Bioinformatics

Bits and Pieces

Sami Khuri
Department of Computer Science
San José State University
San José, California, USA
sami.khuri@sjsu.edu
www.cs.sjsu.edu/faculty/khuri

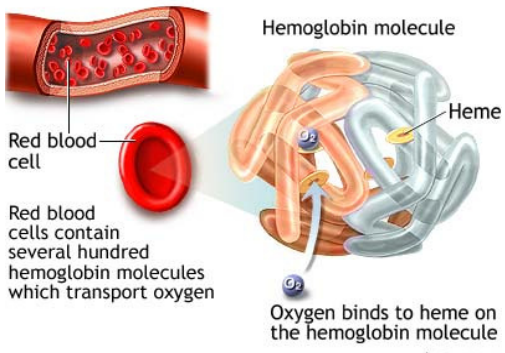
©2013 Sami Khuri

Outline

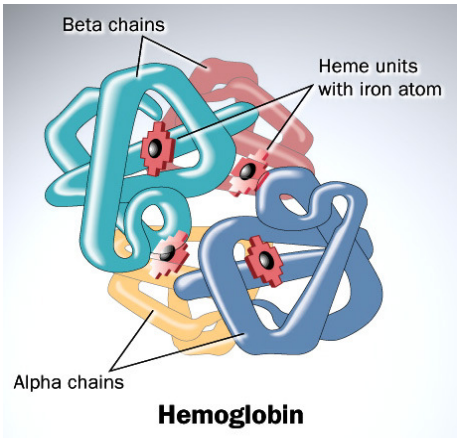
- Hemoglobin
- Hemoglobinopathies
  - Beta Thalassemias
- Micro RNA
- Antagomirs
- Biomarkers
- Direct-to-Consumer
- Lesson Learned [Genome & Biology]

©2013 Sami Khuri

Red Blood Cells

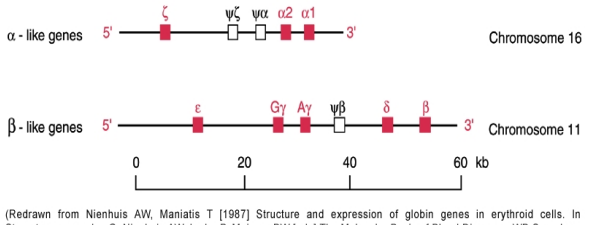


©2013 Sami Khuri



©2013 Sami Khuri

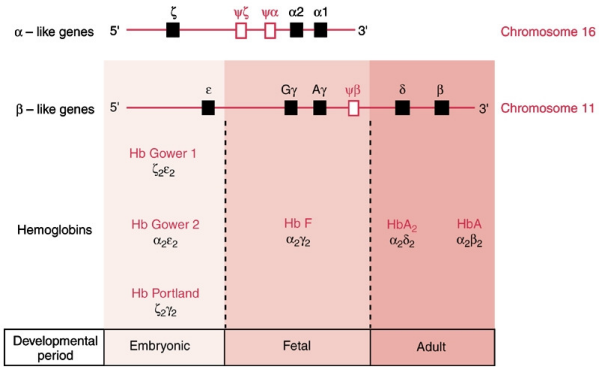
α-like and β-like Globin Genes



(Redrawn from Nienhuis AW, Maniatis T [1987] Structure and expression of globin genes in erythroid cells. In Stamatoyannopoulos G, Nienhuis AW, Leder P, Majerus PW [eds] The Molecular Basis of Blood Diseases. WB Saunders, Philadelphia, pp. 28–65.)

© 2004 Elsevier Inc. All rights reserved.

