

Abstract

Biological
Background

Regression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

Evidence of a Pathway of Reduction in Bacteria: Reduced Quantities of Restriction Sites Impact tRNA Activity in a Trial Set

Oliver Bonham-Carter,
Lotfollah Najjar, Dhundy Bastola

School of Interdisciplinary Informatics
University of Nebraska at Omaha,
Omaha, NE, 68182
USA

Fall 2013

Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

Pathway

Conclusions

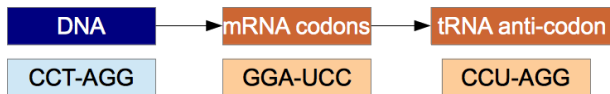
References

Thanks To



- The study's *Pathway of Reduction* describes the phenomenon of missing palindromic DNA content, extending to absent codon and tRNA content.





- Palindromes are generally found in statistically *under-expected* levels in bacterial DNA.
 - Palindromes (*often restriction sites*) are thought to be a danger to the host if they are failed to be methylated.
- DNA makes up triplets (mRNA codons) which are associated with tRNA anti-codons.
- tRNAs serve as the physical link between the nucleotide sequence of nucleic acids (RNA) and the amino acid sequence of proteins.

Abstract

Biological
Background

Regression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

- The understanding of reduced quantities of particular palindromes may help to predict which tRNAs may also be observed in reduced quantities.
- This tRNA usage information may be applied to controlling translation efficiency and to slow protein growth in bacteria. For example, in the case of an infection.

What Are Palindromes?

DNA palindromes are complementary words

Abstract

Biological
Background

Regression
Models

Palindromes

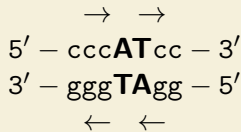
tRNAs

Pathway

Conclusions

References

Thanks To



- $A \implies T, T \implies A$ and $C \implies G, G \implies C$
- The palindrome **AT** is the only common word; it is read the same way from the 5-prime end of both strands.



Restriction Sites are Largely Palindromic

Abstract

Biological
Background

Regression
Models

Palindromes

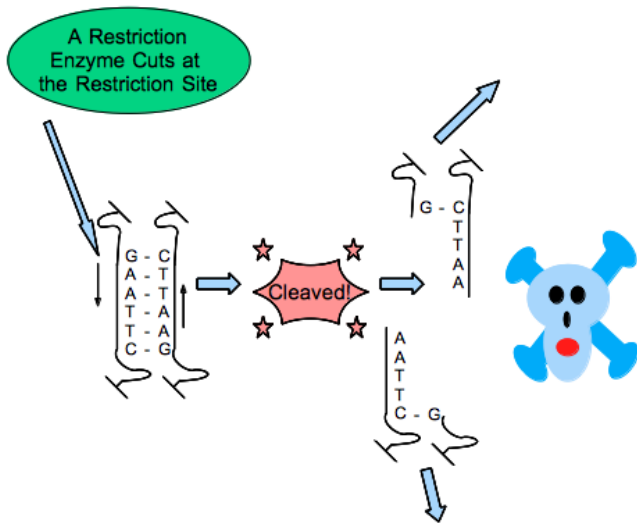
tRNAs

Pathway

Conclusions

References

Thanks To



Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

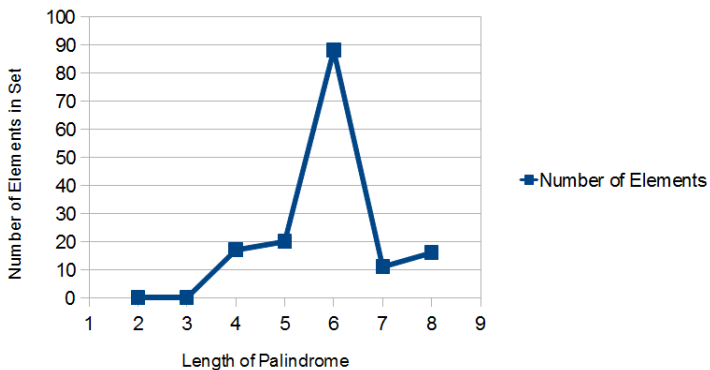


Figure : The most common group of restriction sites is length 6. Prepared using REBASE; R. Roberts, T. Vincze, J. Posfai, and D. Macelis. Rebase; a database for dna restriction and modification: enzymes, genes and genomes. *Nucleic acids research*, 38(suppl 1):D234D236, 2010.

