Gene expression and Cancer: three stories (and a half)

Eytan Domany
Dept of Physics of Complex Systems
Weizmann Institute of Science, Rehovot, Israel

KITP, May 2012

THREE STORIES (and a half)

1. Transcriptional Response of Cells to Stimulus: Discovery of Time-Dependent Transcript Specific Production and Degradation Rates.

2. Outcome Prediction in Breast Cancer: Hope, Hype, Physics and Biology

3. Chromosomal Instabilities in Cancer

3.5 Acute Lymphoblastic Leukemia in Children with Down Syndrome: Causality vs Selection

BIOLOGY vs PHYSICS
UPSIDES listed – *DOWNSIDES* by implication

1. Immediacy of Phenomena – ***Interesting & Exciting*** in a simple, direct way

2. The Beauty of Reductionism - everything stems from ***a few basic laws***, cast in mathematical language. There **is** "***Theoretical Physics***"!

3. An amazing number of basic things that are not known or understood -- new technologies every day -- ***there is a smell of breakthrough in the air***

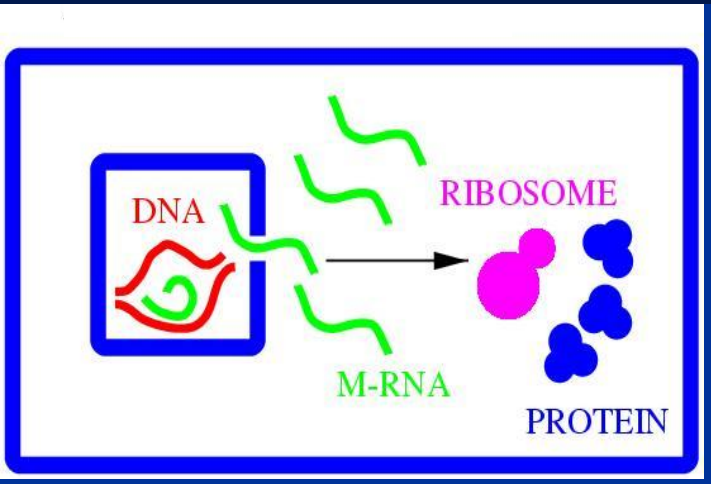
4. Although many papers are "not even wrong" – ***there is an objective truth*** Wrong claims either go unnoticed or have a very short life-time

5. One (I, anyway) is very frequently ***wildly surprised***

6. Submitted papers are reviewed, reviews are very rarely completely idiotic and unfair

7. Selection vs Causality

MEASURING THE TRANSCRIPTOME: ABUNDANCE OF 10,000 – 20,000 mRNA SPECIES



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

Quantitative Real Time PCR – not high throughput

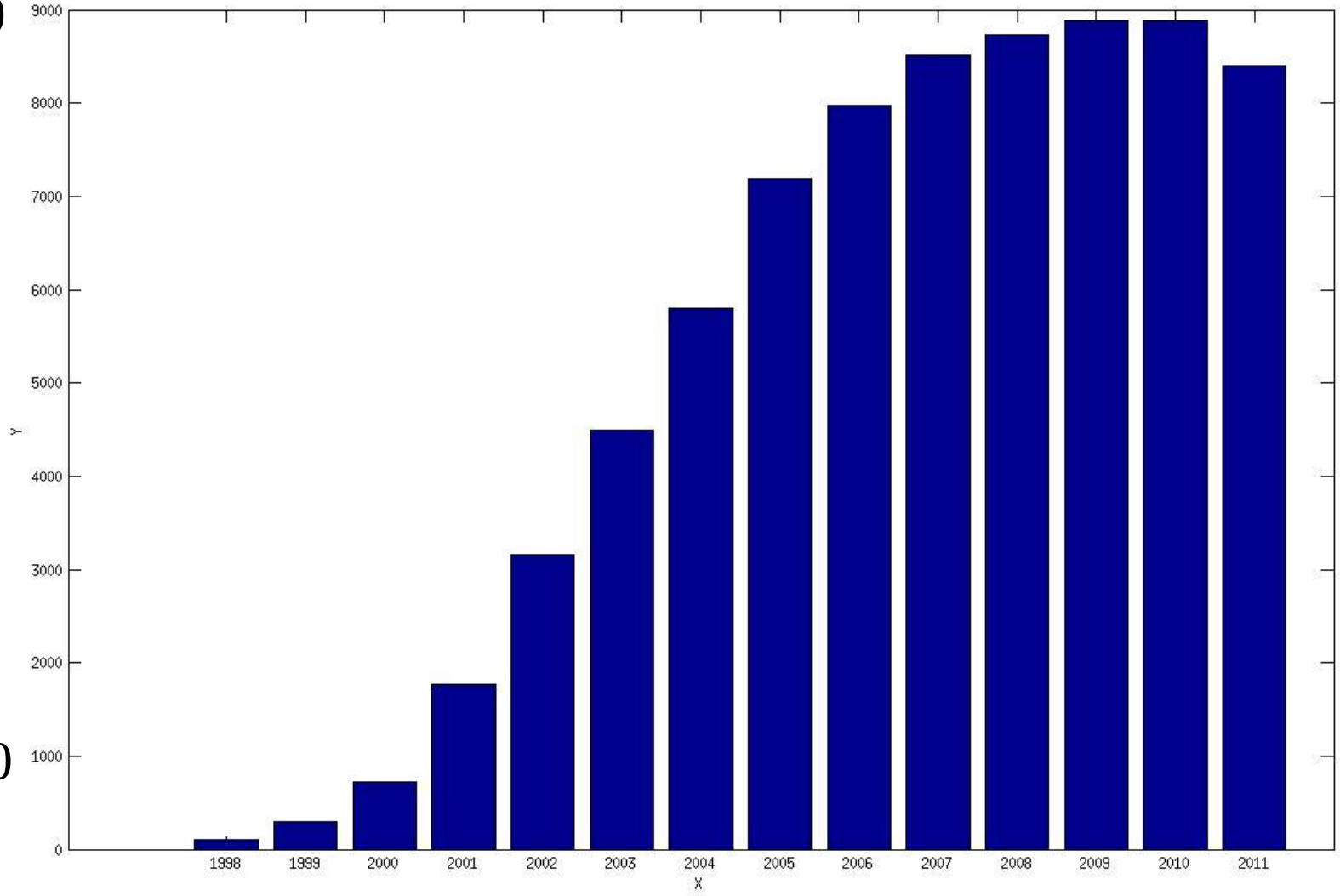
A DNA-CHIP MEASURES CONCENTRATIONS OF THOUSANDS OF DIFFERENT

MESSENGER RNA

HU Exon 1.0 ST – 1,425,647

NUMBER OF PUBLICATIONS (ISI WEB OF KNOWLEDGE) PER YEAR
SEARCH WORD: MICROARRAY

9000



1000

1998

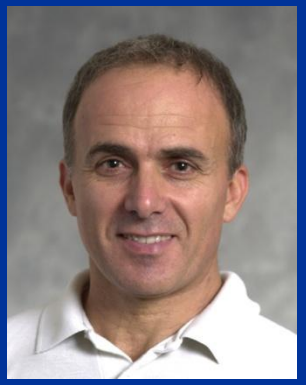
2011

1. DYNAMICS of TRANSCRIPTIONAL RESPONSE of CELLS to STIMULI

THE CAST (IN ORDER OF APPEARANCE):


Amit Zeisel


Wolfgang Köstler


Yossi Yarden


Roni Golan-Lavi


Rita Krauthgamer

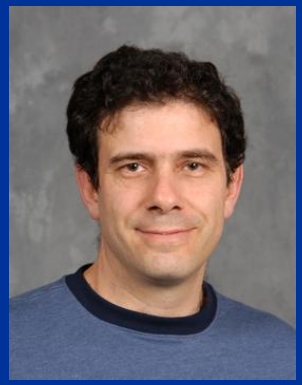

Steffen Jung

THE PLOT:

GENOME-WIDE MEASUREMENTS OF TRANSCRIPTION AND DEGRADATION RATES REVEAL COMPLEX TRANSCRIPT-SPECIFIC TEMPORAL VARIATIONS


Mattia Lauriola


Natali Molotski


Yoav Soen

Molecular Systems Biology 7, 529 (2011)

PHENOTYPIC RESPONSE OF CELLS TO STIMULUS

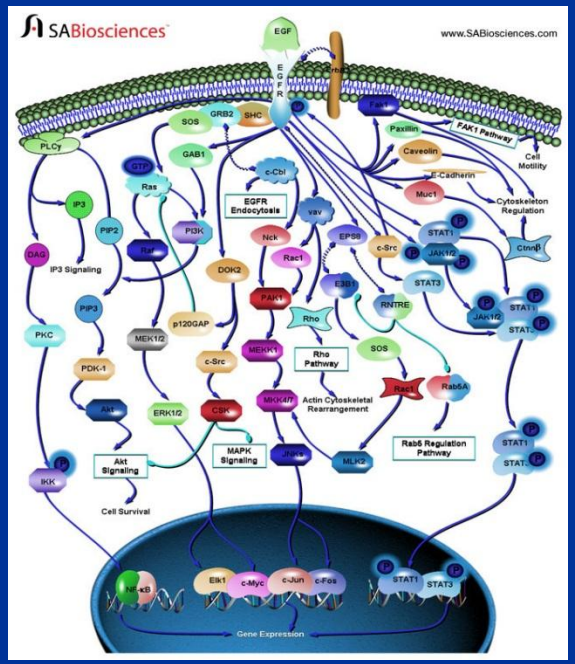
mmary
l cells)

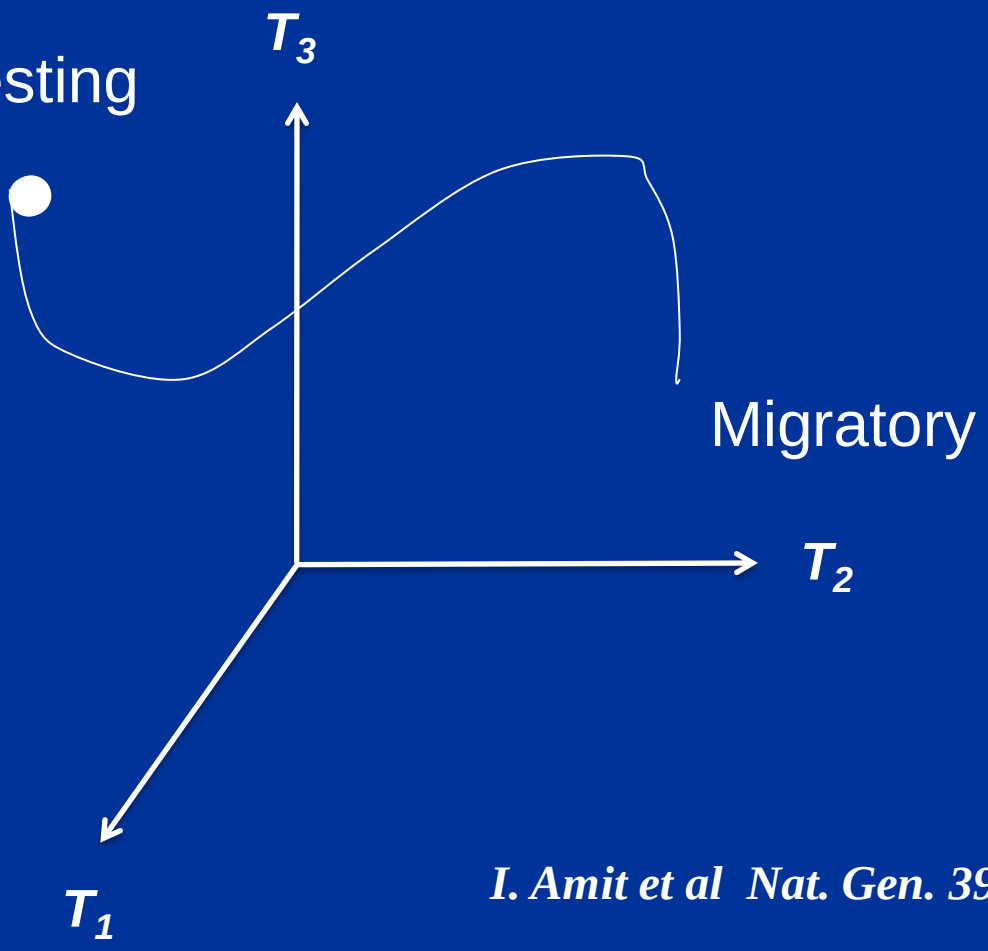
=0)

owth Factor:
ule (peptide)

Phenc
Char

TO CHANGE THEIR PHENOTYPE, CELLS MUST CHANGE THEIR TRANSCRIPTOME (RNA CONTENT)

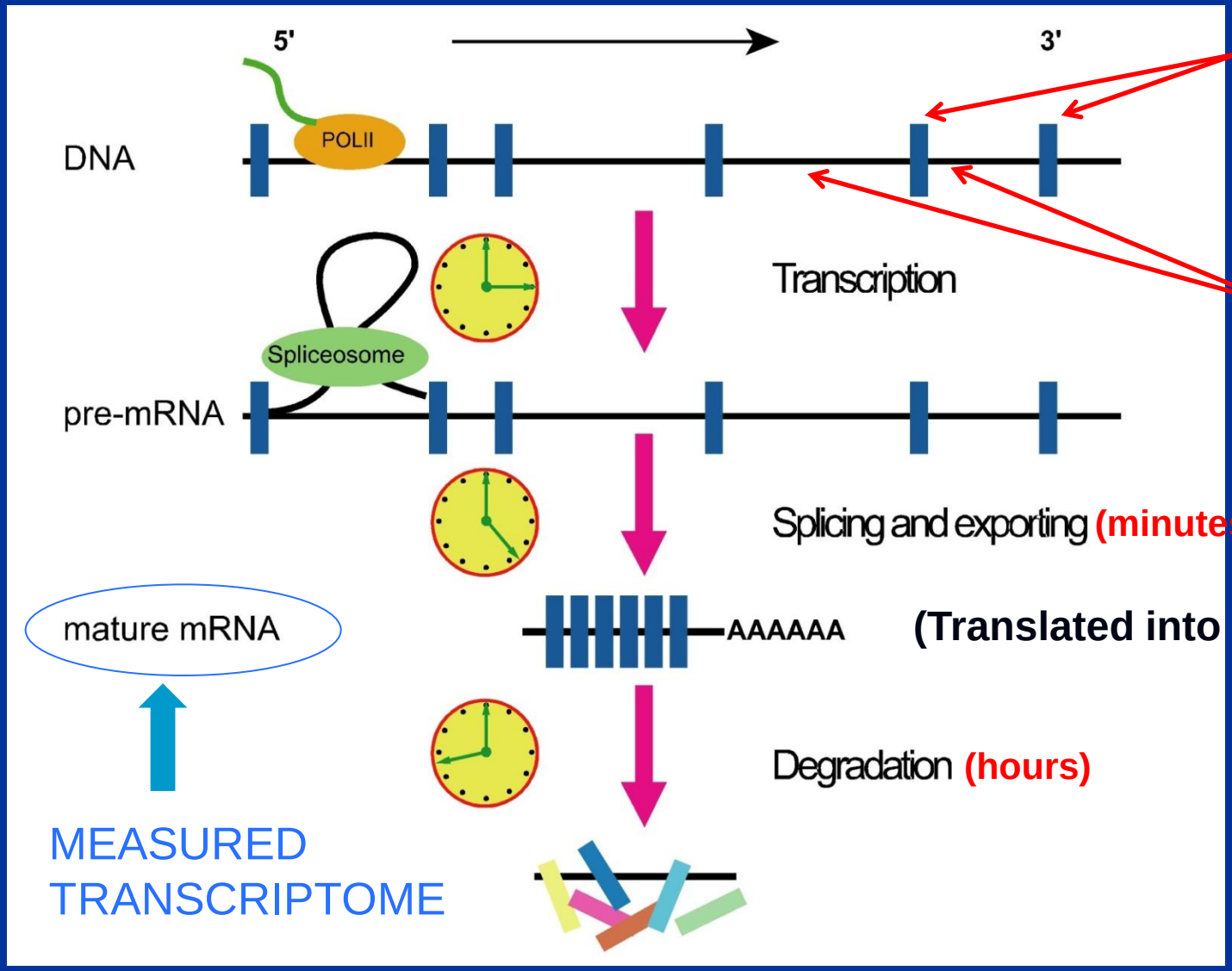




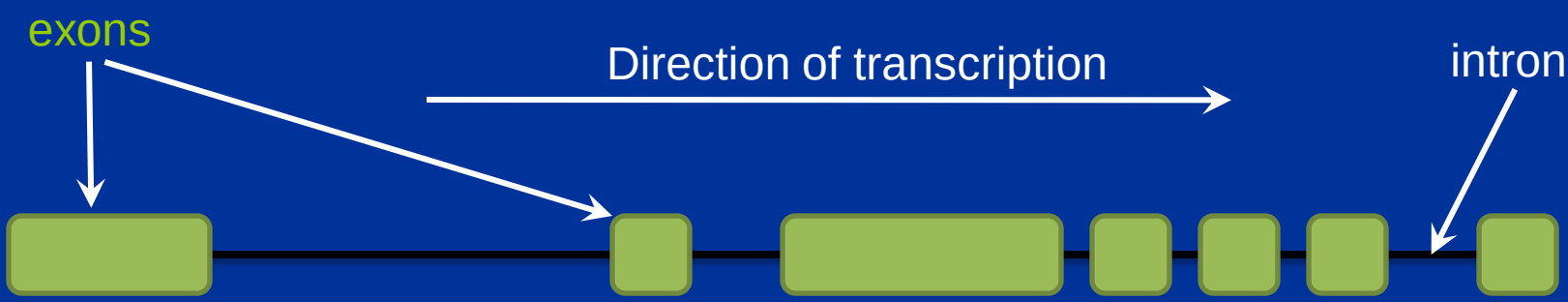
I. Amit et al Nat. Gen. 39, 503 (2007)