©2013 Sami Khuri

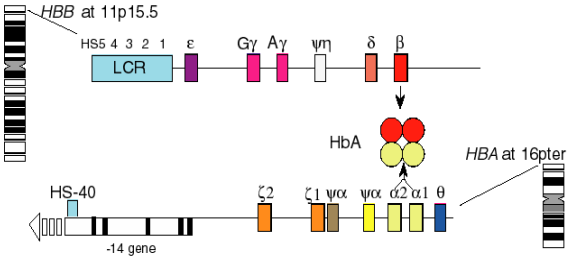


(Redrawn from Stamatoyannopoulos G, Nienhuis AW [1987] Hemoglobin switching. In Stamatoyannopoulos G, Nienhuis AW, Leder P, Majerus PW (eds) The Molecular Basis of Blood Diseases. WB Saunders, Philadelphia, pp. 66–105.)

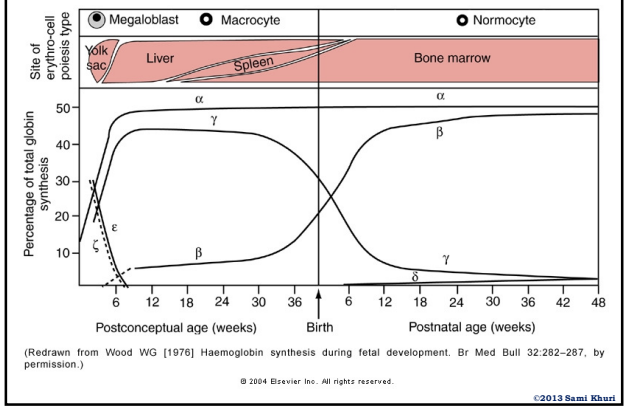
© 2004 Elsevier Inc. All rights reserved.

©2013 Sami Khuri

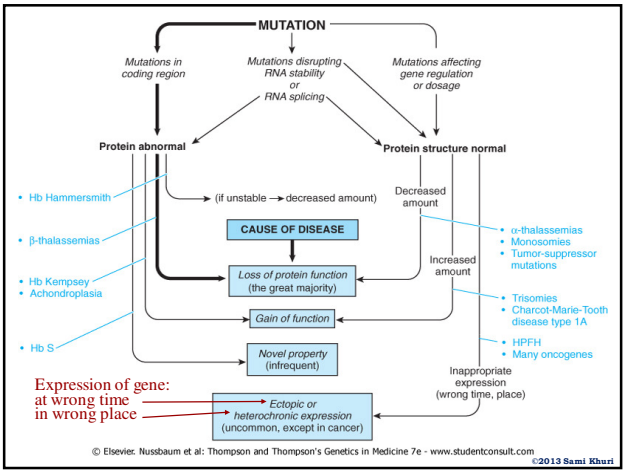
The Human α-Globin and β-Globin Gene Families



©2013 Sami Khuri



©2013 Sami Khuri



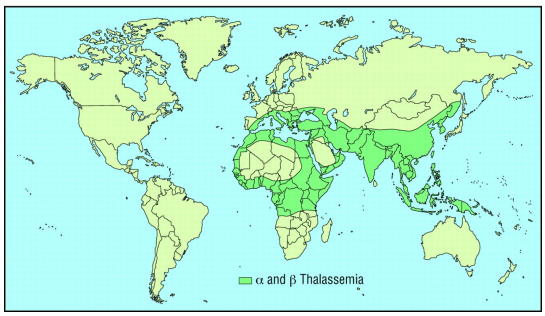
©2013 Sami Khuri

Hemoglobinopathies

- Hemoglobinopathies are the most common inherited disorders in humans, resulting from mutations in the α globin and β globin gene clusters.
- Molecular defects in either regulatory or coding regions of the human α globin, or β globin genes can minimally or drastically reduce their expression, leading to α thalassemia or β thalassemia, respectively.

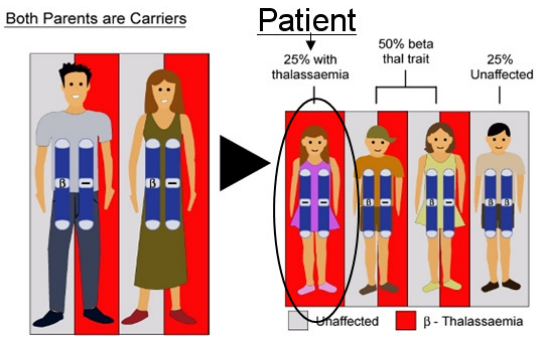
©2013 Sami Khuri

Where is Thalassemia endemic?



