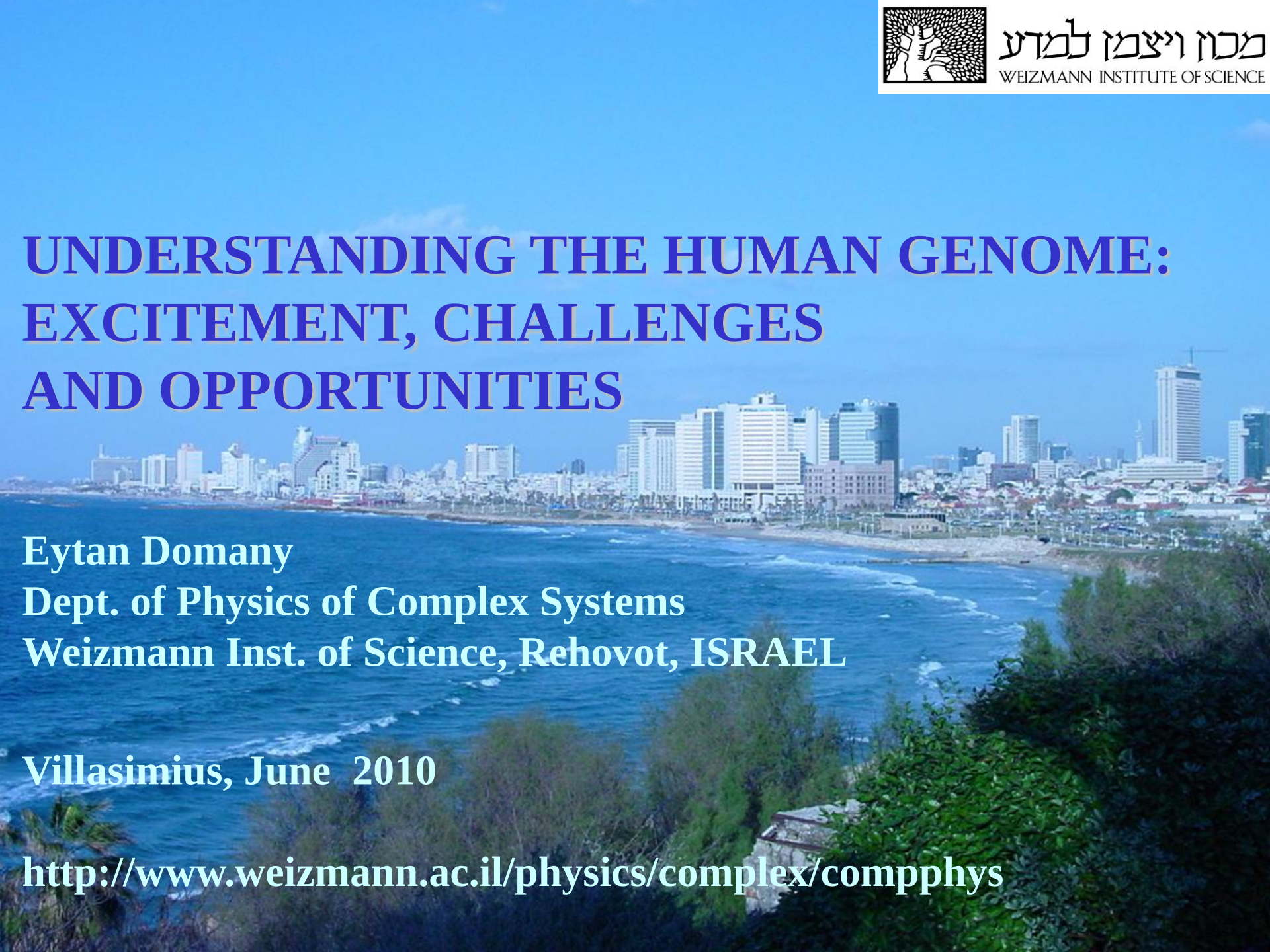


מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

UNDERSTANDING THE HUMAN GENOME: EXCITEMENT, CHALLENGES AND OPPORTUNITIES

Eytan Domany
Dept. of Physics of Complex Systems
Weizmann Inst. of Science, Rehovot, ISRAEL

Villasimius, June 2010

http://www.weizmann.ac.il/physics/complex/compphys

OUTLINE:

1. THESE ARE EXCITING TIMES IN BIOLOGY:
   NEW FRONTIERS, NEW TECHNOLOGIES,
   CENTRAL ROLE FOR COMPUTATIONAL SCIENCE

2. a. INTRODUCTION TO MOLECULAR BIOLOGY OF THE CELL: GENES, GENE EXPRESSION AND ITS MEASUREMENT BY MICROARRAYS
   b. WHAT CONTROLS GENE EXPRESSION?

3. a. THE HALLMARKS OF CANCER;
   b. RESPONSE TO STIMULUS BY A GROWTH FACTOR

4. THE “EPIGENETIC CODE” AND REGULATION OF TRANSCRIPTION

1. THESE ARE ***VERY*** EXCITING TIMES IN BIOLOGY!

articles

Initial sequencing and analysis of the human genome

International Human Genome Sequencing Consortium*

*A partial list of authors appears on the opposite page. Affiliations are listed at the end of the paper.

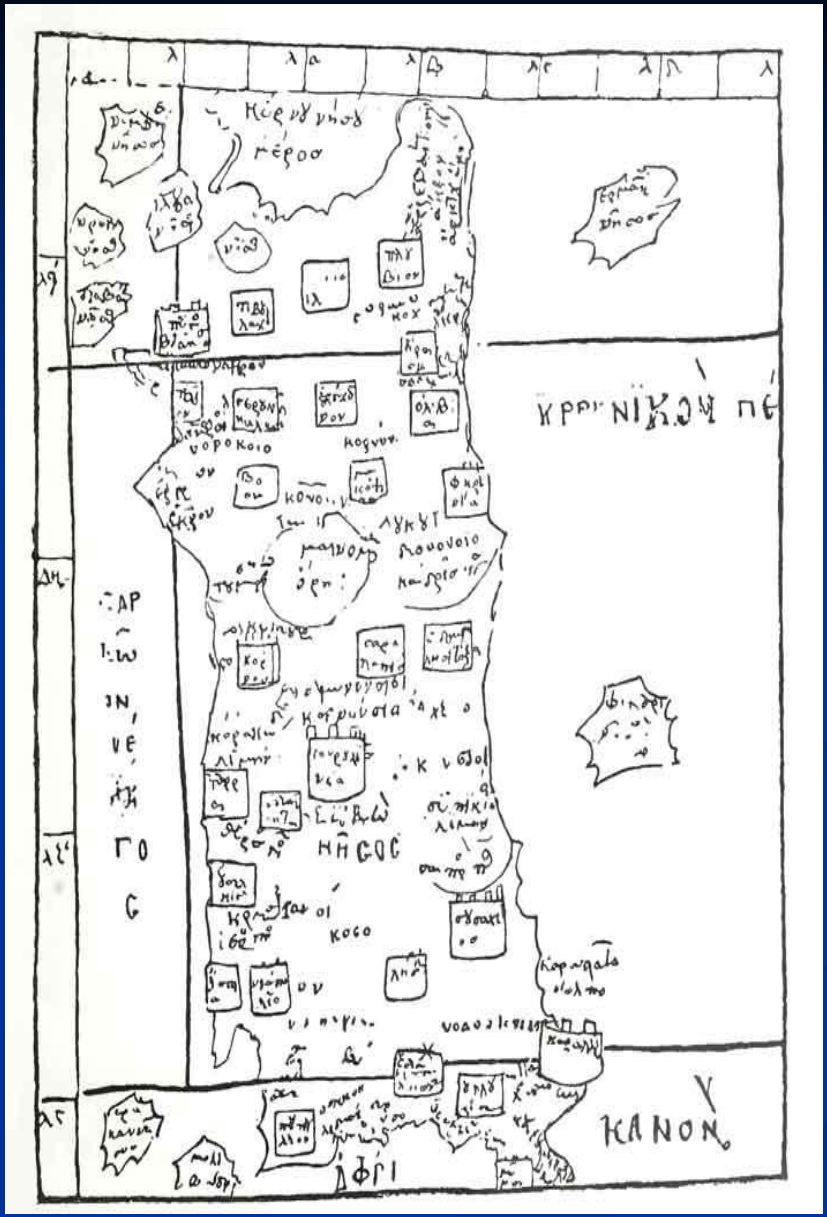
The human genome holds an extraordinary trove of information about human development, physiology, medicine and evolution. Here we report the results of an international collaboration to produce and make freely available a draft sequence of the human genome. We also present an initial analysis of the data, describing some of the insights that can be gleaned from the sequence.

Lander et al, Nature Feb 15, 2001

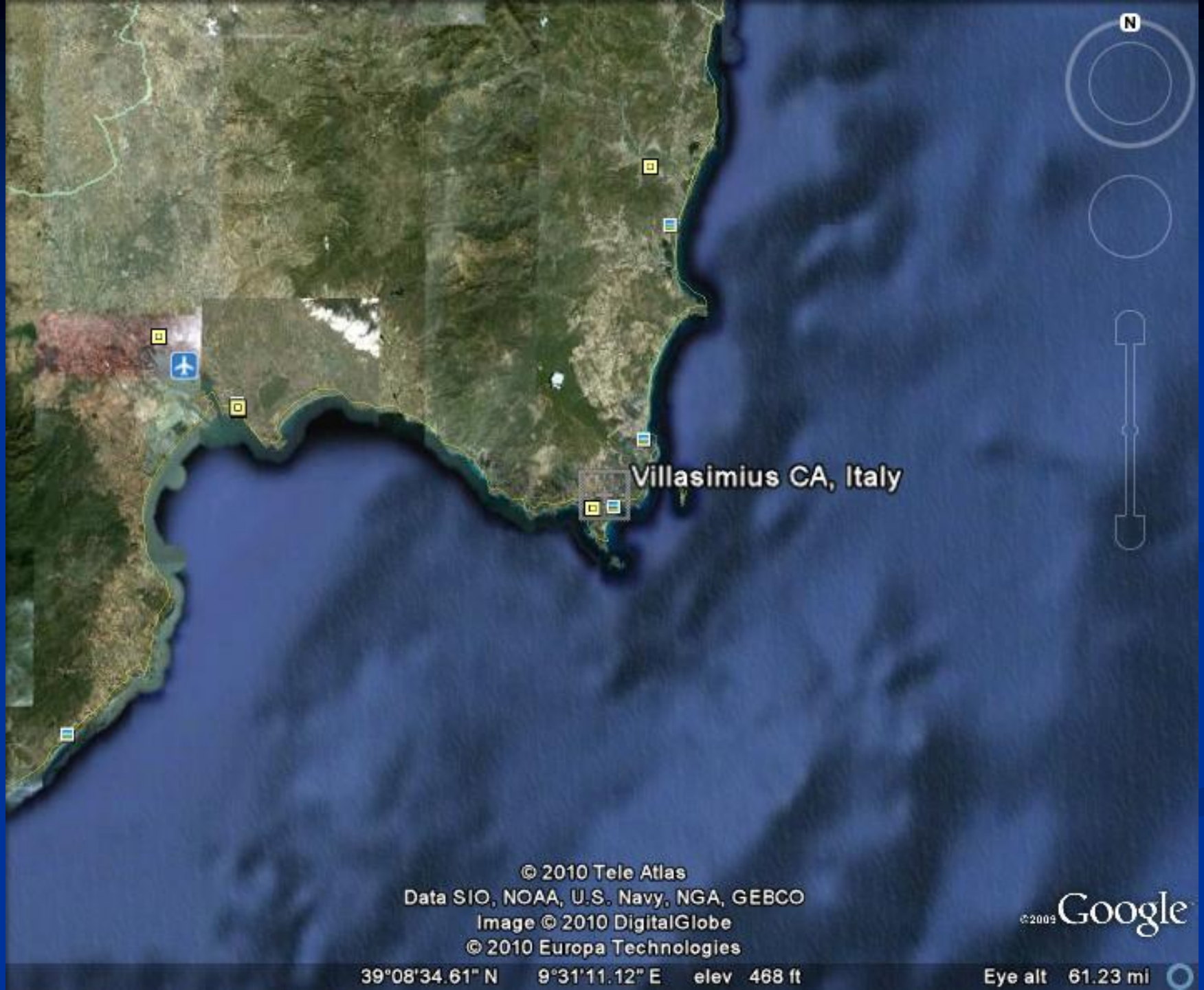
SEQUENCING (A SINGLE) HUMAN GENOME: 15 YEARS, 3 BILLION $, THOUSANDS OF MAN-YEARS

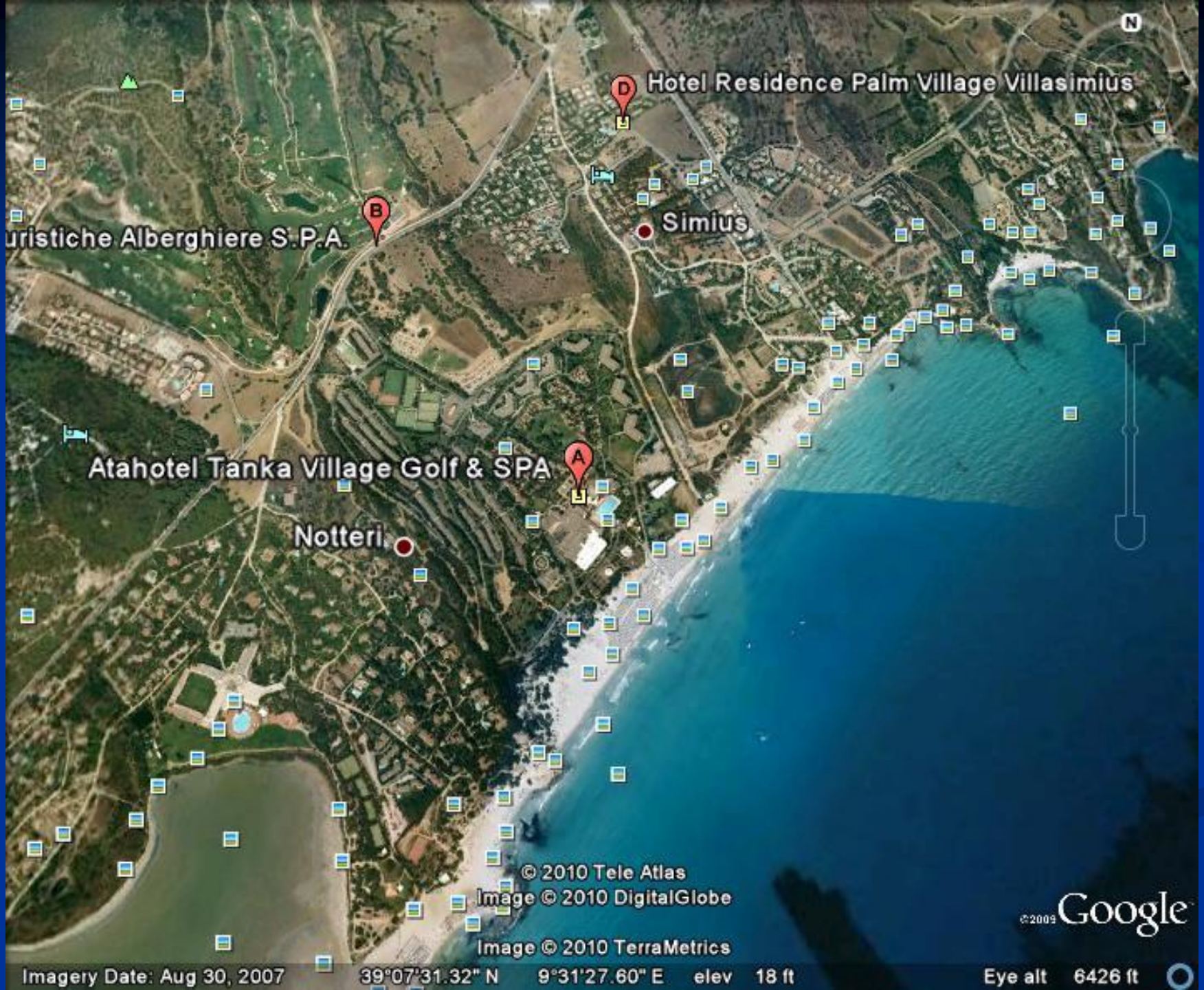
NEXT GENERATION SEQUENCING: TODAY'S COST 50,000$
IN 2 YEARS COST WILL COME DOWN TO 1000$

LANDER COMPARES THE REVOLUTION WE ARE WITNESSING TO THE TRANSITION IN CARTOGRAPHY FROM MEDIEVAL MAPS TO GOOGLE EARTH; IN CHEMISTRY TO BEFORE VS AFTER MENDELEYEV



12th century copy of Ptolemeyan map of Sardinia


N
Villasimius CA, Italy
© 2010 Tele Atlas
Data SIO, NOAA, U.S. Navy, NGA, GEBCO
Image © 2010 DigitalGlobe
© 2010 Europa Technologies
©2009 Google
39°08'34.61" N 9°31'11.12" E elev 468 ft
Eye alt 61.23 mi


