Enzymatic Removal of Ribonucleotides from DNA Is Essential for Mammalian Genome Integrity and Development

Martin A.M. Reijns,1,3 Björn Rabe,1,3 Rachel E. Rigby,1 Pleasantine Mill,1 Katy R. Astell,1 Laura A. Lettice,1 Shelagh Boyle,1 Andrea Leitch,1 Margaret Keighren,1 Fiona Kilanowski,1 Paul S. Devenney,1 David Sexton,1 Graeme Grimes,1 Ian J. Holt,2 Robert E. Hill,1 Martin S. Taylor,1 Kirstie A. Lawson,1 Julia R. Dorin,1 and Andrew P. Jackson1,*

1Medical Research Council Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK
2Medical Research Council Mitochondrial Biology Unit, Cambridge CB2 0XY, UK
3These authors contributed equally to this work
*Correspondence: andrew.jackson@igmm.ed.ac.uk

SUMMARY

The presence of ribonucleotides in genomic DNA is undesirable given their increased susceptibility to hydrolysis. Ribonuclease (RNase) H enzymes that recognize and process such embedded ribonucleotides are present in all domains of life. However, in unicellular organisms such as budding yeast, they are not required for viability or even efficient cellular proliferation, while in humans, RNase H2 hypomorphic mutations cause the neuroinflammatory disorder Aicardi-Goutières syndrome. Here, we report that RNase H2 is an essential enzyme in mice, required for embryonic growth from gastrulation onward. RNase H2 null embryos accumulate large numbers of single (or di-) ribonucleotides embedded in their genomic DNA (>1,000,000 per cell), resulting in genome instability and a p53-dependent DNA-damage response. Our findings establish RNase H2 as a key mammalian genome surveillance enzyme required for ribonucleotide removal and demonstrate that ribonucleotides are the most commonly occurring endogenous nucleotide base lesion in replicating cells.

INTRODUCTION

DNA is believed to have evolved from an ancestral RNA world as a more stable store of genetic information (Alberts et al., 2002; Cech, 2011). Ribonucleotides differ from deoxynucleotides by the presence of a single reactive hydroxyl group at the 2'-position of the ribose sugar, rendering RNA ~100,000-fold more susceptible to spontaneous hydrolysis under physiological conditions (Li and Breaker, 1999). The presence of ribonucleotides in genomic DNA is therefore undesirable, as it renders DNA more sensitive to strand breakage. It has long been thought that such misincorporation is prevented by the stringent selectivity of replicative DNA polymerases, favoring deoxynucleoside triphosphate (dNTP) over ribonucleoside triphosphate (rNTP) substrates (Joyce, 1997). However, recent in vitro experiments have demonstrated that, under physiologically relevant conditions in which rNTPs substantially exceed dNTPs, such DNA polymerases may incorporate a ribonucleotide base every few thousand base pairs (Nick McElhinny et al., 2010a). Budding yeast expressing a less selective replicative polymerase only displayed widespread ribonucleotide incorporation when ribonuclease (RNase) H2 activity was genetically abolished (Nick McElhinny et al., 2010b). This directly implicated RNase H2 in the removal of such ribonucleotides.

RNase H enzymes hydrolyze the RNA strand of RNA/DNA hybrids (Stein and Hausen, 1969). Such hybrids form during many cellular processes, including DNA replication (Machida et al., 1977), telomere elongation (Förstemann and Lingner, 2005), and transcription (Huertas and Aguilera, 2003; Li and Manley, 2005). Eukaryotes have two types of RNase H with distinct biochemical properties and substrate specificity (reviewed in Cerritelli and Crouch, 2009). RNase H1 is a processive monomeric enzyme that requires interaction with 2'-OH groups from four consecutive ribonucleotides for efficient substrate cleavage (Nowotny et al., 1997). Mammalian RNase H1 has two isoforms: a nuclear isoform of undefined function and a mitochondrial isoform that is essential for mitochondrial DNA replication (Cerritelli et al., 2003). However, the predominant source of RNase H activity in mammalian cells is RNase H2 (Büsen, 1980). Like RNase H1, it digests the RNA strand of RNA/DNA hybrids in a processive manner (Chon et al., 2009), but it also recognizes single ribonucleotides in a DNA duplex and cleaves the 5'-phosphodiester bond of the ribonucleotide (Eder et al., 1993). In eukaryotes, RNase H2 is a multimeric complex consisting of three subunits: RNASEH2A, RNASEH2B, and RNASEH2C (Crow et al., 2006a; Jeong et al., 2004). The RNASEH2A subunit contains the catalytic center, whereas the closely intertwined auxiliary RNASEH2B and C subunits are likely involved in interactions with other proteins (Fiegel et al., 2011; Reijns et al., 2011; Shaban et al., 2010). A PIP box motif at the C terminus of the RNASEH2B subunit guides the interaction of RNASEH2A with the catalytic subunit RNASEH2B (Eder et al., 1993).
between RNase H2 and PCNA (Chon et al., 2009) and its localization to replication foci (Bubeck et al., 2011), consistent with a role for the RNase H2 enzyme in DNA replication and/or repair.

