THE GENOMICS EDUCATION PARTNERSHIP
ASSESSING AND IMPROVING A COURSE-BASED
UNDERGRADUATE RESEARCH EXPERIENCE
(CURE)

Karim A. Sharif, PhD

Holy Quran & Science Conference, Canada

November 25, 2018
OUR BELOVED IMAM’S EXPECTATION

NOBEL PRIZE IS MINIMUM TARGET

(April 17, 2008, BAGH-E-AHMAD, GHANA)
There is not an animal that crawls in the earth, nor a bird that flies on its two wings, but they are communities like you. We have left out nothing in the Book. Then to their Lord shall they be gathered together. (Al-Ana’am 39)
GEP OVERVIEW

Primary goals:

- Incorporate genomics / bioinformatics into the undergraduate curriculum
- Engage undergraduates in genomics research
- Assess the impact of Undergraduate student involvement in research on their academic performance.
Membership:

- Over **100 faculty**
  - Mostly from primarily undergraduate institutions
- Includes **4-year and 2-year institutions**
- More than **1000 undergraduates** participate annually
RESEARCH PROBLEM: COMPARATIVE ANALYSIS OF THE DROSOPHILA DOT CHROMOSOME (F ELEMENT)

Research Questions:

- Identify conserved F element features across ~40 million years of evolution
- Elucidate features that enable F element genes to be expressed in a heterochromatic domain
- Annotate genes on various species of Drosophila using bioinformatics tools.
- Assess how involvement in scientific research impacts student learning outcome?
WHAT IS GENE ANNOTATION?

Books of Human Genome

AAACAAACAATCATAAATAGAGGAAGTTTTCCGGAATATACGATAATGGAATATCGTTCT
TAAAAAAGACAGCAAGACAGTTAAAACATGGAAAACAAGATAATTTCGCAATAGGGTAAGA
GTTTCTTTATTGTAATGGAATAGGATGCTAGTCGCAGATGATTAGACATCGGAGCTGGGAA
AATAGGAATCGGAAAGCTATGCAGCTTAACATCAATGGAACATTTGAGTATACGATAG
GATCGGCGGAGGCCGACTTCTCTCAGGATGGGTGGGAGTGCGACGACGATTTATCAGGAT
ACCAGGATGAGGTGCCCAAGTCTCAATCTAGATGCGGCACATTAGGACCATCAGGAA
GCCGCTTTATGTGGAATGCAACAAATATTAAACCATAGGCAAACGGATTGTGGGAATTCAAT
TAGGAAACCAGGCCTAGCCACCCGCTCAACAAGTGGCAAGACCATCTTTGGGGCATAAGC
CCTCATCAAAATTTGGGCACGACATTGGGGGAGACGATGATGGGCCGATAGGACCCACAG
CGTTGGACGGGTCTAATCCATATCCACATACATGAAAAGCTGCTGTTGCTGACGATGGTGCAG
TACGCTCTGGCCGTTGGCGCCTGCTGTCCGATAATGGGACACGTTTGCTGTGAGCTGGGAA
TGTTGGAGTGCGGATTTGCTCTAAACTGACTGGAATAATCAAATCGGCCGCGGACAGGACG
CCCTGCTCCGGTCGGTCGGCTCGGAATGTGGGGAATCTCATCTCCAGATGCTCTCAATCG
ATCGGCGGAAATGNTAANAGAATATCAAATATAATTGGCGCAGATAATTGNGCAGATTCAG
ACGTATATCAAAAAATATGCGCCCGTGTTAGTGTAGACAGGCTAAATATCGGAGCT
CAAATATTGGCCACACCGGTGGTCGCTCGGATATTGCTGCGGCAATAGGGGAGG
GCCACACAGGCCGCTGGGGGCGGCACAGTTTTCCAAGCAATCACTGGATGAGGAGA

Start codon
Coding region
Stop codon
Splice donor
Splice acceptor
UTR
GENE ANNOTATIONS USING BIOINFORMATICS

- Projects organized using a UCSC Genome Browser mirror
- Sequence similarity, gene predictions, and RNA-Seq evidence tracks – can be contradictory!
- Students use public databases (e.g., NCBI, FlyBase) to gather additional evidence
- Students construct the best gene model – and defend!
BLASTX search of each D. melanogaster CDS against the contig

Reading frame

Alignment

Contig

Identify the exact coordinates of each CDS using the Genome Browser

Reading frame

Use the Gene Model Checker to verify the final CDS coordinates

Gene model

Coordinates: 1245-1383, 1437-1678, 1740-2081, 2159-2337, 2397-2511
GENES ARE ANNOTATED BASED ON THEIR RELATEDNESS

There is not an animal that crawls in the earth, nor a bird that flies on its two wings, but they are communities like you.