Hotel Residence Palm Village Villasimius
D
B
uristiche Alberghiere S.P.A.
Simius
Atahotel Tanka Village Golf & SPA
A
Notteri
© 2010 Tele Atlas
Image © 2010 DigitalGlobe
Image © 2010 TerraMetrics
©2009 Google
Imagery Date: Aug 30, 2007
39°07'31.32" N
9°31'27.60" E
elev 18 ft
Eye alt 6426 ft


lage Golf & SPA
A
N
© 2010 Tele Atlas
©2009 Google
Imagery Date: Aug 30, 2007
39°07'28.99" N
9°31'30.64" E
elev 13 ft
Eye alt 715 ft

KNOWLEDGE OF THE HUMAN GENOME

1. OPENED NEW HORIZONS IN BIOLOGICAL DATA ACQUISITION, FOLLOWED BY SUPER-EXPONENTIAL PROGRESS IN DEVELOPMENT OF NEW TECHNOLOGIES

2. IDENTIFIED PROTEIN-CODING GENES (SURPRISINGLY FEW ~20,000); REVEALED NEW REGULATORY REGIONS AND MECHANISMS

3. COMPARATIVE GENOMICS: HUMANS VS MOUSE, RAT, ….. (CONSERVED ELEMENTS ARE IMPORTANT)

4. HUMAN-HUMAN DIFFERENCES – SNPs AND THEIR ASSOCIATION WITH DISEASE

5. SYSTEMATIC APPROACH TO DISCOVER & UNDERSTAND THE MOLECULAR MECHANISMS OF DISEASE (CANCER!)

MAJOR DISCOVERIES OF THE LAST DECADE:

1. THE RNA WORLD: microRNA, Long NonCoding RNA

2. CHROMATIN MODIFICATION PLAYS A CENTRAL REGULATORY ROLE; EPIGENETICS

3. MOLECULAR STRATIFICATION OF CANCER

MAJOR NEW PROJECTS:

1. THE CANCER GENOME ATLAS (TCGA)

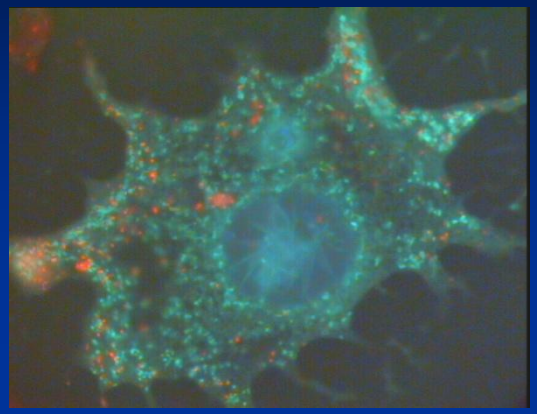
2. THE 1000 GENOME PROJECT

3. THE 10k VERTEBRATES GENOME PROJECT

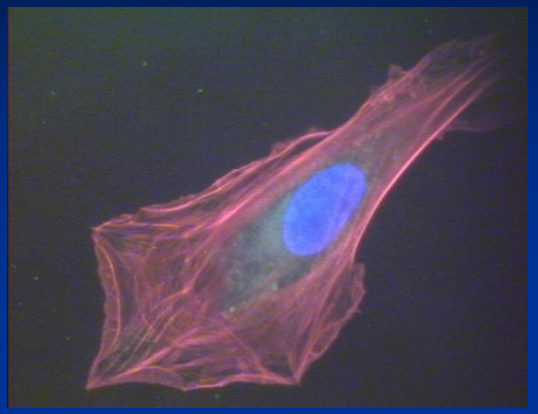
4. GENOME-WIDE ASSOCIATION STUDIES OF MANY DISEASES AND HUMAN TRAITS

2a. INTRODUCTION TO MOLECULAR BIOLOGY OF THE CELL

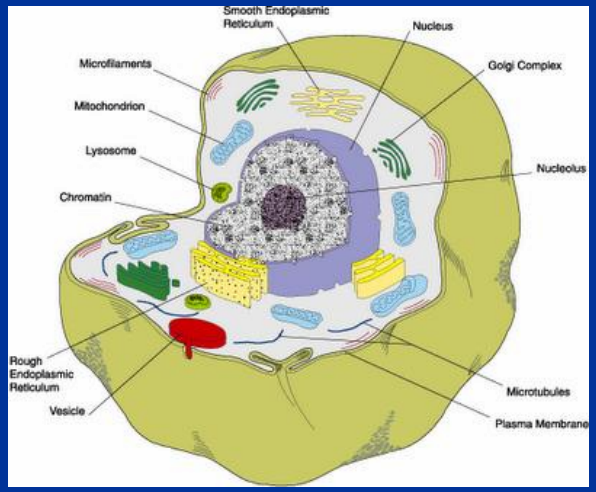
EUKARYOTIC CELLS:



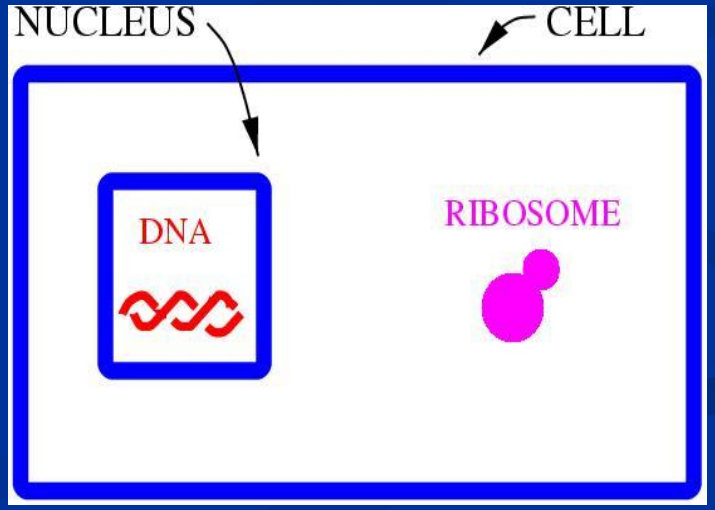
Chinese hamster ovary (CHO) cell
www.kent.edu/projects/cell/



Rat aortic smooth muscle cells
www.kent.edu/projects/cell/

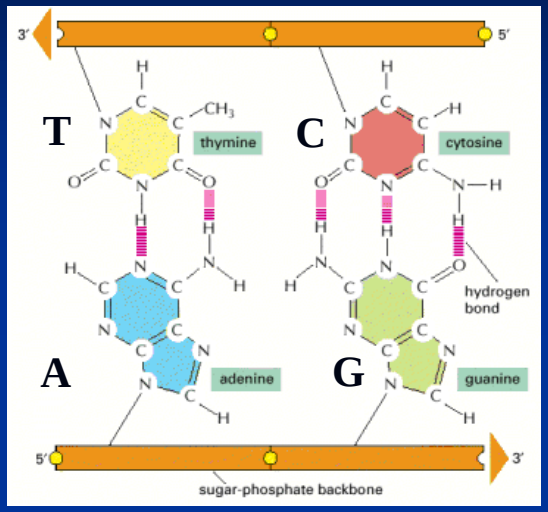


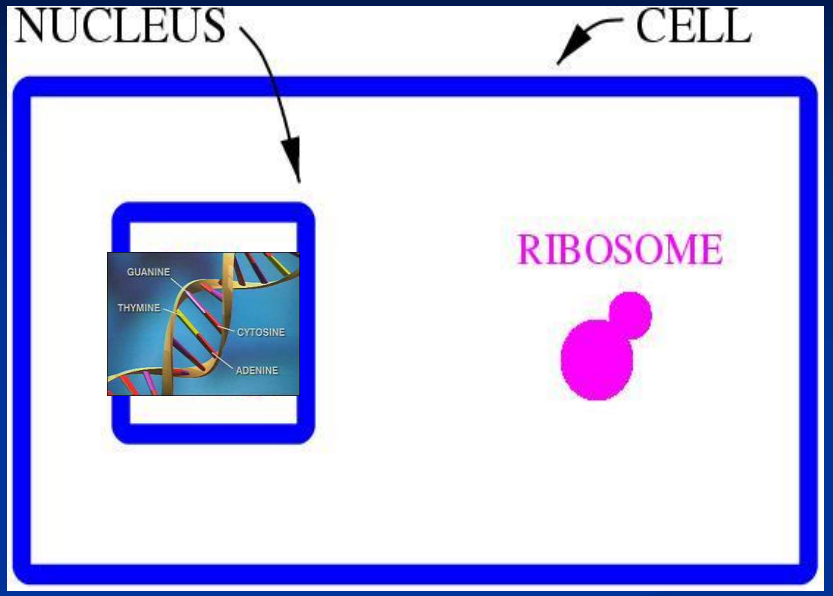
CARICATURE (FOR BIOLOGISTS)



CARICATURE (FOR PHYSICISTS)

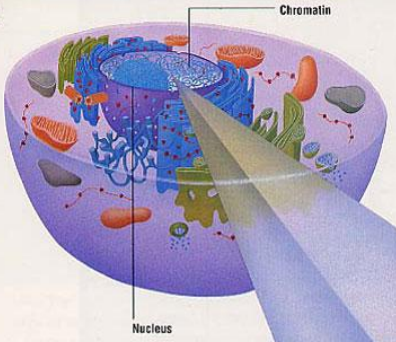
GENE EXPRESSION: OVERVIEW





INTERESTING COMPLEX MULTI-SCALE STRUCTURE

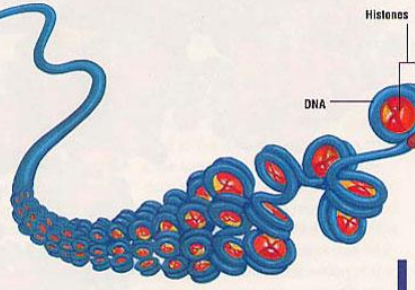
DNA – Structure: From chromosome to chromatin to solenoid



A THRONG OF CHROMATIN. Inside the cell nucleus, 46 strands of chromatin form a tangled mass, not unlike a bowl of microscopic spaghetti—except that each "noodle" has an ornate internal structure that can be fully appreciated only when seen greatly magnified.

A SINGLE STRAND. Shown alone for simplicity, this bit of chromatin bends around and back on itself many times inside the cell nucleus. In actuality, other chromatin strands weave—randomly, as far as anyone knows—between the folds of this one and others.

COILS WITHIN COILS. A close look at a strand of chromatin reveals it to be a coil created from a compactly folded strand of material, which is itself an even finer coil less than a millionth of an inch in diameter *(above, right)*.



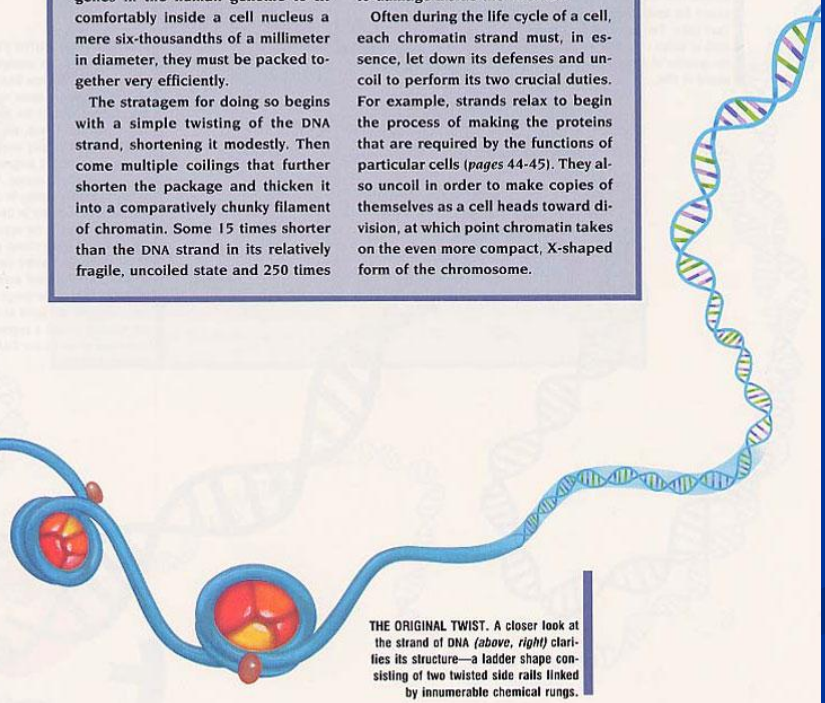
INSIDE THE SMALLEST COIL. The minuscule DNA filament from which chromatin is made can be seen here as it coils twice around a series of beadlike cores consisting of eight protein molecules called histones. Together with a single histone outside the core, these molecules exert the forces that coil the beads together.

COILED AND COMPACT FOR EFFICIENCY

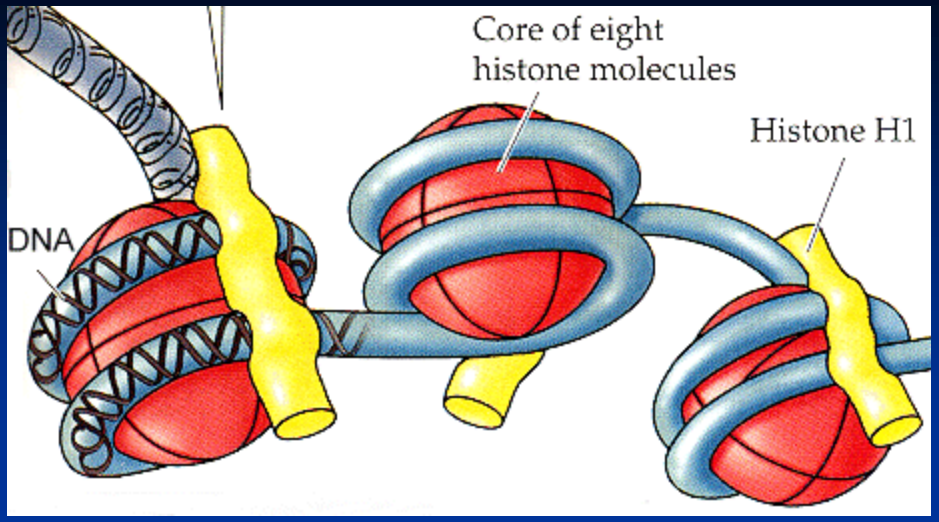
In order for the tens of thousands of genes in the human genome to fit comfortably inside a cell nucleus a mere six-thousandths of a millimeter in diameter, they must be packed together very efficiently.

The stratagem for doing so begins with a simple twisting of the DNA strand, shortening it modestly. Then come multiple coilings that further shorten the package and thicken it into a comparatively chunky filament of chromatin. Some 15 times shorter than the DNA strand in its relatively fragile, uncoiled state and 250 times thicker, chromatin is less susceptible to damage inside the nucleus.

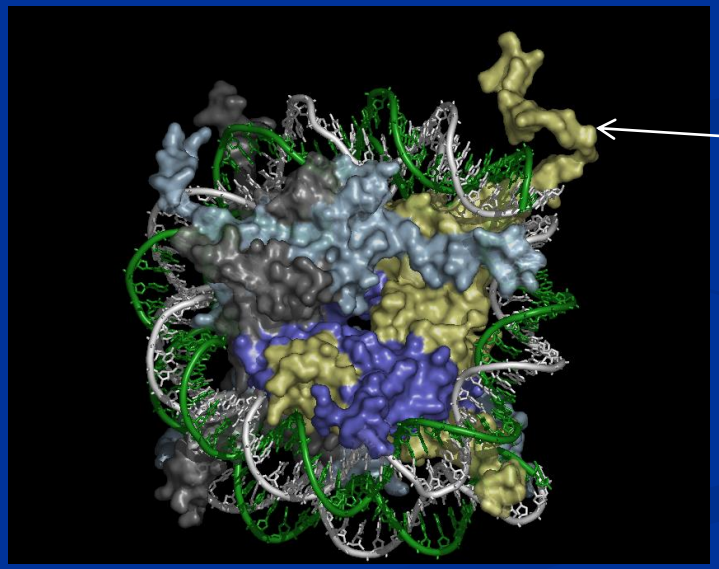
Often during the life cycle of a cell, each chromatin strand must, in essence, let down its defenses and uncoil to perform its two crucial duties. For example, strands relax to begin the process of making the proteins that are required by the functions of particular cells (*pages* 44-45). They also uncoil in order to make copies of themselves as a cell heads toward division, at which point chromatin takes on the even more compact, X-shaped form of the chromosome.



THE ORIGINAL TWIST. A closer look at the strand of DNA *(above, right)* clarifies its structure—a ladder shape consisting of two twisted side rails linked by innumerable chemical rungs.