THE AIM OF OUR STUDY: CHARACTERIZE THE DYNAMICS OF TRANSCRIPTIONAL RESPONSE TO STIMULATION BY ***EGF***

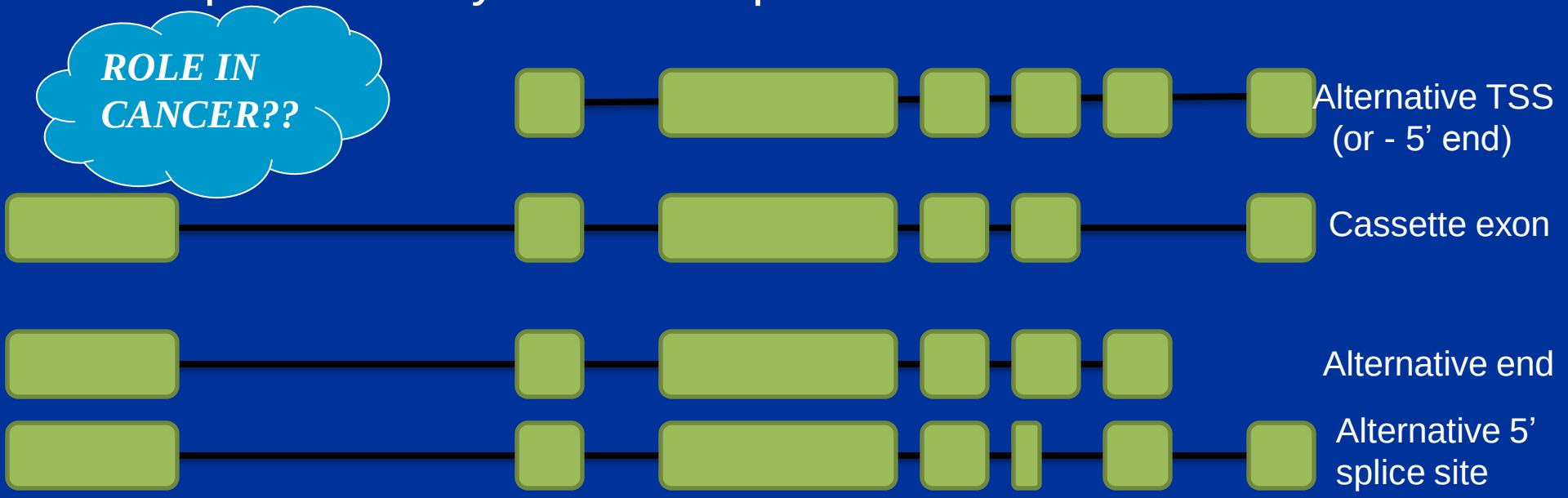
DETAILED LOOK AT TRANSCRIPTION: GOVERNED BY MULTIPLE DYNAMIC PROCESSES



IN FACT, THE AIM WAS TO STUDY ROLE OF ***TRANSCRIPT ISOFORM VARIATION*** IN THE PHENOTYPE

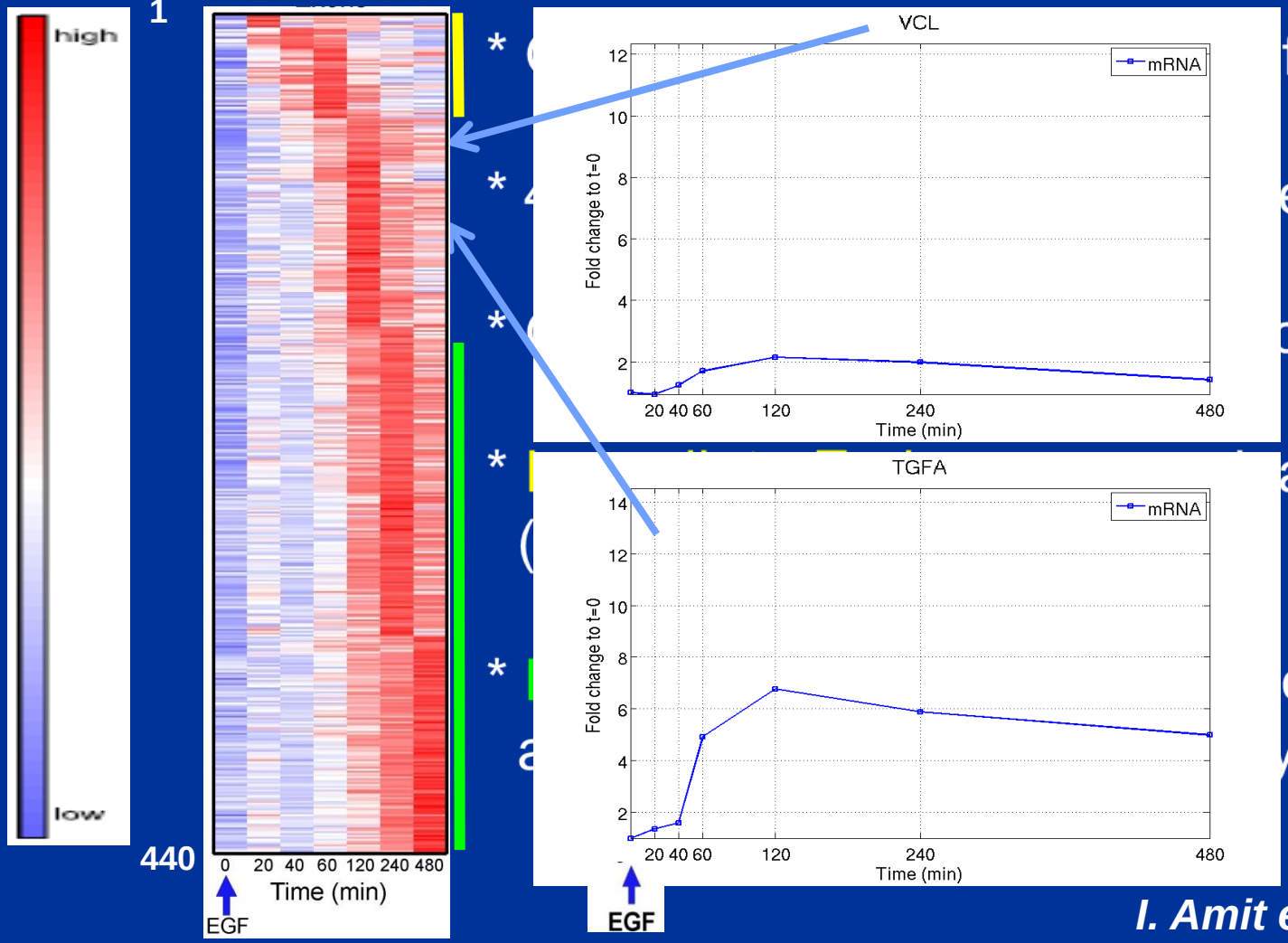


In principle all possible combinations of exons may exist.
In practice only few are expressed and observed.



TRANSCRIPTIONAL RESPONSE OF CELLS TO STIMULUS: THE TRANSCRIPTOME (mRNA)

The transcript levels (mRNA concentration) of genes change:



filing genome wide??

enes (FC>2)

of mRNA

at < 1 hour

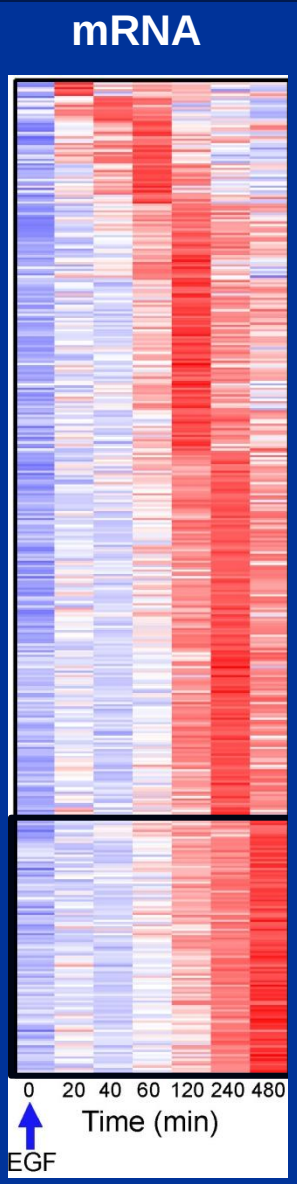
ours –

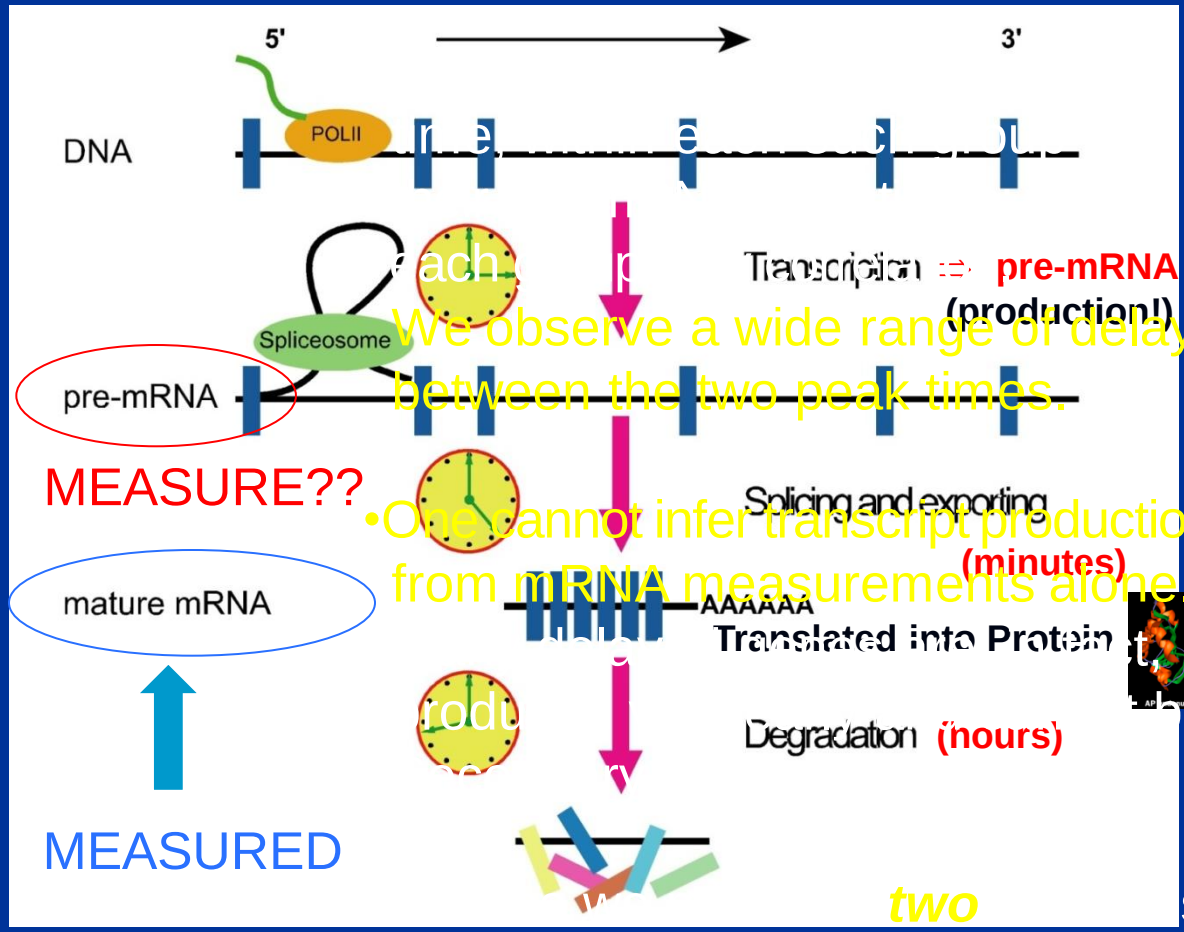
y Response?

I. Amit et al Nat Gen 2007

THE NOVELTY OF THIS STUDY:

INFER NETWORK FROM TRANSCRIPT DYNAMICS





We observe a wide range of delays between the two peak times.

•One cannot infer transcript production from mRNA measurements alone.

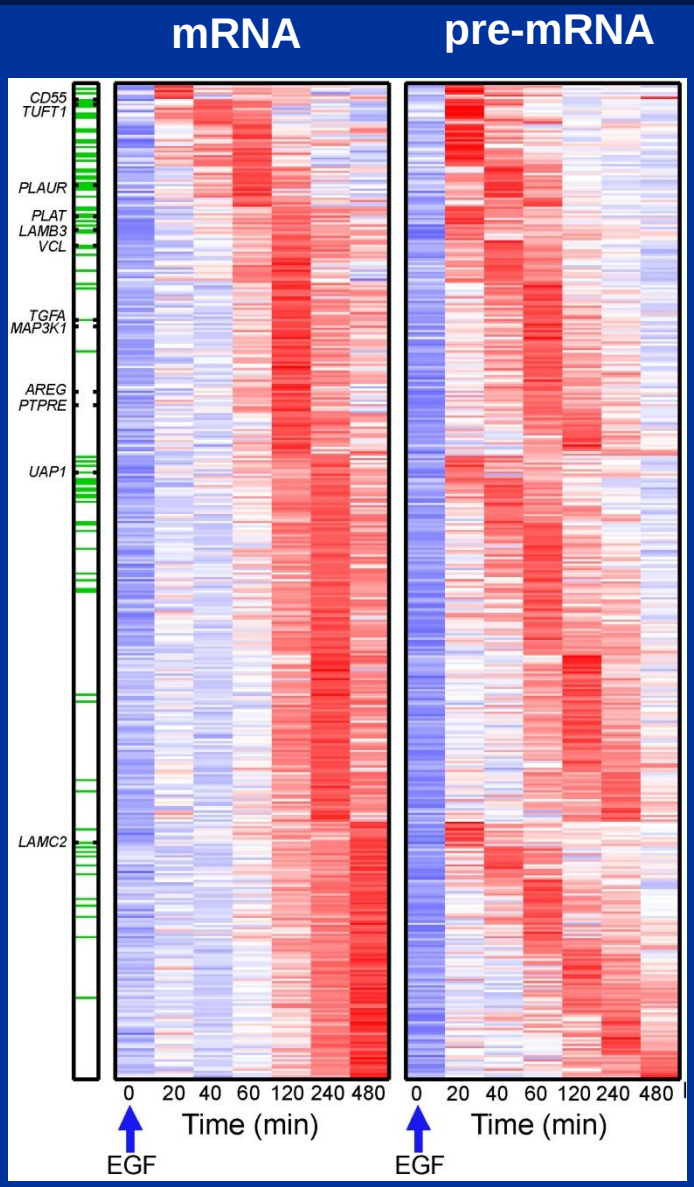
two

we can infer both production and degradation

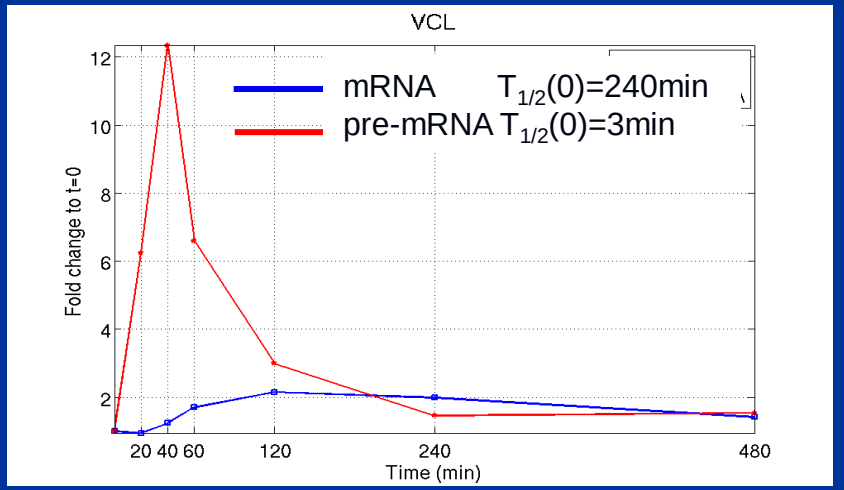
ANOTHER SURPRISE: PRODUCTION OVERSHOOT

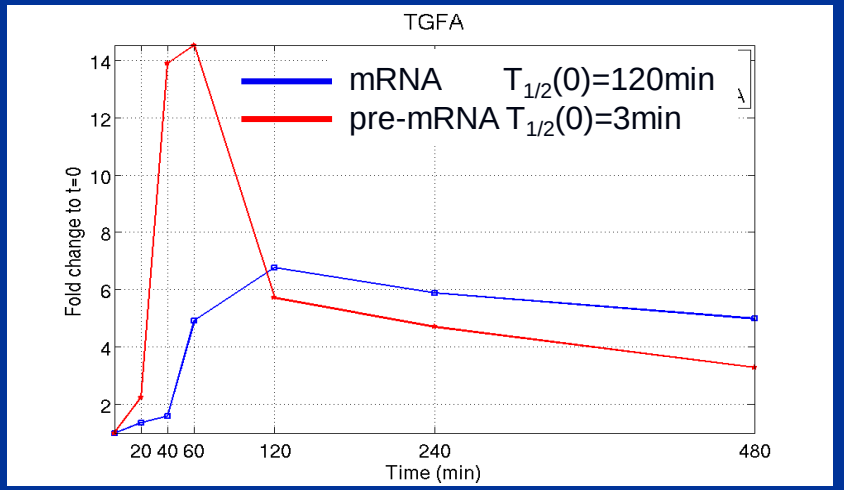
80 Genes exhibit production overshoot

Genes for qRT-PCR



Production overshoot:
Peak pre-mRNA Fold Change > > 2 X Peak mRNA Fold Change

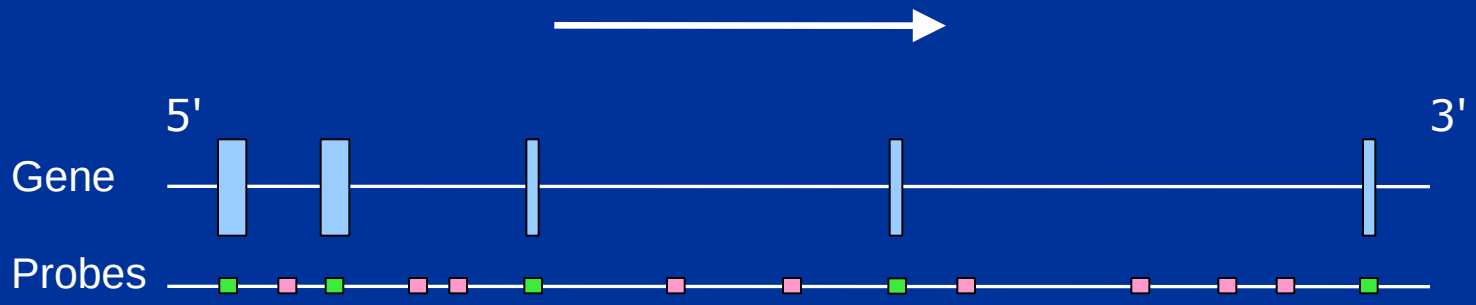




MEASURING *GENOME-WIDE* pre-mRNA and mRNA FOLD CHANGES USING EXON ARRAYS

3.measurements

Affymetrix exon arrays measure expression of gene regions encoding exons and/or introns