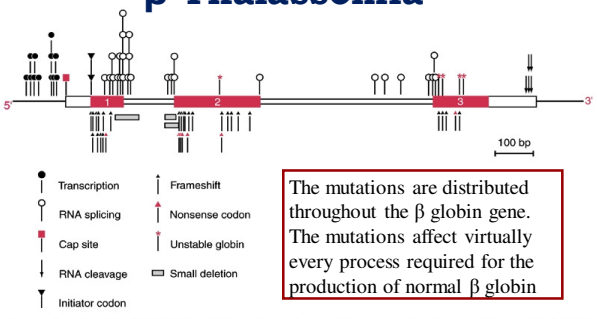
©2013 Sami Khuri

β Thalassemia from Parents



©2013 Sami Khuri

Point Mutations Causing β-Thalassemia



The mutations are distributed throughout the β globin gene. The mutations affect virtually every process required for the production of normal β globin

(Redrawn from Kazazian HH [1990] The thalassemia syndromes: Molecular basis and prenatal diagnosis in 1990. In Miescher LA, Jaffé ER [eds] Seminars in Hematology 27. WB Saunders, Philadelphia, pp. 209–228.)

© 2004 Elsevier Inc. All rights reserved.

©2013 Sami Khuri

Molecular Defects in β Thalassemia (I)

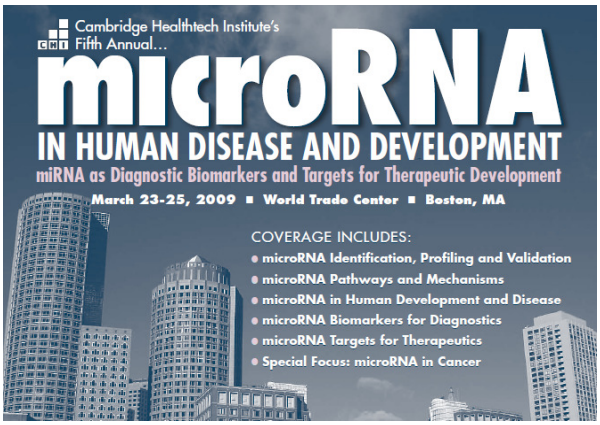
- Large Deletions and Insertions
- Non-deletional forms of β thalassemia
- β-globin Mutations Affecting Transcription
  - Promoter Mutations
    - TATA-Box, CAT-Box, CACCC-TFBS
  - Mutations of the 5' UTR
- β-globin Mutations Affecting mRNA Processing
  - Junctional mutations [GT – AG]
  - Consensus-sequence mutations
  - Cryptic splice-site mutations in introns
  - Cryptic splice-site mutations in exons
  - 3'UTR and Polyadenelation site mutations

©2013 Sami Khuri

Molecular Defects in β Thalassemia (II)

- β-Globin Mutations Affecting mRNA Translation
  - Start and Stop codon mutations
    - Several mutations of AUG have been found all of which produce beta-zero thalassemia.
  - Missense (frameshift) and nonsense mutations
    - Around half of the β thalassemia alleles are characterized by premature β-chain termination (mainly in exons 1 and 2), produced by frameshift or nonsense mutations.
    - **Examples**:
    - **Nonsense mutation**:
      - cd39 CAG → TAG (Mediterranean).
    - **Frameshift mutation**:
      - cd17 AAG → TAG (Chinese, Japanese).

©2013 Sami Khuri



©2013 Sami Khuri

Micro RNAs

- **Micro RNAs** are a class of non-coding RNA gene whose products are nucleotide sequences (about 22 nucleotides long) that play important roles in regulation of translation and degradation of mRNAs through base pairing of partially complementary sites in the untranslated regions (UTRs) of the message.
- **miRNAs** are a class of small, evolutionarily conserved RNA molecules that regulate gene expression at the post-transcriptional level.

