Title

Temporal and spatial rearrangements of a repetitive element array on C57BL/6J mouse genome

Journal

Experimental and Molecular Pathology, 98(3)

ISSN

0014-4800

Authors

Lee, Kang-Hoon
Yee, Lisa
Lim, Debora
et al.

Publication Date

2015-06-01

DOI

10.1016/j.yexmp.2015.03.037

Peer reviewed

HHS Public Access

Author manuscript

Exp Mol Pathol. Author manuscript; available in PMC 2016 June 01.

Published in final edited form as:

Exp Mol Pathol. 2015 June ; 98(3): 439–445. doi:10.1016/j.yexmp.2015.03.037.

Temporal and spatial rearrangements of a repetitive element array on C57BL/6J mouse genome

Kang-Hoon Lee, **Lisa Yee**, **Debora Lim**, **David Greenhalgh**, and **Kiho Cho***

Department of Surgery, University of California, Davis and Shriners Hospitals for Children Northern California, Sacramento, CA 95817

Abstract

Repetitive elements (REs) make up the vast majority of the mammalian genomes. We identified species-specific genomic libraries of RE arrays. The non-random configurations of RE arrays suggest their functions. We tested whether RE arrays undergo age- and tissue/cell-specific rearrangements. An RE array of C57BL/6J mice, containing tandem repeats of a mosaic of transposable REs, was selected to examine rearrangements in different ages and tissues. There were marked changes in the array configuration in the genomes of skin and brain in all mice of six weeks and older, whereas heart and liver had alterations at 29 weeks. The temporal variations were confirmed by identifying putative rearrangement junctions. Temporal and spatial rearrangements of certain RE arrays may contribute to the acquired characteristics of the genome information system.

Keywords

RE array$_{CHR7.32}$; temporal rearrangement; spatial rearrangement; acquired genome characteristics; dynamic genome; four-dimensional genome unit

Introduction

The sum of conventional exon sequences is reported to occupy ~1.2 % of the human and mouse genomes, while a plethora of both characterized and uncharacterized repetitive elements (REs) represent the greater fraction of the rest (more than 50 % of the genome) (International Human Genome Sequencing, 2004; Lander et al., 2001; Mouse Genome Sequencing et al., 2002). Some REs are transposable in the genome via “copy and paste” or “cut and paste” functions in response to environmental stress (Deininger and Batzer, 2002; Lee et al., 2013b; Munoz-Lopez and Garcia-Perez, 2010; Young et al., 2014). Certain transposable REs (TREs), such as endogenous retroviruses (ERVs), which make up ~10 %

*Corresponding author: Kiho Cho, DVM, PhD, Department of Surgery, University of California, Davis and Shriners Hospitals for Children Northern California, 2425 Stockton Blvd. Sacramento, CA 95817, Tel: 916-453-2284, Fax: 916-453-2288, kcho@ucdavis.edu.

of the human genomes, are reported to respond to specific stress signals (*e.g.*, injury, infection), resulting in the induction of their expression in conjunction with production of functional proteins (Belshaw et al., 2005; Bhat et al., 2014; Cho et al., 2008; Kwon et al., 2009; Lee et al., 2013b; Lee et al., 2014). Recently, it has been demonstrated that the size and structure of the C57BL/6J inbred mouse genomes change temporally and spatially in conjunction with the differential activities of various families of TREs, such as ERVs, long-interspersed nuclear elements (LINEs), and/or short-interspersed nuclear elements (SINEs) (Lee et al., 2012a; Shukla et al., 2013). In addition, we recently identified species-specific libraries of complex, but highly-ordered RE arrangement structures, named RE arrays, from the human and mouse reference genomes (Chung et al., 2011; Kim et al., 2012; Lee et al., 2013a; Lee et al., 2011; Lee et al., 2012b).

Interestingly, tandem repeat units of certain RE arrays consist of a mosaic of various TREs, suggesting the inherent potential for changes in the structure of the RE arrays due to acquired activity of the TREs (Lee et al., 2012b). Furthermore, the species-specific (human vs. mouse) genome-wide profiles of RE arrays and the ordered complexity of individual RE arrays suggest that RE arrays contribute to a host of biologic processes, some of which are species-specific. In this study, we investigated whether the structural configuration of a tandem repeat cluster, which consists of TRE mosaics, within an RE array locus on chromosome 7 of C57BL/6J inbred mice, remains static or is rearranged temporally (age-dependent) and spatially (tissue/organ-dependent).

Materials and Methods

Nucleotide sequences of RE array$_{CHR7.32}$ and its surrounding region

The RE array$_{CHR7.32}$ locus and its neighboring region sequences on mouse chromosome 7 were obtained from the C57BL/6J strain-based mouse reference genome database (Annotation Release 104) at NCBI.

Animal experiment

Five age groups (neonate to ~29 weeks) of C57BL/6J female mice were purchased (pregnant mouse for neonates) from the Jackson Laboratory (West Sacramento, CA). The experimental protocol was approved by the Animal Use and Care Administrative Advisory Committee of the University of California, Davis. Three mice from each age group were sacrificed by CO_2 inhalation to collect tissues which were then snap-frozen in liquid nitrogen.

