Damage-Induced Localized Hypermutability

Dmitry Gordenin

Chromosome Stability Section
Laboratory of Molecular Genetics

National Institute of Environmental Health Sciences
Damage-Induced Localized Hypermutablety (LHM) in *Saccharomyces cerevisiae*

- Finding with small reporters (*Yang et al., PLoS Genetics 2008*)
- Large size of a region
- Genome-wide scale
  - Multiple regions in a genome
  - Vast mutation clusters caused by chronic mutagenesis
Mutations Can Alter Gene Function
Multiple Mutations are More Likely to Alter Gene Function

Alleles with high fitness:
(enhanced or even new function)
- Evolution
- Ig-variants (SHM)
- Cancer (oncogenes)

Inactivation or distortion of function:
- Genetic disease
- Cancer
Localized Hypermutability (LHM) Can Produce Alleles with Multiple Changes without Overloading the Rest of the Genome with Mutations

Genome-wide hypermutability

LHM
Ig/SHM – programmed

Catastrophic mutation load

Mechanism for non-programmed?

Few additional mutations
**Hypothesis:**

Error-prone Translesion Synthesis (TLS) in Damaged Long ssDNA can be a Source of Localized Hypermutability and Multiple Mutations

- **5'→3' resection at DSB**
  - Mre11/Rad50/Xrs2
  - Dna2/Sgs1 or Exo1

- **5'→3' resection at uncapped telomere**
  - Cdc13
  - Stn1
  - Tel-cap

- **uncoupled replication fork**
Checkpoints Are Triggered by Long ssDNA

How Efficient Would be Recovery and Repair of Damaged Long ssDNA?

- Repair; Recovery
- Non-rearranged genome; TLS $\rightarrow$ multiple mutations
- Viable cell;
- Senescence; Apoptosis; Necrosis
- Rearrangements; Deletions
- Adaptation
Hypermutablety of Damaged Single-Strand DNA Formed at Double-Strand Breaks and Uncapped Telomeres in Yeast *Saccharomyces cerevisiae*

Yong Yang, Joan Sterling, Francesca Storici*, Michael A. Resnick*, Dmitry A. Gordenin*

Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, North Carolina, United States of America

1. Create a region of persistent ssDNA *in vivo* (put ssDNA “on hold”)

2. Apply DNA damage

3. Restore to dsDNA at will

4. Count mutations in a reporter
Damage-Induced Localized Hypermutability (LHM) in *Saccharomyces cerevisiae*

- Finding with small reporters *(Yang et al., PLoS Genetics 2008)*

- Large size of a region

- Genome-wide scale
  - Multiple regions in a genome
  - Vast mutation clusters caused by chronic mutagenesis
1. Create a region of persistent ssDNA
2. Apply DNA damage
3. Restore to dsDNA
4. Count mutations in a reporter

_Yeast telomere_

\[ cdc13-1 \, (ts) \, yeast \, (G1) \]

\[ \downarrow \text{Shift to non-permissive conditions (37ºC)} \]

Uncapped tel., G2 arrest, 5’→3’ resection

\[ \downarrow \text{UV damage} \]

\[ \downarrow \text{Shift to permissive conditions (23ºC)} \]

Score Lys– mutants

_cdc13-arrest puts ssDNA “on hold”_
Sub-Telomeric LYS2 Reporter

- **Score genome-wide mutability**
- **Score sub-telomeric mutability**

**UV-mutagenesis (x 10⁴)**

<table>
<thead>
<tr>
<th>Tel. ssDNA (cdc13-arrest)</th>
<th>Genome-wide (Lys⁻, LYS2⁺)</th>
<th>Sub-telomeric (lys2⁻)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>+</td>
<td>5</td>
<td>182</td>
</tr>
</tbody>
</table>

UV – 45 J/m²; survival ~50%
Mutagenesis in Transient ssDNA Around a DSB

Inducible DSB
5’→3’ resection

DNA damage (UV or MMS)