Mutations in all three genes that encode the RNase H2 subunits cause the autosomal-recessive disorder Aicardi-Goutières syndrome (AGS) (Crow et al., 2006a). This early-onset neuroinflammatory condition mimics congenital viral infection and has immunological similarities to the autoimmune disease systemic lupus erythematosus (Ramantani et al., 2010). RNase H2 mutations that cause AGS result in partial rather than absolute loss of RNase H2 enzyme function (Reijns et al., 2011; Rice et al., 2007). Two further enzymes have been implicated in AGS: the 3'→5' DNA exonuclease TREX1 (Crow et al., 2006b) and a dNTP triphosphohydrolase, SAMHD1 (Rice et al., 2009). Innate immune-mediated inflammation is thought to result from the accumulation of endogenous nucleic acid species that are usually degraded by these enzymes, e.g., during DNA replication/repair (Yang et al., 2007) or suppression of endogenous retroelements (Manel and Littman, 2011; Stetson et al., 2008).

The nucleic acids that may accumulate as a consequence of impaired RNase H2 function are yet to be defined, and although RNase H2 enzyme activity has been studied for more than 40 years (Stein and Hausen, 1969), its cellular functions are poorly understood. Initially, RNase H2 was proposed to act in removal of the RNA oligonucleotides that prime Okazaki fragment synthesis during lagging-strand replication. In vitro biochemical studies indicate that sequential action of RNase H2 and FEN1 are sufficient to complete this process (Goulian et al., 1990; Turchi et al., 1994). However, primer removal through flap processing by FEN1/DNA2 has become the predominant model for Okazaki fragment maturation (Burgers, 2009; Rossi and Bambara, 2006). RNase H2 may also be important for the resolution of R loops that arise during transcription (El Hage et al., 2010) or for the repression of endogenous retroelements (Cerritelli and Crouch, 2009). Finally, the distinctive property that allows RNase H2 to recognize and cleave single ribonucleotides that are embedded in DNA duplexes suggests a role in the removal of such nucleotides from genomic DNA (Ryndberg and Game, 2002).

Here, we performed targeted mutagenesis of the mouse Rnaseh2b gene to gain insight into the in vivo role of the mammalian RNase H2 enzyme. Ablation of Rnaseh2b in mice leads to early embryonic lethality due to elevated DNA damage and reduced cellular proliferation during gastrulation. We establish that the growth arrest is the consequence of a p53-dependent DNA damage response associated with the accumulation of single ribonucleotides in genomic DNA. Thus, we demonstrate that removal of ribonucleotides to preserve genome integrity is an essential cellular function of RNase H2 in mammals.

**RESULTS**

**Rnaseh2b Is a Developmentally Essential Gene**

Rnaseh2b<sup>E202X</sup> embryonic stem (ES) cells with a premature stop codon in exon 7 of Rnaseh2b were generated by targeted homologous recombination (Figure 1A). Correct recombination of both arms of the targeting cassette was confirmed by Southern blotting and long-range PCR (Figures 1A and 1B) and the presence of a nonsense mutation at codon 202 (E202X) established by sequencing (Figure 1C). ES cells were injected into C57BL/6J host blastocysts to generate germline chimeras and subsequently heterozygous Rnaseh2b<sup>E202X/+</sup> mice. Intercrossing of Rnaseh2b<sup>E202X/+</sup> animals failed to yield live-born homozygous null mice (p < 0.001; Figures 1D and 1E). Similarly, homozygotes were not present (p < 0.001) in litters of a second independent line, Rnaseh2b<sup>em1a</sup>, derived from EUCOMM “knockout-first” ES cells.

As no viable homozygous animals were obtained from either of these alleles, we concluded that Rnaseh2b was likely to be an essential gene that is required for embryonic viability. At embryonic day 6.5 (E6.5), Rnaseh2b<sup>E202X/E202X</sup> embryos were present at Mendelian ratios and were almost indistinguishable in size from wild-type littersmates, suggesting normal progression of early embryogenesis (Figures 2A and 2B). However, by E7.5, there was a 23% and 32% decrease in embryonic height and width, respectively (p < 0.005; Figure 2B), suggesting a failure to increase the rate of proliferation in the epiblast that normally occurs at the onset of gastrulation (Mac Auley et al., 1993; Snow, 1977). Though all mutants proceeded through gastrulation, the embryos remained reduced in size and were developmentally retarded (Figure 2A). By E9.5 (n = 23), they were frequently truncated (57%) with few or very small postcervical somites (52%) and defects in allantois development (48%). At E10.5, histology demonstrated increased numbers of cells with condensed or fragmented nuclei, and at E11.5, a terminal phenotype was evident with loss of tissue morphology and integrity (data not shown).

**Rnaseh2b<sup>E202X</sup> Results in Absolute Loss of RNase H2 Complex Function**

To examine the effect of the Rnaseh2b<sup>E202X</sup> mutation on the RNase H2 enzyme, immunoblotting was performed with affinity-purified antibodies. As the premature stop codon is not in the final Rnaseh2b exon, it is expected to cause nonsense-mediated decay of the transcript. Consistent with this, we did not detect a truncated RNASEH2B<sup>E202X</sup> protein (Figure S1B available online). Also, full-length RNASEH2A, B, and C subunits were all undetectable in E9.5 Rnaseh2b<sup>E202X/E202X</sup> embryo lysates (Figure 2C), indicating that RNASEH2B is required for in vivo stability of the heterotrimeric complex. Furthermore, type 2 RNase H activity against a DNA duplex oligonucleotide with an embedded ribonucleotide was undetectable in Rnaseh2b<sup>E202X/E202X</sup> embryos (Figure 2D). In contrast, general RNase H activity was retained, albeit at <10% of wild-type embryo levels (Figure 2D), consistent with retention of normal RNase H1 activity. Given that RNase H2 is absent from Rnaseh2b<sup>E202X/E202X</sup> embryos, we hereafter refer to these embryos and their derivative cells as Rnaseh2b<sup>−/−</sup> and their genotype as Rnaseh2b<sup>−/−</sup>. Significantly, such embryos entirely lack detectable enzyme activity against isolated ribonucleotides embedded in DNA.

