

## GENOME REMODELING IN DEVELOPMENTAL TIME: ALGORITHMS FOR CILIATES

MARION SCHEEPERS

Genome remodeling occurs in cells as either a slow evolutionary process, or a rapid programmed developmental process. Intriguing examples of each type permeate the tree of life. This summer REU will feature research projects related to rapid programmed developmental processes. Algorithmic details of the developmental program of an organism allow one to investigate the mathematical and technological consequences of being able to control individual program steps.

[10] surveys programmed genome remodeling in ciliates. Ciliates have two types of nuclei, micro- and macro- nuclei. Micronuclei are encrypted versions of macronuclei. Towards illustration of the relation between the micronuclear and macronuclear genomes, consider this diagram:

Micronuclear gene	IE0	M3	IE1	M5	IE2	M2	IE3	M1	IE4	M4	IE5
Macronuclear gene						M1	M2	M3	M4	M5	

In this depiction of a hypothetical gene of a ciliate the functional macronuclear form of the gene is a single DNA sequence composed of five segments denoted M1, M2,  $\dots$ , M5, concatenated in this order. The micronuclear precursor of this gene is a molecule that contains these five DNA segments, dispersed in some order and orientation, and separated by interstitial DNA segments denoted as numbered IE's. These micronuclear configurations are species and gene dependent.

Conjugation and micronuclear gamete exchange triggers decryption of the zygote micronucleus to form a new macronucleus while prior micro- and macro- nuclei are degraded. Putative decryption operations proposed in text [3] have the capability to:

- switch the orientation of a segment of DNA<sup>1</sup>:  $A B C \longrightarrow A \overline{B} C$
- permute positions of two separated segments of DNA:  $A B C \longrightarrow C B A$

These two operations resemble operations in permutation groups and braid groups. It is plausible that the ciliate computing system is capable of solving computational problems in braid groups and symmetric groups.

Results in [8] demonstrated that the micronuclear decryption program can be reliably manipulated to induce in living ciliates a specified decryption different from the canonical one. This programmability suggests using the ciliate decryption apparatus to execute various algorithms. [7] and [8] demonstrated that intermediate results of the decryption process can be read. Thus ciliates constitute a living computing system capable of hosting an input program and input data, and of providing output data.

---

<sup>1</sup>An overline indicates opposite orientation

Why investigate this computing environment? First, it is a DNA computing environment. In a visionary experiment and subsequent report [1] in *Science*, Adleman points out the power of computing using biomolecular processes involving DNA and enzymes: The energy efficiency of this technology is approaching the limits set by the second law of thermodynamics. Information storage in DNA requires a density of around 1 bit per cubic nanometer, while electronic storage media (in 1994) accommodated a density of around 1 bit per  $10^{12}$  cubic nanometer. Adleman writes [1] “*It is a research problem of considerable interest to elucidate the kinds of algorithms that are possible with the use of molecular methods and the kinds of problems that these algorithms can efficiently solve.*” Second, it is a programmable environment with a ready made operating system for DNA computing. The energy intensive part of the experiment was the laboratory work that went into orchestrating the various algorithmic steps of the computation. Using this *in vivo* system transfers the orchestration of the algorithmic steps from the laboratory technician to the built in ciliate “operating system” and program.

In programmed genome remodeling projects students will learn basics of:

- the ciliate genome remodeling process,
- symmetric groups and the structure of permutations,
- braid groups and the relation with permutation groups,
- elementary graph theory,
- computational complexity, and
- computational problems of significant scientific interest.

In their research projects students will

- design algorithms for use by the ciliate computing environment to solve mathematical problems (see for example [11]),
- simulate mini versions of their designs using computers,
- investigate which mathematical problems are in principle solved during the ciliate micro nuclear decryption process,
- investigate the difficulty of decrypting a micro-nuclear precursor of a gene on the basis of what information is available about the target macro nuclear gene,
- compare the ciliate computing environment with other proposed *in vitro* or *in vivo* computing environments: Bacterial computing as described in [2], [4] and [9] is an example of an *in vivo* computing environment, or *in vitro* computing as described in [1] and in [11].
- explore potential practical applications of the ciliate micro nuclear decryption technology.