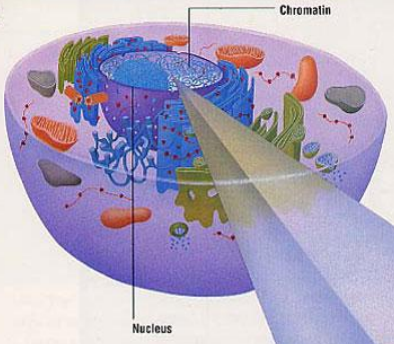


Physicist's nucleosome = basketball

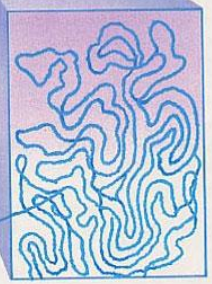


Biochemist's nucleosome

From chromosome to chromatin to solenoid to DNA

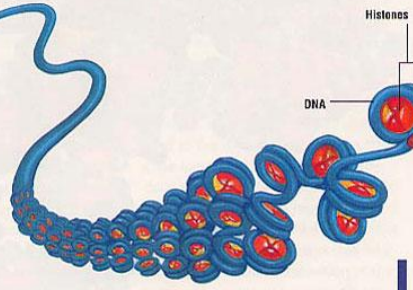


A THRONG OF CHROMATIN. Inside the cell nucleus, 46 strands of chromatin form a tangled mass, not unlike a bowl of microscopic spaghetti—except that each "noodle" has an ornate internal structure that can be fully appreciated only when seen greatly magnified.



A SINGLE STRAND. Shown alone for simplicity, this bit of chromatin bends around and back on itself many times inside the cell nucleus. In actuality, other chromatin strands weave—randomly, as far as anyone knows—between the folds of this one and others.

COILS WITHIN COILS. A close look at a strand of chromatin reveals it to be a coil created from a compactly folded strand of material, which is itself an even finer coil less than a millionth of an inch in diameter *(above, right)*.



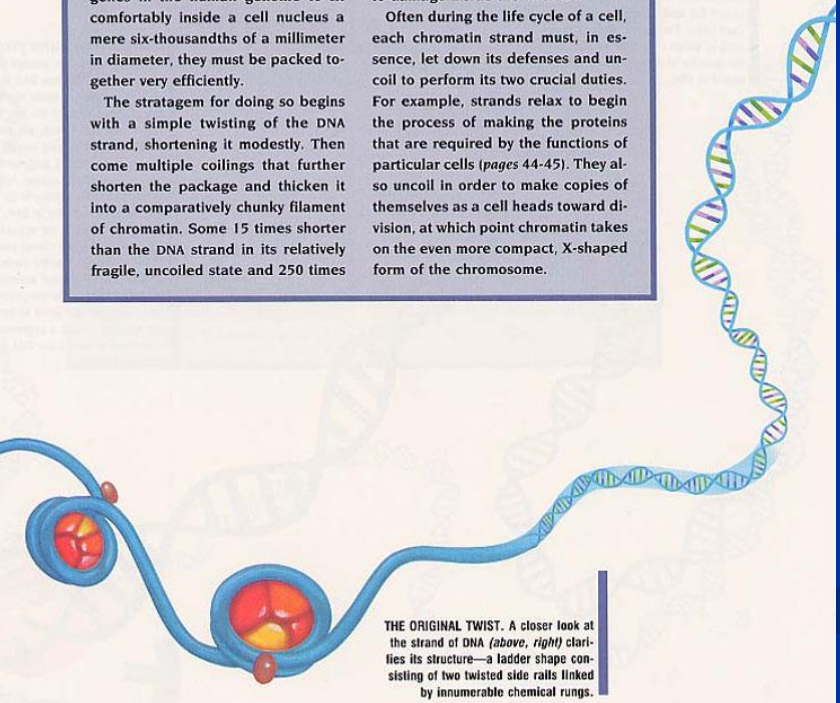
INSIDE THE SMALLEST COIL. The minuscule DNA filament from which chromatin is made can be seen here as it coils twice around a series of beadlike cores consisting of eight protein molecules called histones. Together with a single histone outside the core, these molecules exert the forces that coil the beads together.

COILED AND COMPACT FOR EFFICIENCY

In order for the tens of thousands of genes in the human genome to fit comfortably inside a cell nucleus a mere six-thousandths of a millimeter in diameter, they must be packed together very efficiently.

The stratagem for doing so begins with a simple twisting of the DNA strand, shortening it modestly. Then come multiple coilings that further shorten the package and thicken it into a comparatively chunky filament of chromatin. Some 15 times shorter than the DNA strand in its relatively fragile, uncoiled state and 250 times thicker, chromatin is less susceptible to damage inside the nucleus.

Often during the life cycle of a cell, each chromatin strand must, in essence, let down its defenses and uncoil to perform its two crucial duties. For example, strands relax to begin the process of making the proteins that are required by the functions of particular cells (*pages* 44-45). They also uncoil in order to make copies of themselves as a cell heads toward division, at which point chromatin takes on the even more compact, X-shaped form of the chromosome.

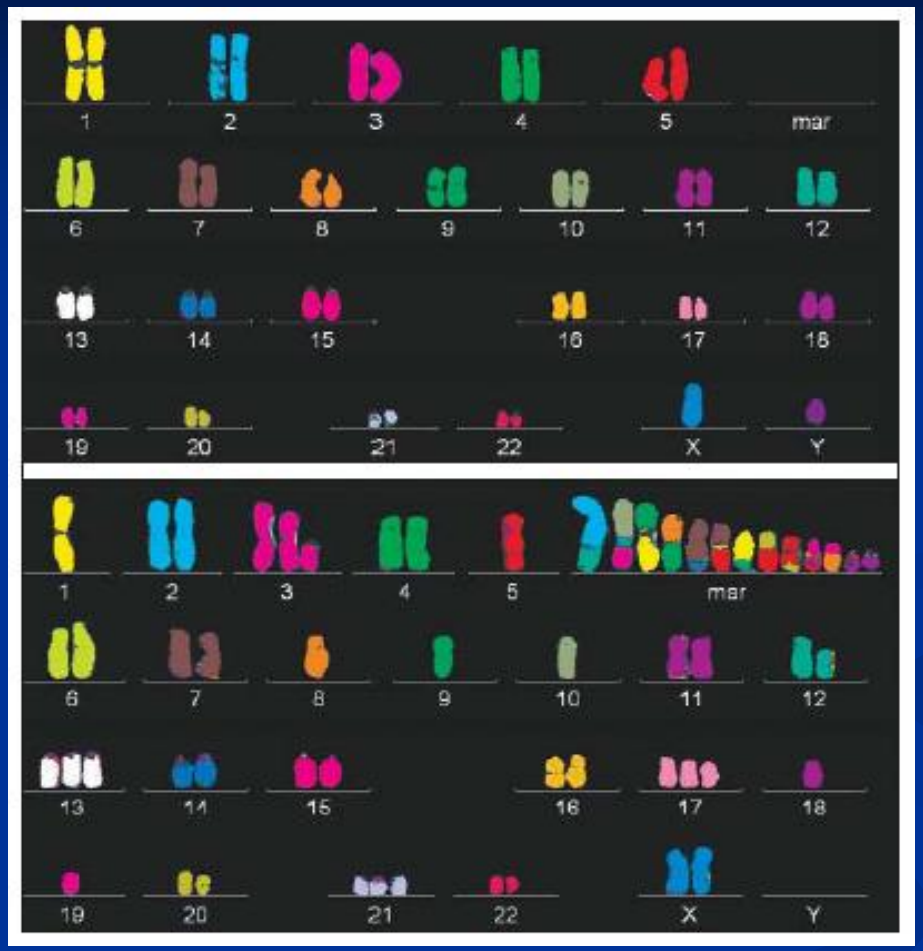


THE ORIGINAL TWIST. A closer look at the strand of DNA *(above, right)* clarifies its structure—a ladder shape consisting of two twisted side rails linked by innumerable chemical rungs.

HUMAN GENOME – 23 PAIRS OF CHROMOSOMES:

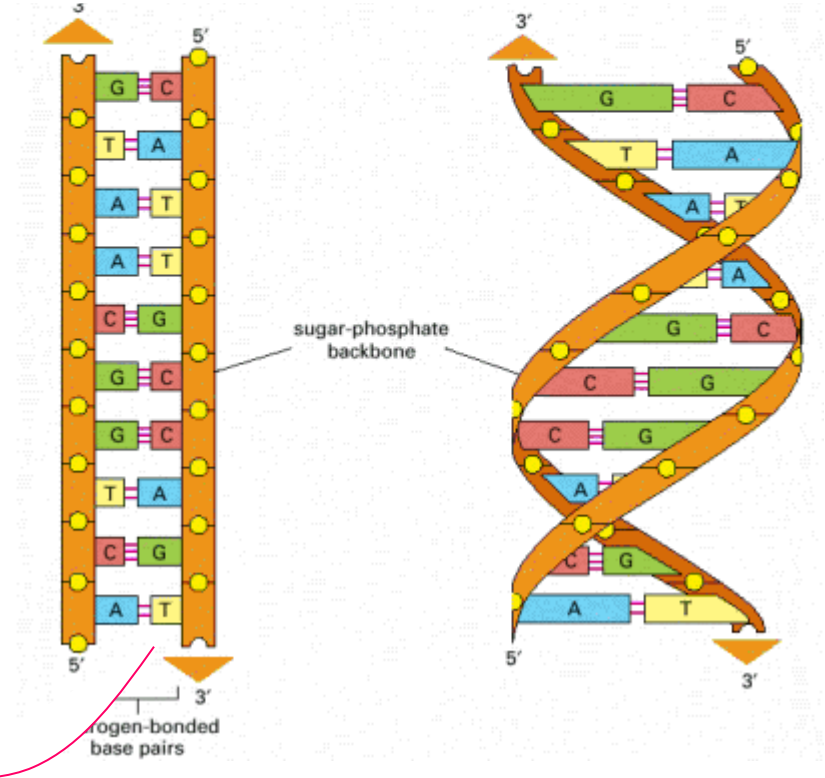
NORMAL:

CANCER:
(LEUKEMIA)



A CENTRAL THEME OF MY RESEARCH: DO CHROMOSOMAL ABERRATIONS PLAY A ***CAUSAL*** ROLE IN CANCER?

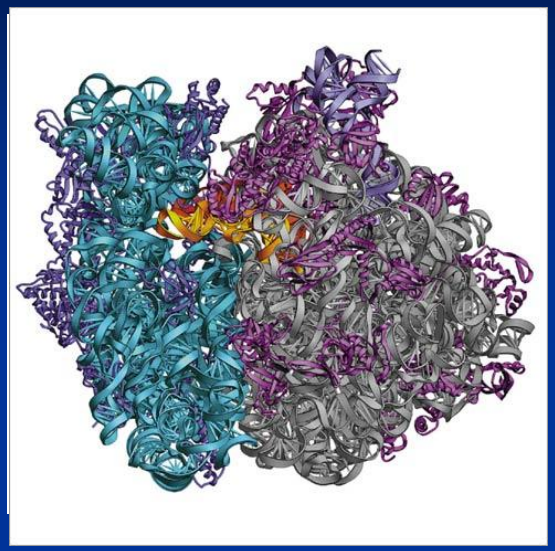
DNA CONTAINS INFORMATION:

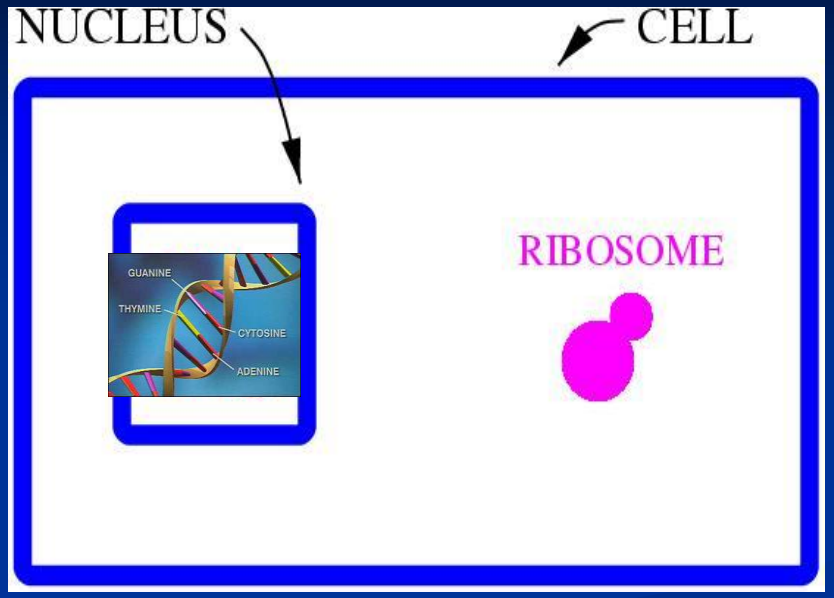


UNIVERSAL GENETIC CODE

TGACCGTTACATGGTGATGGTGCACCTGACTCCTGAGGAGAAGTCT

GENE EXPRESSION: OVERVIEW





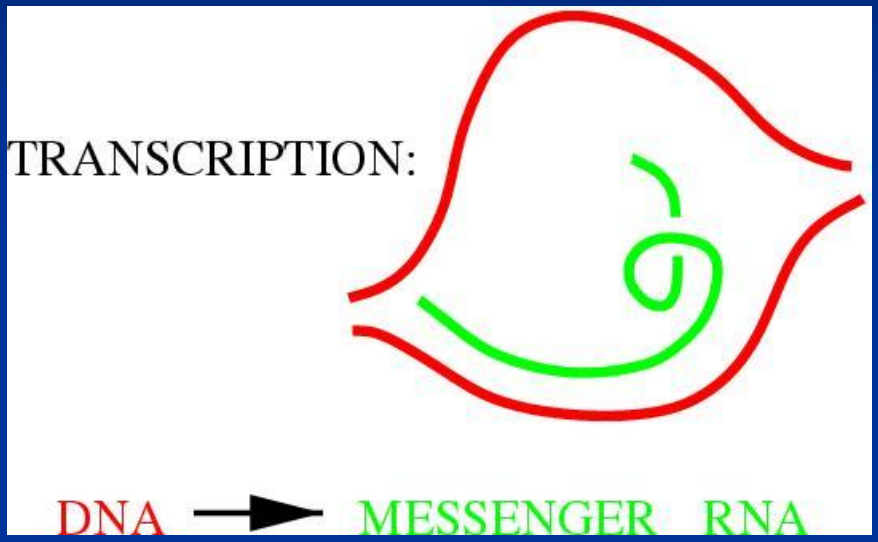
- GENE =SEGMENT OF **DNA** = BLUEPRINT FOR A PROTEIN (or RNA….)
- WHEN A GENE IS EXPRESSED, THE PROTEIN IT CODES FOR IS SYNTHESIZED
- EACH CELL CONTAINS ALL GENES !!!
- NOT ALL GENES ARE "EXPRESSED" (DIFFERENTIATION,TIME)

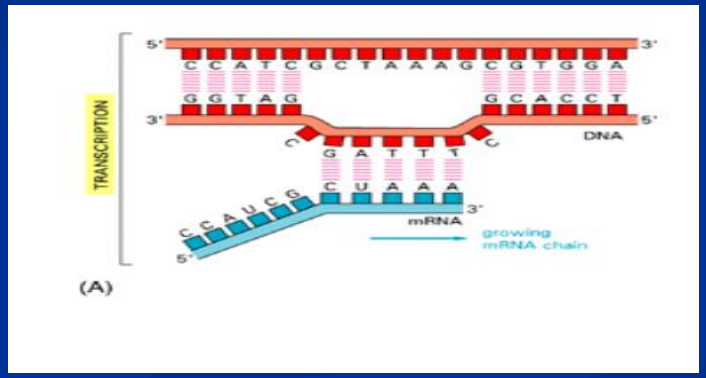
PROTEIN SYNTHESIS TAKES PLACE AT **RIBOSOMES**

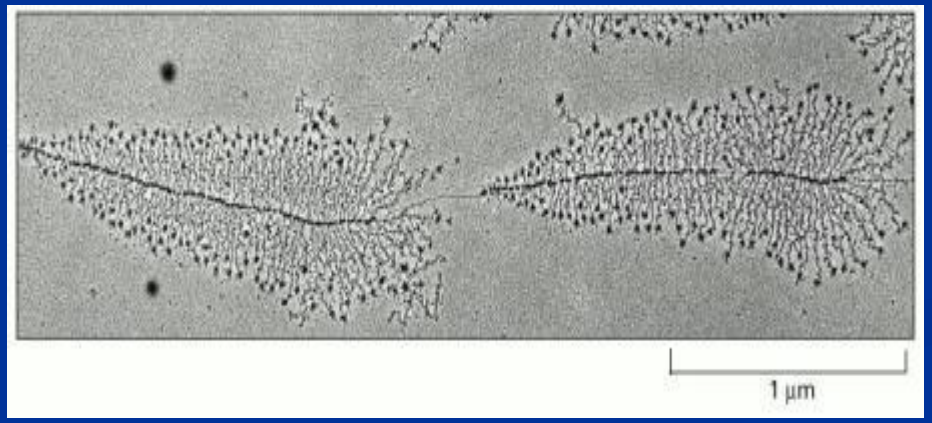
(…2009 Nobel Prize…) *LOGISTIC PROBLEM!*