**RNase H2 Is Highly Expressed in Proliferating Cells**

The anti-RNase H2 antibody specifically detected RNase H2 with an embedded ribonucleotide was undetectable in Rnaseh2b<sup>E202X/E202X</sup> embryos, consistent with RNase H2null and their genotype as Rnaseh2b<sup>−/−</sup>. Endogenous
Figure 1. Targeted Inactivation of the Rnaseh2b Gene Causes Embryonic Lethality

(A) Schematic depicting targeted mutagenesis of exon 7 of the Rnaseh2b gene. (Top) A 7 kb segment of the Rnaseh2b genomic locus; exons 6 (ex6) and 7 (ex7) are indicated by black boxes, flanked by EcoRI sites (R). (Middle) NotI (N)-SalI (S) restriction fragment of the final targeting construct, comprising 4.5 kb of genomic DNA and a Neomycin selection cassette (Neo) flanked by Cre recombinase loxP sites (triangles). (Bottom) Successfully targeted endogenous locus containing the mutagenized exon 7 (ex7*). Red arrowheads indicate primers used to amplify arm I and arm II to confirm correct targeting.

(B) Southern blotting and long-range PCR confirm successful targeting by homologous recombination. The introduction of an additional EcoRI site results in a restriction fragment of 4.1 kb detected on Southern blotting with the 400 bp probe (red bar in A) for the targeted ES cells (E202X/+), that is not present in parental DNA (+/+). Arm I (4.7 kb) and arm II (2.2 kb) fragments are amplified by PCR in correctly targeted ES cells only.

(C) Sequencing traces for Rnaseh2b+/+, Rnaseh2bE202X/+, and Rnaseh2E202X/E202X DNA show the introduced nonsense mutation (red box).

<table>
<thead>
<tr>
<th>Mutant/Stage</th>
<th>+/+</th>
<th>+/m</th>
<th>m/m</th>
<th>Total</th>
<th>χ² test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liveborn pups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rnaseh2bE202X</td>
<td>34% (104)</td>
<td>66% (206)</td>
<td>0% (0)</td>
<td>310</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Rnaseh2bE202Xm1a</td>
<td>41% (22)</td>
<td>59% (32)</td>
<td>0% (0)</td>
<td>54</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Staged dissections</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rnaseh2bE202X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E5.5</td>
<td>31% (12)</td>
<td>46% (18)</td>
<td>23% (9)</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>E7.5</td>
<td>42% (24)</td>
<td>39% (22)</td>
<td>19% (11)</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>E8.5</td>
<td>20% (13)</td>
<td>57% (38)</td>
<td>23% (15)</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>E9.5</td>
<td>24% (19)</td>
<td>46% (36)</td>
<td>30% (24)</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>E10.5</td>
<td>29% (9)</td>
<td>52% (16)</td>
<td>19% (6)</td>
<td>31</td>
<td></td>
</tr>
</tbody>
</table>

(B) Southern blotting and long-range PCR confirm successful targeting by homologous recombination. The introduction of an additional EcoRI site results in a restriction fragment of 4.1 kb detected on Southern blotting with the 400 bp probe (red bar in A) for the targeted ES cells (E202X+) that is not present in parental DNA (+/+). Arm I (4.7 kb) and arm II (2.2 kb) fragments are amplified by PCR in correctly targeted ES cells only.

(C) Sequencing traces for Rnaseh2b+/+, Rnaseh2bE202X/+, and Rnaseh2E202X/E202X DNA show the introduced nonsense mutation (red box).
RNase H2 exhibited nuclear localization in early embryos, consistent with proposed roles in DNA replication and repair. Expression was observed in blastocysts (Figure 3 A) and in all three embryonic layers during gastrulation (Figure 3B), reflecting a ubiquitous presence at early stages of development. Later in embryogenesis and postnatally, expression became more restricted to highly proliferative tissues, such as the subventricular zone during neurogenesis and perinatal hair follicles (Figures S2C–S2F). In adults, RNase H2 was present in proliferative tissues, including intestinal crypt epithelium, and testes (Figures S2G and S2H). Expression correlated most closely with the proliferation marker Ki67, suggesting that RNase H2 is

(D) Multiplex PCR for mouse genotyping. (Top) A 221 bp PCR product (a/c) is present in wild-type mice (+/+); mice containing the Rnaseh2bE202X allele (also) give a 460 bp product (a/b). (Bottom) Schematic indicating position of forward (a) and reverse primers (b and c).

(E) Mice with null mutations (m) in Rnaseh2b are not postnatally viable, whereas E6.5–E10.5 embryos are present at Mendelian ratios. Genotype frequencies for offspring at weaning (and for embryos at indicated stages) derived from Rnaseh2bE202X/E202X and RNaseh2b^tm1a^/E202X intercrosses respectively. p values, \( \chi^2 \) test; n.s., not significant; m, mutant allele. Rnaseh2b accession number: NM_026001.2.