This work will add to a broader ongoing research program that is investigating the algorithmic properties of fundamental biomolecular processes that occur in nature. The summer 2011 REU research team investigated, and simulated *in-silico*, the use of features of the ciliate micronuclear decryption technology to determine phylogenetic relationships among several species of fruitflies [5]. In the summer 2012 REU we will feature two options for the research team:

- (1) **Project 1:** Computing the inverse of an element of a finite permutation group: The genome remodeling operations of ciliates seem to be perfectly

capable of simulating the operations of a finite symmetric group. By a classical theorem of Cayley, this capability can be leveraged to do computations in any finite group. In this project the team will investigate the problem of using a “ciliate computer” to do calculations in a finite group.

- (2) **Project 2:** Comparing the “ciliate computer” with other proposed computing environments: The fundamental question that will be addressed in this project is to what extent the “ciliate computer” can accomplish the computational feats of selected proposed *in vitro* or *in vivo* computing platforms: In particular, can a “ciliate computer” accomplish the computations accomplished by bacterial computing as in [2], [4] and [9]? And can a “ciliate computer” accomplish the computations accomplished by applying biomolecular laboratory techniques as in [1] and [11]?

#### REFERENCES

- [1] L.M. Adleman, *Molecular computation of solutions to combinatorial problems*, **Science** 266 (1994), 1021 - 1024.
- [2] J. Baumgardner, K. Acker, O. Adefuye, S.T. Crowley, W. DeLoache, J.O. Dickson, L. Heard, A.T. Martens, L. Heard, A.T. Martens, N. Morton, M. Ritter, A. Shoecraft, J. Treece, M. Unzicker, A. Valencia, M. Waters, A. Malcolm Campbell, L.J. Heyer, J.L. Poet and T.T. Eckdahl, *Solving the Hamiltonian path problem with a bacterial computer*, **Journa of Biological Engineering** 3:11 (2009),
- [3] A. Ehrenfeucht, T. Harju, I. Petre, D.M. Prescott and G. Rozenberg, *Computation in living cells: gene assembly in ciliates*, **Springer** (2004).
- [4] K.A. Haines, M.L. Broderick, A.D. Brown, T.L. Butner, J.O. Dickson, W.L. Harden, L.H. Heard, E.L. Jessen, K.L. Malloy, B.J. Ogden, S. Rosemond, S. Simpson, E. Zwack, A. Malcolm Campbell, T.T. Eckdahl, L.J. Heyer and J.L. Poet, *Engineering bacteria to solve the burnt pancake problem*, **Journal of Biological Engineering** 2:8 (2008),
- [5] J.L. Herlin, A. Nelson and M. Scheepers, *Can a ciliate compute evolutionary distance?* (in progress)
- [6] J.M. Keil, J. Liu and I. McQuillan, *Algorithmic properties of ciliate sequence alignment*, **Theoretical Computer Science** 411 (2010), 919 - 925
- [7] M. Möllenbeck et al., *The pathway to detangle a scrambled gene*, **Plos One**, 3:6 (2008), 110.
- [8] M Nowacki, V. Vijayan, Y. Zhou, K. Schotanus, T.G. Doak and L.F. Landweber, *RNA-mediated epigenetic programming of a genome-rearrangement pathway*, **Nature** 451 (2008), 153159.
- [9] J.L. Poet, A.M. Campbell, T.T. Eckdahl and L.J. Heyer, *Bacterial computing*, **XRDS** 17:1 (2010), 11 - 15.
- [10] D. M. Prescott, *Genome Gymnastics: Unique modes of DNA evolution and processing in ciliates*, **Nature Reviews** 1 (2000), 191198.
- [11] X. Wang, Z. Bao, J. Hu, S. Wang and A. Zhan, *Solving the SAT problem using a DNA computing algorithm based on ligase chain reaction*, **Biosystems** 91 (2008), 117 - 125.