scenario far more conceivable. First, a circular DNA encompassing 39 kb has been reported in yeast, containing 2 histone genes as well as a centromere and origins of replication.<sup>14</sup> More than 80% of the human genes in the ensembl annotation would fit in circles of this size.<sup>29,30</sup>

Second, numerous human genes are deprived of introns and therefore occupying a limited genomic space. Intronless human genes—presumably generated through retrotransposition using the LINE L1 machinery<sup>31</sup>—are enriched for signal transducing- and regulatory genes, and for tissue-specific expression in brain or testis.<sup>32,33</sup> Clearly, such relatively short genes could be contained in circular DNAs from a more modest size range.

### **Implications of DNA circularization by human transposable elements**

In nitrogen-deprived yeast cultures amplifications and deletions of the GAP1 gene have been associated with the presence of circular DNAs harboring the GAP1 gene and an origin of replication.<sup>5</sup> As nitrogen limitation results in elevated GAP1 expression,<sup>34</sup> this prompted speculations that DNA circularization could provide a means for the cell to increase GAP1 activity through the presence of multiple gene copies residing in DNA circles.<sup>5</sup>

Although the number of cell divisions varies immensely between human tissues,<sup>35</sup> and the specific nature of mammalian replication origins is still not well-understood,<sup>36</sup> circular DNA carrying human genes could increase in numbers, providing additional gene copies to the cell. Yet, affecting gene copy numbers may not be the only impact of circular DNAs. Transcriptional regulation in mammalian cells often works at a regional scale, where genetic promoter regions physically interact with enhancer structures located at a considerable distance.<sup>37</sup> Transcription is further regulated by epigenetic modifications, which can travel across larger genomic regions<sup>38</sup>—the spread of which intriguingly can be stopped by the presence of transcriptionally active transposable elements.<sup>39</sup> Transcription initiated at one genomic site may also drive or repress transcription at neighboring sites. This is referred to as transcriptional interference,<sup>40</sup> a phenomenon that may even shape genome architecture.<sup>41</sup>

A gene residing in a circular DNA should hence evade such regional regulatory effects, and one may therefore speculate if the major impact of circular DNAs is not primarily in gene copy number variations, but in taking genes out of their regulatory context. It is also conceivable that gene copies residing in circular DNAs bind regulatory proteins, thereby potentially changing the dynamics and kinetics at

other chromosomal loci. Similarly, transcripts from genes in circular DNAs may act as sinks by binding regulatory RNAs as well as proteins. Such an effect may obviously manifest itself without the entire gene sequence being present in circular DNAs, and circular DNAs harboring incomplete gene sequences may serve as previously described regulatory pseudogenes.<sup>42</sup>

In summary, we have previously shown that circular DNAs are generated from genomic copies of yeast transposable elements.<sup>7</sup> It is currently unknown to which extent this is happening in mammalian genomes rich in transposable element sequences. Circular DNAs have a huge potential in altering gene copy numbers,<sup>5</sup> and – as speculated here – in harboring genes that will be taken out of their regulatory context. Whether the generation of circular DNAs has an adaptive aspect, and may add novel functionality to the cell is of course an entirely different question. As circular DNAs from ribosomal genes are found in aging yeast cells,<sup>43</sup> circularization may simply reflect a genomic deterioration associated with cellular decay. The true extent to which human transposable elements are a source for the generation of circular DNAs will be determined in coming experiments. But crucially, such experiments need to be tailored toward circular structures both in experimental design and in downstream analysis.

### Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

### ORCID

Tobias Mourier  <http://orcid.org/0000-0002-8522-7583>

### References

- [1] von Schwedler U, Jack HM, Wabl M. Circular DNA is a product of the immunoglobulin class switch rearrangement. *Nature* 1990; 345:452-6; PMID:2111465; <http://dx.doi.org/10.1038/345452a0>
- [2] Horowitz H, Haber JE. Identification of autonomously replicating circular subtelomeric Y' elements in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1985; 5:2369-80; PMID:3915542; <http://dx.doi.org/10.1128/MCB.5.9.2369>
- [3] Moller HD, Parsons L, Jorgensen TS, Botstein D, Regenberg B. Extrachromosomal circular DNA is common in yeast. *Proc Natl Acad Sci U S A* 2015; 112:E3114-22; PMID:26038577; <http://dx.doi.org/10.1073/pnas.1508825112>
- [4] Dillon LW, Kumar P, Shibata Y, Wang YH, Willcox S, Griffith JD, Pommier Y, Takeda S, Dutta A. Production of extrachromosomal MicroDNAs is linked to mismatch repair pathways and transcriptional activity. *Cell Rep* 2015; 11:1749-59; PMID:26051933; <http://dx.doi.org/10.1016/j.celrep.2015.05.020>
- [5] Gresham D, Usaite R, Germann SM, Lisby M, Botstein D, Regenberg B. Adaptation to diverse nitrogen-limited environments by deletion or extrachromosomal element formation of the GAP1 locus. *Proc Natl Acad Sci U S A* 2010; 107:18551-6; PMID:20937885; <http://dx.doi.org/10.1073/pnas.1014023107>
- [6] Cohen S, Segal D. Extrachromosomal circular DNA in eukaryotes: possible involvement in the plasticity of tandem repeats. *Cytogenetic Genome Res* 2009; 124:327-38; PMID:19556784; <http://dx.doi.org/10.1159/000218136>
- [7] Moller HD, Larsen CE, Parsons L, Hansen AJ, Regenberg B, Mourier T. Formation of extrachromosomal circular DNA from long terminal repeats of Retrotransposons in *Saccharomyces cerevisiae*. *G3* 2016; 6:453-62; <http://dx.doi.org/10.1534/g3.115.025858>
- [8] Ewing AD. Transposable element detection from whole genome sequence data. *Mobile DNA* 2015; 6:24; PMID:26719777; <http://dx.doi.org/10.1186/s13100-015-0055-3>
- [9] Mourier T, Willerslev E. Large-scale transcriptome data reveals transcriptional activity of fission yeast LTR retrotransposons. *BMC Genomics* 2010; 11:167; PMID:20226011; <http://dx.doi.org/10.1186/1471-2164-11-167>
- [10] Kim JM, Vanguri S, Boeke JD, Gabriel A, Voytas DF. Transposable elements and genome organization: A comprehensive survey of retrotransposons revealed by the complete *saccharomyces cerevisiae* genome. *Genome Res* 1998; 8:464-78; PMID:9582191
- [11] Jordan IK, McDonald JF. Tempo and mode of Ty element evolution in *Saccharomyces cerevisiae*. *Genetics* 1999; 151:1341-51; PMID:10101161
- [12] Carr M, Bensasson D, Bergman CM. Evolutionary genomics of transposable elements in *Saccharomyces cerevisiae*. *PLoS One* 2012; 7:e50978; PMID:23226439; <http://dx.doi.org/10.1371/journal.pone.0050978>
- [13] Moller HD, Bojsen RK, Tachibana C, Parsons L, Botstein D, Regenberg B. Genome-wide purification of extrachromosomal circular DNA from Eukaryotic cells. *J Vis Exp* 2016:e54239; PMID:27077531; <http://dx.doi.org/10.3791/54239>
- [14] Libuda DE, Winston F. Amplification of histone genes by circular chromosome formation in *Saccharomyces cerevisiae*. *Nature* 2006; 443:1003-7; PMID:17066037; <http://dx.doi.org/10.1038/nature05205>
- [15] Mieczkowski PA, Lemoine FJ, Petes TD. Recombination between retrotransposons as a source of chromosome rearrangements in the yeast *Saccharomyces cerevisiae*. *DNA Repair* 2006; 5:1010-20; PMID:16798113; <http://dx.doi.org/10.1016/j.dnarep.2006.05.027>
- [16] Farnet CM, Haseltine WA. Circularization of human immunodeficiency virus type 1 DNA in vitro. *J Virol* 1991; 65:6942-52; PMID:1834863
- [17] Hong T, Drlica K, Pinter A, Murphy E. Circular DNA of human immunodeficiency virus: analysis of circle