See also Figure S1.
significant reduction was evident (Figures 4B and 4C). Reduced cells undergoing active DNA synthesis; however, by E7.5, a of less than 6 hr (Snow, 1977).

A p53-Dependent Damage Response Is Evident in RNaseH2null Embryos

RNaseH2null embryonic growth failure could be the consequence of inefficient DNA replication or activated DNA damage signaling. Embryonic tissues were therefore immunostained for histone H2A.X phosphorylated at serine 139 (pH2AX), a marker of DNA damage response to double-strand breaks and arrested replication (Rogakou et al., 1998; Ward and Chen, 2001). Though no difference in pH2AX levels was observed between mutant and control blastocysts, at E6.5, there was a substantial increase in nuclear pH2AX staining in epiblast cells of RNaseH2null embryos (Figure 4A), coinciding with a period of rapid cell cycles of less than 6 hr (Snow, 1977).

At E6.5, there was no alteration in the percentage of epiblast cells undergoing active DNA synthesis; however, by E7.5, a significant reduction was evident (Figures 4B and 4C). Reduced embryo growth appeared to be the result of arrested cell proliferation, rather than cell death, as no widespread increase in apoptosis was observed at E7.5 or E9.5 by activated caspase 3 immunostaining (data not shown).

To investigate the molecular basis of this growth arrest, whole-genome expression analysis was performed using Illumina microarrays. Transcript levels of 197 genes were significantly upregulated in E9.5 mutant embryos, whereas 115 genes were downregulated when compared to age-matched controls (p < 0.05; > 1.5-fold change). Of note, the four genes with the greatest fold increase in expression (Figure 4D), nonsense-mediated decay of this transcript (Figure 4E). Although RNase H2 could have a role in suppressing expression of endogenous retroelements (Bhoj and Chen, 2008), no changes in retroelement transcript levels were identified, indicating that there was no widespread dysregulation of retroelements at this stage of embryonic development. Similarly, there was no transcriptional evidence of an immune-mediated process (data not shown).

We postulated that the DNA damage observed in the E6.5 embryos led to p53 activation and reduced cellular proliferation as a consequence of cyclin G1 and p21-mediated cell-cycle arrest (Cazzalini et al., 2010; Kimura et al., 2001). In keeping with this, partial rescue of the RNaseH2null embryonic phenotype was observed in a p53−/− background (Figure S3). Loss of p53 also fully rescued defective cell proliferation (Figure 4F) in primary cultures of E10.5/E11.5 mesodermal tissue, performed to generate mouse embryonic fibroblast (MEF) cell lines.

Ribonucleotide Accumulation in Genomic DNA of RNaseH2null Cells

Rnaseh2b−/−;p53−/− MEFs were next used to investigate the molecular consequences of RNase H2 loss. As expected,
Figure 4. A p53-Dependent DNA-Damage Response Is Activated in RNase H2-Deficient Embryos, Leading to Arrest in Cellular Proliferation

(A) Markedly elevated levels of nuclear pH2AX foci are evident in the epiblast of E6.5 embryos. Confocal projection of transverse cryosections through the decidua of E6.5 mutant RNaseH2\(^{-\text{mut}}\) and control littermate embryos. Scale bar, 10 \(\mu\text{m}\).

(B and C) The number of S phase epiblast cells is reduced in E7.5 embryos.

(D) Significant upregulation of p53 target genes in E9.5 RNaseH2\(^{-\text{mut}}\) embryos is detected by Illumina MouseWG-6 v2.0 Expression BeadChip microarray analysis. Plotted data points correspond to Illumina probes.

(F and G) Cell-cycle arrest in RNaseH2\(^{-\text{mut}}\) embryos is p53 dependent.

See also Figure S3.
Figure 5. RNaseH2null Genomic DNA Is More Sensitive to Alkali Hydrolysis

(A and B) Levels of pH2AX foci are elevated in Rnase2b−/−;p53−/− MEFs.

(A) Representative immunofluorescent images of Rnase2b+/+, +/−, −/− cell lines. Scale bar, 10 μm.

(B) Percentage of cells with more than three strong p-H2AX foci.

(C) MEFs (p53−/−) E10.5 (p53−/−) E9.5 (p53+/+)

(D) Normalized fragment count for DNA fragments of different sizes.

(E) Gel position (kb) vs. normalized fragment count.

(F) Gel position (kb) vs. densitometric intensity.

(G) Gel position (kb) vs. densitometric intensity.
type 2 RNase H activity and the RNase H2 protein complex were absent in these cell lines (Figures S4A and S4B). Significantly elevated levels of pH2AX foci were also present (Figures S5A and S5B), indicative of DNA double-strand breaks and/or replication fork arrest. In addition, we examined polyADPriboseylation (PAR) as another early marker of DNA-damage activation, which occurs in response to DNA breaks (Caldecott, 2008; Satoh and Lindahl, 1992). Levels of PAR were substantially raised in RNaseH2null MEFs, as shown by immunoblotting (Figure S4C), confirming the presence of cellular DNA-strand breaks.