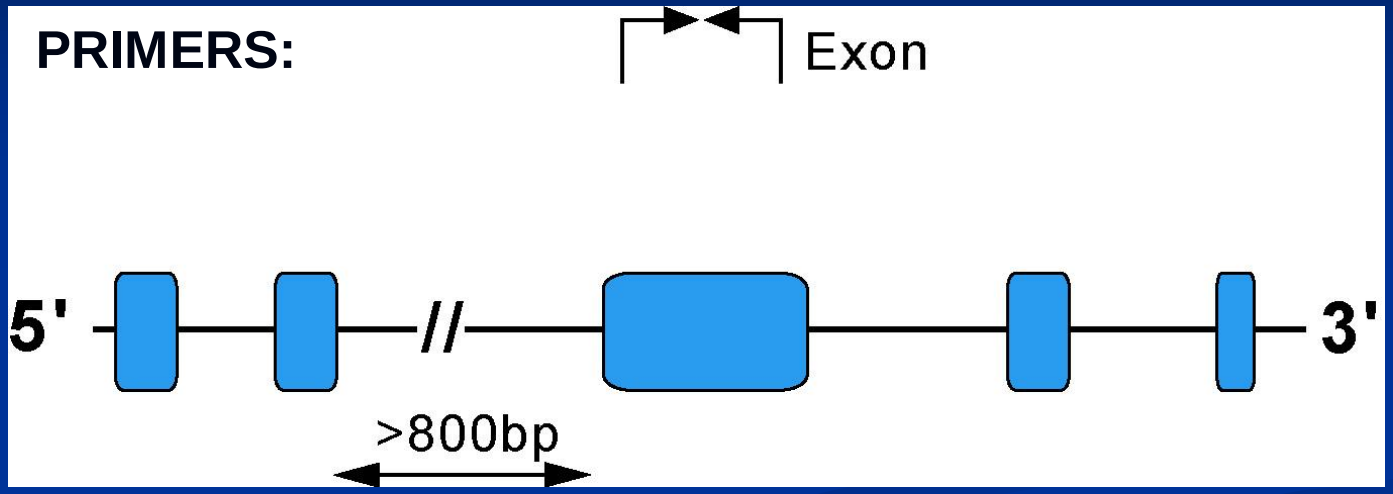
- Intronic probes measure pre-mRNA
- Exonic probes measure pre-mRNA A expression (but pre-mRNA < mRNA, so t NA)

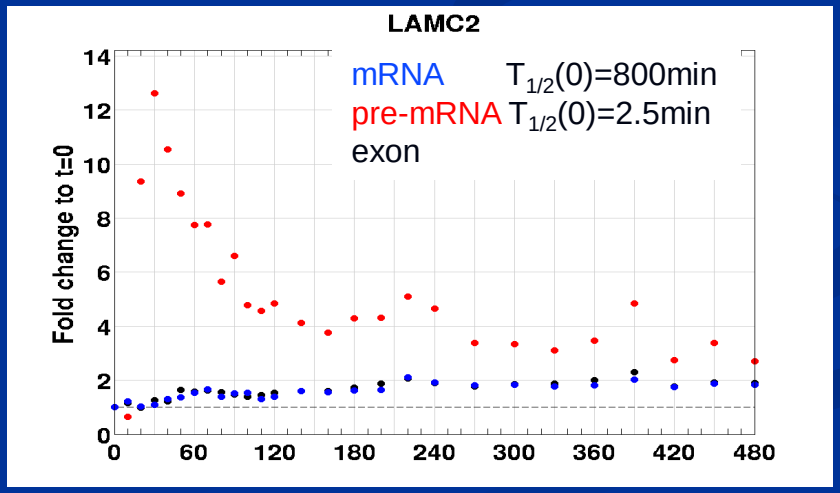
HU Exon 1.0 ST – 1,425,647

For selected transcripts we measured pre-mRNA, mRNA and exonic signals directly by q-RTPCR

FOR SELECTED TRANSCRIPTS: RT-PCR MEASUREMENTS AT CLOSELY SPACED TIMEPOINTS AFTER STIMULUS

3.measurements

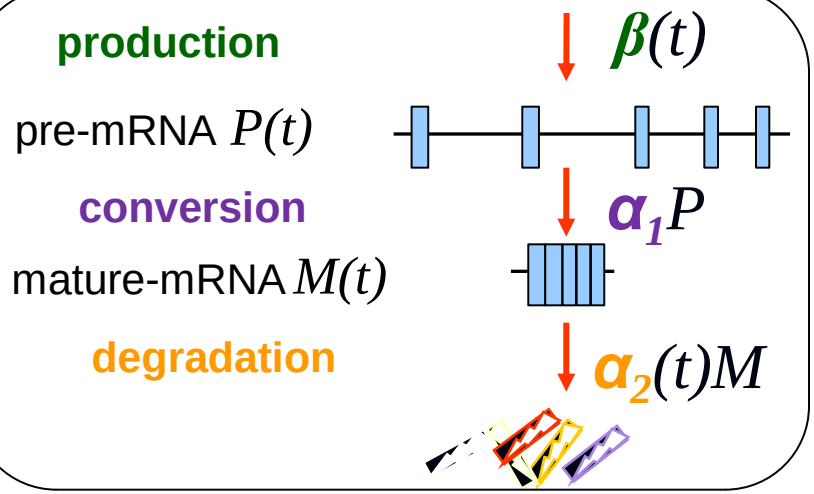




“REFINED MODEL” for DYNAMICS of TRANSCRIPTIONAL RESPONSE TO STIMULUS

$$\frac{dP(t)}{dt} = \beta(t) - \alpha_1 P(t)$$

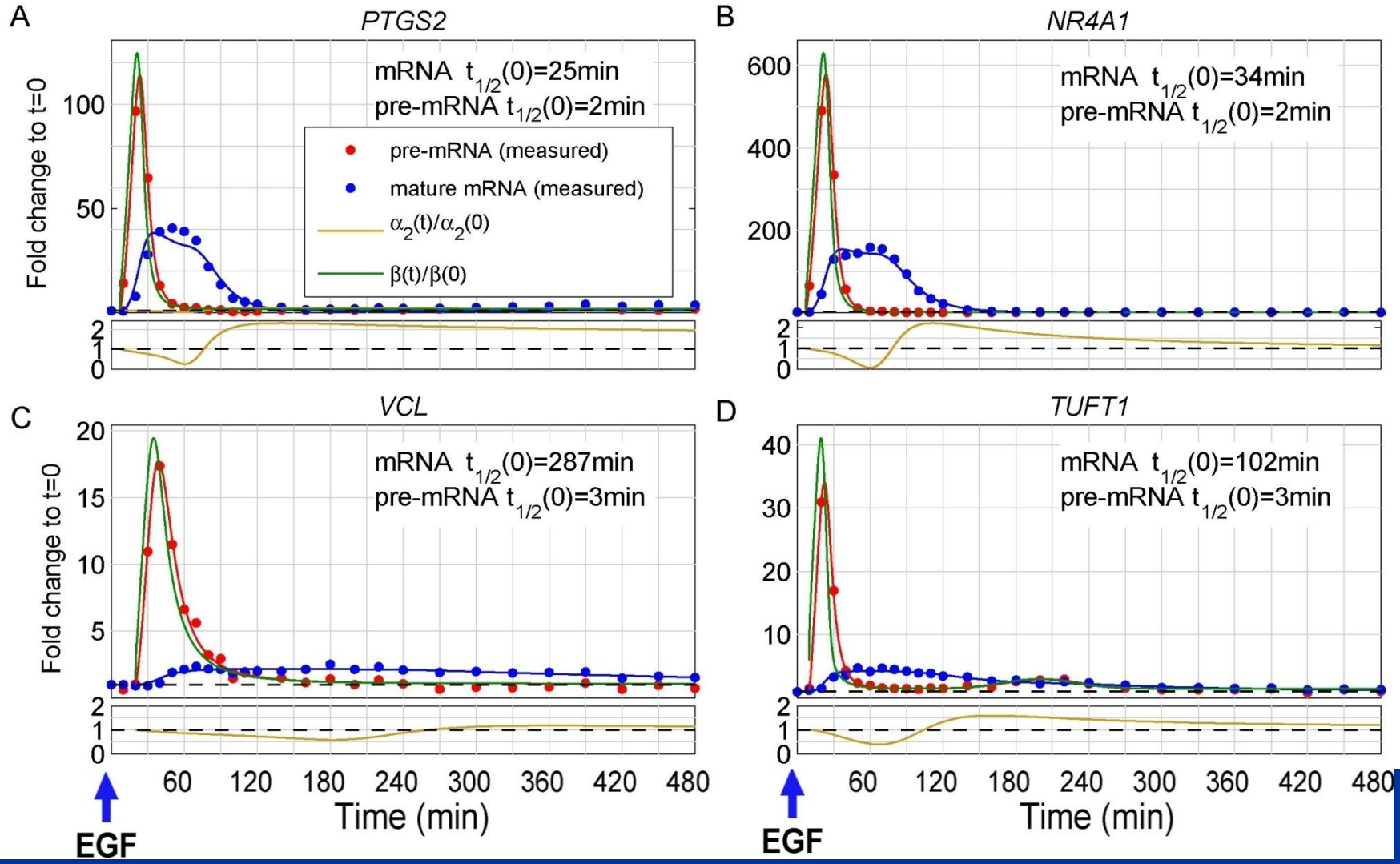
$$\frac{dM(t)}{dt} = \alpha_1 P(t) - \alpha_2(t) M(t)$$



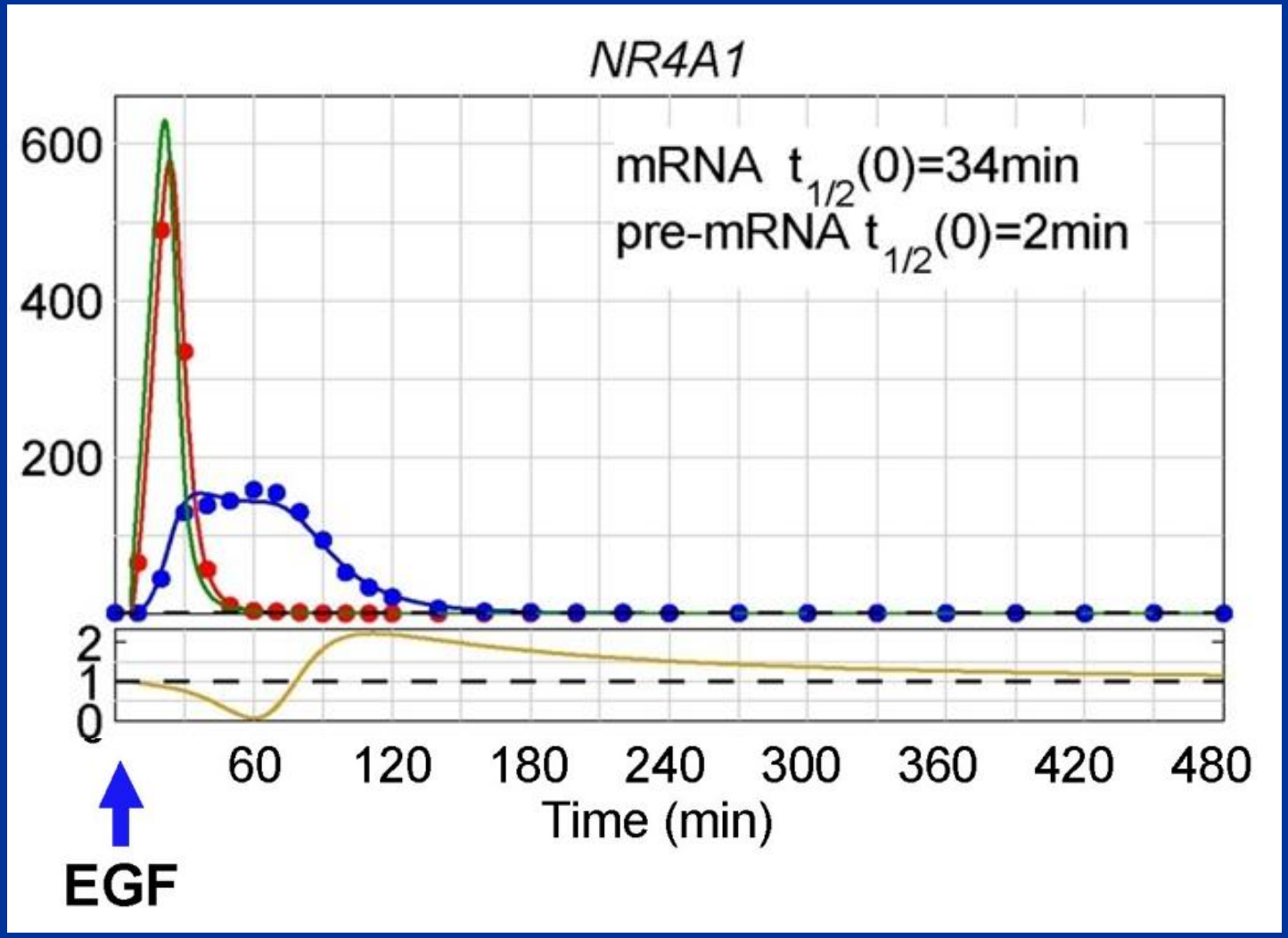
WE MEASURE $P(t)$ and $M(t)$; THEN USE THE EQUATIONS TO INFER TIME-

- DEPENDENT PRODUCTION $\beta(t)$ **and** TIME-DEPENDENT DEGRADATION $\alpha_2(t)$

REVERSE THE EQUATIONS: FROM pre-mRNA and mRNA DATA INFER PRODUCTION $\hat{\beta}(t)$ and DEGRADATION $\hat{\alpha}_2(t)$



WE DISCOVERED PRONOUNCED TRANSCRIPT-DEPENDENT COMPLEX TEMPORAL VARIATION OF DEGRADATION!!



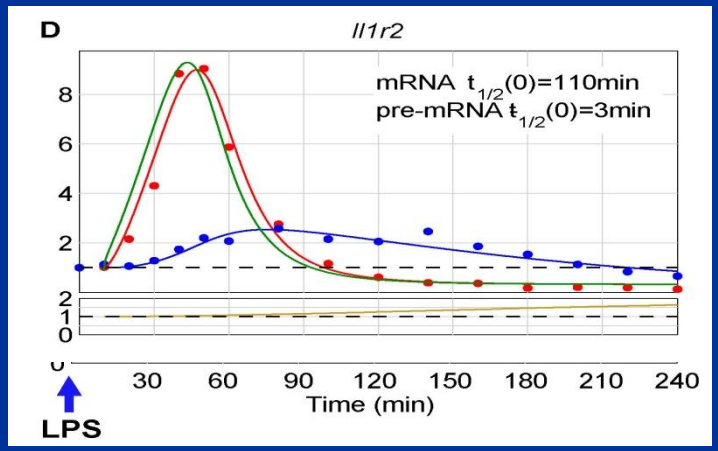
OPERATIONAL STRATEGY:

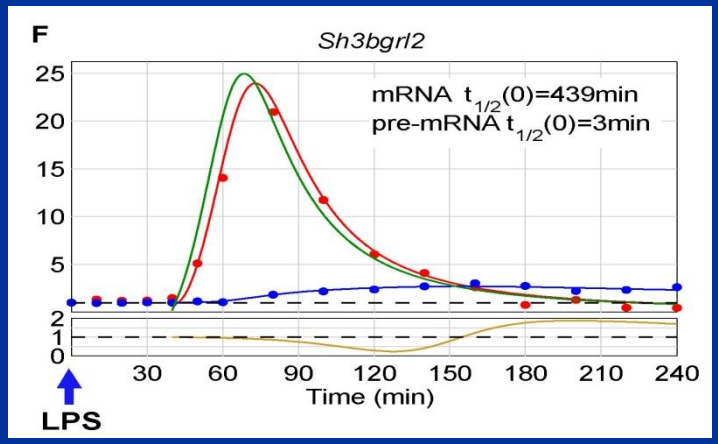
STRATEGY: PRODUCTION OVERSHOOT and TRANSIENT STABILIZATION ACCELERATE INDUCTION OF mRNA:

"SLAM DOWN" ALL THE WAY FOR A SHORT TIME, TO BRING THE TRANSCRIPT FAST TO THE DESIRED VALUE

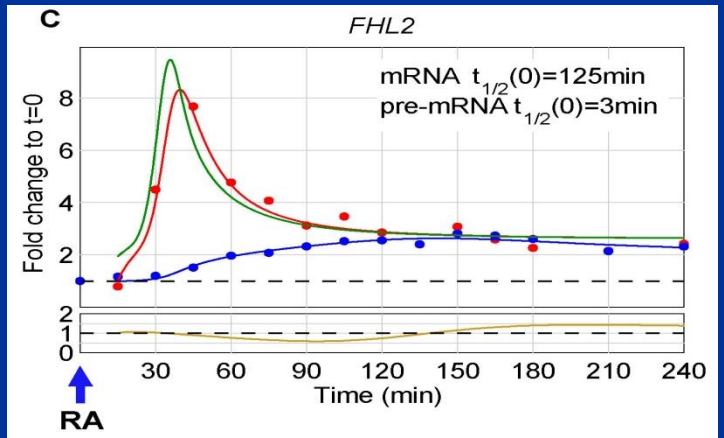
UNIVERSALITY: IS THIS STRATEGY USED BY OTHER TYPES OF MAMMALIAN CELLS AND STIMULI ?

