



microRNA Studies

Chen-Hanson Ting

SVFIG

June 23, 2018



Summary

- **MicroRNA (miRNA)**
- **Species and organisms studied**
- **miRNA in mitochondria**
- **Huge genome files**
- **miRNA in human Chromosome 1**
- **miRNA in bacteria**
- **Tools used**



MicroRNA

MicroRNA (miRNA) are small RNA molecule (~22 nucleotides) found in plants, animals and some viruses, that functions in RNA silencing and post-transcriptional regulation of gene expression.



MicroRNA

- **miRBase Database**
 - **48,885 miRNA's**
 - **27,790 unique miRNA's**
 - **5312 human miRNA's**



Species and Organisms

- **Mitochondrion (human)**
- **Nasuia (*Nasuia deltocephalinicola*)**
- **Ruddii (*Candidatus Carsonella ruddii*)**
- **Mgen (*Mycoplasma genitalium*)**
- **Ecoli (*Escherichia coli*)**
- **Nematode (*Caenorhabditis elegans*)**
- **Mouse (*Mus musculus*)**
- **Human (*Homo sapiens*)**



Species and Organisms

<u>Organsims</u>	<u>Genome</u>	<u>Genes</u>	<u>Chromosome</u>
<u>Mitochondrion</u>	16,569 bp	37	1 Circular Chromosome
<u>Nasuia</u>	112,091	137	1 Circular Chromosome
<u>Ruddii</u>	159,662	182	1 Circular Chromosome
<u>Mgen</u>	580,070	525	1 Circular Chromosome
<u>Ecoli</u>	4,639,221	4377	1 Circular Chromosome
Nematode	100,258,171	21,733	6 Chromosomes
Mouse	2.7 Gb	23,000	21 chromosomes
Human	3.289 Gb	21,000	23 chromosomes



miRNA in Mitochondrion

- **Mitochondrion is an independent organism supplying energy to cells.**
- **Cell controls mitochondrion by sending miRNA to it.**
- **There must be sites in mitochondrion genome to receive miRNA signal.**



miRNA in Human Mitochondrion

Location	miRNA	Bases
624	hsa-miR-4284	GGGCTCACATCACCCCAT
2538	hsa-miR-4485-5p	ACCGCCTGCCCAGTGA
2562	hsa-miR-4485-3p	TAACGGCCGCGGTACCCTAA
2581	hsa-miR-1973	ACCGTGCAAAGGTAGCATA
5747	ppy-miR-4484	AAAAAGGCGGGAGAAGCCCCG



miRNA in All Mitochondria

Many miRNAs appear in human mitochondrion.

miRNAs generally appear in mitochondria of all species.

They must serve important (yet to be found) functions in mitochondria.

miRNA in All Mitochondria

miRNA	Freq	miRNA Name	Code Sequence
M23835	2525	cgr-miR-1973 <i>Cricetulus griseus</i>	TGACCGTGCAAAGGTAGC
M24383	679	ppy-miR-4484 <i>Pongo pygmaeus</i>	AAAAAGGCGGGAGAAGCCCCG
L09448	523	hsa-miR-1973 <i>Homo sapiens</i>	TATGCTACCTTTGCACGGT
L18987	272	hsa-miR-4463 <i>Homo sapiens</i>	GGCCCCACCCCAGTCTC
M27339	237	mmu-miR-6481 <i>Mus musculus</i>	CACTGAAAATGCTTAGATG
M09448	153	hsa-miR-1973 <i>Homo sapiens</i>	ACCGTGCAAAGGTAGCATA
M32116	62	hsa-miR-4485-5p <i>Homo sapiens</i>	ACCGCCTGCCCAGTGA
L20652	47	rgl-miR5141 <i>Rehmannia glutinosa</i>	TTATCTGTTCAGTCGCGTCGGGTCT
M22175	41	osa-miR5539a <i>Oryza sativa</i>	AAGAAAACGGATGCGCGTGCTA
M19019	37	hsa-miR-4485-3p <i>Homo sapiens</i>	TAACGGCCGCGGTACCCCTAA
L21013	35	gma-miR4995 <i>Glycine max</i>	CCCTTAACCAAGCCACTGCCT
M20652	35	rgl-miR5141 <i>Rehmannia glutinosa</i>	AGACCCGACGCGACTGACAGATAA
L22175	21	osa-miR5539a <i>Oryza sativa</i>	TAGCACGCGCATCCGTTTTTCTT
M21013	21	gma-miR4995 <i>Glycine max</i>	AGGCAGTGGCTTGGTTAAGGG
M22174	20	osa-miR5538 <i>Oryza sativa</i>	ACTGAACTCAATCACTTGCTGC
L22174	16	osa-miR5538 <i>Oryza sativa</i>	GCAGCAAGTGATTGAGTTCAGT
L23835	16	cgr-miR-1973 <i>Cricetulus griseus</i>	GCTACCTTTGCACGGTCA
M17325	14	mmu-miR-466i-5p <i>Mus musculus</i>	TGTGTGTGTGTGTGTGTGTG
L45455	13	pla-miR11602 <i>Paeonia lactiflora</i>	GATCCAATAGCGTTCGGTTAGA
L24793	12	hbr-miR6173 <i>Hevea brasiliensis</i>	AGTATCCATCGTTTTACGGC