A Mechanism for Protein Translation

Abstract

Biological
Background

Regression
Models

Palindromes

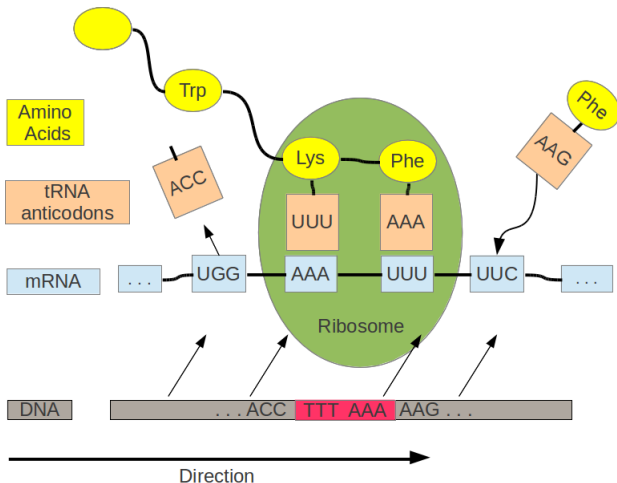
tRNAs

Pathway

Conclusions

References

Thanks To



The ribosome translates the RNA to amino acid residues.

Abstract

Biological
Background

Regression
Models

Palindromes

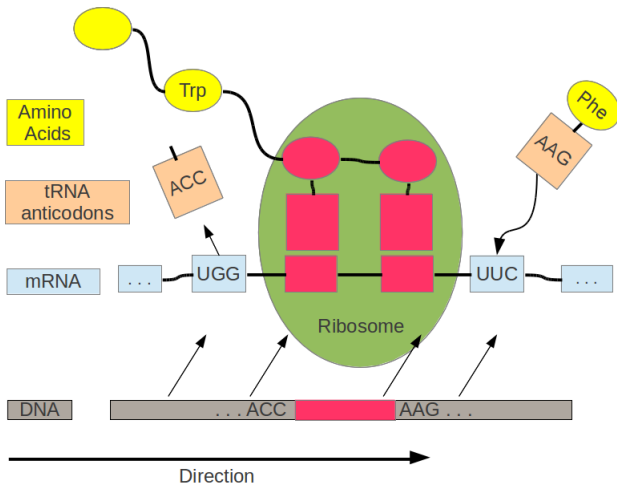
tRNAs

Pathway

Conclusions

References

Thanks To



DNA affects tRNA mechanisms.

Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

Organism Name	Locus	Abbrev.	Material
<i>Bacillus subtilis</i>	NC_000964	BS	genome
<i>Escherichia coli</i>	NC_010498	EC	genome
<i>Haemophilus influenzae</i>	NC_000907	HI	genome
<i>Methanococcus jannaschii</i>	NC_000909	MJ	genome
<i>Mycoplasma genitalium</i>	NC_000908	MG	genome
<i>Synechocystis sp.</i>	NC_000911	Ssp	genome
<i>Marchantia polymorpha</i>	NC_001660	Mit-MP	Mito
<i>Marchantia polymorpha</i>	Z98094.1	Chlor-MP	Chloro

The organisms, their abbreviations and the type of data used in our study. We note that "Mito" and "Chloro" indicate "mitochondria" and "chloroplasts," respectively.

Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

cellular organisms

Bacteria

⊕ Synechocystis sp.⊕ Mycoplasma genitalium⊕ Gammaproteobacteria⊕ Haemophilus influenzae⊕ Escherichia coli⊕ Bacillus subtilis⊕ Methanocaldococcus jannaschii (Methanococcus jannaschii)⊕ Marchantia polymorpha

Table : Graphic prepared using the taxonomy tool by;
D. Wheeler, T. Barrett, D. Benson, S. Bryant, K. Canese, V. Chetvernin, D. Church, M. DiCuccio, R. Edgar,
S. Federhen, *et al.* Database resources of the national center for biotechnology information. *Nucleic acids
research*, 35(suppl 1): D5D12, 2007.

Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

Next steps

- → **What do the palindromic avoidance trends look like?**
- What are the tRNAs which are absent in the organismal data **AND** also in the palindromic DNA content?

Stepwise Regression Model Building

Why use stepwise regression models?

Abstract

Biological
Background

Regression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

- A regression model may be prepared when there is a statistically significant relationship between its variables.
 - Modeling provided a good way to determine which variables may have similar palindromic content trends.
- Since many models would have to be tested, we utilized stepwise regression software to automate the building procedure.

Stepwise Regression Model Analysis

Preliminary results: significant avoidance relationships were found

Abstract

Biological
Background

Regression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

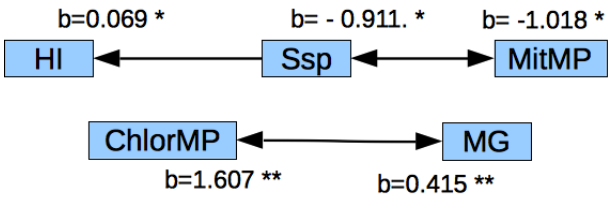
- We applied this palindromic avoidance data (lengths 4, 5 and 6) to stepwise regression models.
- The regression models showed that there were similar trends of palindromic avoidance (all lengths) between the organisms.
- This implies that avoidance is a common trend in the organisms of our data.

Avoided palindromes of length 4.

Significant likeness across *most* of the data

- Abstract
- Biological Background
- Regression Models
- Palindromes**
- tRNAs
- Pathway
- Conclusions
- References
- Thanks To

Organism Name	Locus	Abbrev.	Material
<i>Bacillus subtilis</i>	NC_000964	BS	genome
<i>Escherichia coli</i>	NC_010498	EC	genome
<i>Haemophilus influenzae</i>	NC_000907	HI	genome
<i>Methanococcus jannaschii</i>	NC_000909	MJ	genome
<i>Mycoplasma genitalium</i>	NC_000908	MG	genome
<i>Synechocystis sp.</i>	NC_000911	Ssp	genome
<i>Marchantia polymorpha</i>	NC_001660	Mit-MP	Mito
<i>Marchantia polymorpha</i>	Z98094.1	Chlor-MP	Chloro



The α -value significance of each relationship is indicated by * or ** for $\alpha = 0.01$ or $\alpha = 0.05$, respectively.

Avoided palindromes of length 5.

Significant likeness across *all* of the data

Abstract

Biological
Background

Regression
Models

Palindromes

tRNAs

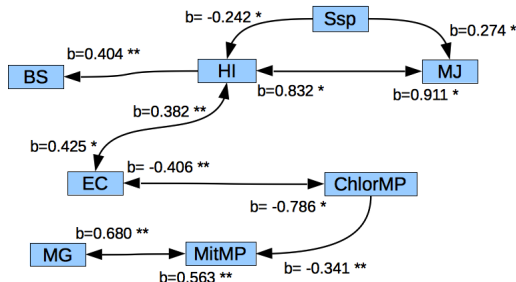
Pathway

Conclusions

References

Thanks To

Organism Name	Locus	Abbrev.	Material
<i>Bacillus subtilis</i>	NC_000964	BS	genome
<i>Escherichia coli</i>	NC_010498	EC	genome
<i>Haemophilus influenzae</i>	NC_000907	HI	genome
<i>Methanococcus jannaschii</i>	NC_000909	MJ	genome
<i>Mycoplasma genitalium</i>	NC_000908	MG	genome
<i>Synechocystis sp.</i>	NC_000911	Ssp	genome
<i>Marchantia polymorpha</i>	NC_001660	Mit-MP	Mito
<i>Marchantia polymorpha</i>	Z98094.1	Chlor-MP	Chloro



The α -value significance of each relationship is indicated by * or ** for $\alpha = 0.01$ or $\alpha = 0.05$, respectively.