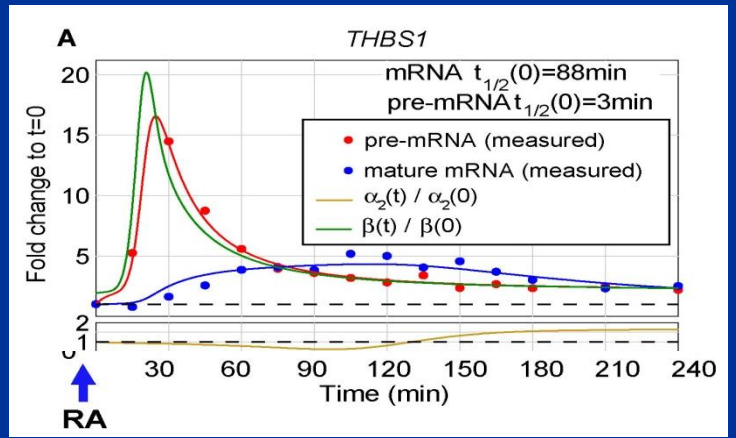
1. MURINE DENDRITIC CELLS (FROM BONE MARROW) RESPOND TO STIMULATION BY ***L***ipo***P***oly***S***accharide – INFLAMMATION & MATURATION





2. HUMAN EMBRYONIC STEM CELLS DIFFERENTIATE INTO NEURAL PROGENITORS IN RESPONSE TO RETINOIC ACID.



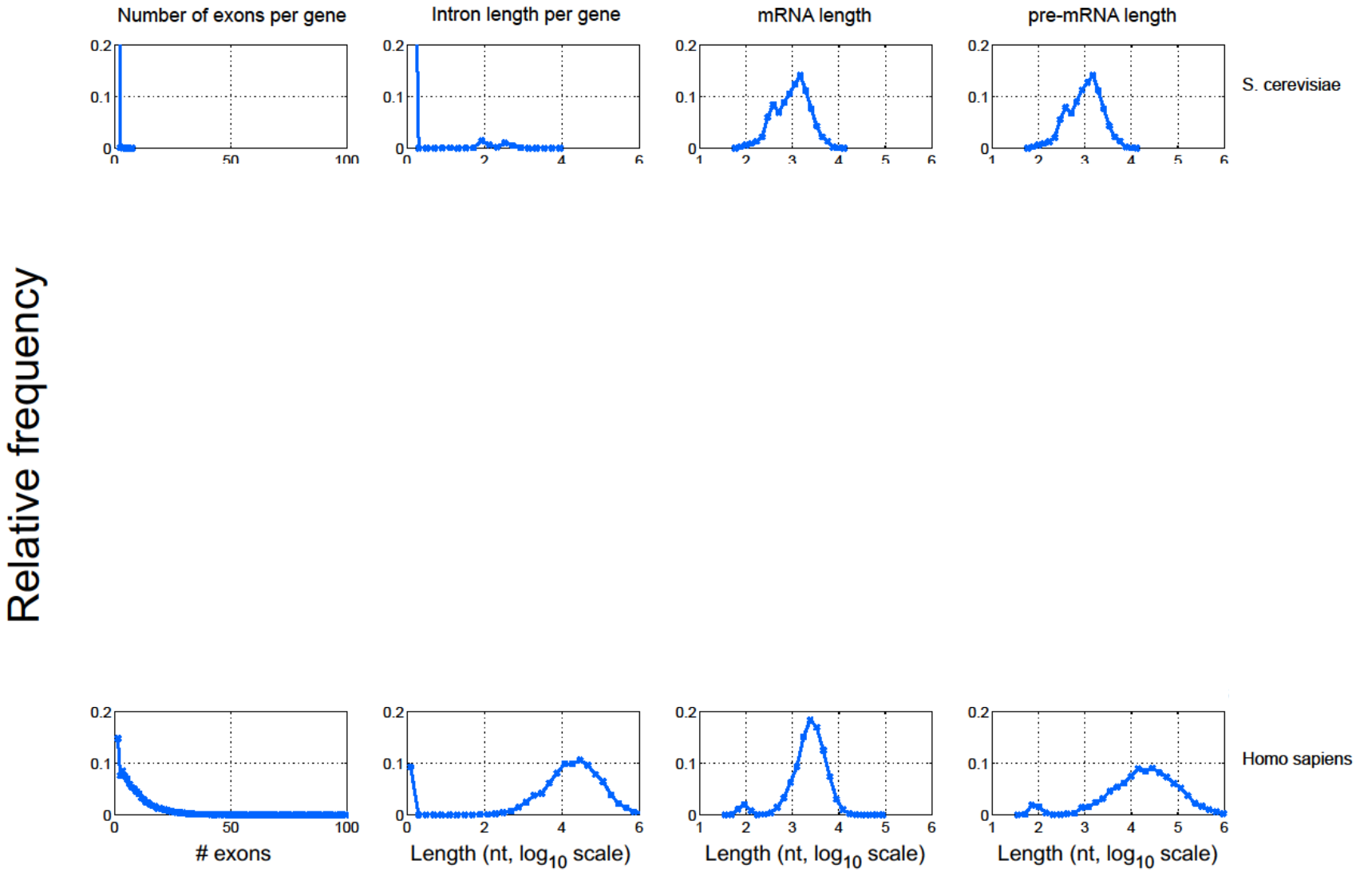


SUMMARY

Molecular Systems Biology 7, 529 (2011)

- Genome-wide measurement of pre-mRNA dynamics reveals lack of correspondence between *mRNA* and *pre-mRNA* profiles. *Must* measure *both* to infer production *and* degradation.
- A simple model for the coupled dynamics of pre-mRNA and mRNA allows inference of time-dependent production and degradation during transcriptional response to stimuli.
- An operational strategy involves *Production Overshoot* together with *Transient Stabilization*: it accelerates mRNA response, allows transcript-specific *control* of the timing and amplitude of the mRNA abundance profile
- This strategy is used by several cellular mammalian systems, in response to different stimuli.

OTHER ANIMALS

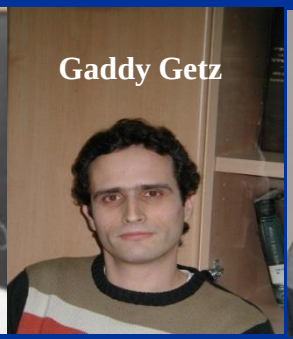


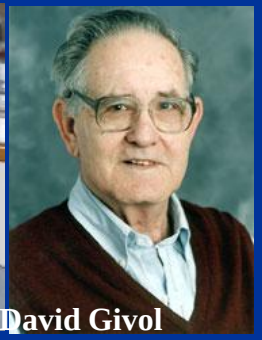
RELEVANCE TO CANCER (SPECULATION)

WE FOUND CONSIDERABLE EVIDENCE FOR GENE SPECIFIC CONTROL OF ***DEGRADATION*** (SO FAR THE CHANGE OF TRANSRIPTOME WAS ATTRIBUTED TO CHANGES IN PRODUCTION). SEARCH FOR ***DEGRADATION FACTORS***

2. OUTCOME PREDICTION IN BREAST CANCER: HOPE, HYPE, PHYSICS AND BIOLOGY

THE CAST:



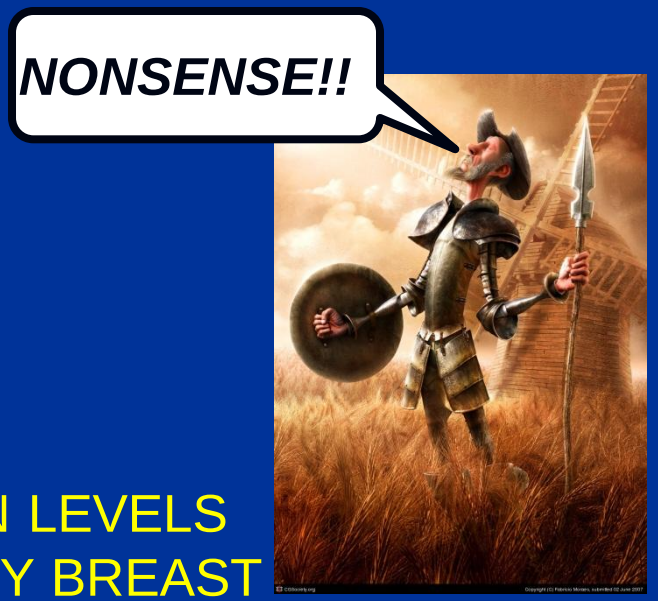


THE PLOT:

Gene expression profiling predicts clinical outcome of breast cancer

Laura J. van 't Veer*†, Hongyue Dai†‡, Marc J. van de Vijver*†, Yudong D. He‡, Augustinus A. M. Hart*, Mao Mao‡, Hans L. Peterse*, Karin van der Kooy*, Matthew J. Marton‡, Anke T. Witteveen*, George J. Schreiber‡, Ron M. Kerkhoven*, Chris Roberts‡, Peter S. Linsley‡, René Bernards* & Stephen H. Friend‡

Nature 2002



CLAIM: FOUND 70 GENES, WHOSE EXPRESSION LEVELS CAN PREDICT WHETHER AN EARLY DISCOVERY BREAST TUMOR IS AGGRESSIVE OR LOW RISK

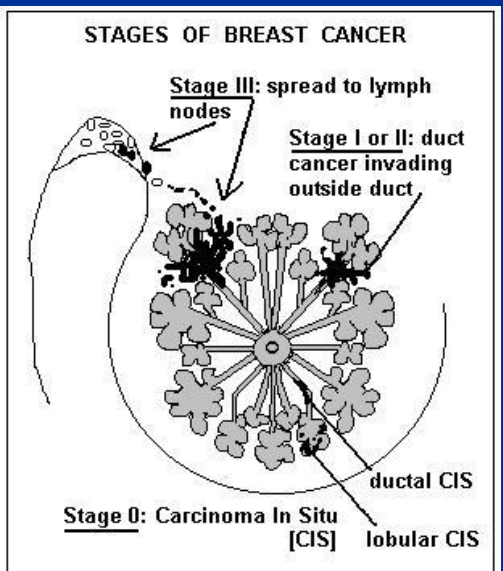
BREAST CANCER:

DEATH RATE 30/100,000 per year

INCIDENCE: ABOUT 1 OUT OF 9 WOMEN AFFECTED.

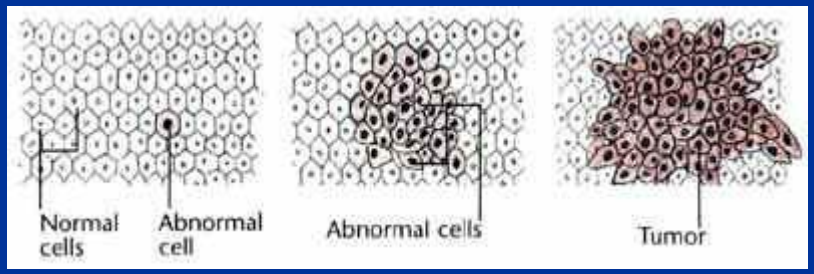
EARLY DISCOVERY: SMALL TUMOR (< 2cm), HAS NOT SPREAD TO LYMPH NODES, LOWEST GRADE, STAGE

TREATMENT: SURGICAL REMOVAL OF TUMOR + RADIOTHERAPY + HORMONAL THERAPY IF ER+ (or PgR+) +Herceptin



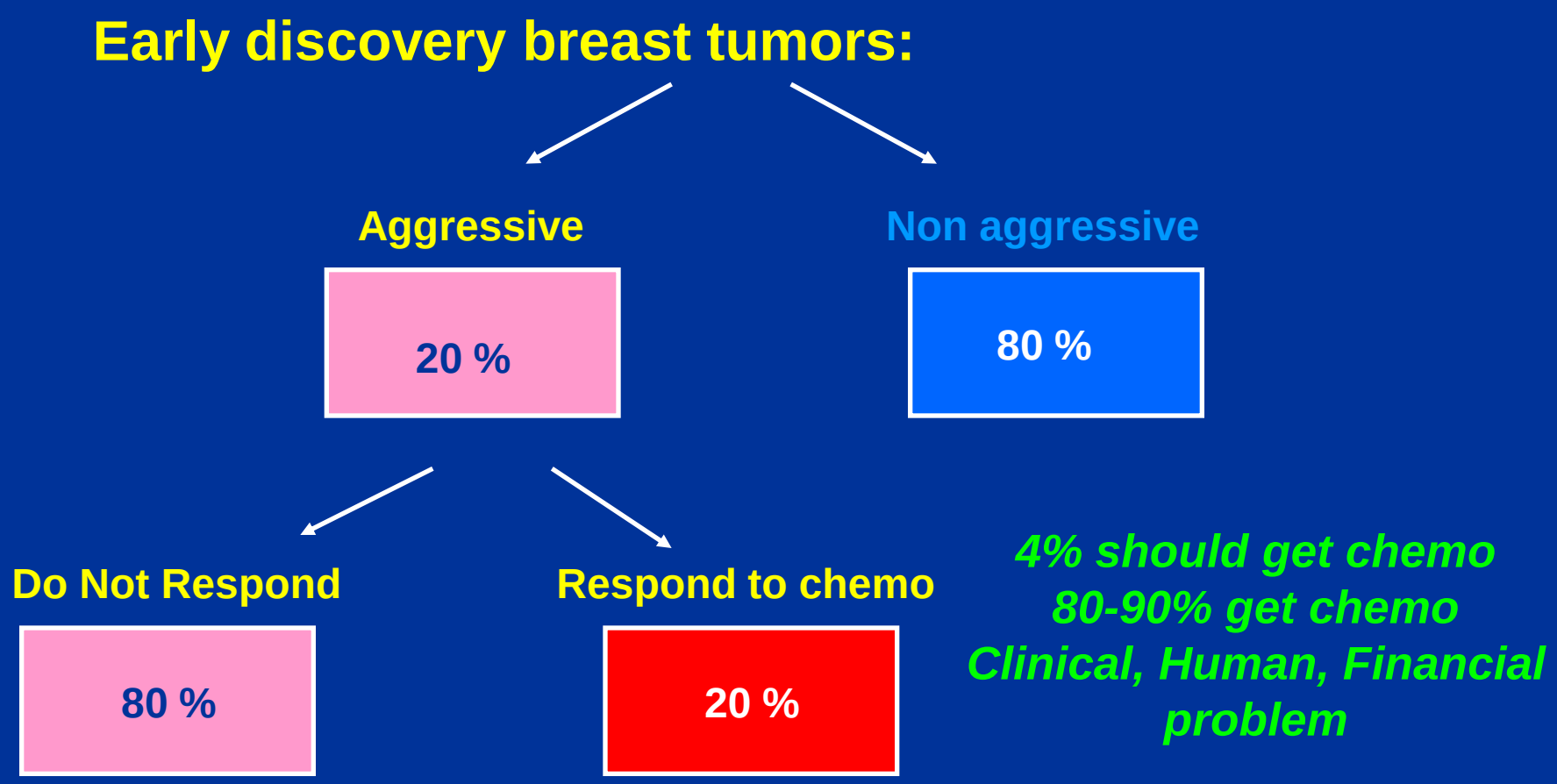
CHEMOTHERAPY ??? No CHEMO if Low Risk

DECISION Yes/No -TAKEN ON THE BASIS OF CLINICAL PARAMETERS: NIH, St Gallen, NPI CRITERIA



GRADES 1,2,3

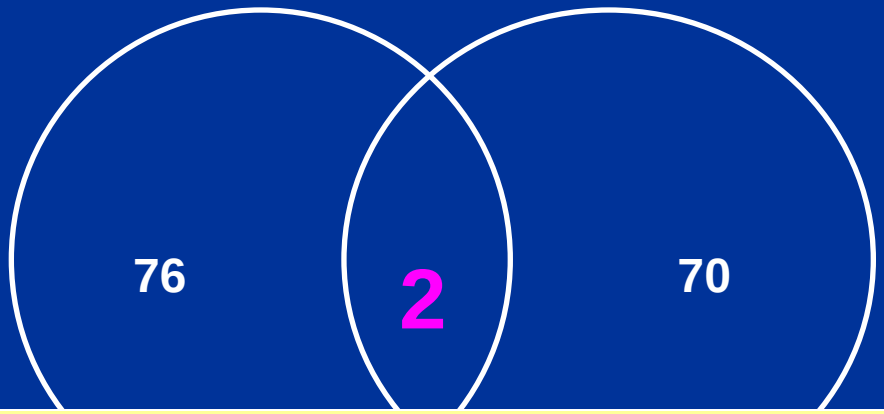
NIH, St GALLEN; HOW WELL DO THESE CRITERIA WORK?