DNA damage in RNaseH2null cells might arise as the consequence of undegraded Okazaki RNA primers, transcription-coupled incorporation of ribonucleotides by DNA polymerases, or transcriptionally induced R loops. To discriminate between these possibilities, total nucleic acids from the MEFs were separated by gel electrophoresis under alkaline conditions. Substantially increased mobility of genomic DNA from both RNaseH2null cell lines after treatment with alkali was observed relative to control genomic DNA (Figures 5C–5E), whereas no significant difference was evident upon electrophoresis under neutral conditions (Figure S4D). Given that phosphodiester bonds 3' of ribonucleotides, but not deoxyribonucleotides, are sensitive to alkali hydrolysis through nucleophilic attack by the 2' hydroxyl group (Lipkin et al., 1954), such increased fragmentation (termed alkali sensitivity) likely indicates the incorporation of ribonucleotides into genomic DNA. Given that alkali treatment also denatures DNA, increased electrophoretic mobility could also be the consequence of increased nicking, or gaps, in genomic DNA. To address this possibility, gel electrophoresis of total cellular nucleic acids was performed after formamide denaturation, which denatures DNA (Tang et al., 1989) without hydrolyzing ribonucleotide phosphodiester bonds (Figure 6A). Under these conditions, RNaseH2null DNA did not demonstrate increased mobility, in contrast to DNA treated with Nb.BspQI nicking endonuclease, which on average introduces nicks every ~11 kb. Therefore, the observed alkali-sensitive sites were consistent with covalently incorporated ribonucleotides, rather than gaps or nicks in genomic DNA.

Treatment with recombinant RNase H2 enzyme also led to widespread fragmentation of RNaseH2null genomic DNA, as shown by increased electrophoretic mobility after formamide denaturation. Significantly, the resulting fragmentation pattern was essentially identical to that seen after alkali treatment (Figure 6B), whereas no increased mobility was observed when inactive recombinant RNase H2 was used. In distinct contrast, recombinant E. coli type-1 RNase H (RNase HI) had no visible impact on the mobility of RNaseH2null genomic DNA (Figure 6C). RNase HI efficiently digests substrates with four or more ribonucleotides but can also hydrolyze double-stranded nucleic acids with three embedded ribonucleotides, albeit at substantially lower rates (Hogrefe et al., 1990). Although the biological significance of this activity is unclear, oligonucleotide substrates with three ribonucleotides were cleaved efficiently under our assay conditions (Figures S5C and S5D). This activity was fully preserved in the presence of RNaseH2null genomic DNA, ruling out any inhibitory effects that may be present within the nucleic acid preparation (Figures S5A–S5C). Thus, the differential sensitivity of genomic DNA from RNaseH2null cells to RNase HI and RNase H2 activities established that the alkali and RNase H2 cleavable sites consist of one or, at most, two consecutive covalently incorporated ribonucleotides.

Quantitative analysis of the alkali-induced fragmentation permitted us to estimate the frequency of embedded ribonucleotides. Determination of DNA fragment distributions (Figure 5E) from the densitometry data (Figure 5D) predicted a rate of ribonucleotide incorporation of ~1 in 7,600 nucleotides (nt) (analytical method described in Figure S4F). Additionally, the fragmentation pattern of hydrolyzed RNaseH2null DNA lies between that generated by two nicking endonucleases, one that cuts, on average, once every 11 kb (7-cutter Nt.BspQI) and the other once every 3.7 kb (6-cutter Nb.BstI) in the mouse reference genome (Figures 5F and S5G), supporting the computational estimate. Early and late passage MEFs exhibited similar levels of alkali sensitivity (data not shown). The incorporation of 1 ribonucleotide every 7,600 nucleotides during each round of replication would maintain such a steady-state level, so this is a minimum estimate for in vivo ribonucleotide misincorporation by polymerases.

Taken together, these results demonstrate the widespread presence of incorporated ribonucleotides in genomic DNA of RNaseH2null MEFs. Ribonucleotide misincorporation was also evident in RNaseH2null embryos (Figures 5C and 5E), consistent with this molecular defect underlying the developmental
phenotype. Finally, we employed a chemical genetic approach analogous to previous yeast genetic experiments, in which a mutated Pol ε with an enhanced propensity for ribonucleotide incorporation was used (Nick McElhinny et al., 2010b). Cells were treated with a low dose of hydroxyurea (HU) to reduce cellular dNTP:nNTP ratios through partial inhibition of ribonucleotide reductase activity. This would then increase ribonucleotide incorporation by DNA polymerases. RNaseH2null cells were observed to be hypersensitive to such a low dose of HU, accumulating in S phase (Figure S5G). Most significantly, such treatment resulted in additional alkaline and RNase H2-sensitive sites in genomic DNA from RNaseH2null cells (Figures 6D–6F, SSE, and SSF), consistent with such sites resulting from ribonucleotide misincorporation by DNA polymerases.

**CHROMOSOME INSTABILITY IN RNASEH2null CELLS**

Substantial levels of micronuclei, indicative of chromosomal breakage (Norppa and Falck, 2003), were observed in RNaseH2null MEFs (Figure 7A), suggesting that the excessive presence of ribonucleotides in DNA causes large-scale genome instability. Likewise, large-scale cytogenetic anomalies were evident in DAPI-stained metaphase chromosomes (Figure 7B). Using satellite FISH probes, chromosomal rearrangements were present in virtually all metaphases from Rnaseh2b−/−;p53−/− MEFs, whereas they were only occasionally seen in control MEFs. Both minutes (marker chromosomes) and interchromosomal translocations were frequently observed, with translocations confirmed by FISH and chromosomal painting (Figures 7, S6A, and S6B).