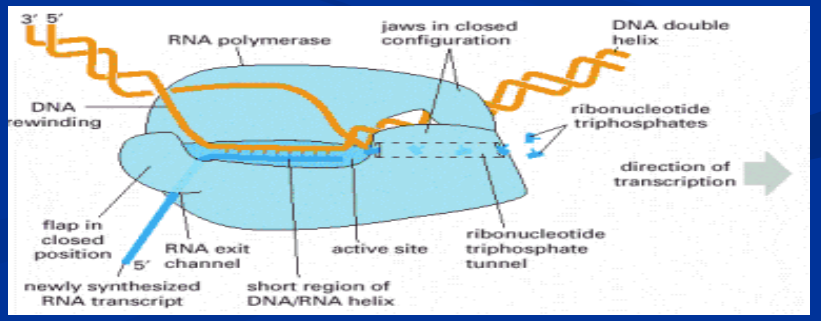
INFORMATION TRANSFER FROM NUCLEUS TO RIBOSOME

TWO STEPS: 1. TRANSCRIPTION



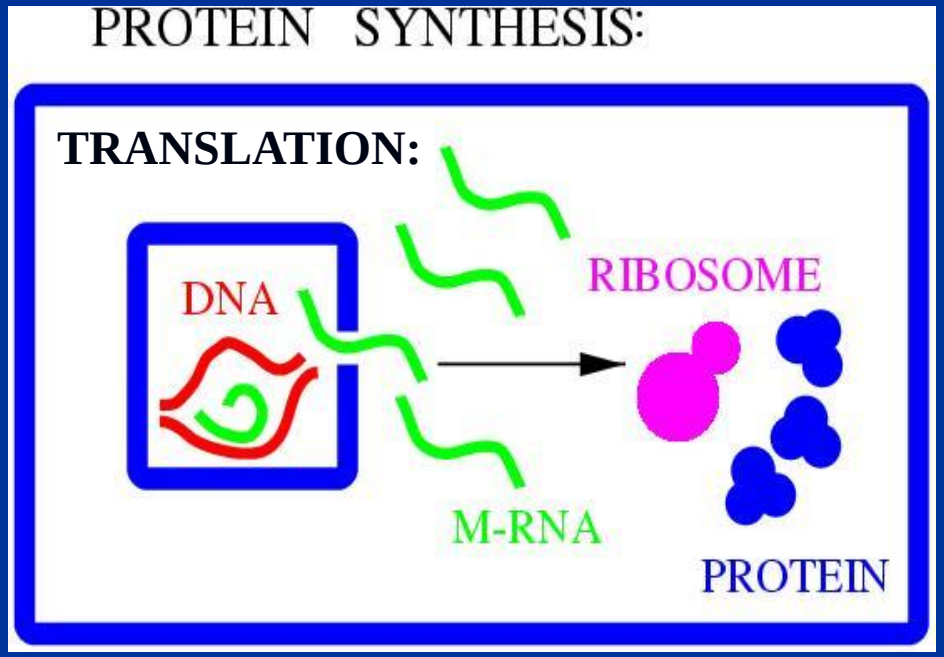


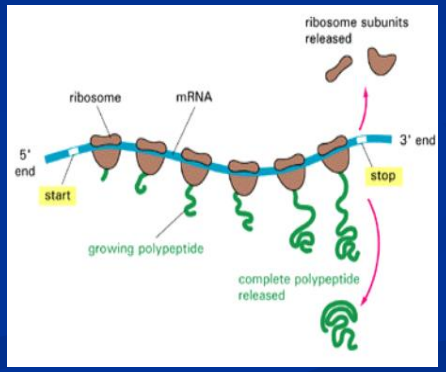


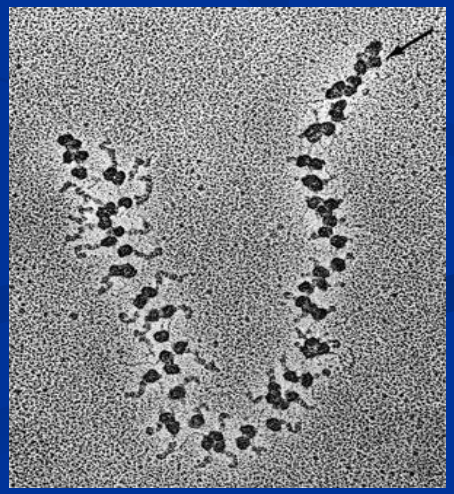


INFORMATION TRANSFER TO RIBOSOME

TWO STEPS: 2. TRANSLATION

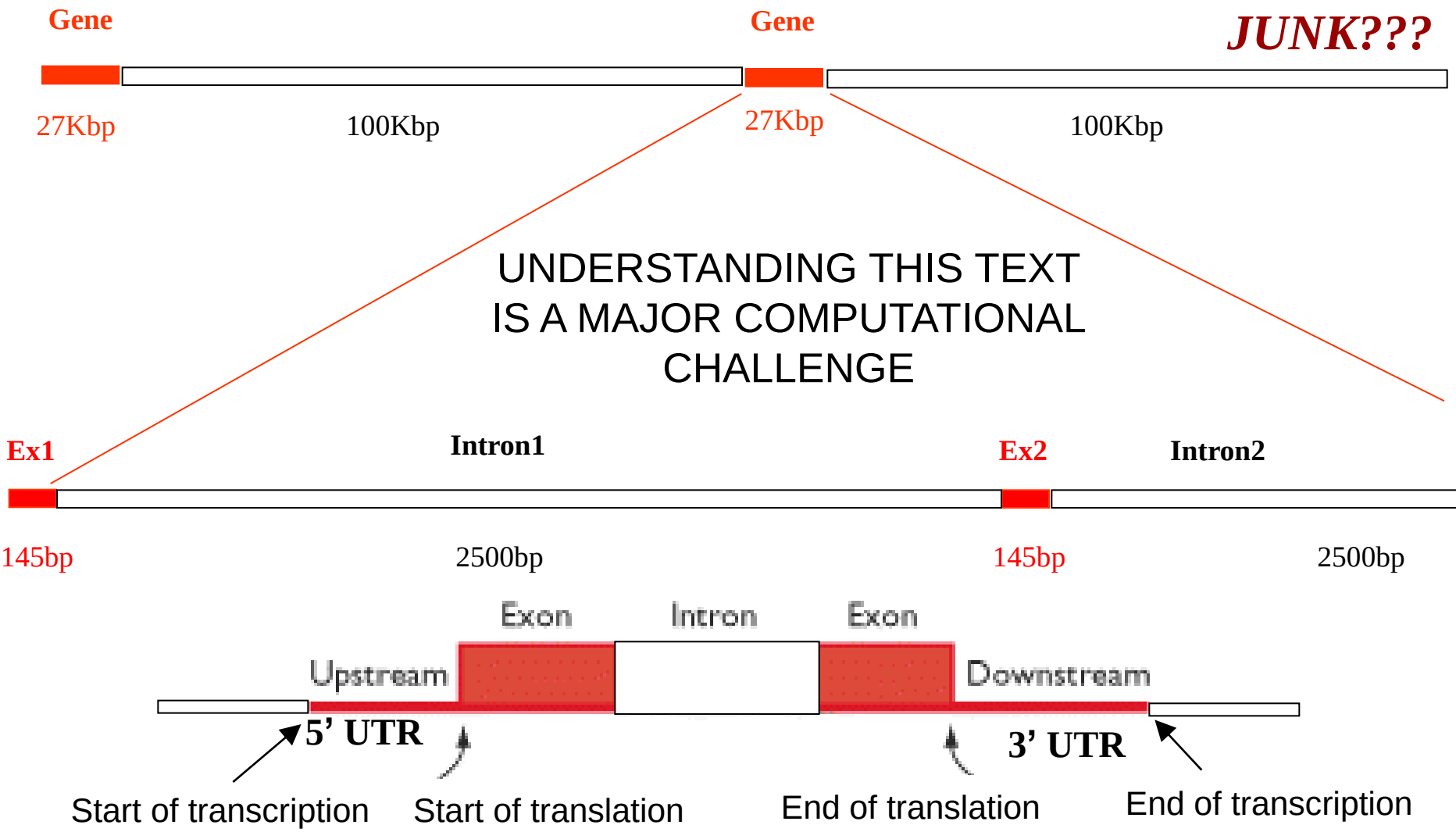




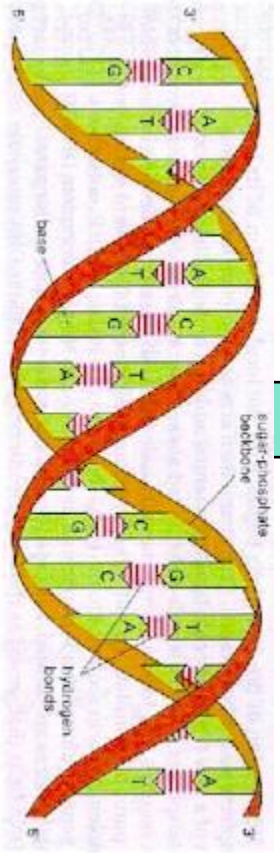


WALK ALONG THE HUMAN GENOME:

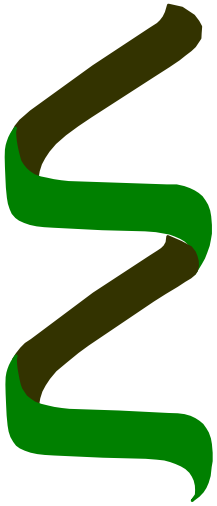
3 10^9 base pairs; ~23,000 protein coding genes ~25% of genome; actual coding ~1.5%



Central Dogma

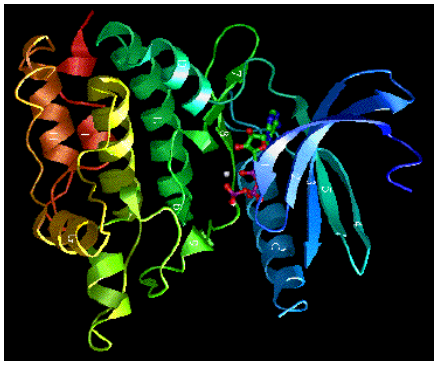


Transcription



mRNA

Translation



Protein

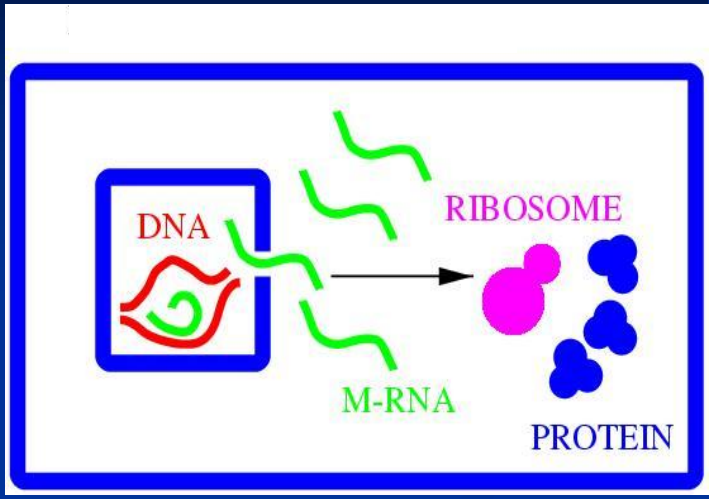
Information stored
In Gene (DNA)

A gene is **expressed** when the mRNA and protein it codes for are produced

Cells express **different** subset of the genes in different tissues and under different conditions

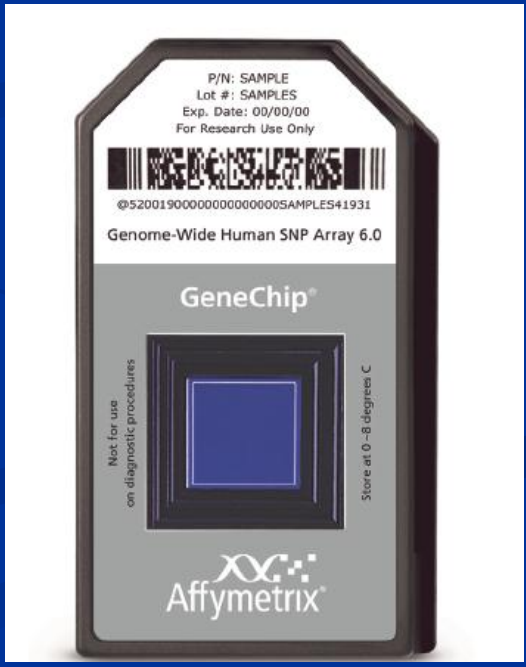
MEASURING GENE EXPRESSION PROFILE ~ 20-30,000 NUMBERS

WHEN A PARTICULAR GENE IS EXPRESSED, THE CONCENTRATIONS OF ITS CORRESPONDING MESSENGER RNA AND PROTEIN ARE HIGH.



A DNA-CHIP MEASURES CONCENTRATIONS OF THOUSANDS OF DIFFERENT

MESSENGER RNA

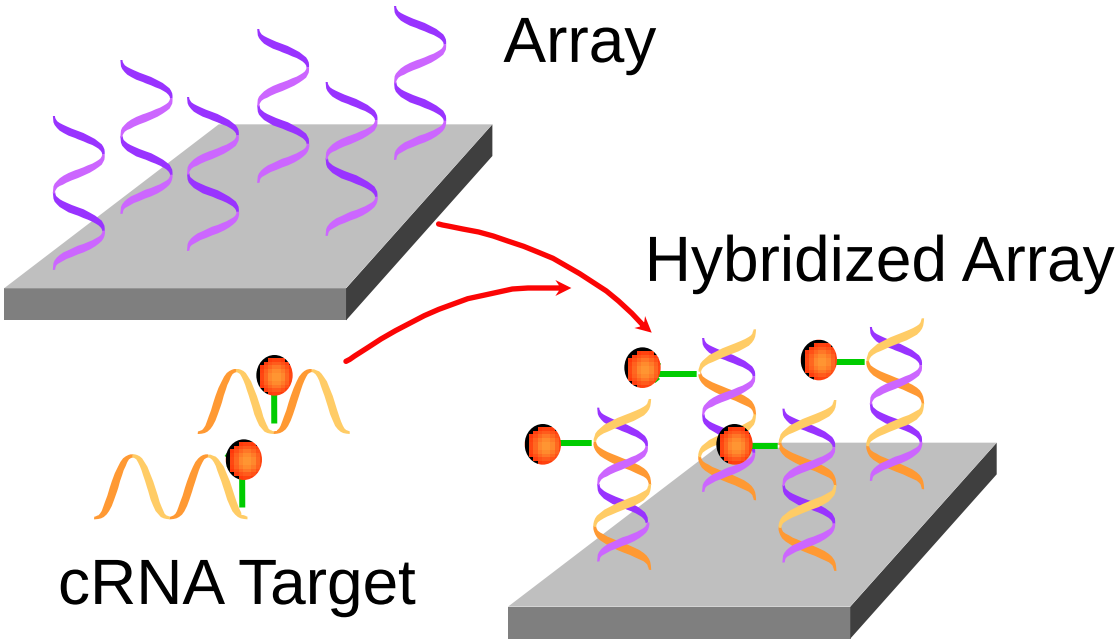


LATEST AFFYCHIPS: Human Gene 1.0 ST Array measures concentration of 20-30,000 genes, using 764,885 "probes" (26 probes per gene)

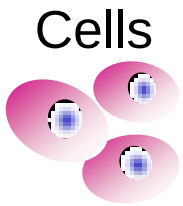
HU Exon 1.0 ST – 1,425,647

Genome-wide Human SNP array 6.0; 1.8M probes (*DNA* CONTENT)