Can we do better in identifying patients at high risk –
and avoid chemotherapy for low-risk? Use expression profiling of tumors

A SUCCESSFUL GENE EXPRESSION BASED ANALYSIS: ANOTHER ONE:

Wang et al. Lancet 2005, List = 76 genes (Rotterdam Signature)



Van't Veer. et al. Nature 2002, List = 70 genes (Amsterdam signature)

VERY SMALL OVERLAP!!! POOR TRANSFERABILITY!!!
WHY ???

Different Platforms !

Different Populations of Patients !

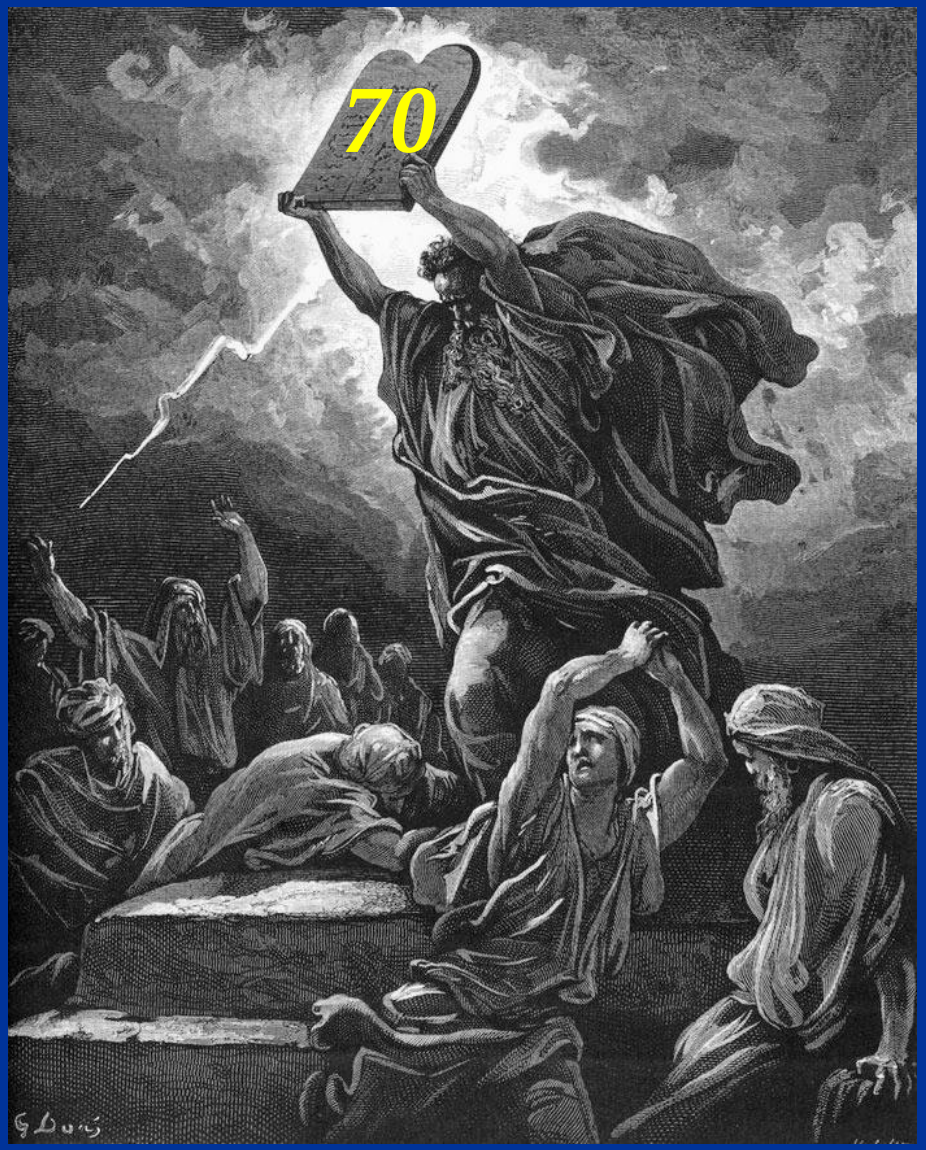
Different Types of Analysis!

NO!!

FOCUS ON THE 70-GENE "AMSTERDAM SIGNATURE" (*MammaPrint*)

1. WHY ***70*** GENES?
2. WHY ***THESE*** 70 GENES?
3. ***HOW WELL*** DOES THE PROGNOSTIC CLASSIFIER ***WORK***?
4. CAN DIFFERENT SIGNATURES BE INTERPRETED IN TERMS OF SAME ***CLINICAL/BIOLOGICAL PROCESSES***?
5. HAS ANY NEW ***KNOWLEDGE*** BEEN GAINED BY THESE 70 GENES?

1. WHY *70* GENES? (OUT OF 5000 CANDIDATES)



1. WHY *70* GENES? (OUT OF 5000 CANDIDATES)

19 left for test set

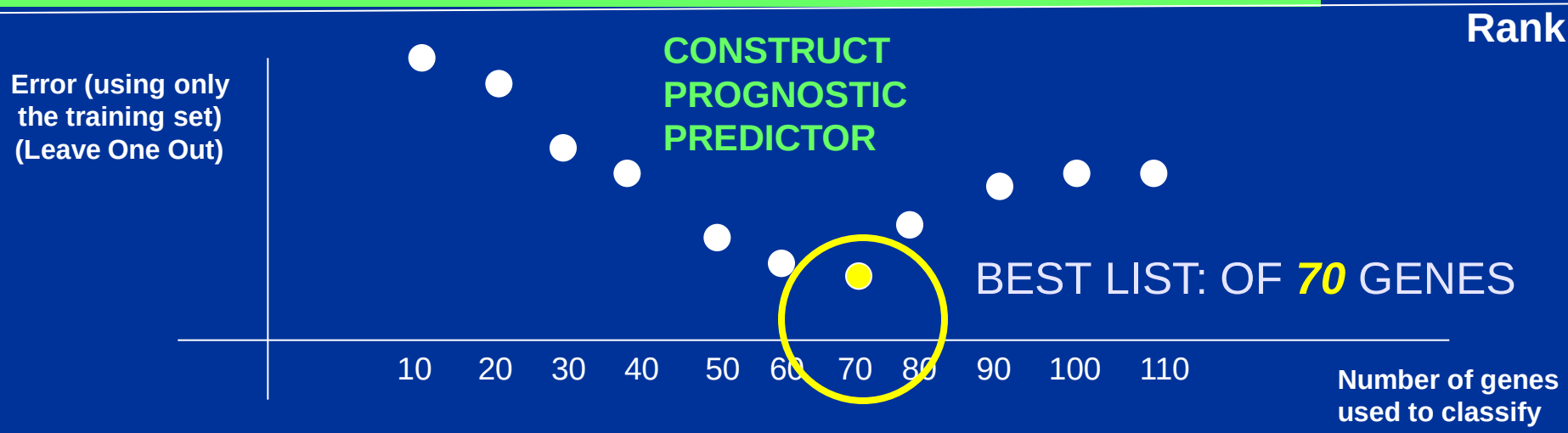
MEASURED EXPRESSION FOR ***97*** TUMORS; WAIT 5 YEARS FOR OUTCOME
78 OF THESE WERE SELECTED ***AT RANDOM*** AS "TRAINING SET",
TO CALCULATE

C(g) = PREDICTIVE POWER OF EACH GENE ***g*** FOR THESE ***78*** TUMORS
= CORRELATION OF EXPRESSION OF GENE ***g*** WITH
OUTCOME OVER THESE 78 TUMORS

RANK THE GENES *g=1,2,…5000* ON THE CHIP BY THEIR ***C(g):***

1 . . . 10 ….20 ….30 ….40 ….50 ….60 ….70 ….80 ….90 ….100 ….110, ….., 5000

Rank



FOCUS ON THE 70-GENE "AMSTERDAM SIGNATURE" (*MammaPrint*)

1.WHY ***70*** GENES? ***FOR NO GOOD REASON (78 SAMPLES)***

THE NUMBER OF "PROGNOSTIC GENES" IS GOVERNED BY THE NUMBER OF SAMPLES THAT WERE USED FOR GENE SELECTION, (i.e. WERE AVAILABLE IN 2002).

2. WHY ***THESE*** 70 GENES?

3. ***HOW WELL*** DOES THE PROGNOSTIC CLASSIFIER ***WORK***?

4. CAN DIFFERENT SIGNATURES BE INTERPRETED IN TERMS OF SAME ***CLINICAL/BIOLOGICAL PROCESSES***?

5. HAS ANY NEW ***KNOWLEDGE*** BEEN GAINED BY THESE 70 GENES?

2. WHY ***THESE*** 70 GENES?

1 . . . 10 ….20 ….30 ….40 ….50 ….60 ….70 ….80 ….90 ….100 ….110, ….., 5852

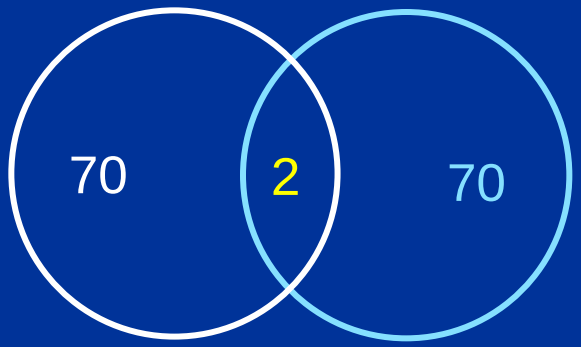
Rank

THESE 70 GENES WERE ***TOP RANKED*** ON THE BASIS OF THEIR PREDICTIVE POWER, CALCULATED USING EXPRESSION DATA FROM ***78*** SAMPLES. THESE WERE ***RANDOMLY SELECTED*** (OUT OF 97).

IS THE RESULTING PROGNOSTIC LIST ROBUST/REPRODUCIBLE?

REPEAT THE PROCESS WITH ANOTHER SET OF ***78***/97 "AMSTERDAM" SAMPLES -- GET A DIFFERENT GROUP OF 70 "TOP--RANKED" GENES.

TO GET TWO SIMILAR TOP-RANKED GROUPS OF 70, ONE MUST USE 2400 SAMPLES FOR RANKING THE GENES



SET #1 OF TOP 70 GENES

SET #2 OF TOP 70 GENES

BY RANDOMLY PICKING (FROM THE SAME POOL!) A DIFFERENT GROUP OF SAMPLES FOR "TRAINING", WE GET A VERY DIFFERENT LIST OF TOP-RANKED GENES

THE TWO LISTS WERE GENERATED USING THE SAME PATIENTS, PLATFORM, METHOD!

PHYSICS/MATH

Ein-Dor et al PNAS (2006)

Ein-Dor et al Bioinformatics (2005)

Michiels et al Lancet (2005)

FOCUS ON THE 70-GENE "AMSTERDAM SIGNATURE" (*MammaPrint*)

1.WHY ***70*** GENES? ***FOR NO GOOD REASON (78 SAMPLES)***

2. WHY ***THESE*** 70 GENES? ***RANDOMLY SELECTED (out of ~ 1000)***

RANKING ON THE BASIS OF A SMALL NUMBER OF SAMPLES IS AN EXTREMELY NOISY UNSTABLE PROCESS.
HENCE -- THESE ARE 70 ***RANDOMLY SELECTED*** GENES.

PERHAPS THESE 70 RANDOM GENES GIVE BEST PROGNOSIS? (BETTER THAN OTHER RANDOMLY CHOSEN SETS OF 70 GENES)

3. ***HOW WELL*** DOES THE PROGNOSTIC CLASSIFIER ***WORK***?

4. CAN DIFFERENT SIGNATURES BE INTERPRETED IN TERMS OF SAME ***CLINICAL/BIOLOGICAL PROCESSES***?

5. HAS ANY NEW ***KNOWLEDGE*** BEEN GAINED BY THESE 70 GENES?

MAYBE THESE 70 RANDOM GENES GIVE BEST PROGNOSIS?

Van't Veer

SELECTED 70 GENES IS SIMILAR TO THAT
OF MOST OTHER RANDOM SETS OF 70 GENES

Time to distant metastasis (months)

Ein-Dor et al Bioinformatics 21:171 (2005)

FOCUS ON THE 70-GENE "AMSTERDAM SIGNATURE" (*MammaPrint*)

1. WHY ***70*** GENES? ***FOR NO GOOD REASON (78 SAMPLES)***

2. WHY ***THESE*** 70 GENES? ***RANDOMLY SELECTED***

PERHAPS THESE 70 RANDOM GENES GIVE BEST PROGNOSIS? (BETTER THAN OTHER RANDOMLY CHOSEN SET OF 70 GENES) ***No!***

3. ***HOW WELL*** DOES THE PROGNOSTIC CLASSIFIER ***WORK?***

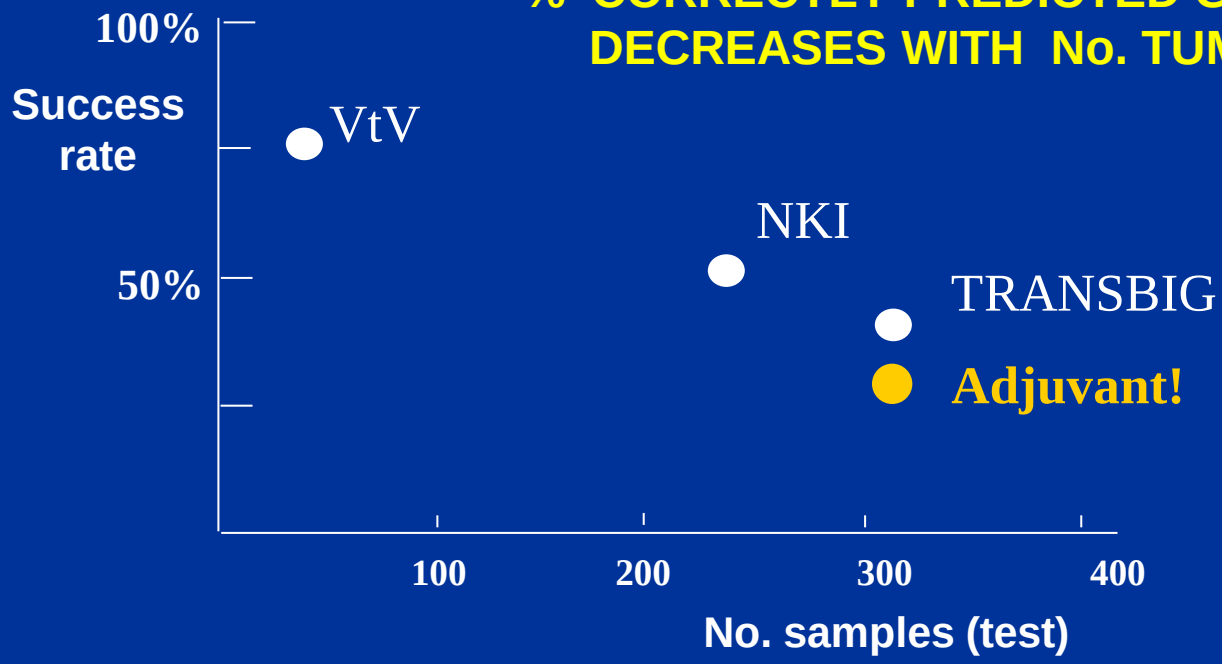
4. CAN DIFFERENT SIGNATURES BE INTERPRETED IN TERMS OF SAME ***CLINICAL/BIOLOGICAL PROCESSES***?

5. HAS ANY NEW ***KNOWLEDGE*** BEEN GAINED BY THESE 70 GENES?

3. HOW WELL DOES THE PROGNOSTIC CLASSIFIER *WORK?*

REQUIRE : % CORRECTLY PREDICTED BAD OUTCOME **> 90%**
AND MEASURE % OF CORRECTLY PREDICTED GOOD OUTCOME,

% CORRECTLY PREDICTED GOOD OUTCOME DECREASES WITH No. TUMORS TESTED



Michiels et al Lancet 2005
Dupuy and Simon JNCI 2007

SUMMARY OF THE 70-GENE "AMSTERDAM SIGNATURE"

1. WHY ***70*** GENES? ***FOR NO GOOD REASON (78 SAMPLES)***

2. WHY ***THESE*** 70 GENES? ***RANDOMLY SELECTED***

3. ***HOW WELL*** DOES THE PROGNOSTIC CLASSIFIER ***WORK***?
   PERFORMANCE GETS WORSE AND APPROACHES "CLASSICAL" AS MORE TUMORS ARE TESTED
4. CAN DIFFERENT SIGNATURES BE INTERPRETED IN TERMS OF SAME ***CLINICAL/BIOLOGICAL PROCESSES***? ***ONLY PROLIFERATION***
   (GENE SET ENRICHMENT ANALYSIS) ***MITOTIC INDEX!!***
5. HAS ANY NEW ***KNOWLEDGE*** BEEN GAINED BY THESE 70 GENES? ***NO***

1,2 Ein-Dor et al Bioinformatics 21:171 (2005)
2 Ein-Dor, Zuk, Domany, PNAS 103:5923 (2006)
4 Drier & Domany PLoS ONE 6:e17795 (2011)

"GO" BIOLOGICAL PROCESSES – NESTED??