©2013 Sami Khuri

First Micro RNAs

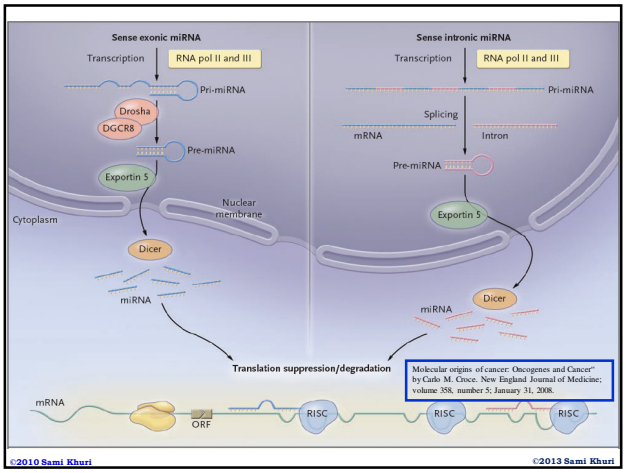
- Micro RNAs (miRNA) were first discovered by Chalfie et al. through genetic studies in the nematode *Caenorhabditis elegans* as essential regulators of development.
  - *lin-4* and *let-7* seemed to be involved in controlling the timing of larval development
- Since then, numerous microRNAs have been found in different species:
  - miRBase (release 13.0) contains 9,499 microRNA entries from 103 species, among which 706 are human microRNAs.
  - many microRNA gene families are conserved among diverse species.

©2013 Sami Khuri

Producing Micro RNAs

- MiRNAs gene encode precursor RNAs that undergo processing to form miRNAs of length approximately 22 nucleotides.

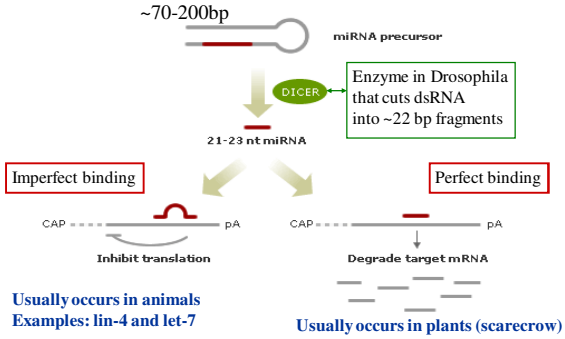
©2013 Sami Khuri



©2010 Sami Khuri

©2013 Sami Khuri

MicroRNA Binding



Usually occurs in animals
Examples: lin-4 and let-7

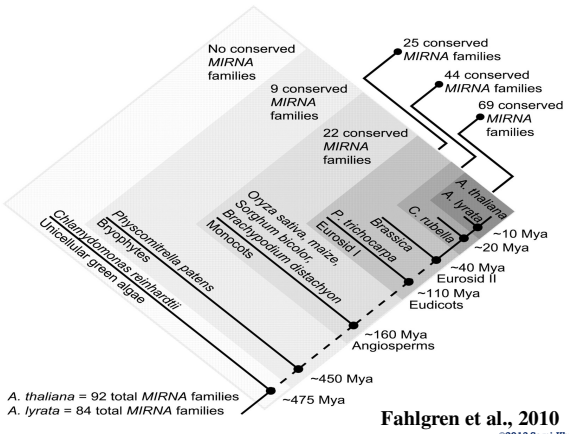
Usually occurs in plants (scarecrow)

©2013 Sami Khuri

Animal versus Plant Targets

- In plants:
  - microRNAs bind almost perfectly to their target mRNAs
  - targets have been found anywhere on the mRNA
  - relatively few targets because microRNA-mRNA binding requires near-perfect complementarity
- In animals:
  - partial base-pairings with the target mRNAs
  - targets are typically found in the 3'-UTR, where the silencing machinery can easily interact with the initiation complex.
  - multiple targets on the same mRNA and often multiple microRNAs target the same mRNA