- junction nucleotide sequences. *J Virol* 1991; 65:551-5; PMID:1985217
- [18] Garfinkel DJ, Stefanisko KM, Nyswaner KM, Moore SP, Oh J, Hughes SH. Retrotransposon suicide: formation of Ty1 circles and autointegration via a central DNA flap. *J Virol* 2006; 80:11920-34; PMID:17005648; <http://dx.doi.org/10.1128/JVI.01483-06>
- [19] Kilzer JM, Stracker T, Beitzel B, Meek K, Weitzman M, Bushman FD. Roles of host cell factors in circularization of retroviral dna. *Virology* 2003; 314:460-7; PMID:14517098; [http://dx.doi.org/10.1016/S0042-6822\(03\)00455-0](http://dx.doi.org/10.1016/S0042-6822(03)00455-0)
- [20] Telesnitsky A, Goff SP. Reverse Transcriptase and the Generation of Retroviral DNA. In: Coffin JM, Hughes SH, Varmus HE, eds. *Retroviruses*. (NY): Cold Spring Harbor; 1997; PMID: 21433342
- [21] Mourier T. Potential movement of transposable elements through DNA circularization. *Curr Genetics* 2016;1-4; PMID:26045103; <http://dx.doi.org/10.1007/s00294-016-0592-4>
- [22] Izsvak Z, Wang J, Singh M, Mager DL, Hurst LD. Pluripotency and the endogenous retrovirus HERVH: Conflict or serendipity? *BioEssays* 2016; 38:109-17; PMID:26735931; <http://dx.doi.org/10.1002/bies.201500096>
- [23] Gaubatz JW. Extrachromosomal circular DNAs and genomic sequence plasticity in eukaryotic cells. *Mutation Res* 1990; 237:271-92; PMID:2079966; [http://dx.doi.org/10.1016/0921-8734\(90\)90009-G](http://dx.doi.org/10.1016/0921-8734(90)90009-G)
- [24] Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. Initial sequencing and analysis of the human genome. *Nature* 2001; 409:860-921; PMID:11237011; <http://dx.doi.org/10.1038/35057062>
- [25] Batzer MA, Deininger PL. Alu repeats and human genomic diversity. *Nat Rev Genet* 2002; 3:370-9; PMID:11988762; <http://dx.doi.org/10.1038/nrg798>
- [26] Deininger PL, Batzer MA. Alu repeats and human disease. *Mol Genetics Metab* 1999; 67:183-93; PMID:10381326; <http://dx.doi.org/10.1006/mgme.1999.2864>
- [27] Ramaprasad A, Mourier T, Naem R, Malas TB, Moussa E, Panigrahi A, Vermont SJ, Otto TD, Wastling J, Pain A. Comprehensive evaluation of *Toxoplasma gondii* VEG and *Neospora caninum* LIV genomes with tachyzoite stage transcriptome and proteome defines novel transcript features. *PloS One* 2015; 10:e0124473; PMID:25875305; <http://dx.doi.org/10.1371/journal.pone.0124473>
- [28] Jeffares D, Mourier T, Penny D. The biology of intron gain and loss. *Trends Genetics* 2006; 22:16-22; PMID:16290250; <http://dx.doi.org/10.1016/j.tig.2005.10.006>
- [29] Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Fitzgerald S, Gil L, et al. Ensembl 2016. *Nucleic Acids Res* 2016; 44: D710-6; PMID:26687719; <http://dx.doi.org/10.1093/nar/gkv1157>
- [30] Smedley D, Haider S, Durinck S, Pandini L, Provero P, Allen J, Arnaiz O, Awedh MH, Baldock R, Barbiera G, et al. The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Res* 2015; 43:W589-98; PMID:25897122; <http://dx.doi.org/10.1093/nar/gkv350>
- [31] Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* 2000; 24:363-7; PMID:10742098; <http://dx.doi.org/10.1038/74184>
- [32] Brosius J. Many G-protein-coupled receptors are encoded by retrogenes. *Trends Genetics* 1999; 15:304-5; PMID:10431191; [http://dx.doi.org/10.1016/S0168-9525\(99\)01783-7](http://dx.doi.org/10.1016/S0168-9525(99)01783-7)
- [33] Grzybowska EA. Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. *Biochem Biophys Res Commun* 2012; 424:1-6; PMID:22732409; <http://dx.doi.org/10.1016/j.bbrc.2012.06.092>
- [34] Usaite R, Patil KR, Grotkjaer T, Nielsen J, Regenberg B. Global transcriptional and physiological responses of *Saccharomyces cerevisiae* to ammonium, L-alanine, or L-glutamine limitation. *Appl Environ Microbiol* 2006; 72:6194-203; PMID:16957246; <http://dx.doi.org/10.1128/AEM.00548-06>
- [35] Tomasetti C, Vogelstein B. Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* 2015; 347:78-81; PMID:25554788; <http://dx.doi.org/10.1126/science.1260825>
- [36] Prioleau MN, MacAlpine DM. DNA replication origins—where do we begin? *Genes Dev* 2016; 30:1683-97; PMID:27542827; <http://dx.doi.org/10.1101/gad.285114.116>
- [37] Chepelev I, Wei G, Wangsa D, Tang Q, Zhao K. Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization. *Cell Res* 2012; 22:490-503; PMID:22270183; <http://dx.doi.org/10.1038/cr.2012.15>
- [38] Grewal SI, Moazed D. Heterochromatin and epigenetic control of gene expression. *Science* 2003; 301:798-802; PMID:12907790; <http://dx.doi.org/10.1126/science.1086887>
- [39] Lunyak VV, Prefontaine GG, Nunez E, Cramer T, Ju BG, Ohgi KA, Hutt K, Roy R, Garcia-Diaz A, Zhu X, et al. Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science* 2007; 317:248-51; PMID:17626886; <http://dx.doi.org/10.1126/science.1140871>
- [40] Mazo A, Hodgson JW, Petruk S, Sedkov Y, Brock HW. Transcriptional interference: an unexpected layer of complexity in gene regulation. *J Cell Sci* 2007; 120:2755-61; PMID:17690303; <http://dx.doi.org/10.1242/jcs.007633>
- [41] Mourier T, Willerslev E. Does selection against transcriptional interference shape retroelement-free regions in mammalian genomes? *PloS One* 2008; 3:e3760; PMID:19018283; <http://dx.doi.org/10.1371/journal.pone.0003760>
- [42] Polisenio L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 2010; 465:1033-8; PMID:20577206; <http://dx.doi.org/10.1038/nature09144>
- [43] Sinclair DA, Guarente L. Extrachromosomal rDNA circles—a cause of aging in yeast. *Cell* 1997; 91:1033-42; PMID:9428525; [http://dx.doi.org/10.1016/S0092-8674\(00\)80493-6](http://dx.doi.org/10.1016/S0092-8674(00)80493-6)