

Genome Remodeling in Developmental Time: Algorithms for Ciliates

Christopher Anderson

Lewis and Clark College

AAAS Pacific Division Conference

June 24-27, 2012



In collaboration with:
Marion Scheepers, Marlena Warner and Helen Wauck
representing
**Boise State University, University of Idaho, and Gustavus
Adolphus College**



Properties of Ciliate Nuclei

¹Mochizuki, Kazufumi, *DNA rearrangements directed by non-coding RNAs in ciliates*, **Wiley Interdisciplinary**

Reviews: *RNA* 1 (3): 376387

Properties of Ciliate Nuclei

- Separate germline Micronucleus and transcriptionally active Macronucleus

¹Mochizuki, Kazufumi, *DNA rearrangements directed by non-coding RNAs in ciliates*, **Wiley Interdisciplinary**

Properties of Ciliate Nuclei

- Separate germline Micronucleus and transcriptionally active Macronucleus
 - Micronucleus- few chromosomes, exchanged during conjugation, interrupted by IESs

¹Mochizuki, Kazufumi, *DNA rearrangements directed by non-coding RNAs in ciliates*, **Wiley Interdisciplinary**

Properties of Ciliate Nuclei

- Separate germline Micronucleus and transcriptionally active Macronucleus
 - Micronucleus- few chromosomes, exchanged during conjugation, interrupted by IESs
 - Macronucleus- can have over 20,000 separate chromosomes, often with a single gene¹

¹Mochizuki, Kazufumi, *DNA rearrangements directed by non-coding RNAs in ciliates*, **Wiley Interdisciplinary**

Properties of Ciliate Nuclei

- Separate germline Micronucleus and transcriptionally active Macronucleus
 - Micronucleus- few chromosomes, exchanged during conjugation, interrupted by IESs
 - Macronucleus- can have over 20,000 separate chromosomes, often with a single gene¹
- Removal of IESs, splicing, inversion, and amplification transforms a copy of the micronucleus into the macronucleus

¹Mochizuki, Kazufumi, *DNA rearrangements directed by non-coding RNAs in ciliates*, **Wiley Interdisciplinary**

Ciliate Micronuclear DNA as permutations

Ciliate Micronuclear DNA as permutations

Ignoring IESs, segments of genes in micronuclear DNA can be viewed as permutations of the final macronuclear gene sequence

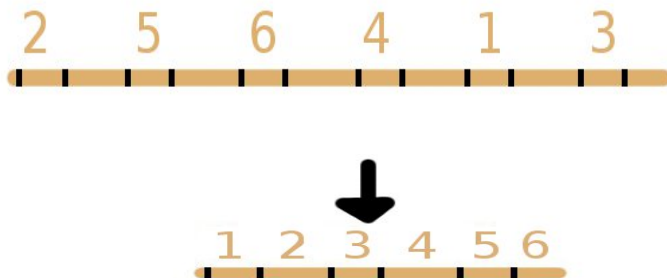
Ciliate Micronuclear DNA as permutations

Ignoring IESs, segments of genes in micronuclear DNA can be viewed as permutations of the final macronuclear gene sequence



Ciliate Micronuclear DNA as permutations

Ignoring IESs, segments of genes in micronuclear DNA can be viewed as permutations of the final macronuclear gene sequence



Other permutation forms

Other permutation forms

This intuitive segment-numbering represents a permutation in inverse-image form.

Other permutation forms

This intuitive segment-numbering represents a permutation in inverse-image form.



Inverse Image Notation: [2 5 6 4 1 3]

Other permutation forms

This intuitive segment-numbering represents a permutation in inverse-image form.



Inverse Image Notation: [2 5 6 4 1 3]

Other forms:

Cauchy's two-line notation:
$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

Other permutation forms

This intuitive segment-numbering represents a permutation in inverse-image form.



Inverse Image Notation: [2 5 6 4 1 3]

Other forms:

Cauchy's two-line notation:
$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

Disjoint Cycle Notation: (152)(36)(4)

Inverting permutations using cds

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$
$$\text{Inverse} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = [5 \ 1 \ 6 \ 4 \ 2 \ 3]$$

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

$$\text{Inverse} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = [5 \ 1 \ 6 \ 4 \ 2 \ 3]$$

Pointer Pair	Unsigned list	Permutation
	[2 5 6 4 1 3]	

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

$$\text{Inverse} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = [5 \ 1 \ 6 \ 4 \ 2 \ 3]$$

Pointer Pair	Unsigned list	Permutation
(1,2)(2,3)	[2 5 6 4 1 3] [5 6 4 1 2 3]	(1 5 4 3 2)(6)

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

$$\text{Inverse} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = [5 \ 1 \ 6 \ 4 \ 2 \ 3]$$

Pointer Pair	Unsigned list	Permutation
	[2 5 6 4 1 3]	
(1, 2)(2, 3)	[5 6 4 1 2 3]	(1 5 4 3 2)(6)
(4, 5)(3, 4)	[1 2 3 4 5 6]	(1 5 2 6 3 4)