**DISCUSSION**

**RIBONUCLEOTIDES ACCUMULATE IN RNASEH2null CELLS AS A CONSEQUENCE OF INCORPORATION BY DNA POLYMERASES**

Here, we report that substantial genome-wide incorporation of ribonucleotides occurs in mammalian genomic DNA and establish that RNase H2 is required for efficient removal of such nucleotides. Recent in vitro biochemical studies with yeast replicative polymerases have shown the misincorporation of ribonucleotides into DNA (Nick McElhinny et al., 2010a). Based on these findings, it was predicted that ribonucleotides may be incorporated into genomic DNA in vivo. Subsequent studies in both fission and budding yeast have established that this is indeed the case using alkali sensitivity assays of yeast genomic DNA (Miyabe et al., 2011; Nick McElhinny et al., 2010b). Here, we show that ribonucleotide incorporation also occurs in metazoans; demonstrate that such ribonucleotide lesions are harmful to mammalian cells; and establish that their removal is required for mouse embryonic development. Previous studies used an elegant genetic approach in which mutator replicative polymerases with increased propensity for ribonucleotide incorporation were used to infer the presence of incorporated ribonucleotides from enhanced alkali sensitivity of RNase H2 null yeast genomic DNA. Our findings provide further characterization of these alkali-sensitive sites by using enzymatic assays to directly substantiate that such lesions are single or diribonucleotides that are covalently incorporated into genomic DNA. Furthermore, we find that such lesions occur at a frequency of at least 1,000,000 sites per cell, establishing them as the most common endogenous base lesions in the mammalian genome.

The presence of such ribonucleotides is most readily explained by misincorporation by the major replicative polymerases, which are estimated to incorporate one ribonucleotide every few thousand nucleotides in vitro (Nick McElhinny et al., 2010a). Alternatively, embedded ribonucleotides could result from failure to remove RNA primers during Okazaki-lagging strand processing. Such primers are ~10 nt in length, much longer than the single/diribonucleotides found. However, retention of single ribonucleotides during this process remains a possibility (Rumbaugh et al., 1997). Theoretically, oxidation of deoxynucleotides that are present in DNA could also result in embedded ribonucleotides (Vengrova and Dalgaard, 2004), although this seems an unlikely explanation for their frequent occurrence.

Genomic DNA from RNase H2 null S. cerevisiae exhibits differential alkali sensitivity that correlates with the propensity of a mutant Pol ε to incorporate ribonucleotides (Nick McElhinny et al., 2010b). Here, we performed an analogous experiment using hydroxyurea to alter dNTP:nNTP ratios, favoring ribonucleotide incorporation. This promoted increased alkali and RNase H2 sensitivity, leading us to conclude that ribonucleotides that are embedded in genomic DNA are most likely the consequence of misincorporation by DNA polymerases. Given their frequent occurrence in genomic DNA, the predominant sources of such ribonucleotides are likely to be Pol ε and Pol δ, the major replicative polymerases that are responsible for leading- and lagging-strand synthesis, respectively (Burgers, 2009; Kunkel and...
However, other DNA polymerases, such as Pol β (Cavanaugh et al., 2010) and lesion bypass polymerases, may also contribute.

Ribonucleotide accumulation in genomic DNA of RNaseH2null mice (Figures 5 and 6) implicates the RNase H2 complex in the maintenance of genome integrity. This DNA repair function was originally suggested by Eder and colleagues (Eder et al., 1993). Ribonucleotides are likely to be harmful, as their ribose 2'-hydroxyl group increases susceptibility of the adjacent phosphodiester bond to hydrolysis. Specific patterns of mutations at nucleotide level have been observed in genomic DNA from RNase H2 null S. cerevisiae (Kim et al., 2011; Nick McElhinny et al., 2010a). Most frequently, these consist of 2–5 bp deletions, which are the result of topoisomerase-I-induced nicks at embedded ribonucleotides (Kim et al., 2011). However, the loss of yeast RNase H2 alone has a relatively small effect on mutation frequency, and the effect of embedded ribonucleotides on large-scale genome stability in yeast has not been reported. Therefore, the frequent occurrence of large-scale genome rearrangements in RNaseH2null MEFs is unanticipated and noteworthy (Figure 7).

We estimate that ribonucleotides are incorporated at a rate of at least 1 every 7,600 nt in RNaseH2null cells, corresponding to ~1,300,000 lesions per cell. This is within the same order of magnitude predicted from in vitro incorporation rates by eukaryotic replicative polymerases (Nick McElhinny et al., 2010a) and is substantially higher than any other endogenous base lesions occurring in the mammalian genome. Even the previously most common lesions, such as abasic sites, 8-hydroxyguanine (8-oxoG), and 7-methylguanine, only occur up to 10,000 times per genome (Ciccia and Elledge, 2010; Lindahl and Barnes, 2000). As misincorporated ribonucleotides occur at least 50-fold higher rates, without an efficient repair mechanism, they would be the most common noncanonical nucleotides present in mammalian DNA. Therefore, defining the processes that remove these ribonucleotides is of substantial interest. Of note, FEN1/Rad27 in conjunction with RNase H2 can excise...
ribonucleotides on an in vitro substrate, generating a single nucleotide gap on which a DNA polymerase and DNA ligase could act directly to repair the lesion (Rydberg and Game, 2002).