Genomic DNA isolation and analysis of structural variations by staggered PCR

Snap-frozen tissue samples (brain, heart, kidney, liver, lung, and skin) were subjected to genomic DNA isolation using a DNeasy Tissue kit (Qiagen, Valencia, CA) followed by normalization to 20 ng/μl. Four staggered primer sets were designed from a mosaic repeat unit of the central tandem repeat cluster within the RE array$_{CHR7.32}$ locus to examine structural changes (Figure 1): Set A (NF1-1A and NF3-2A), Set B (NF1-1A and NF5-2A), Set C (NF1-1A and NF6-2A), and Set D (NF5-1A and NF3-2A). Supplementary Table 1 lists the nucleotide sequences of the primers employed in this experiment. PCR was

performed using 100 ng of genomic DNAs and LA Taq polymerase (Takara Bio, Otsu, Japan) with the following conditions: 30 cycles of 94°C for 30 sec, 53°C for 1 min, and 72°C for 4 min. The GAPDH gene (GAPDH rt-1 For and GAPDH rt-1 Rev primer pair) (Supplementary Table 1) served as a normalization control for the quantity for each DNA sample. The PCR products were resolved by agarose gel electrophoresis followed by visualization using ethidium bromide staining.

Analysis of structural variations by inverse-PCR (I-PCR)

Genomic DNAs (300 ng total) were digested at 37°C for 3 hours with EcoRI (New England Biolab, Ipswich, MA) followed by self-ligation of the restriction fragments using T4 ligase (Promega, Madison, WI) at 4°C overnight. Structural changes in the tandem repeat cluster were examined by amplification of the polymorphic restriction fragments, which surround the mosaic repeat units, using a pair of primers (DiT-1B and DiT-2D). Supplementary Table 1 lists the nucleotide sequences of these primers. PCR was performed using 2 μl of the ligation products and Taq polymerase (Qiagen) with the following conditions: 30 cycles of 94°C for 30 sec, 55°C for 1 min, and 72°C for 4 min. Ten μl each of the PCR products was resolved on an agarose gel followed by ethidium bromide staining for visualization.

Analysis of structural variations by repeat unit-length PCR

Complementary to the I-PCR approach for identifying structural variations in the tandem repeat cluster, a PCR protocol involving a pair of primers (Obox4-1B and NF8r-2B) (Supplementary Table 1), which are capable of amplifying a near full size (~10 Kb of ~11.3 Kb) of the mosaic repeat unit, was employed. PCR was performed using 100 ng of genomic DNAs and LA Taq polymerase (Takara Bio) with the following conditions: 30 cycles of 94°C for 30 sec, 55°C for 1 min, and 72°C for 11 min. The PCR products were resolved by agarose gel electrophoresis followed by staining with ethidium bromide for visualization.

Cloning and sequencing of putative recombinant PCR amplicons

PCR amplicons (both from I-PCR and repeat unit-length PCR) were purified using a Qiaquick Gel Extraction kit (Qiagen) and cloned into the pGEM-T Easy vector (Promega) followed by plasmid DNA preparation using a Qiaprep Spin Miniprep kit (Qiagen). Sequencing was performed at Molecular Cloning Laboratories (South San Francisco, CA) or a DNA sequencing facility at the University of California, Davis.

Characterization of putative recombinant sequences

The sequences of the selected amplicons derived from the I-PCR and repeat unit-length PCR analyses were aligned against the NCBI-based RE array$_{CHR7.32}$ reference sequence to identify shared as well as unshared/deleted regions. For each candidate/putative recombinant sequence (either deletion recombinant or joint-circle as a byproduct of deletion recombination), the unshared/deleted regions, which are mapped on the array's reference sequence, were analyzed further to identify a pair of putative junction-recombination signals at both ends to define a rearrangement event.

Structural configuration of the RE array$_{CHR7.32}$ locus by simulating recombination events

The largest putative deletion recombinant (number 36 of Figure 3) of the tandem repeat cluster isolated in this study was selected for this simulation experiment. Initially, its sequence was aligned to the NCBI-based RE array$_{CHR7.32}$ locus and immediate surrounding regions to delineate the 5′ (3 mismatches out of 1,174 nucleotides)- and 3′ (100 % match out of 260 nucleotides)-sections spanning the deleted sequences. It is expected that the numbers of 5′- and 3′-sections that are identified depend on the degree of sequence matches in individual repeat units. Subsequently, recombination events were introduced *in silico* between each 5′-section and 3′-section combination within the NCBI-based RE array$_{CHR7.32}$ reference sequence region. The individual recombination-simulated RE array$_{CHR7.32}$ region sequences were subjected to self-alignment analyses followed by visualizing the arrangement patterns by dot-matrix presentation of the alignment results using a BLAST program (NCBI).