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

$$\text{Inverse} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = [5 \ 1 \ 6 \ 4 \ 2 \ 3]$$

Pointer Pair	Unsigned list	Permutation
	[2 5 6 4 1 3]	
(1, 2)(2, 3)	[5 6 4 1 2 3]	(1 5 4 3 2)(6)
(4, 5)(3, 4)	[1 2 3 4 5 6]	(1 5 2 6 3 4)

The composition of the two permutations is

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

$$\text{Inverse} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = [5 \ 1 \ 6 \ 4 \ 2 \ 3]$$

Pointer Pair	Unsigned list	Permutation
	[2 5 6 4 1 3]	
(1, 2)(2, 3)	[5 6 4 1 2 3]	(1 5 4 3 2)(6)
(4, 5)(3, 4)	[1 2 3 4 5 6]	(1 5 2 6 3 4)

The composition of the two permutations is
 $(1 \ 5 \ 2 \ 6 \ 3 \ 4) \circ (1 \ 5 \ 4 \ 3 \ 2)(6) =$

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

$$\text{Inverse} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = [5 \ 1 \ 6 \ 4 \ 2 \ 3]$$

Pointer Pair	Unsigned list	Permutation
	[2 5 6 4 1 3]	
(1, 2)(2, 3)	[5 6 4 1 2 3]	(1 5 4 3 2)(6)
(4, 5)(3, 4)	[1 2 3 4 5 6]	(1 5 2 6 3 4)

The composition of the two permutations is

$$(1 \ 5 \ 2 \ 6 \ 3 \ 4) \circ (1 \ 5 \ 4 \ 3 \ 2)(6) =$$

$$(1 \ 2 \ 5)(3 \ 6)(4) =$$

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

$$\text{Inverse} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = [5 \ 1 \ 6 \ 4 \ 2 \ 3]$$

Pointer Pair	Unsigned list	Permutation
	[2 5 6 4 1 3]	
(1, 2)(2, 3)	[5 6 4 1 2 3]	(1 5 4 3 2)(6)
(4, 5)(3, 4)	[1 2 3 4 5 6]	(1 5 2 6 3 4)

The composition of the two permutations is

$$(1 \ 5 \ 2 \ 6 \ 3 \ 4) \circ (1 \ 5 \ 4 \ 3 \ 2)(6) =$$

$$(1 \ 2 \ 5)(3 \ 6)(4) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix}$$

Inverting permutations using cds

$$[2 \ 5 \ 6 \ 4 \ 1 \ 3] = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 6 & 4 & 2 & 3 \end{pmatrix}$$

$$\text{Inverse} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix} = [5 \ 1 \ 6 \ 4 \ 2 \ 3]$$

Pointer Pair	Unsigned list	Permutation
	[2 5 6 4 1 3]	
(1, 2)(2, 3)	[5 6 4 1 2 3]	(1 5 4 3 2)(6)
(4, 5)(3, 4)	[1 2 3 4 5 6]	(1 5 2 6 3 4)

The composition of the two permutations is

$$(1 \ 5 \ 2 \ 6 \ 3 \ 4) \circ (1 \ 5 \ 4 \ 3 \ 2)(6) =$$

$$(1 \ 2 \ 5)(3 \ 6)(4) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 4 & 1 & 3 \end{pmatrix}$$

This is the inverse of the original in disjoint cycle form!

Tracking the Intermediates

Pcr work in Professor James Smith's Lab at BSU will use specific primers to search for certain intermediate sequences

²GENBank Accession Number HM920148

Tracking the Intermediates

Pcr work in Professor James Smith's Lab at BSU will use specific primers to search for certain intermediate sequences

- Current work with *Stylonychia lemnae* focuses on sequences know to be invertible by cds, like the αTBP gene²

²GENBank Accession Number HM920148

Tracking the Intermediates

Pcr work in Professor James Smith's Lab at BSU will use specific primers to search for certain intermediate sequences

- Current work with *Stylonychia lemnae* focuses on sequences know to be invertible by cds, like the αTBP gene²
- Results of pcr so far inconclusive - primers have not exhaustively tested all possible paths.

²GENBank Accession Number HM920148

Biological computing?

Biological computing?

- Remember: CDS operations cannot resolve EVERY permutation of length n

Biological computing?

- Remember: CDS operations cannot resolve EVERY permutation of length n $[n \ 1 \ 2 \ \dots \ n - 1]$

Biological computing?

- Remember: CDS operations cannot resolve EVERY permutation of length n $[n \ 1 \ 2 \ \dots \ n - 1]$
- However, one of our results implies:

Biological computing?

- Remember: CDS operations cannot resolve EVERY permutation of length n $[n \ 1 \ 2 \ \dots \ n - 1]$
- However, one of our results implies:
Any length n permutation $[a_1 \ \dots \ a_n]$ represented in a ciliate by the length $n+1$ sequence

Biological computing?

