Using Games as a platform for Citizen Science in Molecular Biology

Evonne McArthur
COMP 585 Research Presentation
What is citizen science?

- Also called Crowd-sourced science, Civic Science
- Participation of public volunteers, usually amateur participants, in solving scientific problems
- “People Power”
  - Power in numbers

Projects:
- Great Worldwide Star Count
- Hummingbirds at Home
- Big Sleep Survey
How did games come to be used for citizen science?

- SETI@home
- 1999 at University of California, Berkeley

- Harnessed volunteers’ computers to sift through radio telescope data
- “Search for ExtraTerrestrial Intelligence”
How did games come to be used for citizen science?

- Protein Folding - an NP hard problem
  - 20 amino acids, each with different structure and properties
  - Each interacts differently with other AA and its surroundings
  - Want to create the most energetically favorable 3D shape (natural folded protein) from a linear sequence
(a) Normal amino acid sequence

Thr 4  Pro 5  Glu 6  Glu 7

Normal red blood cells

(b) Single change in amino acid sequence

Thr 4  Pro 5  Val 6  Glu 7

Sickled red blood cells

Figure 3-13  Biological Science, 2/e
How did games come to be used for citizen science?

- Folding@home (2000 - Stanford), Rosetta@home (2005 - UW Seattle)
  - Outsources problems and creates giant virtual clusters of computing power with little cost

- 2008 - Created a new interface for Rosetta@home: “Foldit”
  - Allows users to assist in the computation
  - Gives them an incentive to do so by turning it into an online game
What qualities make humans superior at solving complex biological puzzles than computers?

- Human Spatial Reasoning
  - Pattern recognition
  - “Big picture” and 3D visualization
  - Spotting subtleties
- Creativity
  - Multiple different perspectives
What makes a successful citizen science game for biological problems?

- Games that allow users to exercise what humans are good at (spatial reasoning) but allows users to leave the tediousness and details to the computers

- 1. Appropriate simplification and abstraction
- 2. Connection to the science
- 3. Community
- 4. Room for creativity
1. Appropriate simplification and abstraction

- Simplification and abstraction of unnecessary complexity for the layman
- However, cannot be grossly oversimplified
- Many games use colors and simplified shapes to indicate different biological phenomena like complementary DNA bases or hydrogen bonding in a folded protein
Phylo (Multiple Sequence Alignment) versus

Multiple Sequence Alignment of the sequences included:

<table>
<thead>
<tr>
<th>Include</th>
<th>Number</th>
<th>Sequence_id</th>
<th>Reference</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔️</td>
<td>1</td>
<td>P04_37_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td>✔️</td>
<td>2</td>
<td>P04_39_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td>✔️</td>
<td>3</td>
<td>P04_40_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td>✔️</td>
<td>4</td>
<td>P04_41_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>P04_42_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td>✔️</td>
<td>6</td>
<td>P04_43_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>P04_44_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td>✔️</td>
<td>8</td>
<td>P04_45_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>P04_47_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>P04_48_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>P04_50_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>P04_52_051_new5_HKE293_M13</td>
<td>TGAATCTTTGAGCGCGGTTCCGTTAACGCTGAA</td>
<td>CTAAGCCTTTACGCGCCGTTCCGTTAACGCTGAA</td>
</tr>
</tbody>
</table>
Foldit (Protein Folding) versus
“This is the challenging thing for the games. They have to keep the data in the original format so they don’t break the analysis but also make a fun game” – Genes in Space Designer
2. Connection to the science

- Keeping connected to the problem and the science behind it—even if just a little bit—encourages users to play and keep playing the problem

  “Maybe something I do will help contribute an answer to curing cancer or AIDS or the common cold” – FoldIt player on his reason for playing
Phylo (Multiple Sequence Alignment)

Select a disease to get an associated level

Completed puzzles

332,133

Latest achievements

<table>
<thead>
<tr>
<th>User</th>
<th>Puzzle ID</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guest</td>
<td>712</td>
<td>RB1</td>
</tr>
<tr>
<td>Guest</td>
<td>93</td>
<td>Leukencephalop...</td>
</tr>
<tr>
<td>Guest</td>
<td>381</td>
<td>P53-3</td>
</tr>
<tr>
<td>Guest</td>
<td>99</td>
<td>Hemolytic uremic...</td>
</tr>
</tbody>
</table>
EteRNA (RNA folding)

Foldit (protein folding)

“There are different profiles of players in Eterna. Some may be only concerned by their ranking, and then, only their ability to solve puzzles will count towards defining how good they are. Then, there are those like me who are in this for the science and the research.” (Nando)
3. Community

- Leaderboard, Scoring Mechanism, Achievement Icons/Badges
  - Puzzles, Challenges (weekly, monthly, etc)
- Chatroom
- Teams
- Forums
- Wikis

- Encouraging both collaboration and competition
Leaderboard

Foldit (Protein Folding)

EteRNA (RNA folding)

EyeWire (Neuron Mapping)
Foldit Gamers Solve Riddle of HIV Enzyme within 3 Weeks

The online game poses protein-folding puzzles, and participants provided insights recently used to solve the structure of an enzyme involved in reproduction of HIV.

September 20, 2011 | By Michael J. Crenn and Fast Company

When video gamers armed with the world’s most powerful supercomputers take on science and its most vexing riddles, who wins? Sometimes, it’s the gamers.

Humans retain an edge over computers.
4. Room for creativity

- Sandbox mode: create original proteins, molecular structures/switches, etc
  - Users can attempt to solve puzzles or create their own
- Many are connected with a lab which may actually synthesize some of these structures

“The initial appeal was the quick ‘success’ of solving an RNA design puzzle. Later it was the challenge of designing RNA to be folded and tested in vitro.” (Jeff Anderson-Lee)
Sandbox Mode

- Nanocrafter: in Sandbox mode can build synthetic DNA circuits, nanoscale machines
- Focused on allowing users to invent things using DNA as building blocks
- Challenges and tutorials help give users familiarity with how DNA acts
These results demonstrate that human creativity can extend beyond the macroscopic challenges encountered in everyday life to molecular-scale design problems.”
References