We have left out nothing in the Book. Then to their Lord shall they be gathered together.

(Al-Ana’am 39)
TRACKING PROPERTIES OF DOT CHROMOSOME OVER 40 MILLION YEARS OF EVOLUTION

### Muller F Element Wanderer Genes

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CG11077</td>
<td>F</td>
<td>F</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>CG9935</td>
<td>F</td>
<td>F</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>CG1732</td>
<td>F</td>
<td>F</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>CG11076</td>
<td>F</td>
<td>F</td>
<td>A</td>
<td>E</td>
</tr>
<tr>
<td>CG9935</td>
<td>F</td>
<td>F</td>
<td>A</td>
<td>E</td>
</tr>
<tr>
<td>CG11076</td>
<td>F</td>
<td>F</td>
<td>A</td>
<td>F</td>
</tr>
</tbody>
</table>

### Wanderer Genes

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Or13a</td>
<td>A</td>
<td>A</td>
<td>F</td>
<td>E</td>
</tr>
<tr>
<td>CG5367</td>
<td>B</td>
<td>B</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>rho-5</td>
<td>B</td>
<td>B</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>CG4038</td>
<td>C</td>
<td>C</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>CG5262</td>
<td>D</td>
<td>D</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>PRY</td>
<td>chrY</td>
<td>chrY</td>
<td>chrY</td>
<td></td>
</tr>
<tr>
<td>Cyp1_alpha</td>
<td>NA</td>
<td>NA</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>dot</td>
<td>D. virilis</td>
<td>D. melanogaster</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dot</td>
<td>D. virilis</td>
<td>D. melanogaster</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Diagram Description**

- CG9935: D. melanogaster
- CG5367: rho-5
- CG1732: D. melanogaster
- CG11076: D. melanogaster
- CG1038: D. melanogaster
- CG11077: D. melanogaster
- CG5262: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG1732: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
- CG9935: D. melanogaster
- CG11076: D. melanogaster
Drosophila Muller F Elements Maintain a Distinct Set of Genomic Properties Over 40 Million Years of Evolution

Wilson Leung and Participating Students and Faculty of the Genomics Education Partnership

Undergrads power genomics research
With 2014 authors, an article by Leung et al. in the May issue of G3 has the largest author list of any paper published in the journal. More than 900 of those authors were undergraduate students when they performed the research.
GEP STUDENTS SHOW HIGHER LEARNING GAINS THAN SUMMER RESEARCH STUDENTS WHEN SUFFICIENT CLASS TIME IS INVESTED

GEP research resulted in better understanding of gene structure by students.
GEP STUDENTS SHOW HIGHER LEARNING GAINS THAN SUMMER RESEARCH STUDENTS WHEN SUFFICIENT CLASS TIME IS INVESTED

Data separated into quartiles based on the number of hours devoted to the annotation project.

Positive correlation between knowledge quiz scores and self-reported gains.

1. Understanding the research process
2. Knowledge construction
3. Readiness for research
4. Tolerance for obstacles
5. Skill interpreting results
6. Clarifying career choices
7. Integrating theory/practice
8. Tackling real problems
9. Assertions need evidence
10. Ability to analyze data
11. Reading/understanding primary science literature
12. Understanding science
13. Ethical conduct
14. Lab techniques
15. Skill in oral presentation
16. Skill in scientific writing
17. Understanding how scientists think
18. Independence
19. Learning community
20. Teaching potential
GEP STUDENT ALUMNI ATTITUDES CORRELATE WITH THE EXTENT OF THEIR GEP EXPERIENCE

- Asked GEP student alumni to reflect on their GEP experience
- Ratings range from 1 (strongly disagree) to 5 (strongly agree)

Shaffer CD et al. (2014)
CONCLUSIONS

• GEP students show higher levels of learning gains compared to summer research students.
• Genomics projects can benefit from the efforts of “massively parallel” undergraduates.
• Central and community support aid faculty start-up and persistence.
SO GET INVOLVED IN RESEARCH EARLY IF YOU WANT TO FULFILL BELOVED IMAM’S EXPECTATION

NOBEL PRIZE IS MINIMUM TARGET

(April 17, 2008, BAGH-E-AHMAD, GHANA)
ACKNOWLEDGMENTS

Dr. Sarah Elgin – Principal Investigator
Wilson Leung
Christopher D. Shaffer
GEP Partners
Washington University at St. Louis
Massasoit Community College

Howard-Hughes Medical Institute
National Science Foundation