Hybridization

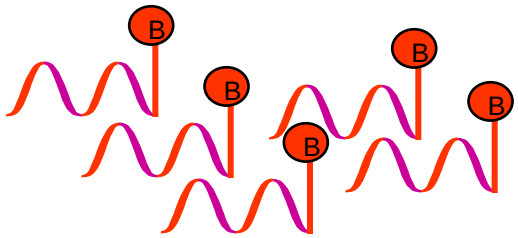


Affymetrix Experimental Design



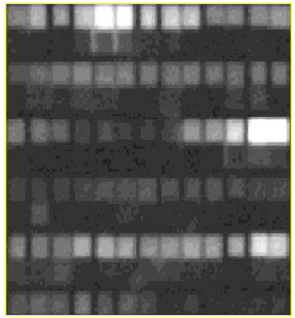
Extract RNA

Amplify, Label, Fragment

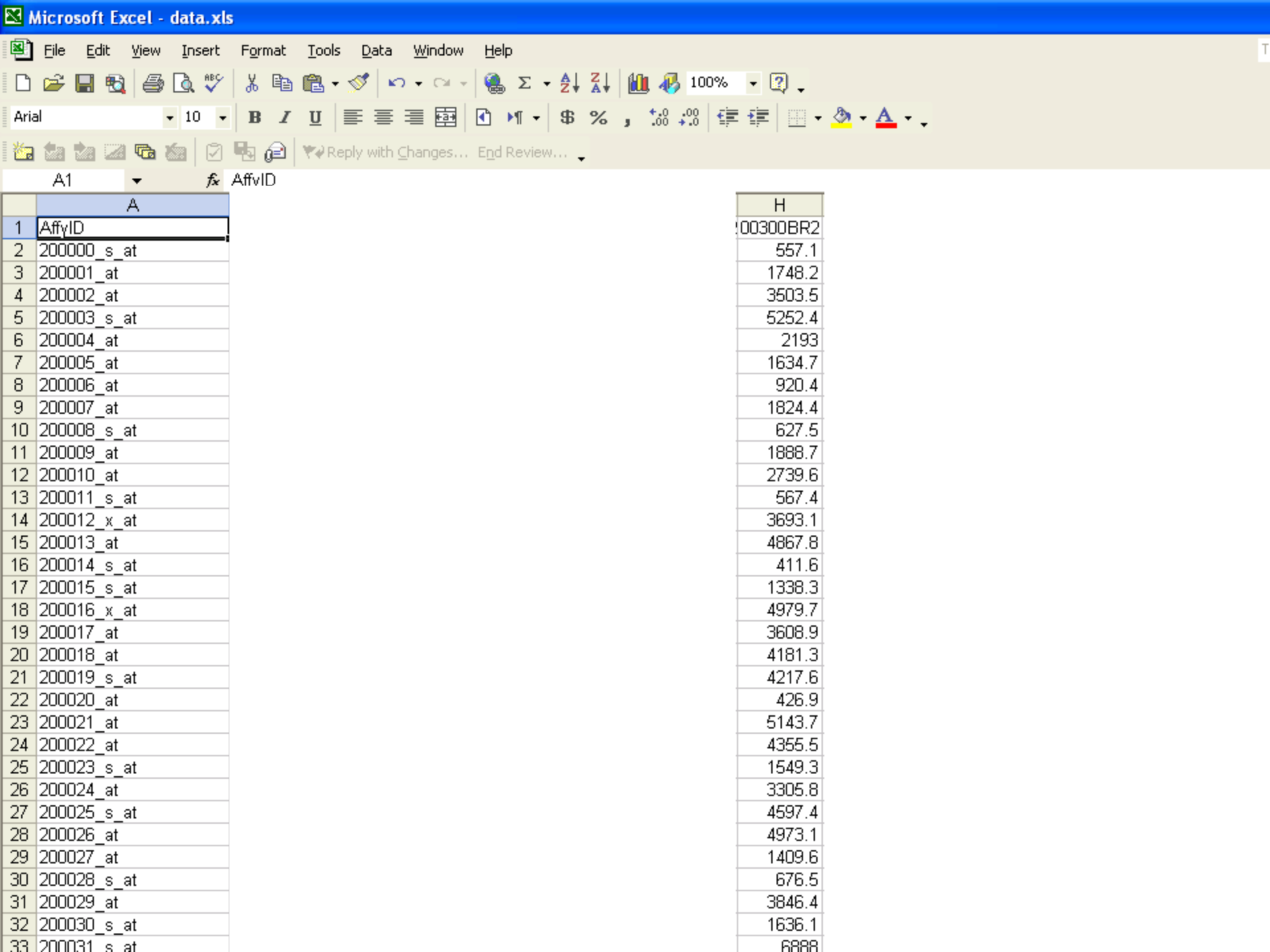


Hybridize to microarray

Scan



MEASURE LIGHT INTENSITY FROM EACH FEATURE , TO COMPOSE EXPRESSION LEVELS OF ~ 30,000 GENES



	A	H
1	AffyID	00300BR2
2	200000_s_at	557.1
3	200001_at	1748.2
4	200002_at	3503.5
5	200003_s_at	5252.4
6	200004_at	2193
7	200005_at	1634.7
8	200006_at	920.4
9	200007_at	1824.4
10	200008_s_at	627.5
11	200009_at	1888.7
12	200010_at	2739.6
13	200011_s_at	567.4
14	200012_x_at	3693.1
15	200013_at	4867.8
16	200014_s_at	411.6
17	200015_s_at	1338.3
18	200016_x_at	4979.7
19	200017_at	3608.9
20	200018_at	4181.3
21	200019_s_at	4217.6
22	200020_at	426.9
23	200021_at	5143.7
24	200022_at	4355.5
25	200023_s_at	1549.3
26	200024_at	3305.8
27	200025_s_at	4597.4
28	200026_at	4973.1
29	200027_at	1409.6
30	200028_s_at	676.5
31	200029_at	3846.4
32	200030_s_at	1636.1
33	200031_s_at	6888

Microsoft Excel - data.xls

File Edit View Insert Format Tools Data Window Help

Arial 10

Reply with Changes... End Review...

A1 AffyID

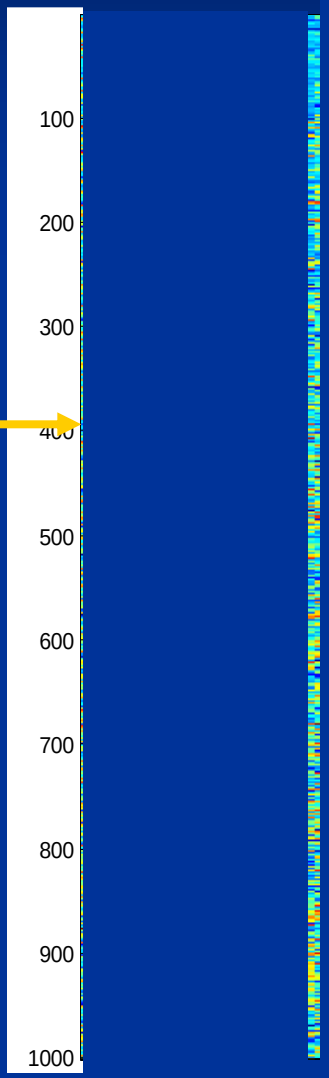
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	AffyID	00450UR2	00485UR2	00823UR2	00838UR2	A5371HR2	C0139KR2	00300BR2	02184AR2	02815AR1	03283AR1	03519AR2	03531AR2
2	200000_s_at	369.3	383.9	477.5	330.9	322.8	348.6	557.1	380.3	529.8	257.5	253.1	596.6
3	200001_at	633.8	806.2	740.7	915.6	1244.1	678.7	1748.2	1217.6	1085.4	1022.7	1364	1152.9
4	200002_at	3400.8	3007.2	3133.7	4032.3	2521.8	1906.7	3503.5	3218.3	2724	3775.7	2550.4	2944.3
5	200003_s_at	4131.6	4439.9	4445.1	4878.7	3257.8	2699.6	5252.4	4450.6	4335.1	4861.3	3428.2	3600.5
6	200004_at	1612.1	1310.8	1211.7	2241.2	2127.4	1525.4	2193	2099.9	1896.4	1750.1	2170.2	1889.7
7	200005_at	1113.7	1323	731.8	1265.7	555.5	482.9	1634.7	818.3	916	1337.3	850.1	1073.8
8	200006_at	1151.1	978.8	1549.7	1632.9	1448	1855.8	920.4	1627.6	1681	2119.9	1973.3	1916.9
9	200007_at	1051.6	988.7	1363.4	1131.3	1716.2	1323.7	1824.4	1646	1862.5	1679.2	1479.7	1691.6
10	200008_s_at	133	294.6	347.2	218.1	782.7	558.9	627.5	753.7	771.3	922	679.3	1088.2
11	200009_at	521.5	904.3	1222.7	820.5	1518.1	1385.1	1888.7	1416.5	1501.9	1691.8	1532.7	1888.8
12	200010_at	1815.9	1483.8	2425.7	2672.4	2578.7	2045.3	2739.6	3015.2	2424.3	3916	1608.3	2970.5
13	200011_s_at	744.8	483.4	451.3	555.3	1018.7	279.1	567.4	438.9	452.2	489.8	718.3	539.6
14	200012_x_at	1931.5	3217.9	3720.1	2565.9	3089.1	3140.6	3693.1	4154	4533.5	4471.1	2629.5	3260.2
15	200013_at	3400.9	3817.9	4032.6	4113.5	2621	2710.6	4867.8	3678.9	3406.5	3718.5	2806.5	3146.7
16	200014_s_at	509	456.8	340.9	625.5	411.8	466.6	411.6	488.3	454.8	651.1	709.6	733.3
17	200015_s_at	1323.1	1181.6	1014.6	1117	830	599.2	1338.3	1115.3	1296.2	1199.6	982.6	977.4
18	200016_x_at	2477.3	2920.4	2832.4	3546.8	3128.6	1770.1	4979.7	3707.6	2895.6	3898.4	2332.7	3923.2
19	200017_at	2997.7	2231.5	3292.5	3659.5	3204.1	2507.2	3608.9	4830.1	4017	4477.7	3380.7	3762.4
20	200018_at	4566.5	4903	3459	4152.2	3491.4	3507.9	4181.3	4629.1	4723	4170.8	3373.4	4067.7
21	200019_s_at	5019.8	4420.3	3201.1	4148.2	2785.2	3126.8	4217.6	3047	2637.9	3361.3	2836.8	3328.3
22	200020_at	627.2	449	593.8	589.6	497.4	340.5	426.9	490.1	510.4	698.8	658.5	669.5
23	200021_at	8773.2	8539.6	6521.3	6543.2	5513.6	3511.7	5143.7	6145.4	5306.4	5737.9	4173.5	4557.7
24	200022_at	3300.9	4028.1	3490.3	4832.2	2944.5	2436.4	4355.5	3040.1	2582.5	3288.9	2736.2	3520.8
25	200023_s_at	1094.8	1004.9	781.1	1098.5	1175.4	830.9	1549.3	1121.4	1429	1250.7	819.4	1216.6
26	200024_at	1075.3	1738.5	1815.3	2710.2	2240.3	1712.6	3305.8	3919.4	3311.2	4087.6	1907.6	2882.6
27	200025_s_at	4743.7	5353	3494.5	4986.7	3120.9	3215.7	4597.4	4309	4109.6	5018.3	3388.1	3753.5
28	200026_at	5905.2	7854.7	4420.4	6630.9	4533.3	3027.7	4973.1	5298.4	4618.8	5537.5	3676.1	4503
29	200027_at	1288.3	1120.6	1017.6	1061.2	1566.3	622.4	1409.6	802	935.5	773.4	1133.7	1970.6
30	200028_s_at	657.2	516	644.9	554.1	639.2	831.7	676.5	716.4	588.7	711.8	426.6	880.6
31	200029_at	4631.1	5165.4	5222.9	4355.2	2452.8	3318.1	3846.4	3363.2	3675.6	4674	3130.1	2966.1
32	200030_s_at	907.1	912.7	970.4	1469.5	2630.1	2293.7	1636.1	1970.7	1910.9	2227	1410	2604.2
33	200031_s_at	5538.1	5656.6	5459.5	5825.3	5401.9	4129.8	6888	6548.4	5618.8	6227.1	4659	4759.2

COLON CANCER DATA:

E_{ij} = EXPRESSION LEVEL OF GENE i IN SAMPLE j

Sample # 127

gene 400



EACH PATIENT IS DESCRIBED BY 30,000 NUMBERS: ITS EXPRESSION PROFILE

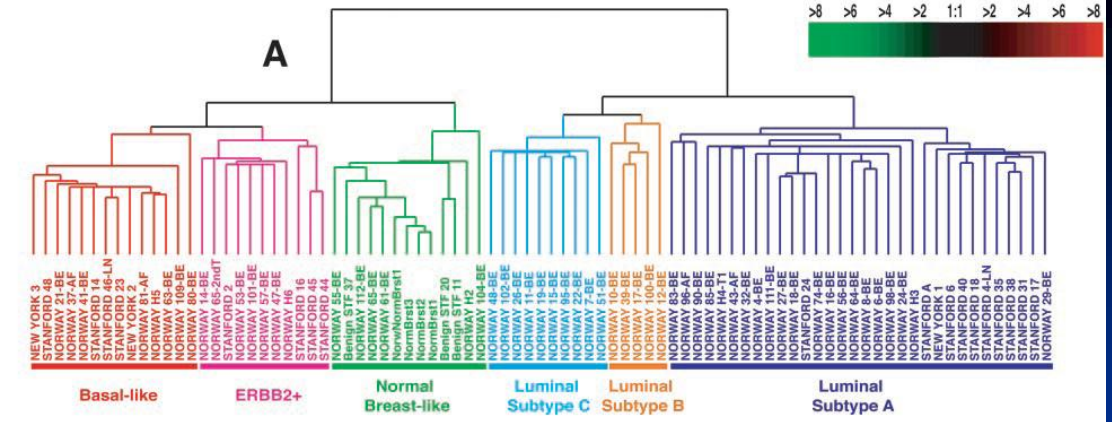
AIMS: **ASSIGN PATIENTS TO GROUPS ON THE BASIS OF THEIR EXPRESSION PROFILES (IDENTIFY CANCER SUBTYPES)**

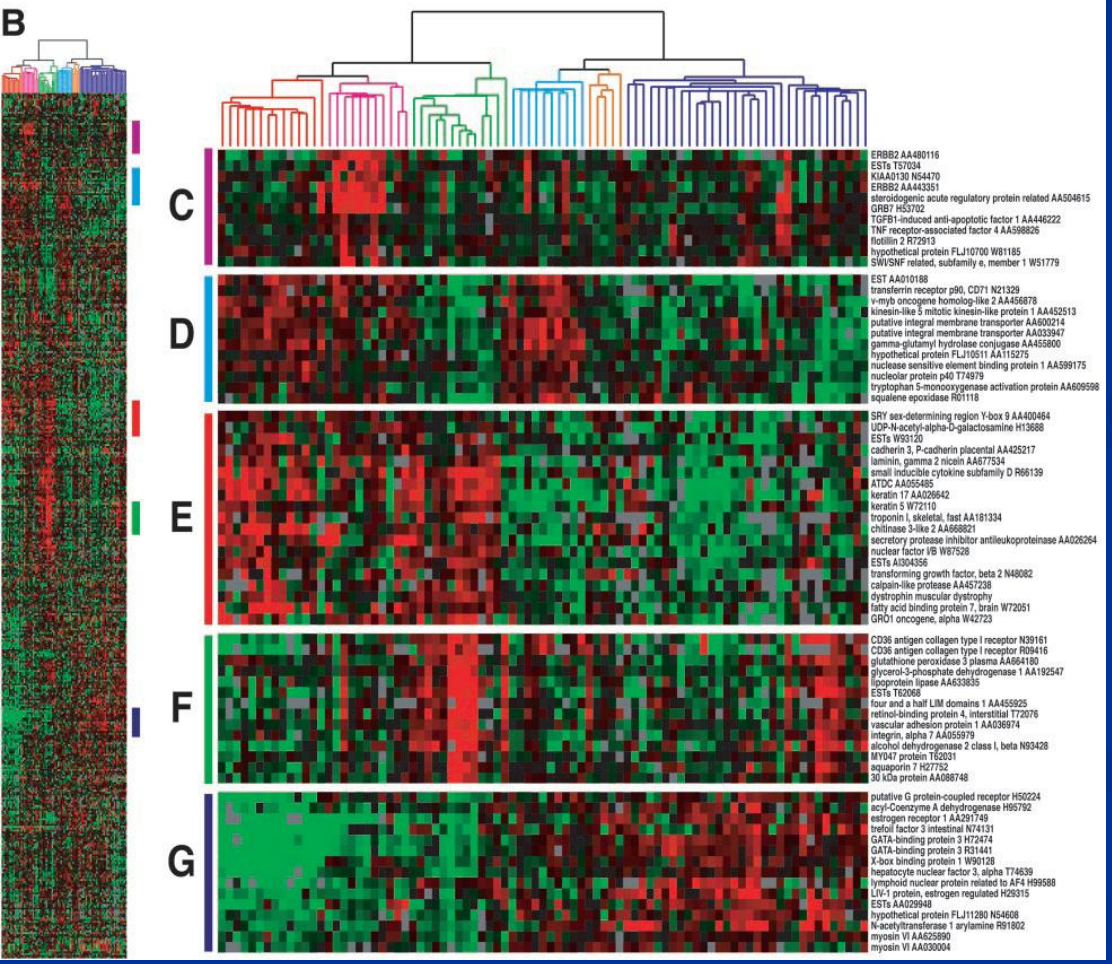
ASSIGN GENES TO FUNCTIONAL GROUPS (DECIPHER MOLECULAR MECHANISMS)

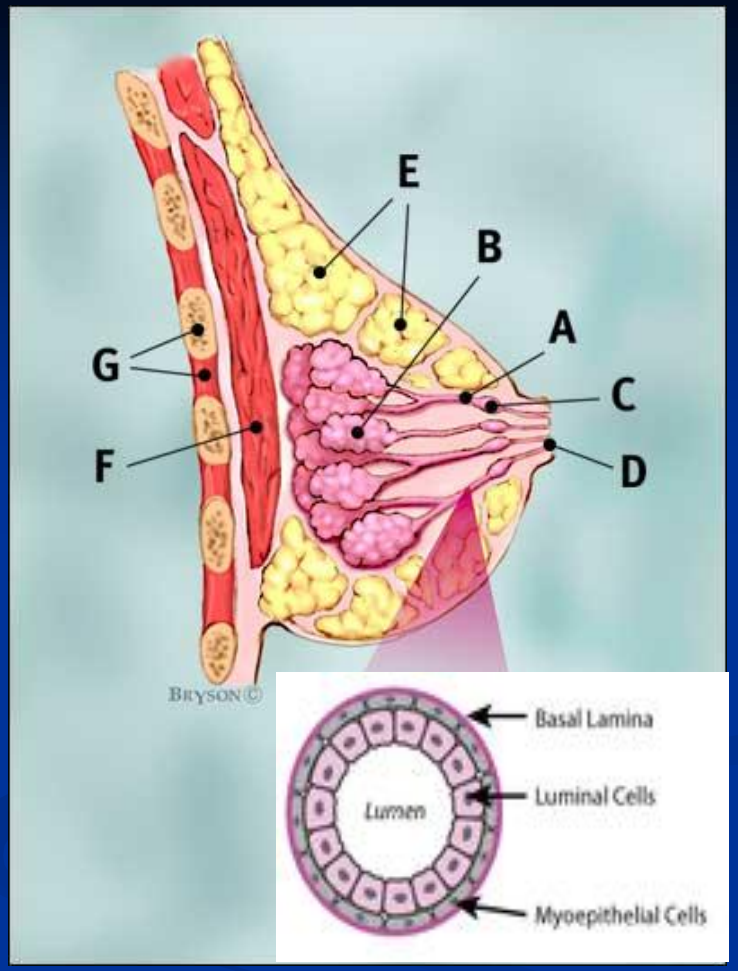
IDENTIFY DIFFERENCES BETWEEN TUMORS AT DIFFERENT STAGES (PERSONALIZED PREDICTIVE MEDICINE)

IDENTIFY GENES THAT PLAY CENTRAL ROLES IN DISEASE PROGRESSION (DRUG DISCOVERY AND DESIGN)

MAJOR COMPUTATIONAL CHALLENGE







BASAL, ERBB2/HER2+,
NORMAL-LIKE
LUMINAL A/B/C

Sorlie et al PNAS 2001

2b. REGULATION OF EXPRESSION:

WHO DECIDES, AND HOW, THAT “IT IS TIME”

FOR A GENE TO BE EXPRESSED?