Results and Discussion

To determine whether there are temporal and spatial rearrangements in certain RE arrays, one RE array on chromosome 7, named RE array$_{CHR7.32}$, was selected among a library of RE arrays identified from the reference C57BL/6J mouse genome (Build 37.1) from the National Center for Biotechnology Information (NCBI) (Lee et al., 2012b). RE array$_{CHR7.32}$ displays a pattern of a “computer microprocessor” and harbors a central tandem repeat cluster of ~146 Kb (Figure 1). It was determined that the ~146 Kb tandem repeat cluster is formed with 13 units and each unit consists of a mosaic of various TRE types, predominantly ERVs, LINEs, and SINEs (Lee et al., 2012b).

To examine whether the structure of the tandem repeat cluster was stationary or dynamically (temporally and spatially) rearranged, four sets of PCR primers (A, B, C, and D), which stagger the mosaic components in different combinations, were designed from a repeat unit of ~11.3 Kb (Figure 1). Five age groups (neonate [0 week], two weeks, six weeks, 12 weeks, and 29 weeks) of female C57BL/6J mice were subjected to this study. Genomic DNAs of six organs/tissues from each age group were surveyed by PCR for structural rearrangements in the tandem repeat cluster. Interestingly, there were marked changes in the profiles of PCR amplicons derived from two (A and B) primer sets in the brain genomes and three (A, B, and C) primer sets in the skin genomes, starting at six weeks of age and thereafter (Figure 2). Variations in amplicon profiles were also observed in the skin genomic DNAs of different mice within the individual age groups of six weeks, 12 weeks, and 29 weeks. Additionally, changes in the amplicon profiles were observed in the liver (primer set B) and heart (primer sets A and B) at 29 weeks of age. In contrast, in the kidney and lung, no significant alterations in the amplicon profiles were observed in all primer sets and age groups except for minor variations among individual mice within certain age groups. Furthermore, the profiles of amplicons derived from primer set C, which is designed to generate the largest amplicon (expected to be ~7.7 Kb from intact repeat units), were highly variable among all six of the organs/tissues examined. In all tissues of different age groups, no significant variations were observed in the PCR amplicons derived from the fourth primer set (D), which is designed to produce the shortest amplicon of ~384 nucleotides. The findings from

this study provide evidence that the tandem repeat cluster within the RE array$_{CHR7.32}$ locus on chromosome 7 undergoes temporal and spatial rearrangements in conjunction with age of mice and/or organ/tissue/cell type.

To verify the PCR amplicon profile-based finding that there are structural alterations in the tandem repeat cluster, we cloned and sequenced putative recombinants which are isolated by two independent PCR analyses: 1) inverse-PCR (I-PCR) and 2) repeat unit-length PCR using a primer set spanning ~10 Kb of the repeat unit of ~11.3 Kb (Figure 3). Two types of recombinants were present among a total of 37 different rearranged sequences identified: 1) 36 deletion recombinants amplified by I-PCR analysis and 2) one putative joint-circle recombinant amplified by repeat unit-length PCR analysis. Thirty-three of the I-PCR amplified sequences had putative recombination signals, which range from two to 16 nucleotides, at their deletion junctions (Figure 3A and Figure 4A). Using the deletion recombinant-36 as an example, the data presented in Figure 4A illustrates a potential mechanism regarding how rearrangement events using recombination signal pairs create putative deletions, ranging from ~9.7 Kb to 301.9 Kb. Based on the sequence identity of the junction, the deletion recombinant-36 can be formed using individual pairs of multiple recombination signals found on the RE array$_{CHR7.32}$ locus (Figure 5A). One of them was a pair of recombination signals found on 7th and 13th repeat units, resulting in the putative recombinant-36 and a deletion of 66,440 nucleotides. In addition, one possibility is that the sequence amplified by repeat unit-length PCR is a putative joint-circle recombinant of 5.62 Kb in length which was presumably derived from an intra-allelic rearrangement event between 5th and 6th repeat units (Figure 3B and Figure 4B). The joint-circle sequence that was identified by repeat unit-length PCR analysis is likely to be a byproduct from a rearrangement event between a pair of putative recombination signals consisting of 24 and 26 nucleotides of which one copy was identified in this joint-circle sequence. The rearranged sequence data obtained from this experiment further confirmed that the structural configuration of the tandem repeat cluster within the RE array$_{CHR7.32}$ locus is highly variable.

In addition to the 13 repeat units forming the tandem repeat cluster, three additional repeat units with high sequence homology were mapped immediately downstream of the RE array$_{CHR7.32}$ locus during a survey of the array region using the repeat unit as a probe. During a search of the RE array$_{CHR7.32}$ for regions homologous to the sequence of the deletion recombinant-36, six and four locations which align to the 5′ and 3′ sections of the deletion recombinant, respectively, were identified (Figure 4 and Figure 5A). We then mapped 16 potential rearrangement events by simulating combinatorial rearrangement at six 5′- and four 3′-junction signals (Figure 5B). Although 24 combinations were expected from six 5′- and four 3′-junctions, the pairs in which 3′-junction signals were located upstream of its 5′-counterparts were invalid for simulating deletion-recombination events. As predicted, some of the putative recombination events resulted in marked changes in the structural configuration of the RE array$_{CHR7.32}$ locus and/or surrounding region. *In vivo* identification of these simulated deletion recombinants will require various structural analyses using a set of PCR primers which specifically target each putative deletion recombinant as well as high-fidelity sequencing efforts.