Initiate DSB repair with exogenous oligonucleotides (restore selection marker)

Strand re-synthesis → mutation

Lack of homology → ssDNA “on hold”
UV-Induced Hypermutagenesis in the Vicinity of a DSB

Mutagenesis enhanced by a DSB 250 to 500-fold (mutant frequency 3-6%)

Strong increase in "spontaneous" mutagenesis may be associated with uncontrolled damage to ssDNA (Hypothesis of Strathern et al.)
Hypothesis:

Error-prone Translesion Synthesis (TLS) in Damaged Long ssDNA can be a Source of Localized Hypermutability and Multiple Mutations

5'→3' resection at DSB

5'→3' resection at uncapped telomere

uncoupled replication fork
Error-Prone Translesion Synthesis (TLS) Relies on Pol ζ/Rev1 and PCNA-monoubiquitination at K164

Hypermutability Requires Same Factors as Error-Prone TLS (TLS)
Hypothesis:

Error-prone Translesion Synthesis (TLS) in Damaged Long ssDNA can be a Source of Localized Hypermutability and Multiple Mutations
Strong Strand Bias to Pyrimidines in the Non-resected Strand Indicates that Mutations Result from UV-photoproducts in ssDNA.

- ssDNA – YES
- Multiple mutations?
UV-induced mutants generated via ss-DNA (DSB and Tel) carry widely separated strand-biased multiple mutations

~ $0.5 \times 10^{-3}$ per nt (1,000-fold compared to genome-wide)
Damage-Induced Localized Hypermutability (LHM) in *Saccharomyces cerevisiae*

- Finding with small reporters (Yang et al., PLoS Genetics 2008)

- Large size of a region

- Genome-wide scale
  - Multiple regions in a genome
  - Vast mutation clusters caused by chronic mutagenesis
G2-arested cells (long subtelomeric ssDNA) – *up to 11 mutations spanning over 15 kb*
G1-cells (no subtelomeric ssDNA) – only one mutation in each of 9 mutants

Mutation bias to pyrimidines marks the areas of damaged ssDNA

<table>
<thead>
<tr>
<th>m15</th>
<th>m88</th>
<th>m27</th>
<th>m18</th>
<th>m45</th>
<th>m36</th>
<th>m42</th>
<th>m33</th>
<th>m22</th>
<th>m83</th>
<th>m78</th>
<th>m46</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Tel-5L</th>
<th>NPR2</th>
<th>LYS2 (4.2 kb)</th>
<th>CIN8</th>
<th>PRB1</th>
<th>PCM1</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>indel; complex</th>
<th>Py</th>
<th>Pu</th>
</tr>
</thead>
<tbody>
<tr>
<td>18% (14)</td>
<td>74% (57)</td>
<td>8% (6)</td>
</tr>
</tbody>
</table>
Large Area with Multiple Mutations Associated with Damaged ssDNA

Areas of multiple mutations could cover:

-- ORFs, large exons or large domains ~ 1 kb+

-- Adjacent exons ~ 10 kb+

-- Small genes (e.g. TP53, EGFR-201, CDK2, p21, RAD6) ~ 20 kb+

-- Large genes ~ 100 kb+
Damage-Induced Localized Hypermutability (LHM) in *Saccharomyces cerevisiae*

- Finding with small reporters (Yang et al., PLoS Genetics 2008)

- Large size of a region
- Genome-wide scale
  - Multiple regions in a genome
  - Vast mutation clusters caused by chronic mutagenesis

*Genome-wide search.*
Collaboration with Piotr Mieczkowsky,
-- UNC High-Throughput Sequencing Center
Identifying and Mapping Polymorphisms

Paired end reads
up to 5 kb

35 nt reads

-- Illumina sequencing can detect as little as 10 damage-induced mutations in a yeast genome

-- CLC Bio – a biologist-friendly software
Count Base Substitutions in Subtelomeric and Internal Regions of Chromosomes

Saccharomyces cerevisiae chromosomes

- 25 kb subtelomeric regions, except Tel-5L (Total - 775 kb ~ 7%)

Total – 12,156 kb

Internal (Total 11,131 kb ~ 93%)
Will it find mutations in subtelomeric regions containing homology blocks?