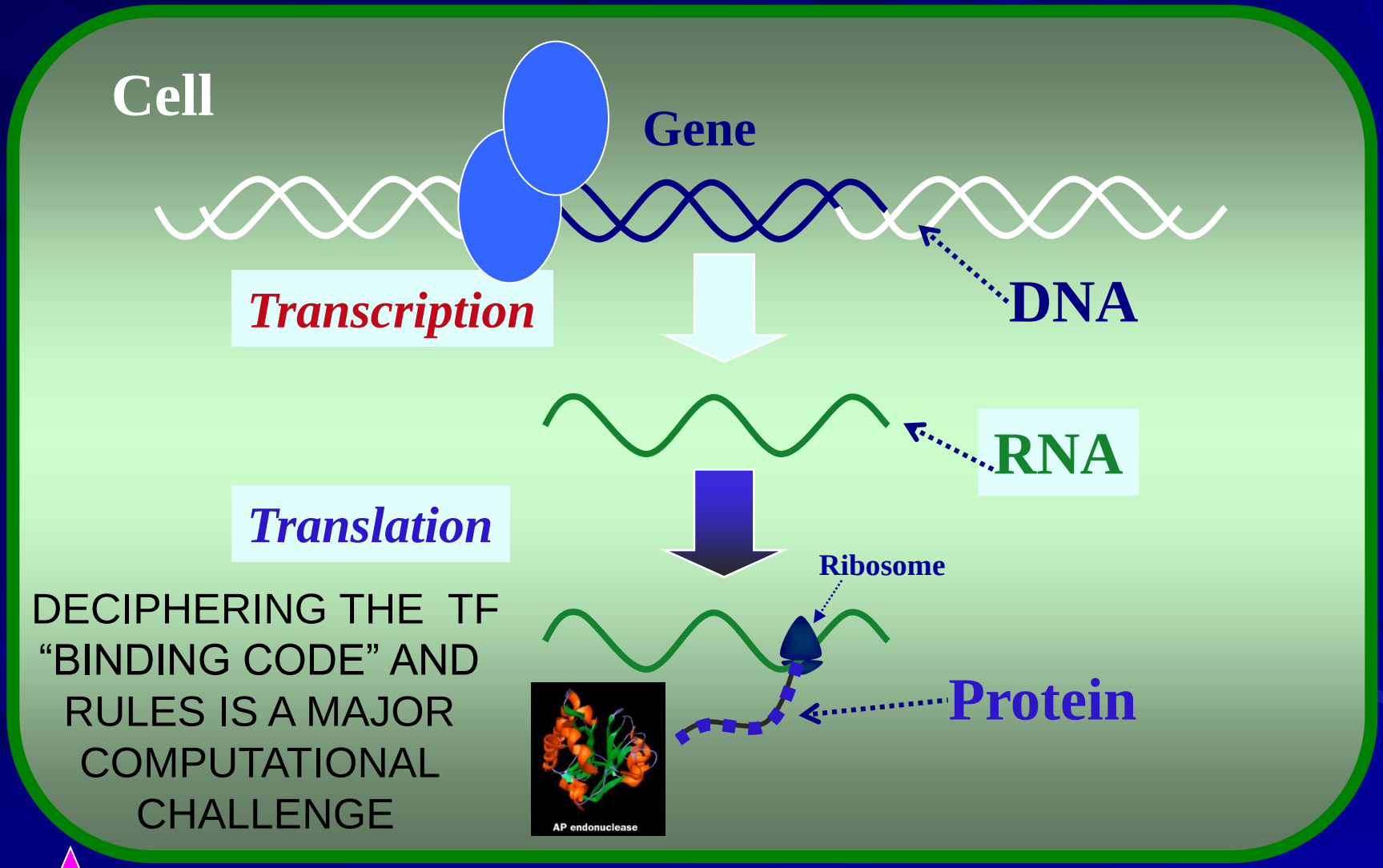
OR - WHAT TURNS A GENE ON?

1. TRANSCRIPTION FACTORS

2. ACCESSIBILITY OF THE DNA

AN ACTIVATOR PROTEIN BINDS TO THE DNA AND INDUCES TRANSCRIPTION

A REPRESSOR PROTEIN BINDS TO THE DNA AND BLOCKS TRANSCRIPTION



DECIPHERING THE TF "BINDING CODE" AND RULES IS A MAJOR COMPUTATIONAL CHALLENGE

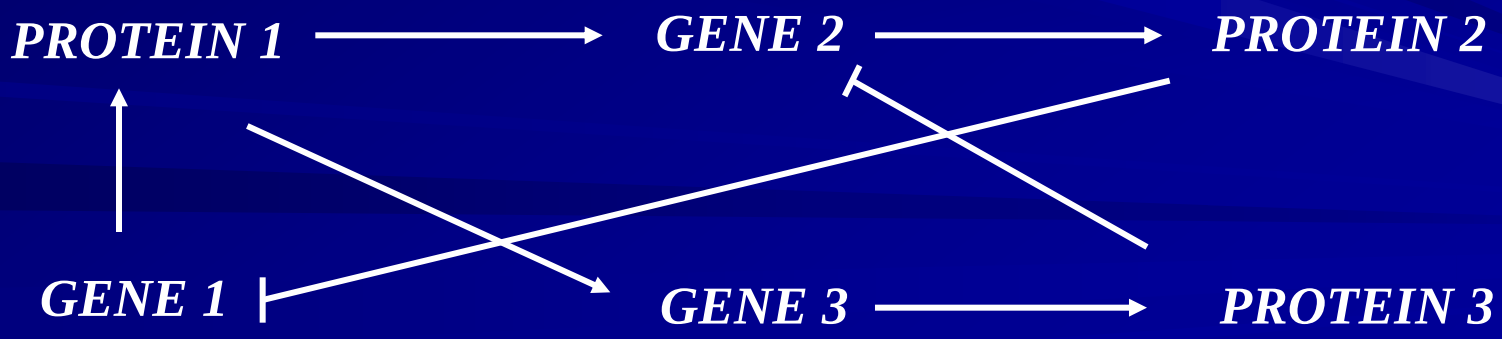
TRANSCRIPTION FACTORS BIND TO THE DNA AT ***BINDING SITES***

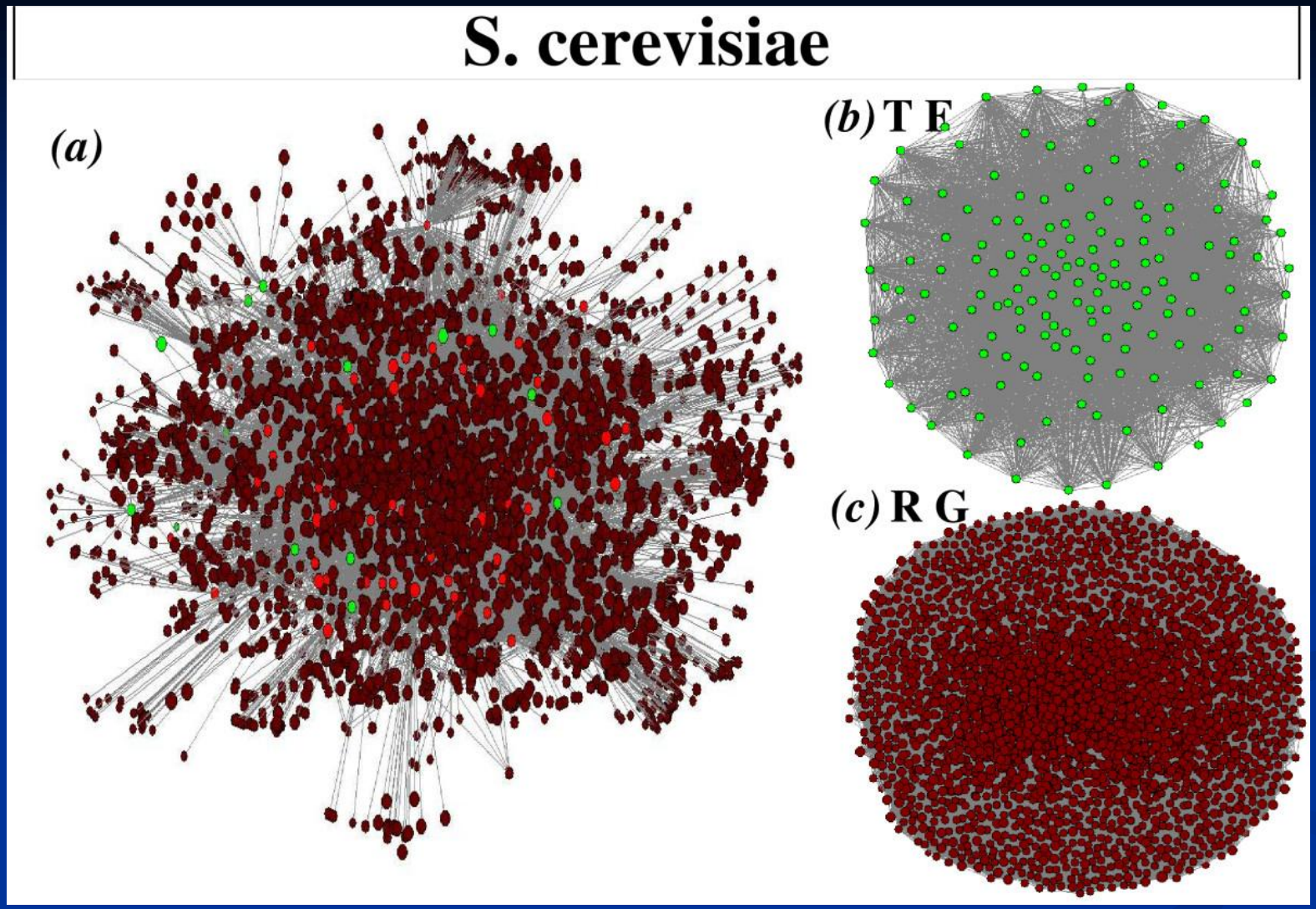
TRANSCRIPTIONAL NETWORKS

GENE CODES FOR PROTEIN

PROTEINS ACTIVATE/SUPPRESS GENE TRANSCRIPTION

GET NETWORK THAT REGULATES THE RNA AND PROTEIN CONTENT OF THE CELL





Guzmán-Vargas aSantillán *BMC Systems Biology* 2008

TRANSCRIPTIONAL REGULATORY NETWORK (Yeast):

Transcription Factors Regulated Genes Both (known)

COMPUTATIONAL CHALLENGES:

1. DEDUCTION OF NETWORK FROM DATA
2. GLOBAL CHARACTERISTICS (HUBS, POWER LAWS…)
3. OVER-REPRESENTED LOCAL MOTIFS
4. DENSE SUBSETS (MODULES)
5. TRANSCRIPTION FACTOR BINDING SITES ON DNA
6. TRANSCRIPTIONAL DYNAMICS
7. NETWORK EVOLUTION

3a. THE HALLMARKS OF CANCER:

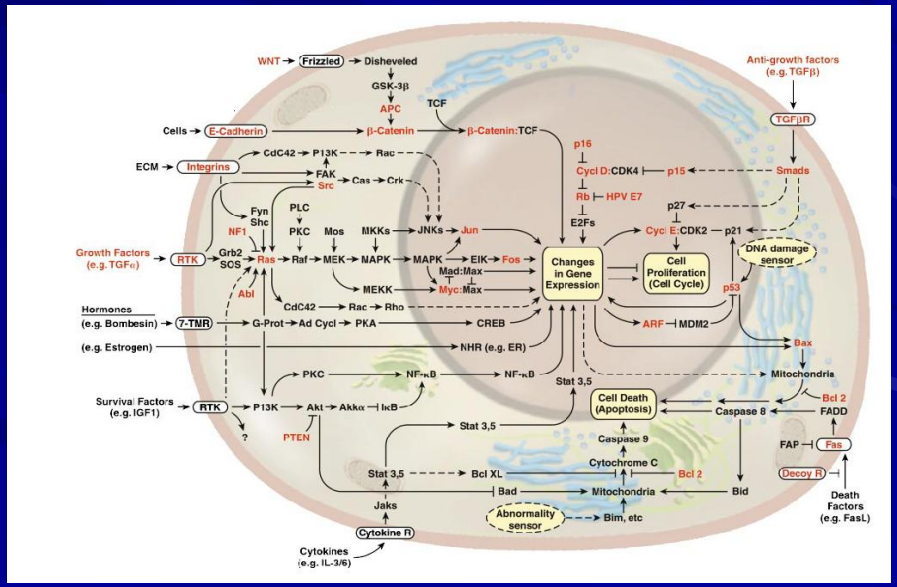
CANCER = UNCONTROLLED GROWTH

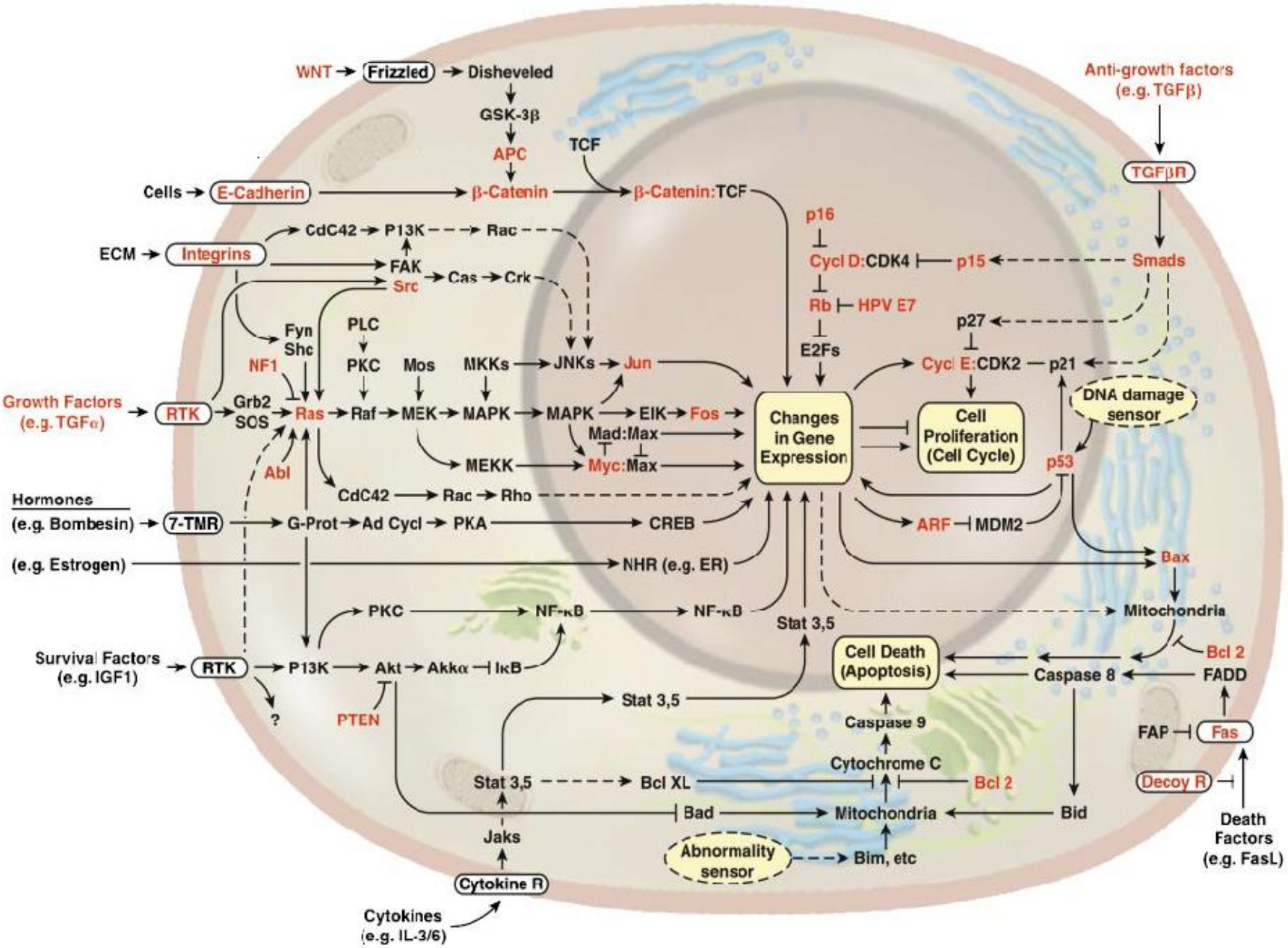
1. VARIOUS REGULATORY NETWORKS PROTECT NORMAL CELLS AGAINST UNCONTROLLED PROLIFERATION