The findings from this study demonstrated that: 1) the tandem cluster of mosaic repeat units within the RE array$_{CHR7.32}$ locus was temporally altered in the genomes of the brain and skin starting when the C57BL/6J inbred mice were six weeks old, 2) in contrast to the brain and skin, the tandem repeat clusters in the genomes of the heart and kidney were determined to be static until 29 weeks when changes started to occur, whereas there were no significant changes in the kidney and lung genomes among the five age groups examined, 3) apparently, there are organ/tissue (perhaps, cell type/cell)-specific polymorphisms in the structure of the tandem repeat cluster within individual C57BL/6J inbred mice, 4) the vast majority of putative recombinants isolated from the tandem repeat cluster and/or its surrounding regions, which share the mosaic repeat units, had deletions of variable sizes between respective pairs of junction-recombination signals, 5) seemingly, identification of putative deletion recombinant sequences spanning the mosaic repeat units parallels the finding that the tandem repeat cluster within the RE array$_{CHR7.32}$ locus of the skin genomic DNAs undergoes temporal rearrangements as mice age, 6) thus, the structural configuration of the tandem repeat cluster within the RE array$_{CHR7.32}$ locus may be highly variable depending on the age of mice as well as organ/tissue/cell type, and 7) an extension of this study into RE arrays residing on different genomic loci may identify additional RE arrays of which structural configurations are dynamically altered within an individual in an age- and/or organ/tissue/cell type-specific manner (Kim et al., 2012).

The outcomes of a future study involving a wide range of tissues, age groups, and mouse strains, which focuses on a comprehensive survey of temporal and spatial rearrangement events occurring at the “local” RE array$_{CHR7.32}$ locus, will provide a foundation for “global” studies examining the role of RE arrays in genome biology. Some of these studies would explore the relationship among the RE arrays’ dynamicity and structural configuration, acquired TRE activity due to environmental stressors (Lee et al., 2011; Lee et al., 2012b), and structure-function of the chromosomal region(s) proximal as well as distal to the locus.

We acknowledge that the data obtained from this study provides no direct clue with regard to the functional significance of the temporal and spatial changes in the tandem repeat cluster within the RE array$_{CHR7.32}$ locus on chromosome 7 of the C57BL/6J inbred mice. However, we can envision that the three-dimensional configuration of certain RE arrays, defined by complexly organized RE interactions (both direct and inverse) in conjunction with their dynamic (temporal and spatial) rearrangements, is an example of a four-dimensional functional unit of the dynamic genome information system.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported, at least in part, by grants from Shriners of North America (No. 86800 to KC, No. 84302 to KHL [postdoctoral fellowship]) and the National Institutes of Health (R01 GM071360 to KC).

Abbreviations

RE	repetitive elements
TREs	transposable repetitive elements
LINEs	long-interspersed nuclear elements
SINEs	short-interspersed nuclear elements
ERVs	endogenous retroviruses
I-PCR	inverse-polymerase chain reaction

References

Belshaw R, et al. Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. J Virol. 2005; 79:12507–14. [PubMed: 16160178]

Bhat RK, et al. Human endogenous retrovirus-K(II) envelope induction protects neurons during HIV/AIDS. PLoS One. 2014; 9:e97984. [PubMed: 24988390]

Cho K, et al. Endogenous retroviruses in systemic response to stress signals. Shock. 2008; 30:105–16. [PubMed: 18317406]

Chung BI, et al. REMiner: a tool for unbiased mining and analysis of repetitive elements and their arrangement structures of large chromosomes. Genomics. 2011; 98:381–9. [PubMed: 21803149]

Deininger PL, Batzer MA. Mammalian retroelements. Genome Res. 2002; 12:1455–65. [PubMed: 12368238]

International Human Genome Sequencing C. Finishing the euchromatic sequence of the human genome. Nature. 2004; 431:931–45. [PubMed: 15496913]

Kim WC, et al. REMiner-II: a tool for rapid identification and configuration of repetitive element arrays from large mammalian chromosomes as a single query. Genomics. 2012; 100:131–40. [PubMed: 22750555]

Kwon DN, et al. Cloning and characterization of endogenous retroviruses associated with postinjury stress signals in lymphoid tissues. Shock. 2009; 32:80–8. [PubMed: 18948849]

Lander ES, et al. Initial sequencing and analysis of the human genome. Nature. 2001; 409:860–921. [PubMed: 11237011]

Lee KH, et al. Age-dependent and tissue-specific structural changes in the C57BL/6J mouse genome. Exp Mol Pathol. 2012a; 93:167–72. [PubMed: 22548960]