©2013 Sami Khuri



Fahlgren et al., 2010

©2013 Sami Khuri

miRNA Challenges and Hope

- The challenges are:
  - Predict the **functions** of the miRNAs
  - Identify the potential **target mRNAs** to which miRNAs will bind
  - Characterize the consequences of their **regulatory interactions**.
- The hope is:
  - RNA interference will be used to inactivate tumor genes or viruses.
    - miRNA-based therapies are under investigation

©2013 Sami Khuri

MicroRNAs and Cancer (I)

- More recently, in the past few years, it has been discovered that some of the 250 to 300 human **miRNA** are linked to cancers, such as leukemia, lung, breast, and colon cancers.
- Mapping of numerous **miRNA** genes has shown that many occur in chromosomal regions that undergo rearrangements, deletions, and amplifications in cancer cells.

Molecular origins of cancer: Oncogenes and Cancer" by Carlo M. Croce. New England Journal of Medicine; volume 358, number 5; January 31, 2008.

©2013 Sami Khuri

MicroRNAs and Cancer (II)

MicroRNAs currently implicated in cancer

MicroRNA	Cancer Role	Cancer Type	Mechanism
miR-15	tumor suppressor	CLL	Bcl-2 inhibition
miR-16	tumor suppressor	CLL	Bcl-2 inhibition
miR-155	oncogene	lymphomas	unknown
let-7	tumor suppressor	lung cancer	ras inhibition
miR-17-92 cluster	oncogene	B cell lymphoma	unknown
miR-372	oncogene	testicular	inhibit p53 pathway
miR-373	oncogene	testicular	inhibit p53 pathway

Chronic Lymphocytic Leukemia (CLL): disease of white blood cells that won't die. It is the most common leukemia.
A **miRNA** can be a tumor suppressor if in a given cell type its target is an oncogene. It can be an oncogene if in a different cell type its target is a tumor-suppressor gene.

Table from "No miR Hype: MicroRNA's Cancer Role Expands" by Ken Garber Journal of the National Cancer Institute, Vol. 98, No. 13, July 5, 2006

©2013 Sami Khuri

miR-15 and miR-16 in CLL

MicroRNAs currently implicated in cancer

MicroRNA	Cancer Role	Cancer Type	Mechanism
miR-15	tumor suppressor	CLL	Bcl-2 inhibition
miR-16	tumor suppressor	CLL	Bcl-2 inhibition
miR-155	oncogene	lymphomas	unknown
let-7	tumor suppressor	lung cancer	ras inhibition
miR-17-92 cluster	oncogene	B cell lymphoma	unknown
miR-372	oncogene	testicular	inhibit p53 pathway
miR-373	oncogene	testicular	inhibit p53 pathway

miR-15 and **miR-16** induce apoptosis by targeting the key survival protein Bcl-2, which is overexpressed in CLL

Table from "No miR Hype: MicroRNA's Cancer Role Expands" by Ken Garber Journal of the National Cancer Institute, Vol. 98, No. 13, July 5, 2006

©2013 Sami Khuri

MiRNAs and Treatment

- Examples of the role **miRNA** plays in cancer pathophysiology involve *miR-15a* and *miR-16-1*, which are deleted or down-regulated in most indolent (slow to develop) cases of chronic lymphocytic leukemia.
- The discovery of the involvement of **miRNAs** in the initiation and progression of human cancer may provide additional targets for anticancer treatments.