Predictions for UV-induced mutations in G2-arrested *cdc13-1* cells:

- Subtelomeric clusters
- Increased probability of subtelomeric mutations
- Strand bias with subtelomeric mutations
**First Indication:** Subtelomeric Clusters – Higher Incidence in G2-arrested \textit{cdc13-1} Cells

### G2-arrested \textit{cdc13-1} cells (37°C)

<table>
<thead>
<tr>
<th>Genome of a mutant</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
<th>XIII</th>
<th>XIV</th>
<th>XV</th>
<th>XVI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
</tr>
<tr>
<td>m15</td>
<td>1</td>
<td></td>
<td>8</td>
<td></td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m33</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m45</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m83</td>
<td>1</td>
<td></td>
<td>8</td>
<td></td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m88</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

### G1 \textit{cdc13-1} cells (23°C)

<table>
<thead>
<tr>
<th>Genome of a mutant</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
<th>XI</th>
<th>XII</th>
<th>XIII</th>
<th>XIV</th>
<th>XV</th>
<th>XVI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
<td>L</td>
<td>R</td>
</tr>
<tr>
<td>m2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m4</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>m7</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Increased Probability of UV-Induced Subtelomeric Mutations in G2-arrested cdc13-1 Cells

![Bar graph showing increased mutations in subtelomeric regions compared to internal regions. The graph indicates a higher number of mutations per $10^6$ nt in the ss-Tel (G2-arrest) condition compared to ds-Tel (G1) condition.]
Reversal of Strand Bias in Subtelomeric Regions

Expect changes in **purines** of the data strand at the **left** end.

Expect changes in **pyrimidines** of the data strand at the **right** end.
Strand Bias as Expected

- **G1 (no ds DNA)**
  - Left: 1
  - Right: 3

- **G2-arrest (ss DNA)**
  - Left: 10
  - Right: 10

Bases mutated in the "data strand"
High Frequency of UV-induced Strand-Biased Multiple Mutations Associated with Transient ssDNA

- Density of UV-induced mutation ~ 0.5 x 10^{-3} per nt (1,000-fold compared to genome-wide)

- Damage-induced hypermutability (UV and MMS) of ssDNA is under Pol ζ control (Rev3, Rev1, PCNA-K164-Ubi)

- Single area of hypermutability spans up to 15 kb

- Cell can tolerate several simultaneous areas of UV-induced LHM
 Genome-Wide Hypermutability – *intolerable mutation load*
 Localized Hypermutability (LHM) – *escape high mutation load*

- in artificial long ssDNA – *YES*
- in "natural" conditions???

Genome-wide mutagenesis "by a textbook"

Catastrophic mutation load -- ~ 1,000 additional mutations

Viable cells ≤10 additional mutations
Damage-Induced Localized Hypermutability (LHM) in *Saccharomyces cerevisiae*

- Finding with small reporters (Yang et al., PLoS Genetics 2008)

- Large size of a region
- Genome-wide scale
  - Multiple regions in a genome
  - Vast mutation clusters caused by *chronic mutagenesis*

Genome-wide search.
Collaboration with Piotr Mieczkowsky,
-- UNC High-Throughput Sequencing Center
MMS-induced hypermutability in ssDNA around an I-SceI-DSB (Yang et al.):

- mostly substitutions in cytosines (also some substitutions in adenines);
- in agreement with 3meC and 1meA (or 3meA) being primary mutagenic lesions.