Lee KH, et al. Large interrelated clusters of repetitive elements (REs) and RE arrays predominantly represent reference mouse chromosome Y. Chromosome Res. 2013a; 21:15–26. [PubMed: 23359013]

Lee KH, et al. Identification of a unique library of complex, but ordered, arrays of repetitive elements in the human genome and implication of their potential involvement in pathobiology. Exp Mol Pathol. 2011; 90:300–11. [PubMed: 21376035]

Lee KH, et al. Injury-elicited stressors alter endogenous retrovirus expression in lymphocytes depending on cell type and source lymphoid organ. BMC Immunol. 2013b; 14:2. [PubMed: 23289855]

Lee KH, et al. Divergent and dynamic activity of endogenous retroviruses in burn patients and their inflammatory potential. Exp Mol Pathol. 2014; 96:178–87. [PubMed: 24509167]

Lee YK, et al. Unique profile of ordered arrangements of repetitive elements in the C57BL/6J mouse genome implicating their functional roles. PLoS One. 2012b; 7:e35156. [PubMed: 22529984]

Mouse Genome Sequencing C et al. Initial sequencing and comparative analysis of the mouse genome. Nature. 2002; 420:520–62. [PubMed: 12466850]

Munoz-Lopez M, Garcia-Perez JL. DNA transposons: nature and applications in genomics. Curr Genomics. 2010; 11:115–28. [PubMed: 20885819]

Shukla R, et al. Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. Cell. 2013; 153:101–11. [PubMed: 23540693]

Young GR, et al. Microarray analysis reveals global modulation of endogenous retroelement transcription by microbes. Retrovirology. 2014; 11:59. [PubMed: 25063042]

Highlights

- Structure of the RE array$_{CHR7.32}$ locus is highly variable depending on tissue types
- RE array$_{CHR7.32}$ of the skin and brain undergoes temporal rearrangements as mice age
- Rearrangements of RE arrays may contribute to the acquired genome characteristics
- Comprehensive studies examining the role of RE arrays in genome biology are needed

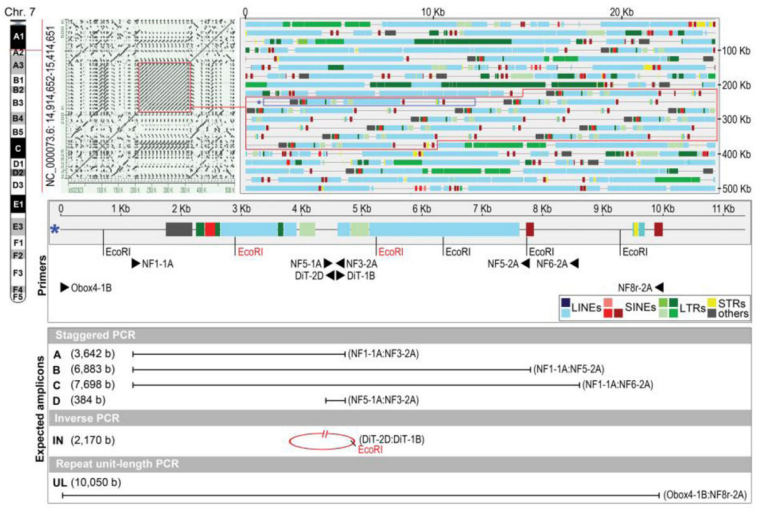

Figure 1.

Three different PCR strategies for structural analyses of the central tandem repeat cluster within the RE array$_{CHR7.32}$ locus on chromosome 7. The tandem repeat cluster (in both forms of dot-matrix structure and linear-visualization of mosaic repeat units) within the RE array$_{CHR7.32}$ locus is outlined with red lines. PCR primers for three different types of structural analyses of the highlighted tandem repeat cluster are mapped: 1) staggered PCR (A/B/C/D amplicons: NF1-1A, NF5-1A, NF3-2A, NF5-2A, and NF6-2A), 2) repeat unit-length PCR (UL amplicon: Obox4-1B and NF8r-2B), and 3) I-PCR (IN amplicon: DiT-1B and DiT-2D). In addition, putative PCR amplicon sizes expected from the individual primer sets without any recombination events are provided as a reference.