2. THE BREAKDOWN OF THESE NETWORKS ARE THE ***HALLMARKS OF CANCER*** ***(4 + 2)***

Hanahan & Weinberg
Cell 2000

SYSTEMATIC APPROACH TO DISCOVER & UNDERSTAND THE MOLECULAR MECHANISMS OF CANCER




WNT
Frizzled
Disheveled
GSK-3β
APC
TCF
β-Catenin
β-Catenin:TCF
Cells
E-Cadherin
Anti-growth factors
(e.g. TGFβ)
TGFβR
Smads
ECM
Integrins
CdC42
P13K
Rac
FAK
Src
Cas
Crk
p16
Cycl D:CDK4
p15
Rb
HPV E7
E2Fs
p27
Cycl E:CDK2
p21
Fyn
Shc
PLC
PKC
Mos
MKKs
JNKs
Jun
NF1
Growth Factors
(e.g. TGFα)
RTK
Grb2
SOS
Ras
Raf
MEK
MAPK
MAPK
EIK
Fos
Mad:Max
Myc:Max
MEKK
Changes
in Gene
Expression
Cell
Proliferation
(Cell Cycle)
DNA damage
sensor
p53
Abl
CdC42
Rac
Rho
Hormones
(e.g. Bombesin)
7-TMR
G-Prot
Ad Cycl
PKA
CREB
ARF
MDM2
(e.g. Estrogen)
NHR (e.g. ER)
Bax
PKC
NF-κB
NF-κB
Stat 3,5
Mitochondria
Survival Factors
(e.g. IGF1)
RTK
P13K
Akt
Akkα
IκB
Cell Death
(Apoptosis)
Caspase 8
Bcl 2
FADD
?
PTEN
Stat 3,5
Caspase 9
FAP
Fas
Cytochrome C
Stat 3,5
Bcl XL
Bcl 2
Decoy R
Bad
Mitochondria
Bid
Jaks
Abnormality
sensor
Bim, etc
Death
Factors
(e.g. FasL)
Cytokine R
Cytokines
(e.g. IL-3/6)

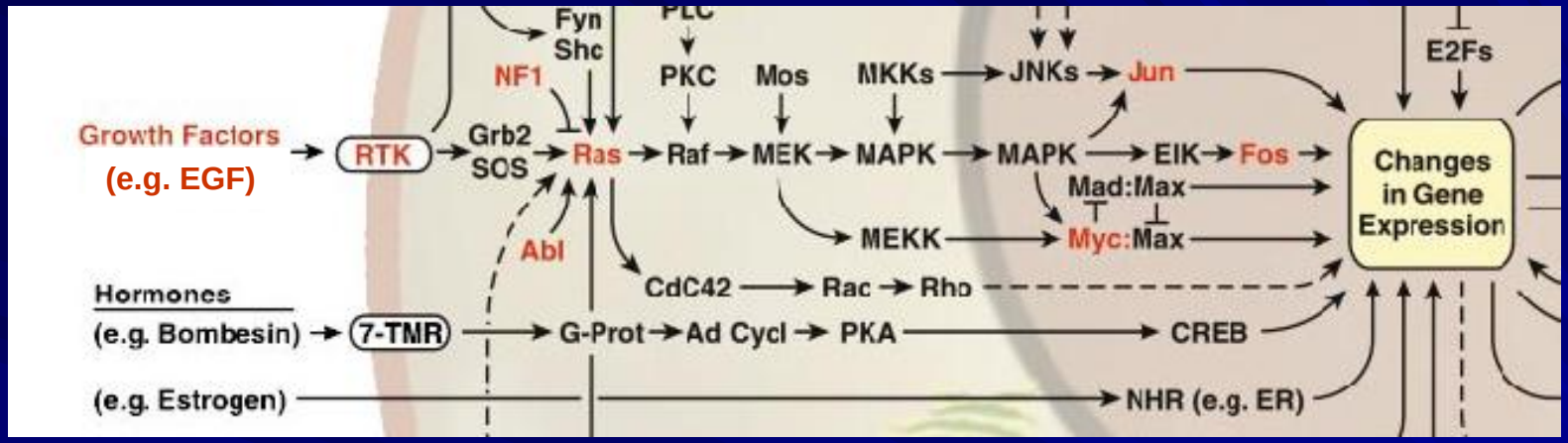
THE HALLMARKS OF CANCER (4 + 2)

1. SELF-SUFFICIENCY IN GROWTH SIGNALS
2. IGNORE ANTI-GROWTH SIGNALS
3. EVADE APOPTOSIS
4. IMMORTALIZATION

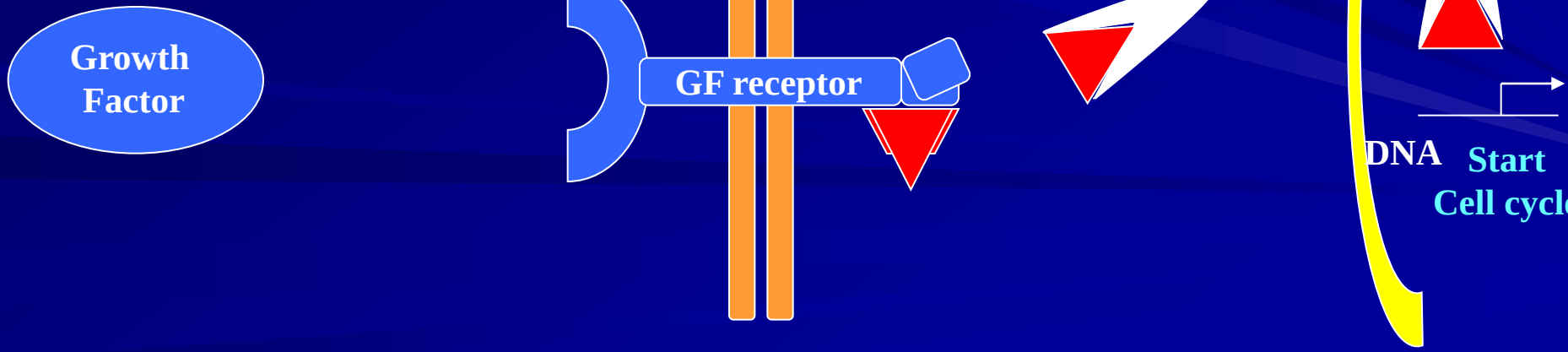
TWO MORE:

5. ANGIOGENESIS – GROWTH OF BLOOD VESSELS
6. METASTASIS – COLONIZATION OF VITAL ORGANS

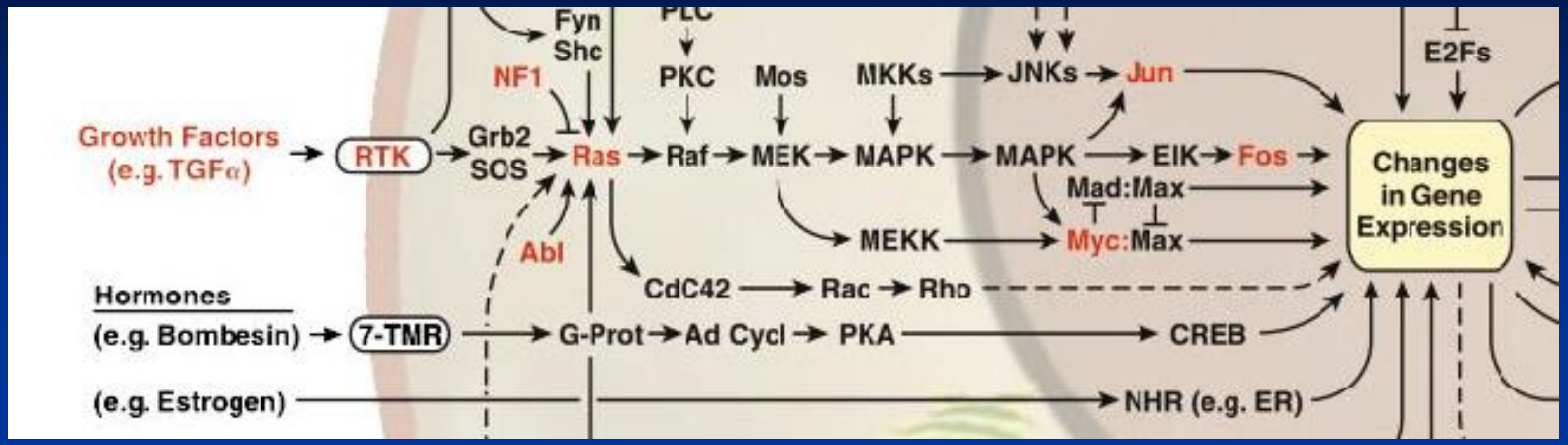
1. SELF SUFFICIENCY IN GROWTH SIGNALS



Normal signaling cascade:



1. SELF SUFFICIENCY IN GROW SIGNALS

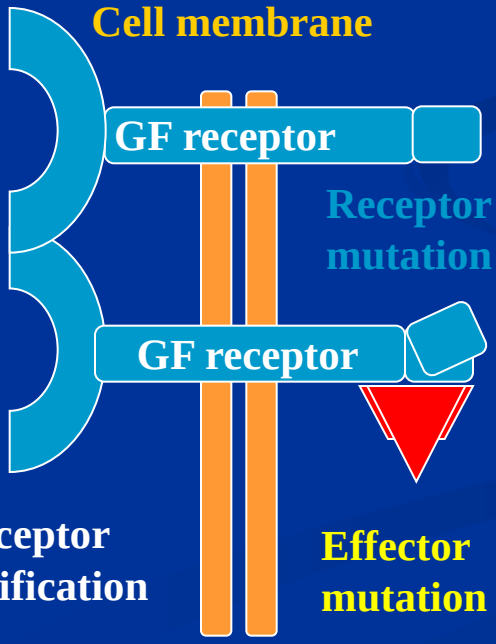


What can go wrong?

Growth Factor

Growth Factor

Autonomous GF production



(ONCOGENE)

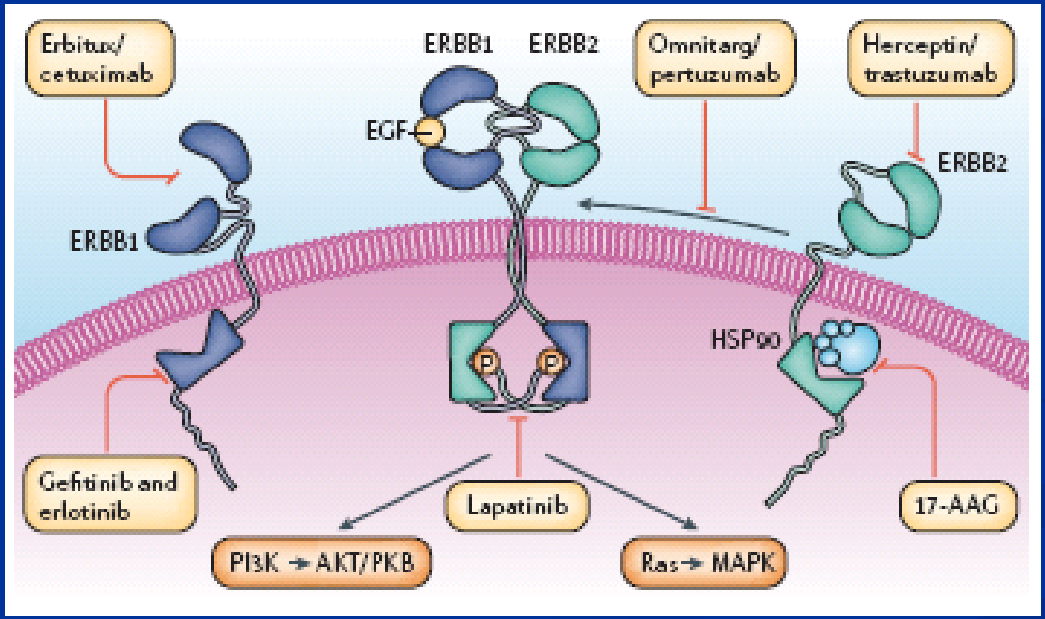
TARGETING THE FAMILY OF EGF RECEPTORS

EGF RECEPTOR ANOMALIES ARE IMPLICATED SEVERAL CANCERS. AMPLIFICATION (GLIOBLASTOMA, BREAST,…)

TARGETING THE EGF RECEPTORS:

ERBITUX

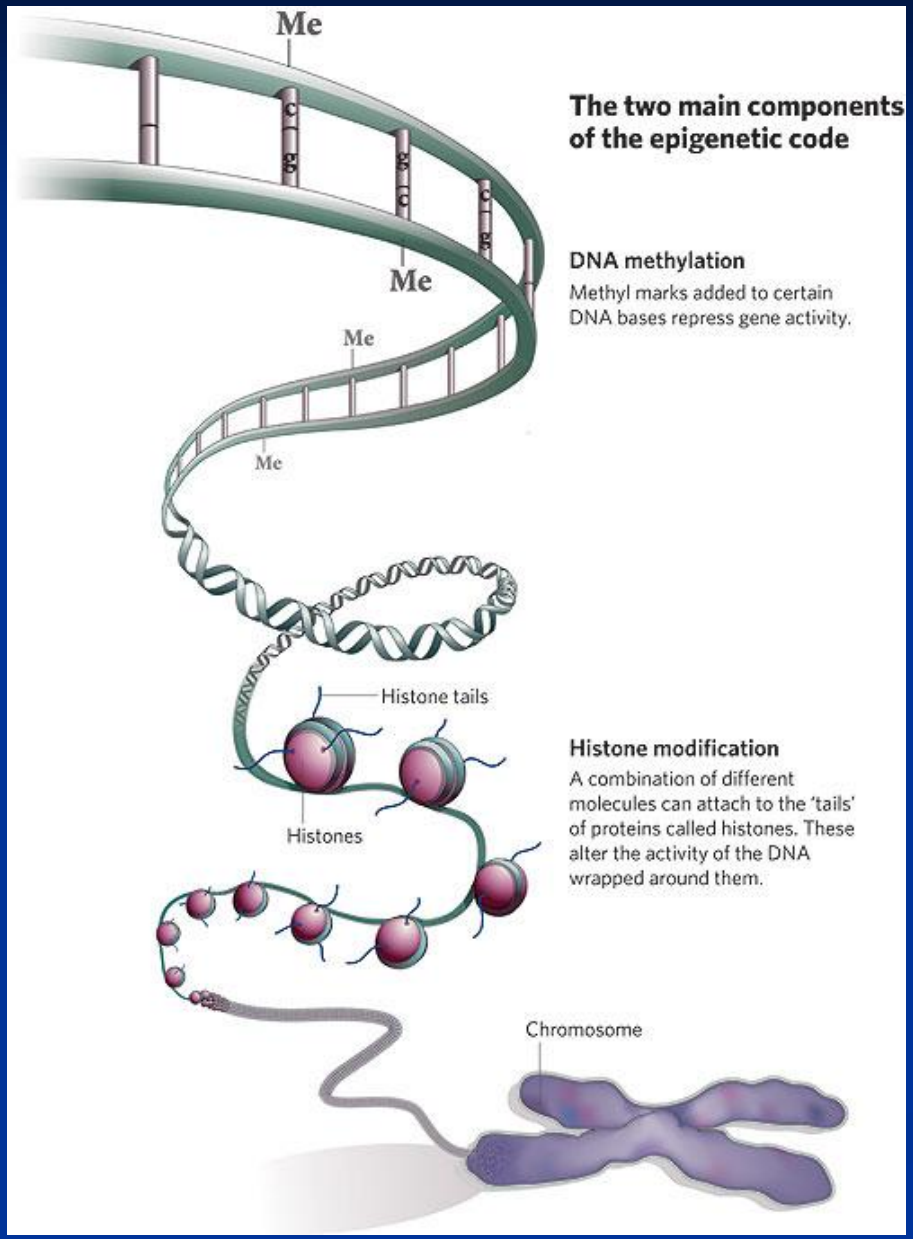
IRESSA, TARCEVA



HERCEPTIN

PERSONAL DRUG SELECTION, DICTATED BY THE ANOMALY

4. THE "EPIGENETIC CODE" AND REGULATION OF TRANSCRIPTION



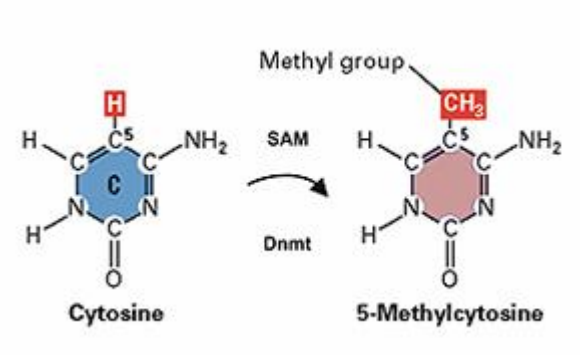
1. DNA METHYLATION
2. HISTONE MODIFICATIONS