SUMMARY OF THE 70-GENE "AMSTERDAM SIGNATURE"

1. WHY ***70*** GENES? ***FOR NO GOOD REASON (78 SAMPLES)***
2. WHY ***THESE*** 70 GENES? ***RANDOMLY SELECTED***
3. ***HOW WELL*** DOES THE PROGNOSTIC CLASSIFIER ***WORK***?
   PERFORMANCE GETS WORSE AND APPROACHES "CLASSICAL" AS MORE TUMORS ARE TESTED
4. CAN DIFFERENT SIGNATURES BE INTERPRETED IN TERMS OF SAME ***CLINICAL/BIOLOGICAL PROCESSES***? ***ONLY PROLIFERATION***
   (GENE SET ENRICHMENT ANALYSIS)
5. HAS ANY NEW ***KNOWLEDGE*** BEEN GAINED BY THESE 70 GENES? ***NO***

K. Solomon et al Ecclesiastes 1:9 (900 B.C.):

מַה־שֶּׁהָיָה הוּא שֶׁיִּהְיֶה, וּמַה־שֶּׁנַּעֲשָׂה, הוּא שֶׁיֵּעָשֶׂה; וְאֵין כָּל־חָדָשׁ תַּחַת הַשָּׁמֶשׁ

What has been will be again, what has been done will be done again; there is nothing new under the sun.

BIOLOGY vs PHYSICS

1. Immediacy of Phenomena – Interesting & Exciting in a simple, direct way

2. The Beauty of Reductionism - everything stems from a few basic laws, cast in mathematical language. There is "Theoretical Physics"!

3. An amazing number of basic things that are not known or understood -- there is a smell of breakthrough in the air

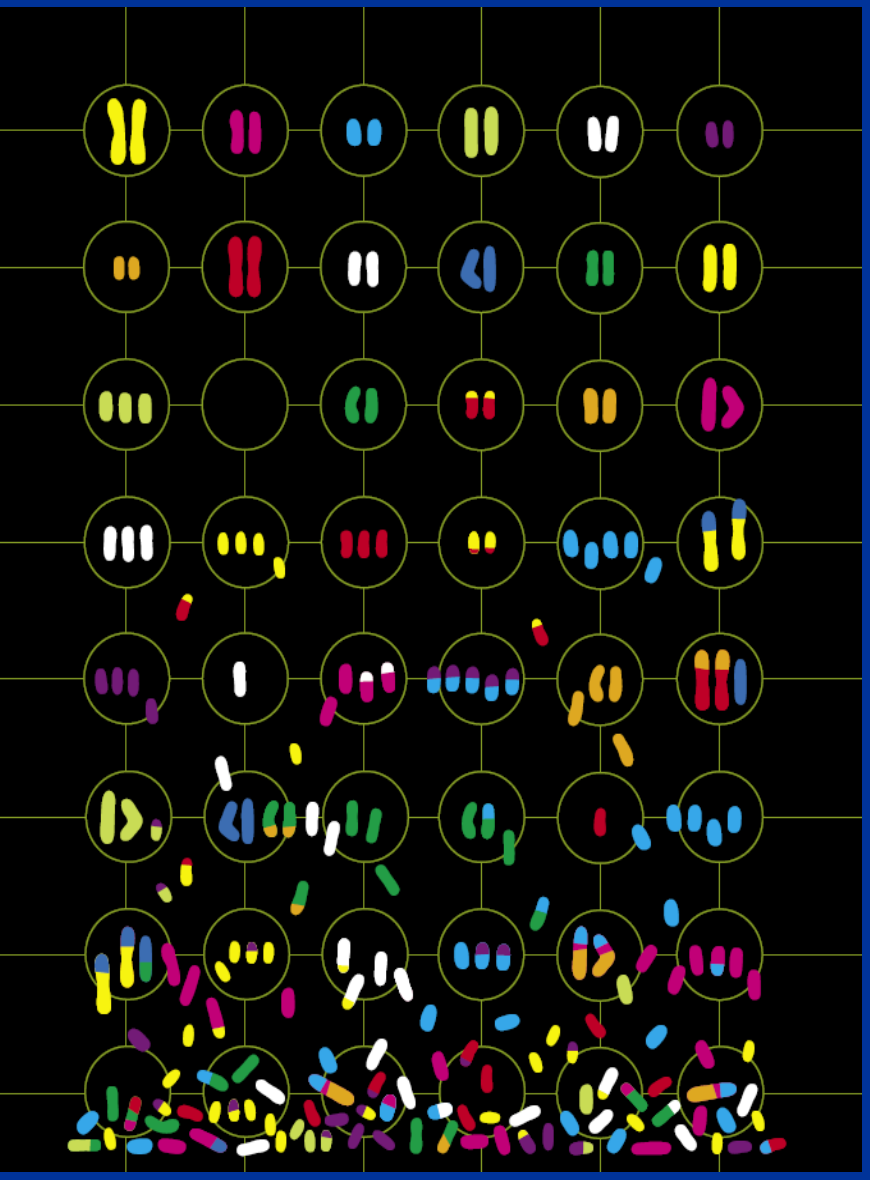
4. Although many papers are "not even wrong" –there is ~~an~~ no objective truth
   Wrong claims either go unnoticed or have a very short life-time
   Leading journals publish many wrong papers that stay around for ages.

5. One (I, anyway) is very frequently wildly surprised

6. Submitted papers are not reviewed, reviews are very ~~rarely~~ often completely idiotic and unfair

7. Selection vs Causality

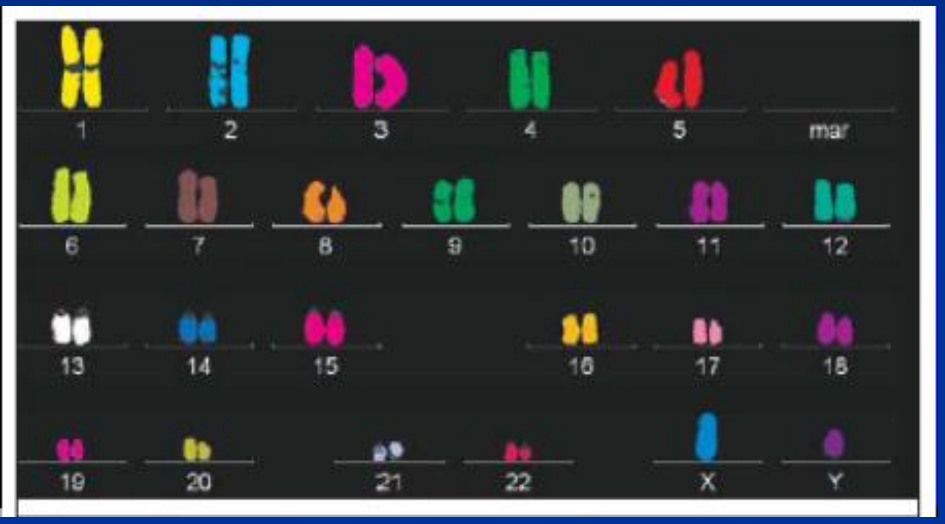
3. CHROMOSOMAL INSTABILITIES (CIN) AND CANCER

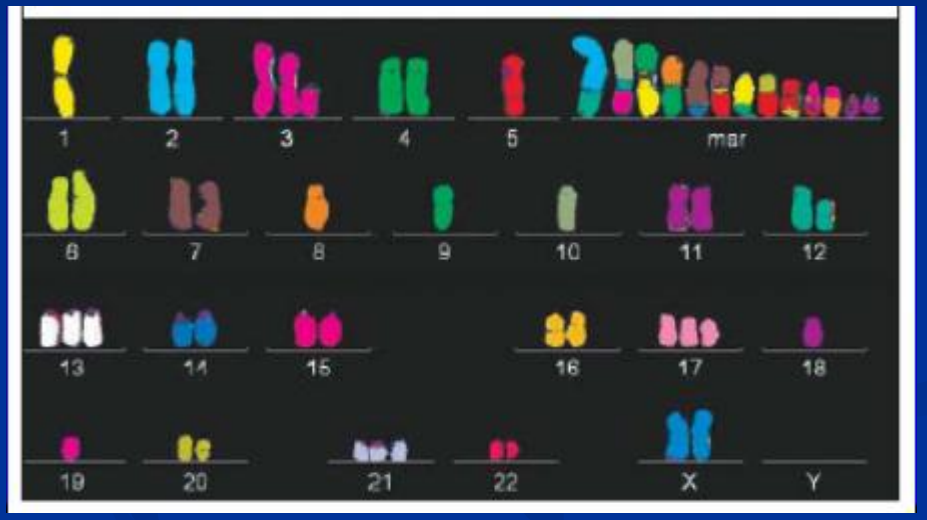


1. “MAPPING” AND INTERPRETING *CIN*
2. *CIN* DYNAMICS AND EVOLUTION
3. DOES *CIN* CAUSE CANCER?

TWO OBSERVATIONS:

1. NORMAL CELLS MAINTAIN A VERY STABLE KARYOTYPE (CHROMOSOMAL SET)





2. CANCER CELLS EXHIBIT ABNORMAL CHROMOSOME COPY NUMBERS (***ANEUPLOIDY***) ***von Hansemann 1890***

SPECULATION: malignant tumours might be the **consequence** of a certain abnormal chromosome constitution, which in some circumstances can be generated by multipolar mitoses (***Boveri***, 1902)

CANCER IS CAUSED BY BREAKDOWN OF REGULATORY NETWORKS THAT PROTECT CELLS AGAINST UNCONTROLLED PROLIFERATION

WHAT CAUSES THIS BREAKDOWN?

DEBATE: (Marx, Science 2002):

1. THE CLASSICAL *Tumor-Suppressor/Oncogene* PICTURE:

BREAKDOWN OF THESE NETWORKS IS CAUSED BY SINGLE-GENE ALTERATIONS . CHROMOSOMAL ABERRATIONS ARE THE EFFECT OF THE MALIGNANT TRANSFORMATION

Weinberg et al Cell 2000,Tomlinson Can Res2001, Dove PNAS 2000

2. AN ALTERNATIVE PICTURE:

CHROMOSOMAL INSTABILITIES PLAY A CENTRAL CAUSATIVE ROLE IN TUMORIGENESIS

Duesberg, Sci. Am. 2007,Science 2005; Weaver, Cancer Cell 2007, Kops, Nature Cancer Reviews 2005

"Genome instability is clearly an enabling characteristic that is causally associated with the acquisition of hallmark capabilities"

Hanahan & Weinberg, Cell 2011

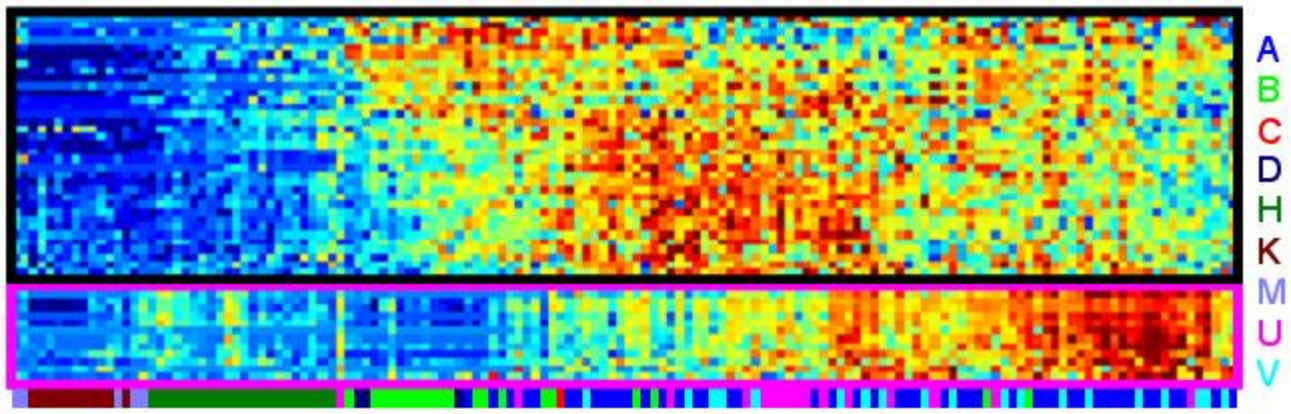
COLON CANCER CONSORTIUM: NCI PPG 2002-2007

SCOPE OF STUDY: 336 PATIENTS
691 TISSUES
EXPRESSION (Affy U133A) 344/264
SNP CHIPS (Affy 50K) 145/84
SNP 309; MSI; METHYLATION
MUTATIONS: BRAF, KRAS, P53

IN 2003 – ONLY 144 (EXPRESSION):
22 NORMAL COLON, 24 POLYPS
47 CARCINOMA, 11 LIVER,16 METS
5 LUNG, 19 LUNG METS

PRINCIPAL INVESTIGATORS:

F. BARANY	Cornell
A. LEVINE	Princeton IAS
D. NOTTERMAN	Princeton U
P. PATY	Memorial SK
W. GERALD	Memorial SK
R. STENGEL	Princeton U
J. OTT	Rockefeller
E DOMANY	Weizmann



32 genes – overexpressed in carcinoma & mets vs normal tissue;
polyps – intermediate

7 out of 32 – from 20q

32 genes – overexpressed in carcinoma & mets vs normal colon; polyps – intermediate

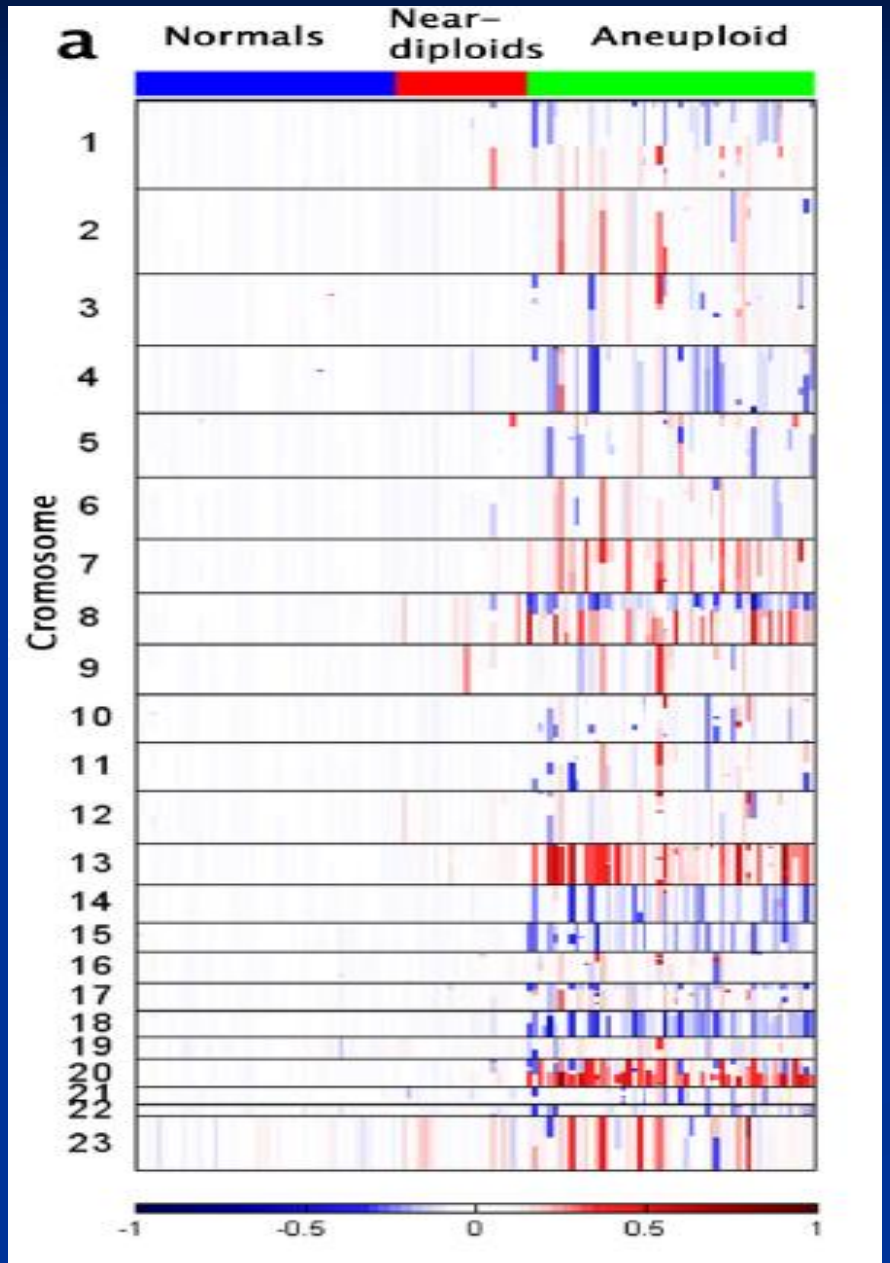
Probe set ID	Title	Gene Symbol	Map location
200903_s_at	S-adenosylhomocysteine hydrolase	AHCY	20cen-q13.1
210052_s_at	chromosome 20 open reading frame 1	C20orf1	20q11.2
218384_at	calcium regulated heat stable protein 1, 24kDa	CARHSP1	16p13.2
202370_s_at	core-binding factor, beta subunit	CBFB	16q22.1
208712_at	cyclin D1 (PRAD1: parathyroid adenomatosis 1)	CCND1	11q13
201326_at	chaperonin containing TCP1, subunit 6A (zeta 1)	CCT6A	7p11.2
203213_at	cell division cycle 2, G1 to S and G2 to M	CDC2	10q21.1
201853_s_at	cell division cycle 25B	CDC25B	20p13
210766_s_at	CSE1 chromosome segregation 1-like (yeast)	CSE1L	20q13
201479_at	dyskeratosis congenita 1, dyskerin	DKC1	Xq28
218435_at	DnaJ (Hsp40) homolog, subfamily D, member 1	DNAJD1	13q14.1
205983_at	dipeptidase 1 (renal)	DPEP1	16q24.3
219787_s_at	epithelial cell transforming sequence 2 oncogene	ECT2	3q26.1-q26.2
203462_x_at,	eukaryotic translation initiation factor 3, subunit 9 eta, 1	EIF3S9	7p22.3
218984_at	hypothetical protein FLJ20485	FLJ20485	7q22.2
201338_x_at,	general transcription factor IIIA	GTF3A	13q12.3-q13.1
218507_at	hypoxia-inducible protein 2	HIG2	7q32.2
206976_s_at	heat shock 105kDa/110kDa protein 1	HSPH1	13q12.3
201601_x_at,	interferon induced transmembrane protein 1 (9-27)	IFITM1	11p15.5
32137_at	jagged 2	JAG2	14q32
212281_s_at,	hypothetical protein MAC30	MAC30	17q11.2
205361_s_at	prefoldin 4	PFDN4	20q13
201558_at	RAE1 RNA export 1 homolog (S. pombe)	RAE1	20q13.31
206918_s_at	RNA binding motif protein 12	RBM12	20q11.21
201063_at	reticulocalbin 1, EF-hand calcium binding domain	RCN1	11p13
204127_at	replication factor C (activator 1) 3, 38kDa	RFC3	13q12.3-q13
201195_s_at	solute carrier family 7 (cationic amino acid transporter, y	SLC7A5	16q24.3
213811_x_at	transcription factor 3 (E2A immunoglobulin enhancer bir	TCF3	19p13.3
201291_s_at	topoisomerase (DNA) II alpha 170kDa	TOP2A	17q21-q22
202954_at	ubiquitin-conjugating enzyme E2C	UBE2C	20q13.12
203797_at	visinin-like 1	VSNL1	2p24.3
213097_s_at	zuotin related factor 1	ZRF1	7q22-q32