Huge Genome Files

- **Human genome is 3.2 Gb long.**
- **Mouse genome is 2.7 Gb long.**
- **I had the privilege to traverse both, looking for microRNAs in them.**



Huge Genome Files

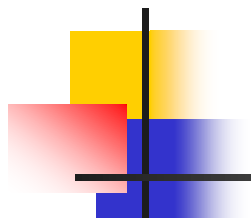
- **A typical genome file, of human Chromosome 1, is 233 Mb long.**
- **7670 miRNAs are found in Chromosome 1.**
- **Prevailing miRNAs are identified.**



miRNA in Chromosome 1

- **Two miRNAs appear most frequently in Chromosome 1, similarly in other chromosomes: miR 916-5p and mir 466i-5p.**
- **They look like the prevailing CALL and RET instructions in a real computer.**

miRNA in Human Chromosome 1



1	M26622	4044	>hsa-miR-619-5p MIMAT0026622 Homo sapiens	GCTGGGATTACAGGCATGAGCC
2	M17325	2768	>mmu-miR-466i-5p MIMAT0017325 Mus musculus	TGTGTGTGTGTGTGTGTGTG
3	M30415	115	>hsa-miR-1273h-5p MIMAT0030415 Homo sapiens	CTGGGAGGTCAAGGCTGCAGT
4	M16192	112	>ppy-miR-1268 MIMAT0016192 Pongo pygmaeus	CGGGCGTGGTGGTGGGCG
5	M23748	42	>cgr-miR-1285 MIMAT0023748 Cricetulus griseus	TGCACTCCAGCCTGGGCAACATAG
6	M03499	40	>efu-miR-9315 MIMAT0034991 Eptesicus fuscus	AAGGAATATGAATCCCATGGGT
7	M22286	36	>hsa-miR-5585-3p MIMAT0022286 Homo sapiens	CTGAATAGCTGGGACTACAGGT
8	M21038	35	>hsa-miR-548ap-3p MIMAT0021038 Homo sapiens	AAAAACCACAATTACTTTT
9	M33692	26	>hsa-miR-8485 MIMAT0033692 Homo sapiens	CACACACACACACACACGTAT
10	M32114	25	>pmi-miR-22-3p MIMAT0032144 Patiria miniata	TCAGCTGCCCGGTGAAGTGTAG
11	M22264	23	>hsa-miR-548aq-3p MIMAT0022264 Homo sapiens	CAAAAAGTCAATTACTTTTGC
12	M09447	19	>hsa-miR-1972 MIMAT0009447 Homo sapiens	TCAGGCCAGGCACAGTGGCTCA
13	M24377	19	>ppy-miR-548h MIMAT0024377 Pongo pygmaeus	AAAAGTAATTGCGGTTTTTG
14	M22723	18	>hsa-miR-548h-3p MIMAT0022723 Homo sapiens	CAAAAACCGCAATTACTTTTGCA
15	M22722	14	>hsa-miR-548g-5p MIMAT0022722 Homo sapiens	TGCAAAAAGTAATTGCAGTTTTTG
16	M34814	14	>efu-miR-9221 MIMAT0034814 Eptesicus fuscus	TCCAAGATCAAGGTGCTGGCA



No miRNA in Bacteria

- **No miRNA was found in single cell bacteria like ruddii, nasuia, and ecoli.**
- **If it is true, my hypothesis of cell computers based on miRNAs as tokens would fail.**



20-b Patterns in Bacterial Genomes

- **There are 156 recurring 20 base patterns in nasuia genome.**
- **There are 498 recurring 20 base patterns in ruddii genome.**



Tools Used

- **F# eForth under Windows 7**
- **Microsoft Words**
- **Microsoft Excel**
- **Microsoft Notepad**
- **Python**



Tools Used

- **F# works well for files up to 300MB. It cannot read bigger files.**
- **All genome files are smaller than 300MB, except the RNA files in human and mouse genome.**
- **I split big files into 200MB pieces with Python. Then F# works fine.**
- **Python is too slow to do anything significant.**

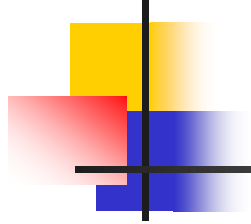


Conclusion

- **I believe each cell is a computer.**
- **I believe the genome contains the complete code of the cell computer.**
- **I believe mRNAs are primitive code to produce proteins.**
- **I believe miRNAs are high level code to activate mRNAs and miRNAs.**



Questions?



Thank You!