**Misincorporated Ribonucleotides Induce DNA Damage**

In itself, ribonucleotide incorporation does not prevent replication: cellular proliferation is seen in both RNase H2 null mouse cells (p53−/−; Figure 4) and RNase H2 null budding yeast, in which p53 signaling is not evolutionarily conserved (Arudchandran et al., 2000; Belyi et al., 2010; Nick McElhinny et al., 2010b). DNA polymerases can tolerate templates containing ribonucleotides (Watt et al., 2011), which may explain why early embryogenesis in RNaseH2null embryos proceeds normally. The absence of grossly perturbed transcriptional profiles later in development (Figure 4) suggests that mammalian RNA polymerases also tolerate ribonucleotide-containing templates.

However, excessive numbers of ribonucleotides do appear to be problematic. Replication fork stalling may occur in regions that contain clustered ribonucleotides, as seen at the S. pombe mating switch locus (Vengrova and Dalgaard, 2006). Incorporation of ribonucleotides in difficult to replicate regions or in close proximity to other lesions may be similarly detrimental. This is likely to explain the activation of DNA-damage response signaling observed in RNaseH2null MEFs and embryos (Figures 4 and 5). Chromosomal rearrangements and micronuclei indicate the occurrence of double-strand DNA breaks. Such breaks may result from subsequent replication fork collapse or may be caused directly by hydrolysis of ribonucleotides on opposing DNA strands (see model in Figures S6C and S6D). Alternatively, the observed increase in PAR (Figure S4C) could suggest the presence of frequent single-strand breaks that would be converted at low frequency to double-strand lesions during replication. The accumulation of ribonucleotides in conjunction with rapid cell cycles in the epiblast (Snow, 1977) probably underlies the marked activation of DNA-damage signaling in the embryo. This then results in a p53-mediated inhibition of proliferation that is likely to substantially contribute to the lethality observed at E11.5 in RNaseH2null embryos.

**Ribonucleotide Incorporation in Health and Disease**

To our knowledge, stable incorporation of ribonucleotides has only been reported to date in two contexts. First, a diribonucleotide has been reported in somatic mammalian cells (Grossman et al., 1973), and we now show these to be mono or diribonucleotides (Figure S5H). These sporadic ribonucleotides appear to be randomly distributed and thus are likely to result from replicative polymerase incorporation. The selectivity of the mitochondrial polymerase γ would be consistent with the presence of 10–30 ribonucleotides in mature mtDNA (Kasi viswanathan and Copeland, 2011). These may be tolerated by the mammalian genome either because of its relatively slow replication rate (Bogenhagen and Clayton, 1977) or owing to different mechanism(s) of genome replication (Clayton, 2003; Holt, 2009). Likewise, ribonucleotide incorporation is well tolerated in RNase H2-deficient S. cerevisiae, with normal viability and efficient cellular proliferation in unperturbed cells (Arudchandran et al., 2000; Nick McElhinny et al., 2010b). As recently reported, template switch and translesion DNA synthesis postreplication repair pathways may permit such ribonucleotides to be tolerated (Lazzaro et al., 2012). In contrast, in mice, we find that ribonucleotide removal is essential early in development. Similarly, mutation of other genes ensuring genome integrity, such as the catalytic subunit of pol ɣ (Rev3), are viable in yeast but cause embryonic lethality at a similar stage in mutant mice (Esposito et al., 2000; Wittschieben et al., 2000). Such lethality may therefore be explained by the much larger size and complexity of the mammalian nuclear genome.

Low levels of ribonucleotide incorporation in the nuclear genome may be tolerated, and this could well be relevant to the autoimmune disorder Aicardi-Goutières syndrome (AGS), in which aberrant nucleic acid substrates are thought to drive an innate immune response (Crow and Rehwinkel, 2009). Reduced RNase H2 activity in AGS may therefore result in a chronic low level of ribonucleotide incorporation that is then processed by alternative (non-RNase H2 dependent) repair pathways. The increased levels of polyADPribosylation (Figure S4C), as well as the enhanced sensitivity to hydroxyurea observed in Rnaseh2b−/− MEFs (Figure S5G), could be consistent with this possibility. Aberrant nucleic acid species generated by such repair could then trigger an innate immune response. Alternatively, embedded ribonucleotides might induce DNA-damage response signaling that itself stimulates interferon production (Bzostek-Racine et al., 2011).

In summary, ribonucleotides are highly deleterious to the mammalian cell when left incorporated in the nuclear genome, causing substantial genome instability. RNase H2 is therefore a critical enzyme for ensuring the integrity of genomic DNA. Defining the pathway(s) that remove these ribonucleotides from genomic DNA, the site and nature of ribonucleotide-induced DNA damage, as well as the genome distribution of ribonucleotides will now be of substantial interest. This will help improve our understanding of the pathological and physiological roles of ribonucleotides in genomic DNA, of significance to both nucleic acid-driven autoimmunity and carcinogenesis.