![Diagram showing MMS-induced mutations in the area of 5'→3' resection at DSB]

- **CAN1 mutation frequency x 10^4**

- **MMS** vs. **MMS+DSB**
Detecting Transient Regions of MMS-Induced LHM

- closely spaced CAN1 and URA3 in chromosome II – to screen for closely spaced multiple mutations
- surrounded by essential genes – to eliminate GCR

**Hypothesis:**
Double inactivation of URA3 (5FOA-R) and CAN1 (Can-R) would be often caused by transient localized hypermutability
MMS-Induced *can1 ura3* Often Carry More than Two Mutations

16 out of 22 *can1 ura3* mutants fit specificity determined for ssDNA -- [C, C+A] or [G, G+T] mutated in the same strand -- *strand-coordinated clusters*

![Diagram showing mutations](image-url)
MMS-Induced Double and Multiple Mutants in Closely Spaced CAN1 URA3 are Enriched with Strand-Coordinated Mutation Pairs

G-G and C-C are the most frequent neighbors at the same strand

\[
\begin{align*}
\text{Observed} & \quad \text{Expected} \\
G \rightarrow G & \quad 8 \quad 15 \\
C \rightarrow C & \quad 7 \\
C \rightarrow C & \quad 8 \\
G \rightarrow G & \quad 1 \\
C \rightarrow C & \quad 1 \\
\end{align*}
\]

\[P_{H0} < 0.03\]
How Large are Strand-Coordinated Mutation Clusters?

Address by whole-genome re-sequencing

- indels; complex
- bps-C
- bps-G
- bps-T
- bps-A

URA3 0.8 kb  CAN1 1.8 kb
How Large are Strand-Biased Mutation Clusters?

- #3: 30 mut./95 kb !!!
- #10: 4 mut./3 kb
- #9: 4 mut./11 kb
- #1: 4 mut./9 kb
Large Cluster Can Arise from a DSB with Abnormally Long Resection
Large Cluster Can Arise from a DSB with Abnormally Long Resection
How Many Mutations in the Rest of the Genome?

30 mut./95 kb !!!

4 mut./3 kb

4 mut./11 kb

4 mut./9 kb
High Mutation Density in a Cluster; Low Mutation Load in the Rest of the Genome

<table>
<thead>
<tr>
<th>Mutations</th>
<th>Clusters 3-95 kb</th>
<th>The rest of genomes ~1,100-1,200 kb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>4-30</td>
<td>9-14</td>
</tr>
<tr>
<td>Density</td>
<td>0.3-0.8 mut/kb</td>
<td>~0.01 mut/kb</td>
</tr>
</tbody>
</table>

Does it happen in nature?
Damage-Induced Localized Hypermutability

- Finding with small reporters
  -- 1000-fold in ss DNA at DSBs and telomeres

- Large size of a region
  -- up to 100 kb

**GENOME-WIDE SCALE**

- Multiple regions in a genome
  -- simultaneous LHM in several uncapped telomeres

- Vast mutation clusters caused by chronic mutagenesis
  -- 4-30 mutations in 3-100 kb
  -- more mutations in a cluster than in the rest of the genome
Damage-Induced Localized Hypermutability -- -- Questions and Perspectives

- Sources and mechanisms
  - Acute and chronic DNA damage
  - Meiosis
  - Uncoupled replication forks
  - ?? dsDNA ?? (chromatin, transcription, etc.)

- Role in evolution and population dynamics

- Impact on human health (cancer, genetic disease)

- Detecting transient stretches of ssDNA in a cell

- Genomic toxicology – use whole genome as a cumulative dosimeter of mutagenic insults
Acknowledgements

Chromosome Stability Section, LMG, NIEHS
- Yong Yang
- Joan Sterling
- Francesca Storici (now at Georgia Tech.)
- Steve Roberts
- Mike Resnick

Molecular Genetics Core, NIEHS
- Nell Burch

UNC High-Throughput Sequencing Center
- Piotr Mieczkowsky