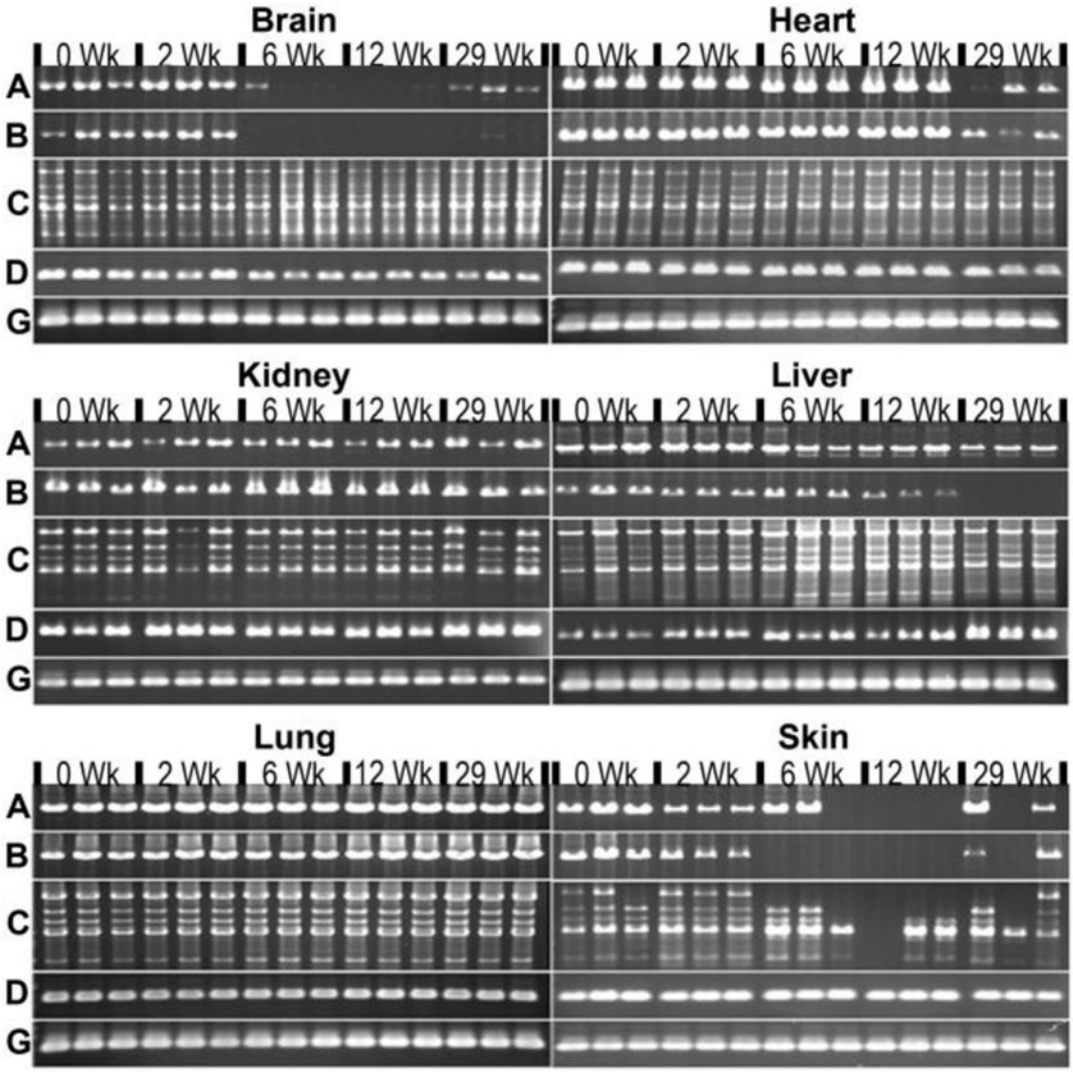

Figure 2.

Structural variations in the central tandem repeat cluster within the RE array$_{CHR7.32}$ locus. Six organs/tissues from five age groups of C57BL/6J female mice were surveyed for structural variations in the tandem repeat cluster within the RE array$_{CHR7.32}$ locus using staggered PCR primer sets. Organ/tissue-specific and age-dependent structural variations in the tandem repeat cluster within the RE array$_{CHR7.32}$ locus were identified in the five age groups (0 week [neonate], 2 weeks, 6 weeks, 12 weeks, and 29 weeks) by examining four sets of the PCR amplicons (A, B, C, and D) which are generated using staggered primer pairs (Figure 1). wk (week); G (GAPDH)

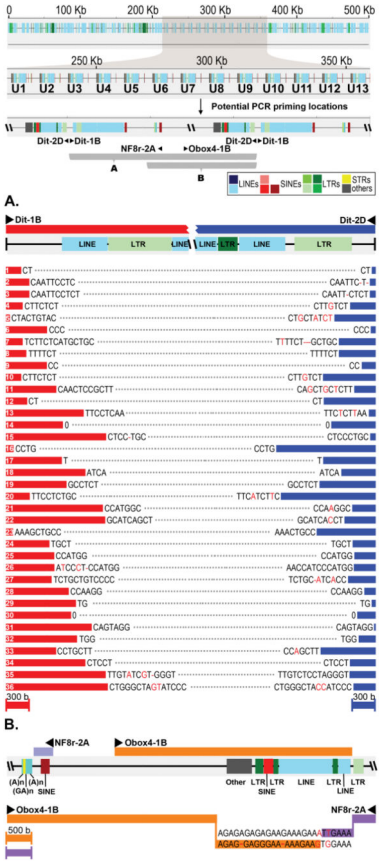

Figure 3.