©2013 Sami Khuri

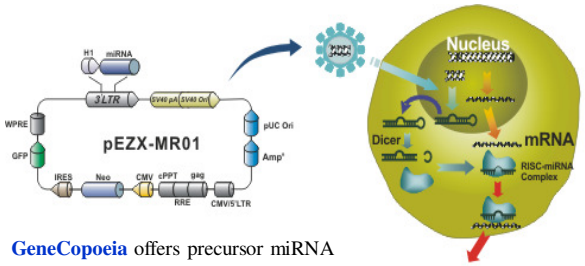
Antagomirs

- Chemically modified antisense oligonucleotides (i.e., short strings of DNA bases complementary in sequence to their targets) injected into mice potentially silenced a target miRNA in the liver.
- The oligonucleotides were dubbed **antagomirs**.
- It is believed that **antagomirs** should be more effective against cancer-causing miRNAs than classic antisense therapy has been against protein-coding mRNAs:
  - **antagomirs** compete with miRNA targets for binding. An easier task than interfering with the protein translation machinery, which is the classic antisense mechanism.

Silencing of microRNAs in vivo with 'antagomirs by Jan Krützfeldt et al. Nature 438, 685-689 (Dec 2005).

©2013 Sami Khuri

GeneCopoeia



GeneCopoeia offers precursor miRNA expression constructs in a feline immunodeficiency virus (FIV) based lentiviral vector system

www.genecopoeia.com/product/mirna

©2013 Sami Khuri

Urine Biomarkers for Kawasaki Disease

- **Boston Children's Mass Spec Study Yields Urine Biomarkers to Diagnose Kawasaki Disease**
- Currently KD is diagnosed by ruling out other disorders using a variety of clinical measures. But this can delay definitive diagnosis. Though the disease is highly curable with early treatment using aspirin and gammaglobulin, it can potentially lead to severe cardiovascular complications if not caught early.

www.genomeweb.com/proteomics/boston-childrens-mass-spec-study-yields-urine-biomarkers-diagnose-kawasaki-disea

©2013 Sami Khuri

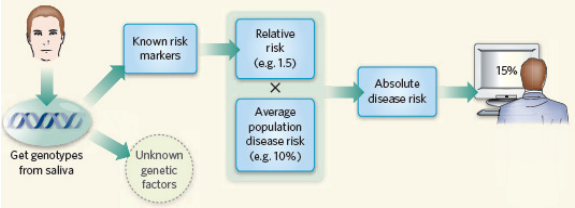
A Step Toward a Universal Cancer Blood Test

A new cancer blood test looks for abnormal chromosomes, such as the rearrangements in these breast cancer cells (different colors on the same chromosome)
Image: Mira Grigorova and Paul Edwards/University of Cambridge

www.wired.com/wiredscience/2012/11/universal-cancer-test/

©2013 Sami Khuri

Direct-to-Consumer Disease Risk

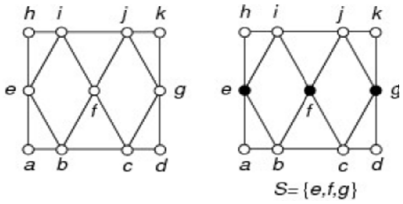


- More than 1,000 DNA variants associated with diseases and traits have been identified.
- Direct-to-consumer (DTC) companies are harnessing these discoveries by offering DNA tests that provide insights

"An agenda for personalized medicine" by P. Ng et al., Nature, October 2009

©2013 Sami Khuri

Minimum Dominating Set Problem



A **dominating set** is a set of nodes S such that every node in the network graph G is a neighbor of at least one element of S. The **Minimum Dominating Set** (MDS) problem is to find a minimum such S for a given network graph.

www.cs.iastate.edu/~chaudhur/cs611/Sp07/notes/lec22.pdf

©2013 Sami Khuri

Lessons Learned: The Genome

- Like living genomes, our understanding of a genome sequence changes and evolves over time.
- By their very nature, genome sequence databases are dynamic and frequently updated.
  - An unknown gene may be annotated and categorized the next time we query the database.

@2002-09 Sami Khuri

©2013 Sami Khuri

Lessons Learned: Biology & Science

- You should be aware, however, that there are probably exceptions to every general statement we have made in this chapter!

  [Understanding Bioinformatics by M. Zvelebil, 2008]
- People outside scientific fields often think of science as a list of facts to be memorized rather than a series of questions to be asked.

  [Discovering Genomes by A. Campbell]

©2013 Sami Khuri