Avoided palindromes of length 6.

Significant likeness across *some* of the data

Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

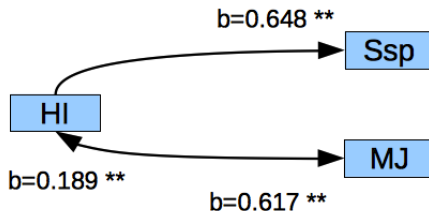
Pathway

Conclusions

References

Thanks To

Organism Name	Locus	Abbrev.	Material
<i>Bacillus subtilis</i>	NC_000964	BS	genome
<i>Escherichia coli</i>	NC_010498	EC	genome
<i>Haemophilus influenzae</i>	NC_000907	HI	genome
<i>Methanococcus jannaschii</i>	NC_000909	MJ	genome
<i>Mycoplasma genitalium</i>	NC_000908	MG	genome
<i>Synechocystis sp.</i>	NC_000911	Ssp	genome
<i>Marchantia polymorpha</i>	NC_001660	Mit-MP	Mito
<i>Marchantia polymorpha</i>	Z98094.1	Chlor-MP	Chloro



The α -value significance of each relationship is indicated by * or ** for $\alpha = 0.01$ or $\alpha = 0.05$, respectively.

Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

REGRESSION

```
/DESCRIPTIVES MEAN STDDEV CORR  
SIG N  
/SELECT=id EQ 2  
/MISSING LISTWISE  
/STATISTICS COEFF OUTS R ANOVA  
/CRITERIA=PIN(.05) POUT(.10)  
/NOORIGIN  
/DEPENDENT {DependentVariable} (i.e.  
MitMP)  
/METHOD=STEPWISE {Predictors}(i.e  
Ssp HI EC MJ BS MG ChlorMP).
```

SPSS software suite (IBM Corp. Released 2010. IBM SPSS Statistics for Windows, Version 19.0. Armonk, NY: IBM Corp.)

A Relation By Avoidance

Concluded from the models

Abstract

Biological
Background

Regression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

- A high degree of palindromic avoidance common across diverse bacterial organisms.
- Arrows between variables suggested significant correlation.
 - There were many cases to suggest that palindromic avoidance was similar across the data (length 5 palindromic data was especially correlated).
- There were few variables of length-4 and 6 shown in the correlation graphs.
 - This may suggest that there was so much palindromic avoidance that there was little non-zero data to graph.

Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

Next steps

- What do the palindromic avoidance trends look like?
- → **What are the tRNAs which are absent in the organismal data AND also in the palindromic DNA content?**

Which tRNAs are Missing in the Organisms?

Finding missing tRNAs that are also found in palindromic DNA

Abstract

Biological
Background

Regression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

We obtained the tRNA's in our set of organisms:

- Isolated the organismal tRNA sequence data from the Genbank records
- Obtained the amino acid anticodons from this sequence data by BLASTing over known tRNA sequence data in other organisms
- Prepared a combined list of all tRNA's taken from the organisms together and found which tRNAs were missing from the list.



Which tRNAs are Missing in the Palindromes?

Finding missing tRNAs that are also found in palindromic DNA

Abstract

Biological
Background

Regression
Models

Palindromes

tRNAs

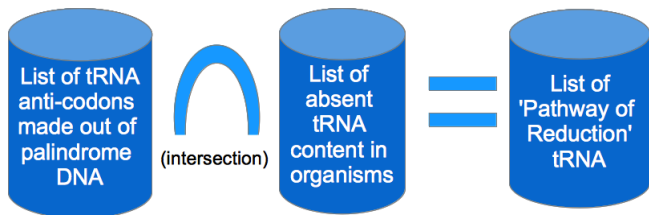
Pathway

Conclusions

References

Thanks To

- Extracted possible codons from palindromic DNA to determine tRNA content
- Prepared a list of tRNAs created from the DNA of the avoided palindromes.
- Determined which tRNAs from palindromic DNA were also found in the organismal data.