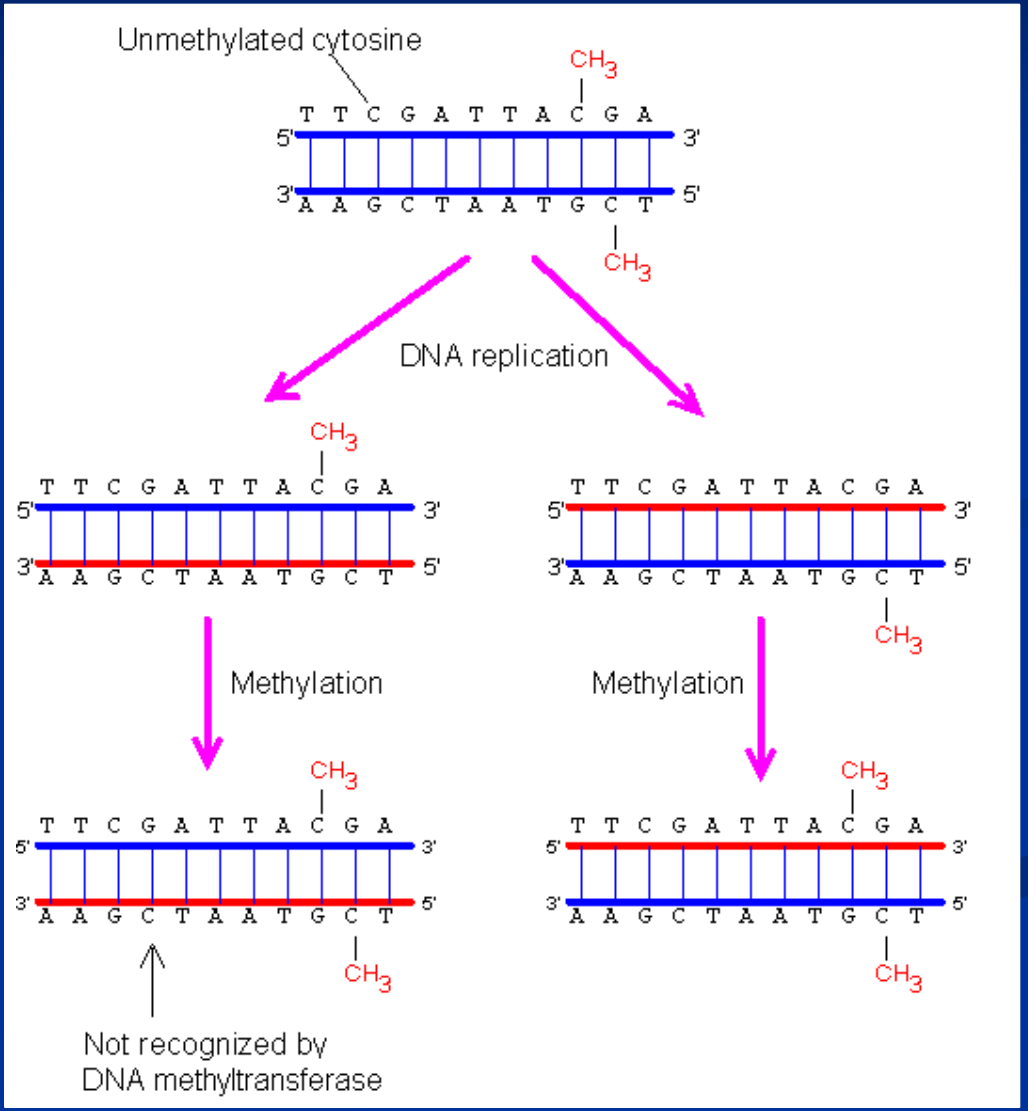
IDENTICAL TWINS DO NOT NECESSARILY GET THE SAME "GENTICALLY DRIVEN" DISEASES

(Esteller, Nature 2006)

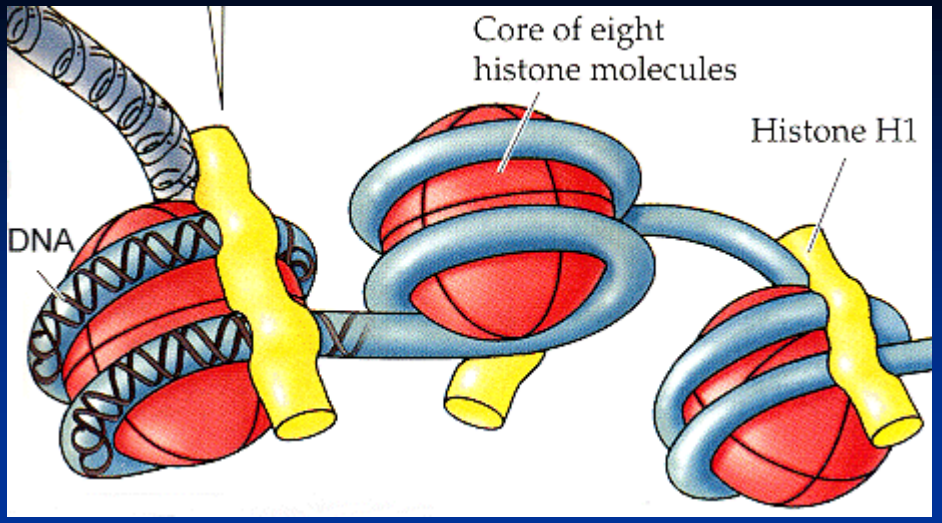
EPIGENETIC SIGNALS CONTROL DNA ACCESSIBILITY ON ALL SCALES

1. DNA METHYLATION: SILENCES GENES, PASSES TO DAUGHTER CELLS DURING DNA REPLICATION





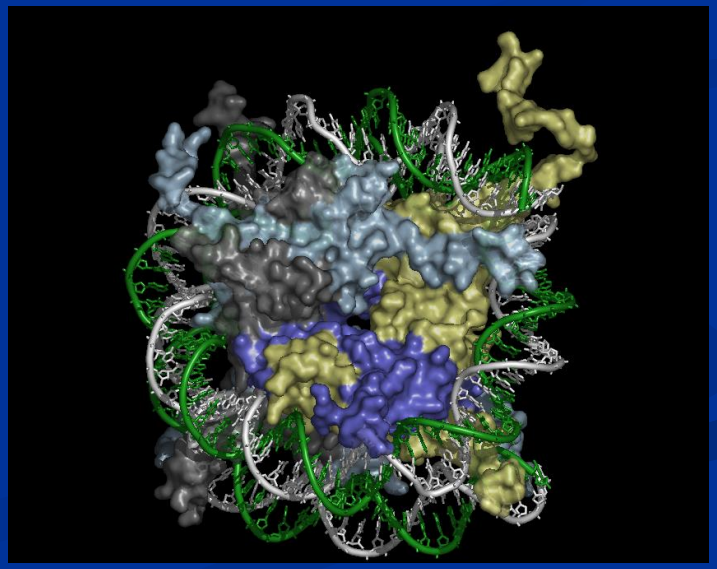
2. HISTONE MODIFICATIONS:



Physicist's nucleosome = basketball

TIGHT CLOSED PACKING OF CHROMATIN DOES NOT ALLOW ACCESS OF TRANSCRIPTION FACTORS AND POLYMERASE TO DNA. OPEN CHROMATIN ALLOWS TRANSCRIPTION.

CHROMATIN PACKING IS CONTROLLED BY HISTONE MODIFICATIONS

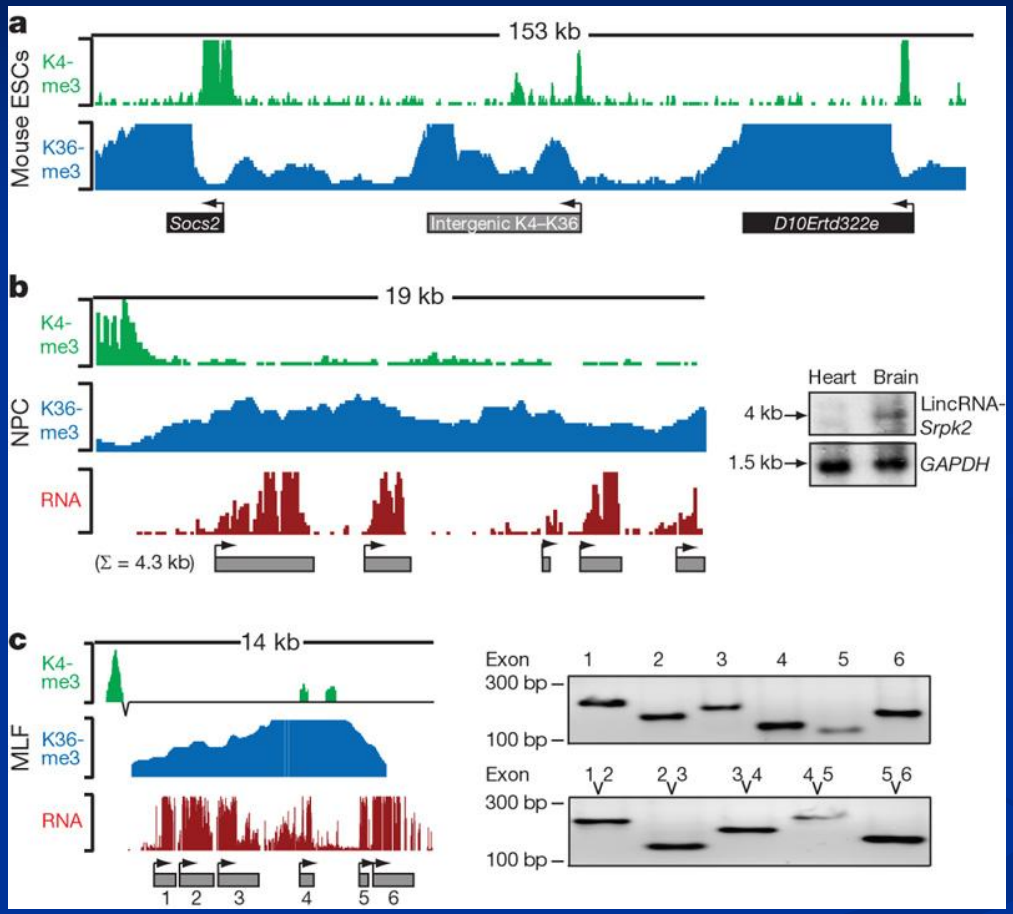


Biochemist's nucleosome

2. HISTONE MODIFICATIONS:

TRI - METHYLATION OF H3K4 AND H3K36 =>

=> OPEN CHROMATIN, TRANSCRIBED DNA

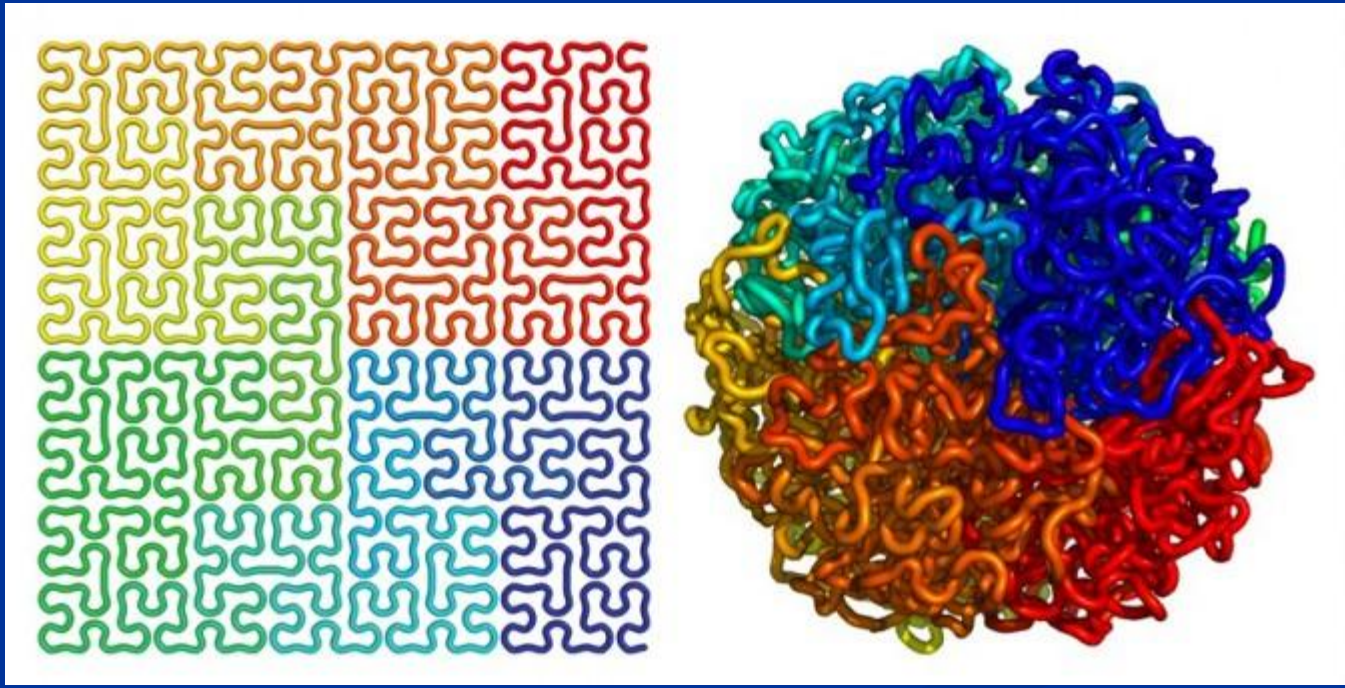


GENOME-WIDE IDENTIFICATION OF TRI-METHYLATED H3K4 – K36 LED TO DISCOVERY OF 1600 LONG INTERVENING NON-CODING RNA IN MOUSE ***(Guttman et al Nature 2009)***

THE GEOMETRY OF CHROMOSOME PACKING

Lieberman-Aiden, Mirny et al Science 2009

AT MEGABASE SCALE, THE CHROMATIN CONFORMATION IS A FRACTAL GLOBULE (~HILBERT CURVE IN 3-d), KNOT-FREE, ALLOWING MAXIMALLY DENSE PACKING WHILE PRESERVING THE ABILITY TO EASILY FOLD AND UNFOLD ANY GENOMIC LOCUS



SUMMARY:

1. THESE ARE EXCITING TIMES IN BIOLOGY:
   NEW FRONTIERS, NEW TECHNOLOGIES, CENTRAL ROLE FOR COMPUTATIONAL SCIENCE

2. a. INTRODUCTION TO MOLECULAR BIOLOGY OF THE CELL: GENES, GENE EXPRESSION AND ITS MEASUREMENT BY MICROARRAYS
   b. WHAT CONTROLS EXPRESSION?

3. a. THE HALLMARKS OF CANCER;
   b. RESPONSE TO STIMULUS BY A GROWTH FACTOR

4. THE “EPIGENETIC CODE” AND REGULATION OF TRANSCRIPTION

Funding and support:

NIH, EC/RTN, EC/6FW, ISF, GIF, Bikura, IMOS/Tashtiyot, Ridgefield, Minerva, Levine, Wolfson Foundations, DIP

Collaborators and teachers:

Weizmann: D. Givol, L. Sachs, E. Canaani, V. Rotter, Y. Yarden, M. Oren, E.Tzahor, D. Lancet, A. Ben-Zeev, B. Geiger, A. Gross, Z. Eshhar, I. Cohen

Memorial SKMC P. Paty

Cornell Med Sch F. Barany

Princeton D. Notterman

CHUV Lausanne: M. Hegi, R. Stupp

Regina Elena, Rome: G. Blandino

Sheba Med C. G. Rechavi, S. Izraeli

Hadassa MC E. Galun, D. Gazit

Students & postdocs:

L. Mirny, M. Blatt, S. Wiseman, M. Vendruscolo, E. Kussell, G. Getz, E. Levine, M. Sheffer, L. Hertzberg, H. Gal, I. Kela, L. Ein-Dor, O. Zuk, Y. Tabach, M. Golan, T. Shay, A. Zeisel, Y. Drier, N. Bossel, M. Koudritzky, Y. Horesh, D. Netanely