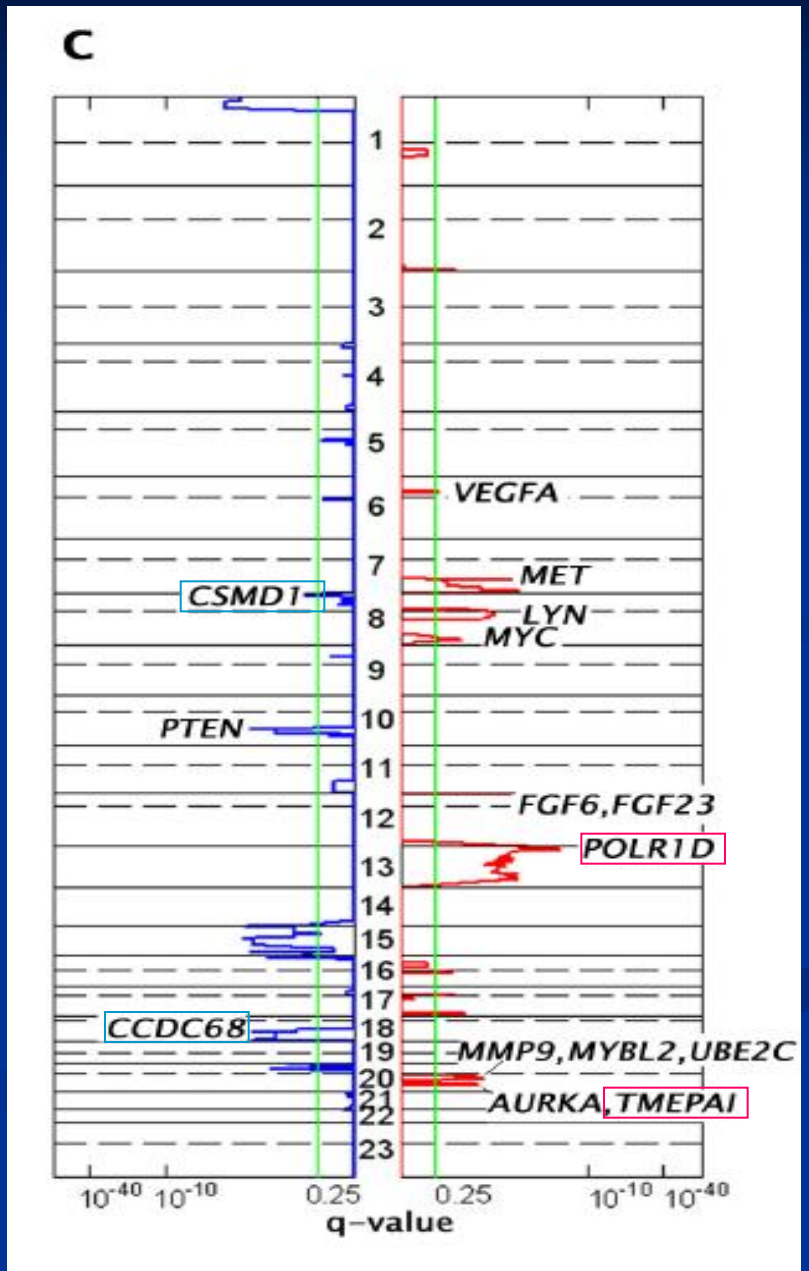
20q GENES – OVER – REPRESENTED

PERHAPS THE INCREASED EXPRESSION OF 20q GENES REFLECTS AMPLIFICATION OF CHROMOSOMAL ARM 20q IN COLON CANCER? KNOWN ***C***HROMOSOMAL ***IN***STABILITY (85% OF COLON CANCER ARE ***CIN***)

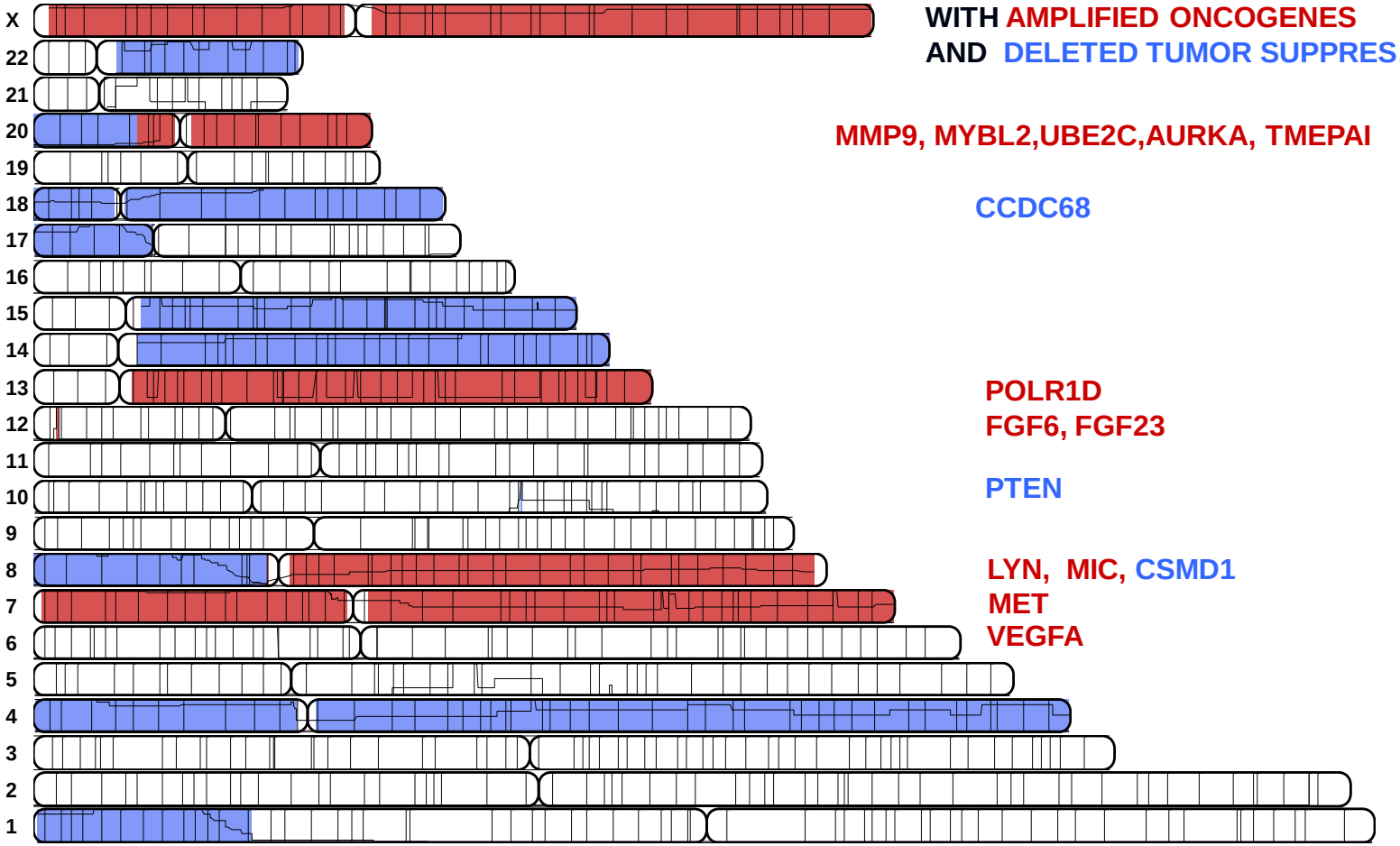
1. GENE (DNA) COPY NUMBER CORRELATES WITH EXPRESSION (mRNA)
2. ONE CAN INFER DNA COPY NUMBER FROM EXPRESSION DATA
3. DEVELOPED A METHOD TO DEDUCE AMPLICONS AND DELETONS FROM aCGH DATA
4. STUDIED COPY NUMBER VARIATIONS IN GLIOBLASTOMA
5. STUDIED COPY NUMBER VARIATIONS IN COLON CANCER

Sheffer et al (PNAS 2009) ANALYSIS OF SNP DATA FOR 130 SAMPLES
using *GISTIC* ***Beroukhim, Getz et al PNAS 2007*** → **45 *CINons***





COMPREHENSIVE CHROMOSOMAL ABERRATION MAP



WITH AMPLIFIED ONCOGENES
AND DELETED TUMOR SUPPRES.

MMP9, MYBL2,UBE2C,AURKA, TMEPAI

CCDC68

POLR1D
FGF6, FGF23

PTEN

LYN, MIC, CSMD1
MET
VEGFA

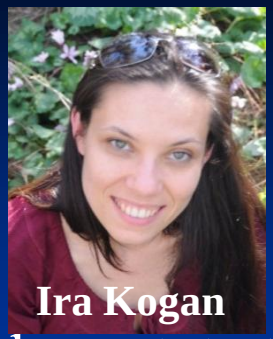
DO CHROMOSOMAL INSTABILITIES CAUSE CANCER?

1. *FOLLOWING THE DYNAMICS OF CIN?*
2. *CLONAL EVOLUTIONARY TREE?*

IN VITRO TRANSFORMATION OF EPITHELIAL CELLS (PROSTRATE)

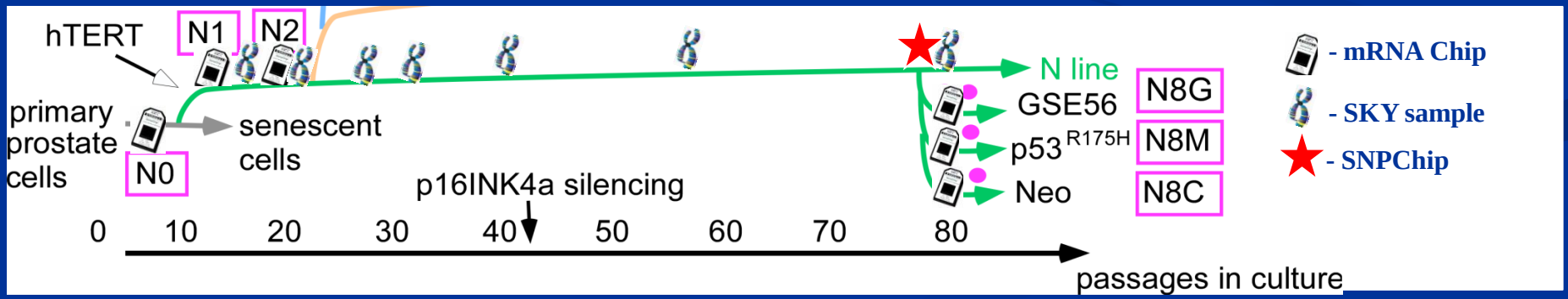
V. Rotter lab: Ira Kogan (expt), Yuval Tabach (analysis)


Varda Rotter

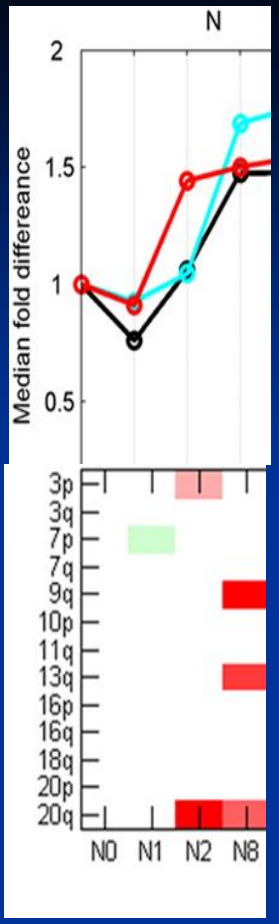

Ira Kogan


Yuval Tabach

1. Isolation of prostate epithelial cells from normal prostate tissue
2. 100K cells plated, immortalized (hTERT), grow.
   Harvest, remove & replant = ONE PASSAGE = 5–6 cycles, 4–10 days.
3. Long term (80 passages, 2 years) *in vitro* culture, leading to "*in vitro* transformation"
4. Measure expression, SKY, SNP, growth rate; P53, P16, Ras mutations at selected time points along the transformation process

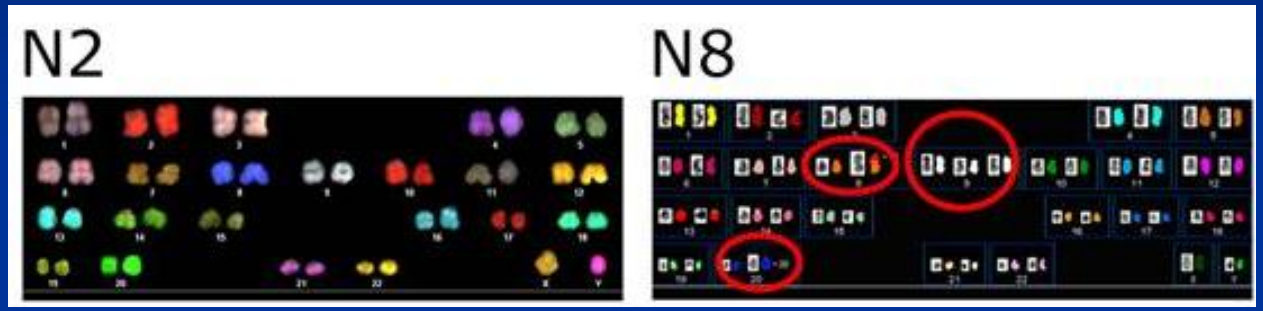


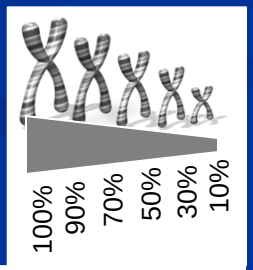
Tabach et al PLoS ONE 2011



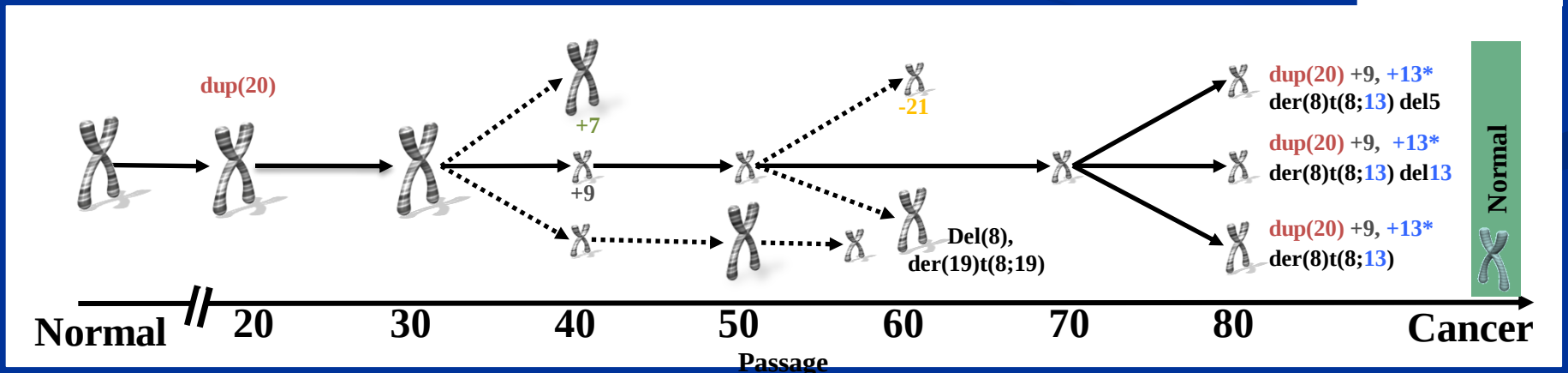
From expression data: 20q duplication starts early, 13q (or 7q, or 9q) follow

From SKY: initially (N2) – normal karyotype, at final time points (N8) - high ANEUPLOIDY





Karyotype evolutionary tree:



FOCUS ON 20q:

BIOINFORMATIC ANALYSIS IDENTIFIED 13
CANCER-INITIATING GENES ON 20q AND THE
PATHWAYS THAT ARE AFFECTED BY THEIR DUPLICATION

MODEL: 20q AMPLIFICATION CAUSED INCREASED EXPRESSION OF GENES – AMONG THEM 13 ***CANCER INITIATING GENES:***
UBE2C, ADRM1, CSE1L, RPN2, C20orf45, MYBL2, TOMM34, AURKA, RAE1, PFDN4, PSMA7,RPS21 and VAPB.
VIA VARIOUS MEDIATORS, THESE ***CIG***s CAUSE INCREASED EXPRESSION OR ACIVATION OF REGULATORS, KNOWN TO PLAY CENTRAL ROLES IN TUMORIGENESIS AND CANCER PROGRESSION.