Structure and size of putative recombinants isolated from the central tandem repeat cluster within the RE array$_{CHR7.32}$ locus. Illustration in the top panel provides information regarding the PCR primer locations, mosaic patterns of repeat units, and repeat unit-length. U (repeat unit) **A**. 36 putative deletion recombinants. A total of 36 putative deletion recombinants were isolated from the I-PCR survey of the tandem repeat cluster within the RE array$_{CHR7.32}$ locus. The positions of the DiT-1B and DiT-2D primer set are mapped on a repeat unit line drawing on which the relevant TRE information is indicated. Apparently, all recombinants except for two contained a pair of putative recombination signals, ranging from two to 16 nucleotides; one had single nucleotide and the other with none. **B**. A putative joint-circle recombinant. One putative recombinant sequence was isolated from the repeat unit-length PCR analysis of the tandem repeat cluster. The positions of the Obox4-1B and NF8r-2A primer set are mapped on a repeat unit line drawing on which the relevant TRE information is indicated. The recombinant sequence is presumed to be derived from a putative joint-circle which is formed during a rearrangement event. A pair of putative recombination signals was identified in the joint-circle clone.

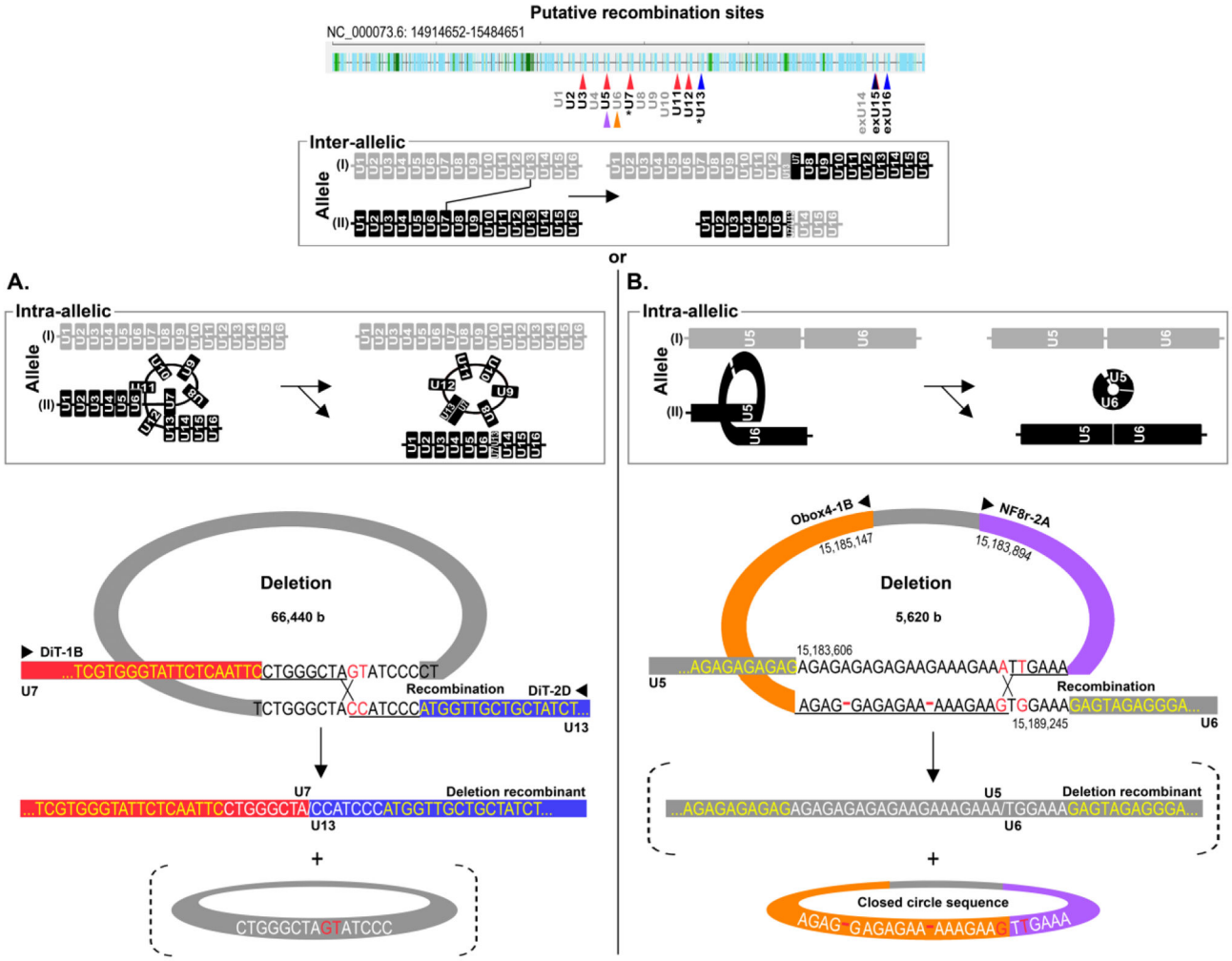

Figure 4.