- Remember: CDS operations cannot resolve EVERY permutation of length n $[n \ 1 \ 2 \ \dots \ n - 1]$
- However, one of our results implies:
Any length n permutation $[a_1 \ \dots \ a_n]$ represented in a ciliate by the length $n+1$ sequence
 $[1 \ a_1 \ \dots \ a_n]$

Biological computing?

- Remember: CDS operations cannot resolve EVERY permutation of length n $[n \ 1 \ 2 \ \dots \ n - 1]$
- However, one of our results implies:
Any length n permutation $[a_1 \ \dots \ a_n]$ represented in a ciliate by the length $n+1$ sequence

$$[1 \ a_1 \ \dots \ a_n]$$

is invertible by a ciliate.

Biological computing?

³M. Nowacki, V. Vijayan, Y. Zhou, K. Schotanus, T.G. Doak and L.F. Landweber, *RNA-mediated epigenetic programming of a genome-rearrangement pathway*, **Nature** 451:10 (2008), 153 - 259

⁴Jeffrey L. Poet, A. Malcolm Campbell, Todd T. Eckdahl, and Laurie J. Heyer. *Bacterial computing*, **XRDS** 17, 1 (September 2010), 10-15. DOI=10.1145/1836543.1836550

Biological computing?

- One can actually "reprogram" the decryption process³ with one's own order, effectively using the ciliate decryption program to invert any permutation of n symbols starting with the first symbol or ending with an $n + 1$ symbol.

³M. Nowacki, V. Vijayan, Y. Zhou, K. Schotanus, T.G. Doak and L.F. Landweber, *RNA-mediated epigenetic programming of a genome-rearrangement pathway*, **Nature** 451:10 (2008), 153 - 259

⁴Jeffrey L. Poet, A. Malcolm Campbell, Todd T. Eckdahl, and Laurie J. Heyer. *Bacterial computing*, **XRDS** 17, 1 (September 2010), 10-15. DOI=10.1145/1836543.1836550

Biological computing?

- One can actually "reprogram" the decryption process³ with one's own order, effectively using the ciliate decryption program to invert any permutation of n symbols starting with the first symbol or ending with an $n + 1$ symbol.
- Math (and computer science) is all about this kind of problem generalization - what else can we phrase as a permutation of n symbols to be solved by ciliates?

³M. Nowacki, V. Vijayan, Y. Zhou, K. Schotanus, T.G. Doak and L.F. Landweber, *RNA-mediated epigenetic programming of a genome-rearrangement pathway*, **Nature** 451:10 (2008), 153 - 259

⁴Jeffrey L. Poet, A. Malcolm Campbell, Todd T. Eckdahl, and Laurie J. Heyer. *Bacterial computing*, **XRDS** 17, 1 (September 2010), 10-15. DOI=10.1145/1836543.1836550

Biological computing?

- One can actually "reprogram" the decryption process³ with one's own order, effectively using the ciliate decryption program to invert any permutation of n symbols starting with the first symbol or ending with an $n + 1$ symbol.
- Math (and computer science) is all about this kind of problem generalization - what else can we phrase as a permutation of n symbols to be solved by ciliates?
 - *E. coli* find Hamiltonian Paths⁴
 - Potential Cryptosystems applications

³M. Nowacki, V. Vijayan, Y. Zhou, K. Schotanus, T.G. Doak and L.F. Landweber, *RNA-mediated epigenetic programming of a genome-rearrangement pathway*, **Nature** 451:10 (2008), 153 - 259

⁴Jeffrey L. Poet, A. Malcolm Campbell, Todd T. Eckdahl, and Laurie J. Heyer. *Bacterial computing*, **XRDS** 17, 1 (September 2010), 10-15. DOI=10.1145/1836543.1836550

Acknowledgments

Acknowledgments

- Prof. Jim Smith, Boise State University, for consultation and use of his laboratory

Acknowledgments

- Prof. Jim Smith, Boise State University, for consultation and use of his laboratory
- Prof. Hans Lipps and Dr. Franziska Jonnson of the University of Witten/Herdecke, Germany, for a gift of DNA of the ciliate *Stylonychia lemnae*

Acknowledgments

- Prof. Jim Smith, Boise State University, for consultation and use of his laboratory
- Prof. Hans Lipps and Dr. Franziska Jonnson of the University of Witten/Herdecke, Germany, for a gift of DNA of the ciliate *Stylonychia lemnae*
- Boise State University Math REU 2012

Acknowledgments

- Prof. Jim Smith, Boise State University, for consultation and use of his laboratory
- Prof. Hans Lipps and Dr. Franziska Jonnson of the University of Witten/Herdecke, Germany, for a gift of DNA of the ciliate *Stylonychia lemnae*
- Boise State University Math REU 2012
- NSF REU funding via grant DMS 1062857

Acknowledgments

- Prof. Jim Smith, Boise State University, for consultation and use of his laboratory
- Prof. Hans Lipps and Dr. Franziska Jonnson of the University of Witten/Herdecke, Germany, for a gift of DNA of the ciliate *Stylonychia lemnae*
- Boise State University Math REU 2012
- NSF REU funding via grant DMS 1062857



- NIH INBRE funding

Thank you!