Tabach et al PLoS ONE 2011

DO CHROMOSOMAL INSTABILITIES

CAUSE CANCER?

PROBABLY YES

3.5 ACUTE LYMPHOBLASTIC LEUKEMIA (*ALL*) IN CHILDREN WITH DOWN'S SYNDROME

THE CAST:

Libi Hertzberg

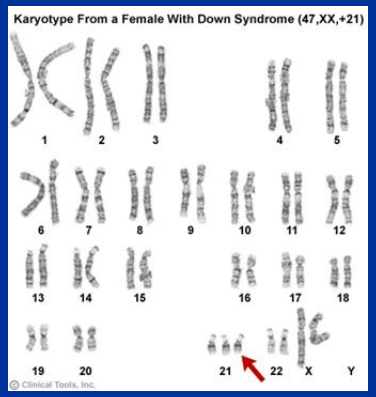
Ithamar Ganmore

Shai Izraeli

THE PLOT:

CHILDREN WITH DOWN'S SYNDROME HAVE 20-FOLD INCREASED RISK FOR ACUTE LYMPHOBLASTIC LEUKEMIA.

WHY??



ANALYSIS OF GENE EXPRESSION DATA: THE CYTOKINE RECEPTOR *CRLF2* IS UPREGULATED IN DOWN'S SYNDROME *ALL* SAMPLES

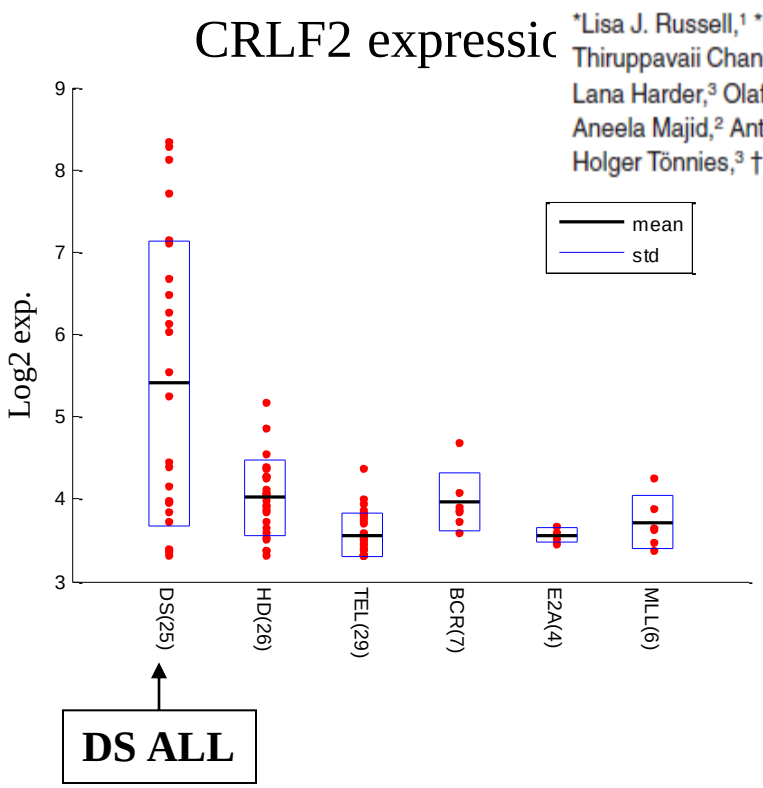
CRLF2 OVER-EXPRESSION IS KNOWN TO BE INVOLVED IN *ALL* !

CRLF2 is upregulated in 60% of DS-ALL (vs 5% of non-DS *ALL*)

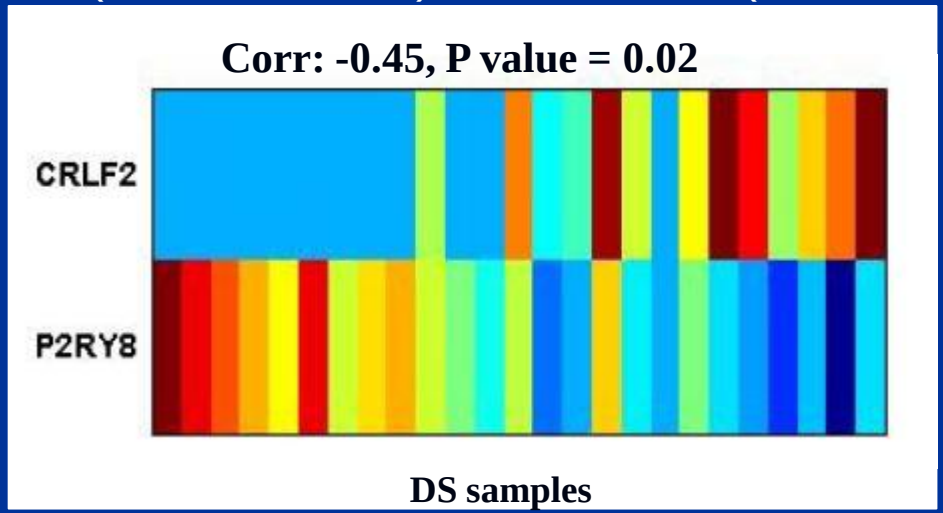
***Blood*, Sept 24, 2009, vol. 114, (13) 2688**

Deregulated expression of cytokine receptor gene, *CRLF2*, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia

*Lisa J. Russell,[1] *Melania Capasso,[2] *Inga Vater,[3] Takashi Akasaka,[2] Olivier A. Bernard,[4] Maria Jose Calasanz,[5] Thiruppavaii Chandrasekaran,[2] Elise Chapiro,[4] Stephan Gesk,[3] Mike Griffiths,[6] David S. Guttery,[2] Claudia Haferlach,[7] Lana Harder,[3] Olaf Heidenreich,[8] Julie Irving,[8] Lyndal Kearney,[9] Florence Nguyen-Khac,[4] Lee Machado,[2] Lynne Minto,[8] Aneela Majid,[2] Anthony V. Moorman,[1] Heather Morrison,[1] Vikki Rand,[1] Jonathan C. Strefford,[10] Claire Schwab,[1] Holger Tönnies,[3] †Martin J. S. Dyer,[2] †Reiner Siebert,[3] and †Christine J. Harrison[1]

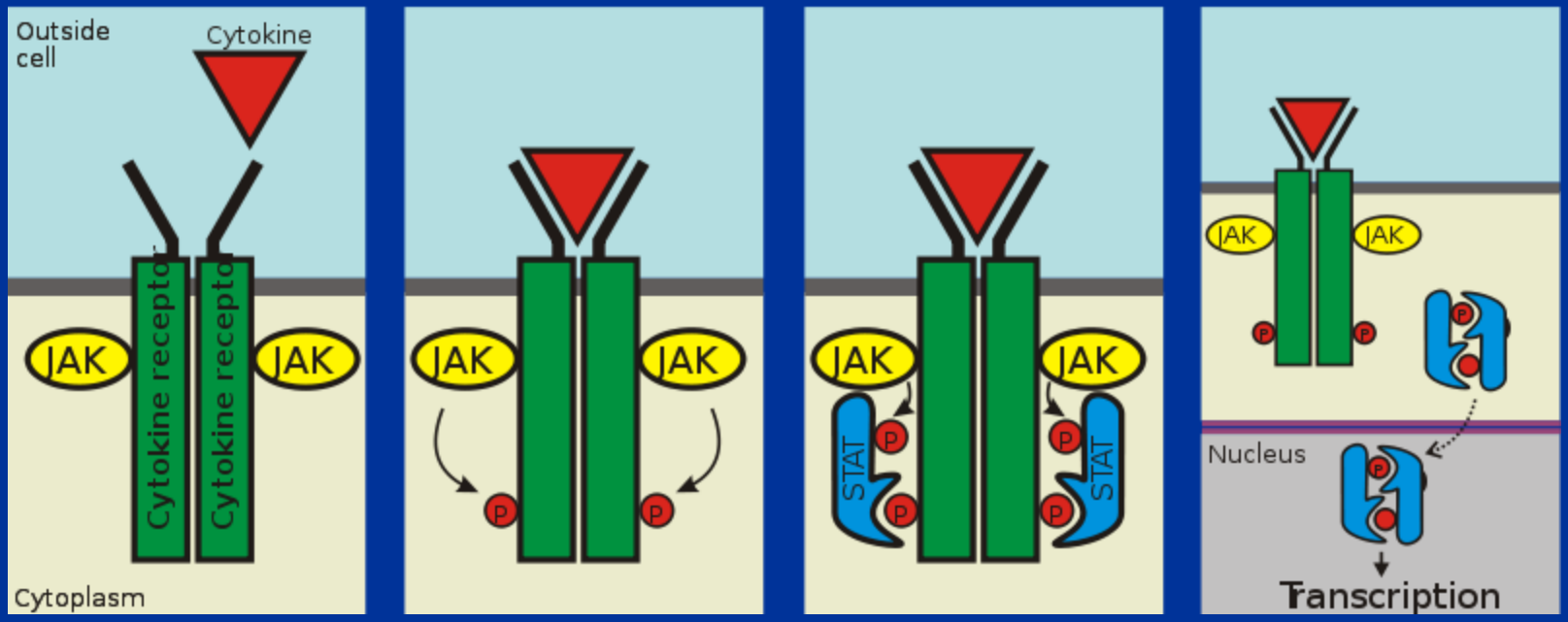


DELETION PLACES *CRLF2* UNDER CONTROL OF A (VERY ACTIVE) PROMOTOR (OF *P2RY8*)

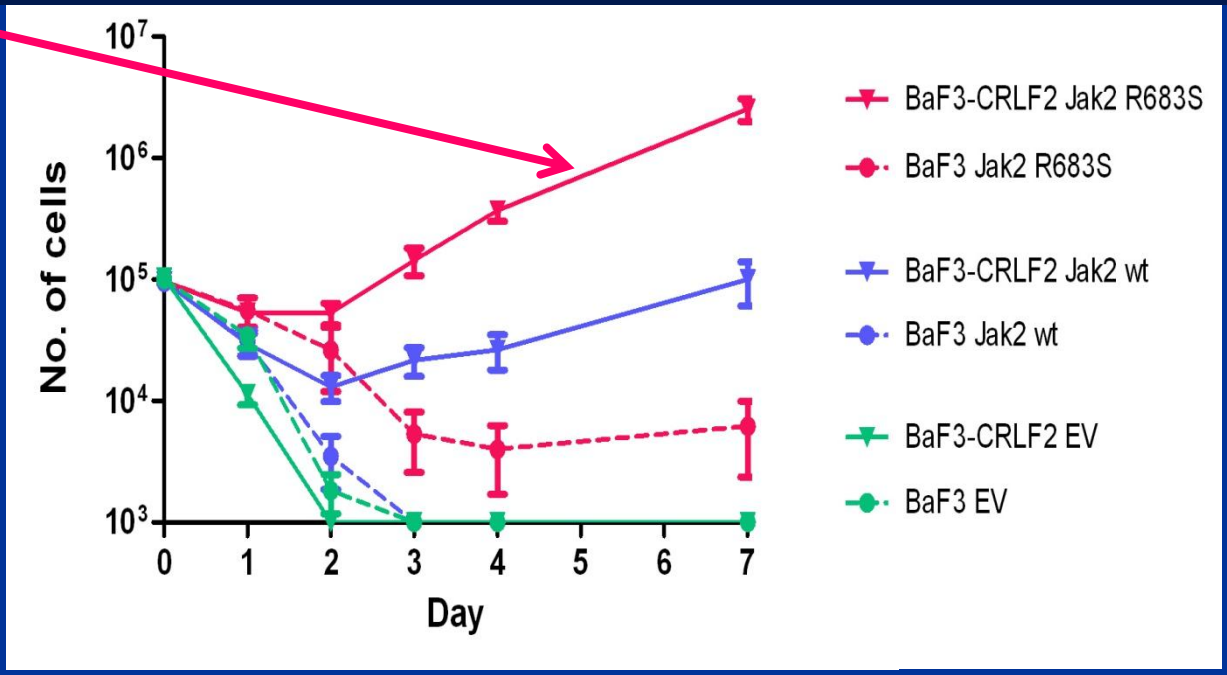


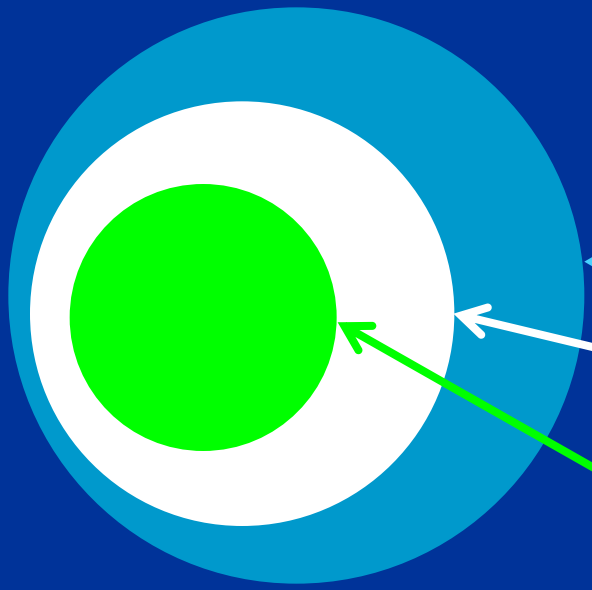
ACTIVATED JAK-STAT PATHWAY INDUCES PROLIFERATION

- JAK2 IS MUTATED IN 20% of DS ALL.
- JAK2 MUTATION CAUSES ACTIVATION OF THE JAK-STAT PATHWAY AND PROMOTES PROLIFERATION.



JAK2 MUTATION AND CRLF2 OVER-EXPRESSION COOPERATE TO INDUCE PROLIFERATION



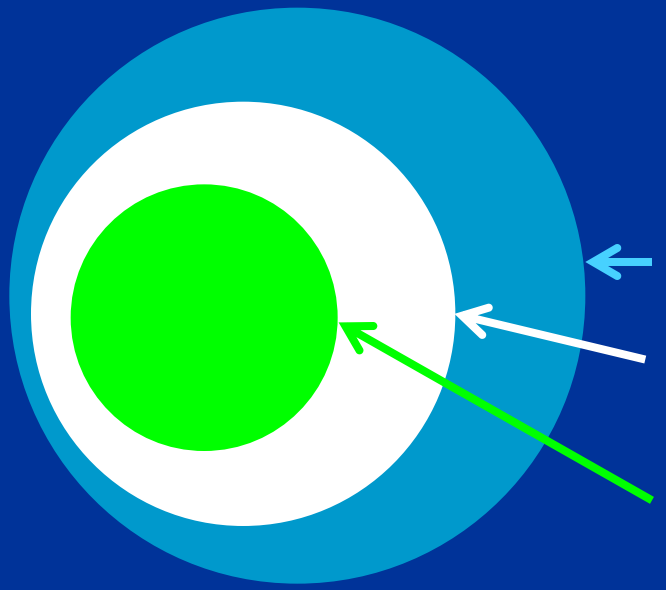


DS *ALL*

***CRLF2* OVEREXPRESSION (60% vs 5% IN non-DS-*ALL*)**

JAK2* MUTATION (20%) all HAVE OVER-EXPRESSED *CRLF2

PHYSICISTS ARE TRAINED TO LOOK FOR CAUSATIVE RELATIONSHIPS



DS *ALL*

***CRLF2* OVEREXPRESSION (60% vs 5% IN non-DS-*ALL*)**

JAK2* MUTATION (20%) all HAVE OVER-EXPRESSED *CRLF2

TRISOMY 21 CAUSES *CRLF2* OVEREXPRESSION

CRLF2 OVEREXPRESSION CAUSES JAK2 MUTATION

BUT HOW?!

BIOLOGY vs PHYSICS

1. Immediacy of Phenomena – Interesting & Exciting in a simple, direct way
2. The Beauty of Reductionism - everything stems from a few basic laws, cast in mathematical language. There is “Theoretical Physics”!
3. An amazing number of basic things that are not known or understood -- there is a smell of breakthrough in the air
4. Although many papers are “not even wrong” – there is an objective truth
5. One (I, anyway) is very frequently wildly surprised
6. Submitted papers are reviewed, reviews are very rarely completely idiotic and unfair
7. Selection vs Causality

MY LESSON: SELECTION vs CAUSALITY

TRISOMY 21 CAUSES *CRLF2* OVEREXPRESSION

CRLF2 OVEREXPRESSION CAUSES JAK2 M UTATION

BUT HOW?!

PHYSICISTS ARE TRAINED TO LOOK FOR *CAUSATIVE RELATIONSHIPS*

BIOLOGY IS GOVERNED BY SELECTION :

MUTATIONS OF JAK2 OCCUR RANDOMLY – WHEN THEY OCCUR IN THE PRESENCE OF OVEREXPRESSED CRLF2, THE CELL WITH THIS CO-OCCURRENCE HAS PROLIFERATIVE ADVANTAGE, IS ***SELECTED*** AND DOMINATES THE CELL POPULATION.

A HUGE HOLE IN OUR EDUCATION!

EPILOGUE

IS THE MICROARRAY TECHNOLOGY GOING TO BE REPLACED (BY NEXT GEN SEQUENCING)?

1997

2001

2011

THANKS FOR LISTENING

&

APOLOGIES FOR RUNNING OVER TIME