Schematic illustration of potential recombination events identified in the central tandem repeat cluster within the RE array$_{CHR7.32}$ locus. **A**. Formation of deletion recombinants. Both inter-allelic and intra-allelic recombinant events are illustrated on the top. The 36th putative recombinant listed on panel A of Figure 3 is employed to explain how a putative intra-allelic recombination event occurred using a pair of recombination signals at the deletion junction. A set of primers (DiT-1B and DiT-2D) used for amplification of the putative recombinant is indicated. U (repeat unit) **B**. Formation of a joint-circle. Both inter-allelic and intra-allelic recombinant events are illustrated on the top. The illustration describes how the joint circle (from panel B of Figure 3) is formed during a putative intra-allelic deletion recombination event. A set of primers (Obox4-1B and NF8r-2A) used for amplification of the putative joint-circle is indicated. U (repeat unit)

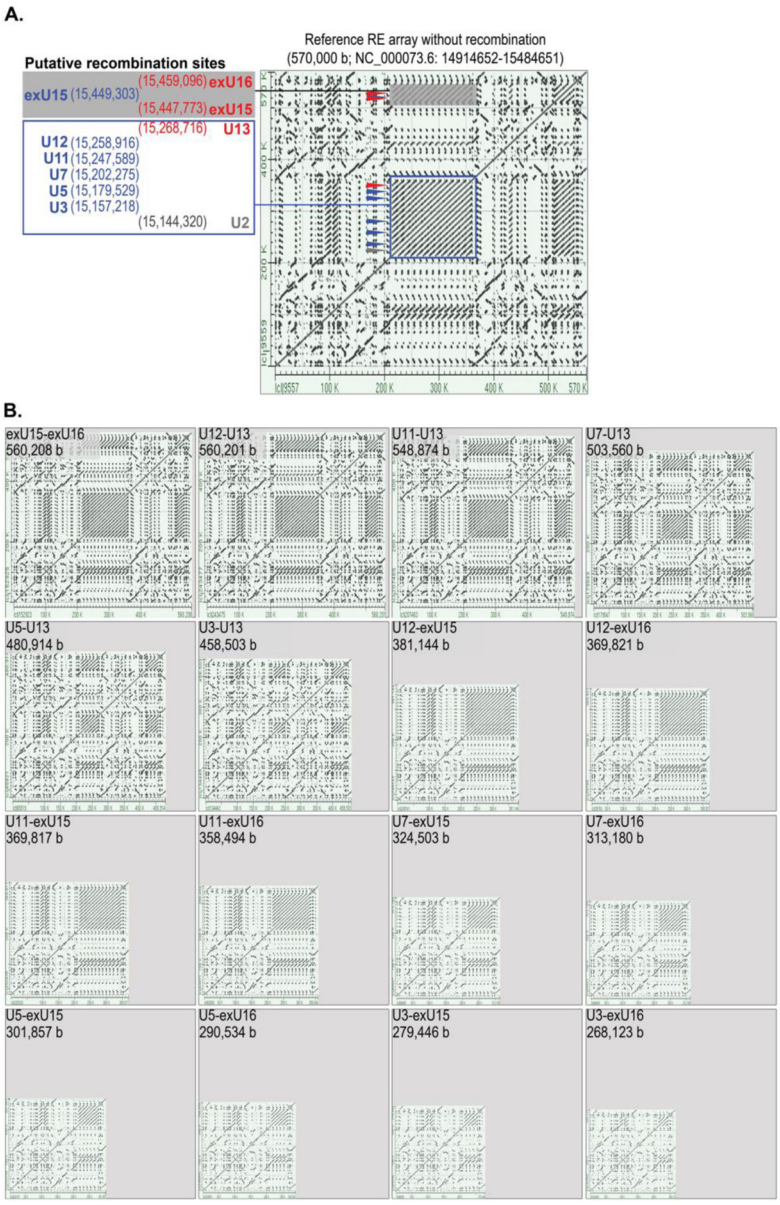

Figure 5.

Simulated rearrangements within the RE array$_{CHR7.32}$ region by combinatorial rearrangement at putative junction signal pairs based on the deletion recombinant-36 sequence. **A**. Putative locations of candidate 5′- and 3′-sections for single-deletion recombinants within the RE array$_{CHR7.32}$ region. A BLAST alignment analysis of the deletion recombinant-36 sequence mapped potential locations of the six 5′- (blue) and four 3′- (red) sections of the putative deletion recombinants within the RE array$_{CHR7.32}$ locus and its surrounding regions. The original RE array$_{CHR7.32}$ dot-matrix pattern without any rearrangements serves as a reference. **B**. Simulated rearrangements of the RE array$_{CHR7.32}$ region. Potential rearrangement events were simulated by combinatorial deletions at six 5′- and four 3′ recombination junction signals followed by visualizing 16 respective RE array structures reflecting the effects of individual recombination events. The junction-

recombination signal pairs in which the 3′ signals are located upstream of its 5′-counterparts are disqualified for simulating deletion-recombination events in this experiment. U (repeat unit); exU (repeat unit existing outside the original RE array$_{CHR7.32}$)