Evidence of the Pathway of Reduction

The number of amino acids possible from codons of palindrome

- Abstract
- Biological Background
- Regression Models
- Palindromes
- tRNAs
- Pathway
- Conclusions
- References
- Thanks To

In Len 4 APs	Count	In Len 5 APs	Count	In Len 6 APs	Count
A	3	A	2	A	15
D	1	E	2	C	2
G	1	D	3	E	3
I	1	G	6	D	5
STOP	1	F	1	G	9
L	1	I	2	F	1
P	1	L	2	I	5
R	3	N	1	H	3
		Q	1	K	1
		P	6	STOP	2
		S	2	M	2
		R	3	L	7
		T	2	N	2
		W	1	Q	2
		V	2	P	10
				S	10
				R	17
				T	5
				W	2
				V	7
				Y	4

Table : A complete listing all codons for amino acids (AAs) that were extracted from the DNA of the avoided palindromes (APs). The columns contain the counts of codons correlating to each extracted amino acid. The gray cells indicate that a triplet from the AP code was also missing a corresponding tRNA according to our analysis using BLAST. These cells are evidence for the pathways of reduction of our study.

Missing tRNAs by BLAST analysis.

As the length of palindromes increases, there are more tRNA missing.

The missing tRNAs: found in the avoided palindromes and also in the organismal tRNA content.

Amino Acid	Codons in Nature	Length 4 missing codons	Length 5 missing codons	Length 6 missing codons
Cysteine (C)	2			tgt
Phenylalanine (F)	2		ttt	ttt
Leucine (L)	6	tta, ttg	tta, ttg	tta, ttg
Asparagine (N)	2		aat	aat
Serine (S)	6		tcc, tcg, tct	tcc, tcg, tct
Tryptophan (W)	1		tgg	tgg
Tyrosine (Y)	2			tat, tac

The tRNAs which were absent across all organisms of our study according to a BLAST analysis. The above tRNAs, by extension from reduced codon content according to avoided palindromic DNA, are the end-points of the pathway of reduction.

Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

- Avoided palindromes imply that its DNA, codons and resulting tRNAs may also be absent.
- The DNA from these palindromes is likely unavailable to form the mRNA that interacts with the tRNA.
- This phenomenon is what we called, the *pathway of reduction*.

Abstract

Biological
Background

Regression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

- Many organisms cannot make the tRNAs they consume for translation.
- Knowledge of these missing tRNAs may be used to study how organisms prepare functional proteins.
- This knowledge of functional proteins may also help to prevent some kind of organisms from growing too quickly: the bacteria during the course of an infection.

Abstract

Biological
BackgroundRegression
Models

Palindromes

tRNAs

Pathway

Conclusions

References

Thanks To

- R. Roberts, T. Vincze, J. Posfai, and D. Macelis. **Rebase; a database for dna restriction and modification: enzymes, genes and genomes.** *Nucleic acids research*, 38(suppl 1):D234D236, 2010.
- D. Wheeler, T. Barrett, D. Benson, S. Bryant, K. Canese, V. Chetvernin, D. Church, M. DiCuccio, R. Edgar, S. Federhen, et al. **Database resources of the national center for biotechnology information.** *Nucleic acids research*, 35(suppl 1): D5D12, 2007.
- M. Gelfand and E. Koonin. **Avoidance of palindromic words in bacterial and archaeal genomes: a close connection with restriction enzymes.** *Nucleic acids research*, 25(12):24302439, 1997.

We would like to thank the support students, faculty and staff in the UNO-Bioinformatics Core Facility. This project has been funded by the grants from the National Center for Research Resources (5P20RR016469) and the National Institute for General Medical Science (NIGMS) (8P20GM103427).

Thank You! Questions?

obonhamcarter@unomaha.edu

IS&T Bioinformatics

<http://bioinformatics.ist.unomaha.edu/>