**EXPERIMENTAL PROCEDURES**

**Generation of Rnaseh2b Null Mice**

While performing targeted homologous recombination to generate 129/Ola ES cells with a c.520G > A (A174T) mutation in exon 7 of Rnaseh2b (that corresponds to the most common, hypomorphic AGS patient mutation, c.529G > A), we identified the Rnaseh2bfl/fl ES clone that had fortuitously acquired an additional mutation resulting in a premature stop codon: c.604G > T, E202X. Further details of gene targeting are available in the Supplemental Information. After karyotyping and blastocyst injection, resulting male chimeras were crossed to C57BL/6J females, giving rise to heterozygous Rnaseh2bE202X/+;p53+/− mice (Clarke et al., 1993) (hereafter referred to as Rnaseh2bfl/fl;E202X). Crosses with p53−/− mice (Clarke et al., 1993) were used to generate Rnaseh2bfl/fl;E202X;p53−/− double mutants. Knockout-first Rnaseh2b mice were generated by blastocyst injection of the Rnaseh2bfl/fl;E202X ES cell clone EPD0087.4_A02 (EUCOMM ID: 24441; elsewhere referred to as Rnaseh2bfl/fl;E202X) (Friedel et al., 2007). The Rnaseh2bfl/fl;E202X allele is designed to prematurely truncate Rnaseh2b transcripts through targeted insertion into

intron 4 of a genetrap cassette that contains a strong splice acceptor and an efficient polyadenylation signal (Testa et al., 2004).

Genotyping

Genotypes of mice and embryos were determined by multiplex PCR. The status of early embryos was determined by r-Rnase H2 immunofluorescence. For further details and primer sequences, see Supplemental Information and Table S1.

RNase H Activity Assays

Enzyme activity assays were performed in triplicate using a FRET-based fluorescent substrate release assay as previously described (Crow et al., 2006a; Reijns et al., 2011) using 100 ng/ml of total protein from whole-cell extracts. RNase H2-specific activity was determined by subtracting the cellular activity against a sequence-matched DNA duplex without ribonucleotides.

Western Blotting, Immunohistochemistry, Immunocytochemistry, and Microscopy

Immunoblotting was performed on whole-cell extracts as described previously (Crow et al., 2006a). For immunohistochemistry, tissues and decidua were dissected into ice-cold PBS and fixed with 4% paraformaldehyde/PBS for 3–16 hr at 4°C with further processing performed by standard methods. Images were collected on Zeiss Axioplan II fluorescence or Nikon A1R confocal microscopes. For full experimental details, see the Supplemental Information; for antibodies and dilutions, see Table S2.

Detection of Ribonucleotides in Genomic DNA

Total nucleic acids were isolated by mechanical disruption of MEFs or yolk sacks in ice-cold lysis buffer (20 mM Tris-HCl [pH 7.5], 75 mM NaCl, and 50 mM EDTA) and subsequent incubation with 100 μg/ml proteinase K, with Sarcosine then added to final 1% concentration. Nucleic acids were sequentially extracted with TE-equilibrated phenol, phenol/chloroform:isoamylalcohol (25:24:1), and chloroform; precipitated with isopropanol; washed with 75% ethanol; and dissolved in water. Mitochondrial DNA (mtDNA) was isolated from sucrose-gradient purified mitochondria as previously described (Pohjoismäki et al., 2010). For alkaline gel electrophoresis, total nucleic acids were heated for 2 hr at 55°C with 0.3 M NaOH and separated on agarose gels (50 mM NaOH, 1 mM EDTA) as previously described (Nick McElhenny et al., 2010b). Control samples were heated with 0.3 M NaCl and separated on 0.5× TBE agarose gels. Alternatively, nucleic acids were treated with RNase H enzymes and heated for 30–60 min in 90% formamide/20 mM EDTA at 37°C before separation on 0.5× TBE agarose gels. Digestions with RNase HII were carried out in 100 μl of 1× reaction buffer (NEB) with 5 U of enzyme for 1 hr at 37°C. Digestions with RNase H2 were carried out for 1 hr at 37°C in 100 μl reaction buffer (50 mM Tris [pH 8], 60 mM KCl, 10 mM MgCl2, 0.01% BSA, 0.01% Triton) using 10 nM of purified recombinant human RNase H2 (Reijns et al., 2011). Nucleic acids were ethanol precipitated and dissolved in 90% formamide/20 mM EDTA. After electrophoresis, gels were stained with SYBR Gold (Invitrogen) or ethidium bromide.

ACCESSION NUMBERS

Microarray data for E9.5 Rnaseh2b−/− and wild-type embryos have been deposited in the Gene Expression Omnibus under the accession number GSE37419.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and two tables and can be found with this article online at doi:10.1016/j.cell.2012.04.011.

ACKNOWLEDGMENTS

We thank I. Adams, N. Gilbert, N. Hastie, and I. Jackson for commenting on the manuscript; W. Bickmore, M. Ansari, D. Fitzpatrick, and T. Kunkel for helpful discussions; A. Ross, G. Waugh, J. Young, A. Hart, and C.A. Martin for technical assistance; A. Pearce, E. Maher, and E. Freyer for assistance with cytogenetic and FACS analysis; and C. Graham for statistical advice. We also thank Aurelio Reyes and Stuart Wood for providing mouse mitochondrial DNA and the Wellcome Trust Clinical Research Facility, Edinburgh for performing microarrays. This work was supported by funding from the MRC and Lister Institute for Preventative Medicine. A.P.J. is a MRC Senior Clinical Fellow and Lister Institute Prize Fellow.

Received: November 9, 2011
Revised: February 21, 2012
Accepted: April 23, 2012
Published online: May 10, 2012

REFERENCES


