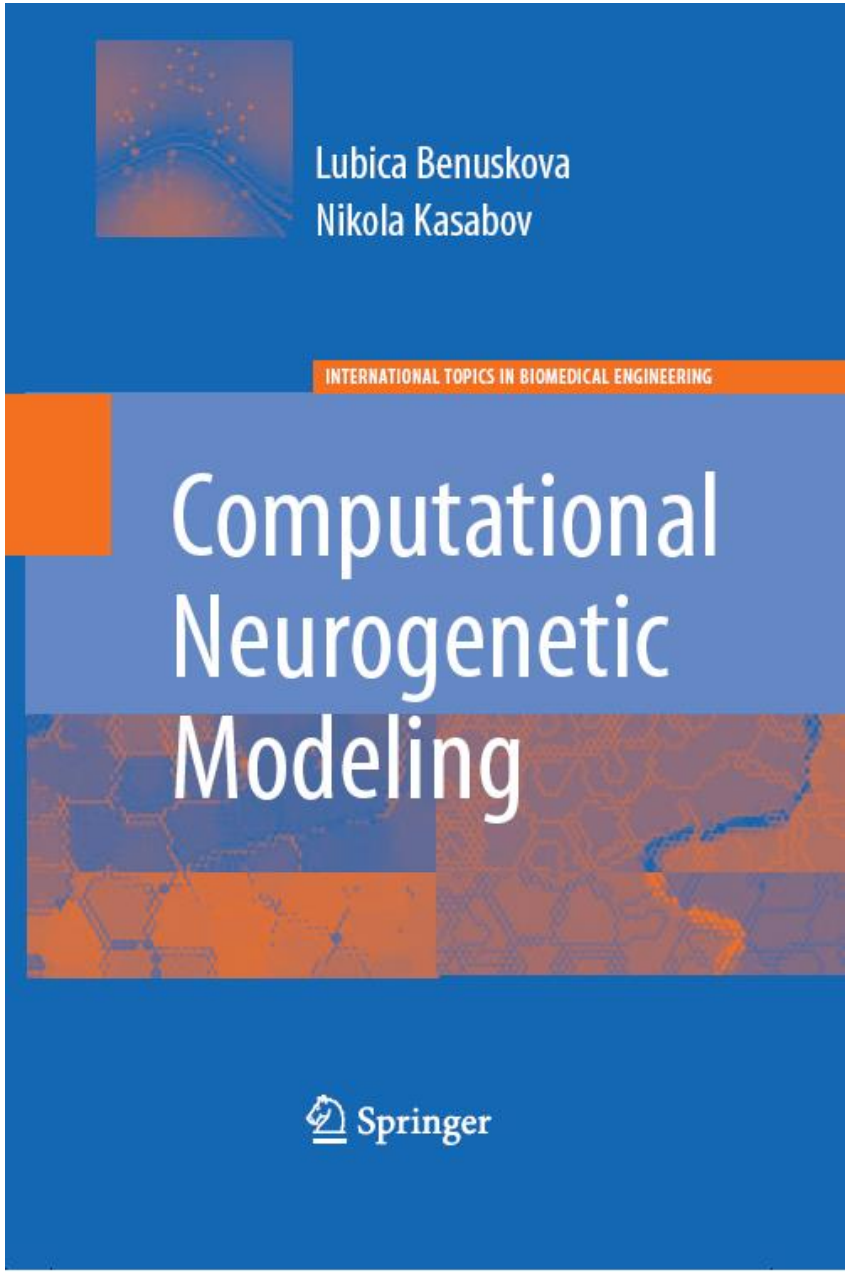
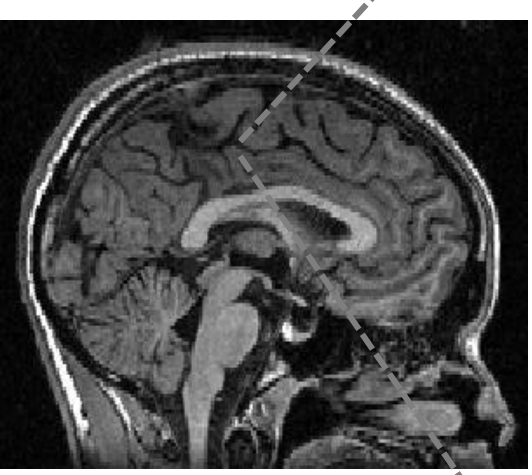
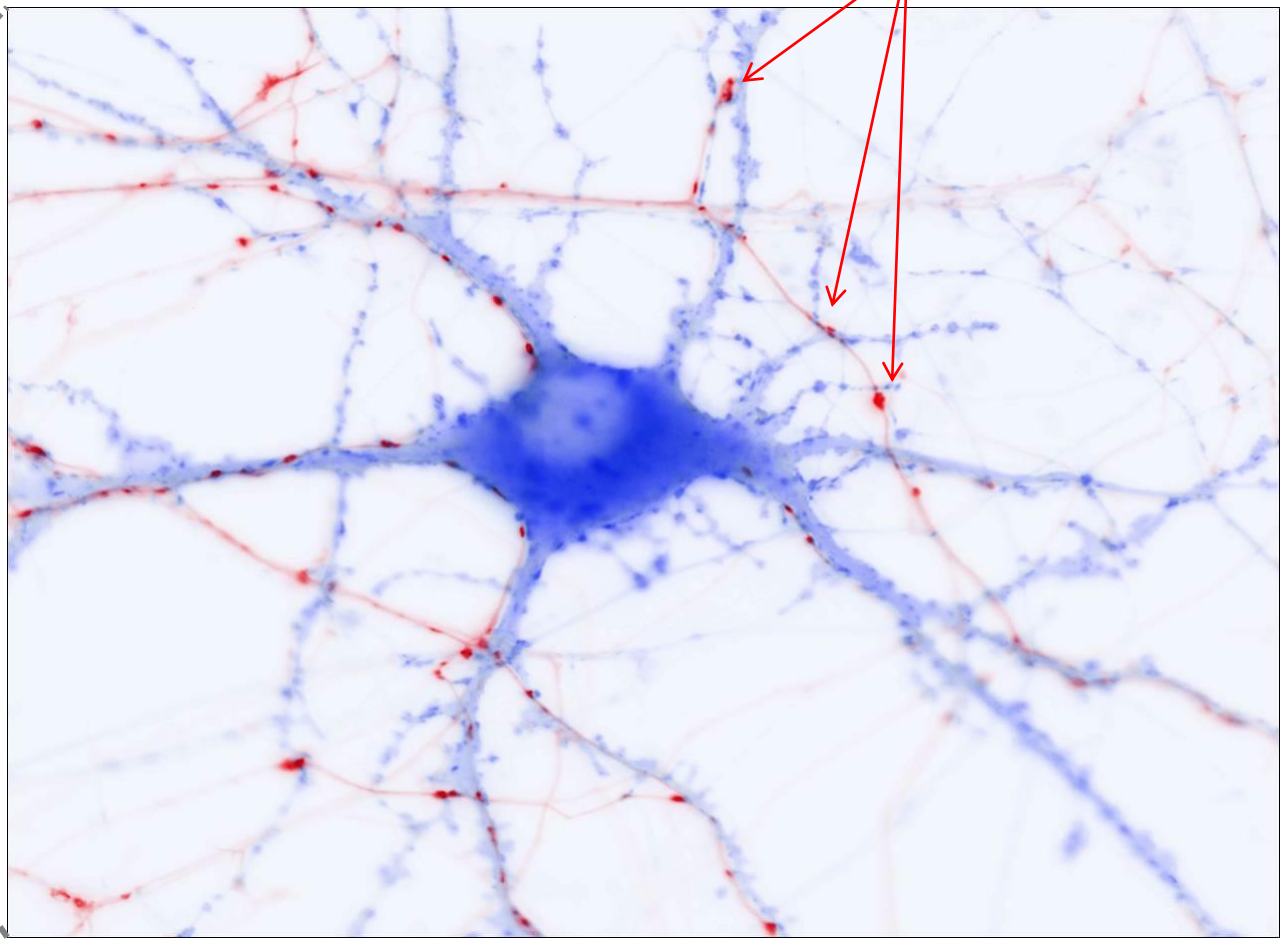

Lubica Benuskova
Nikola Kasabov
INTERNATIONAL TOPICS IN BIOMEDICAL ENGINEERING
Computational Neurogenetic Modeling
Springer

Brain is comprised of networks of neurons connected and communicating by means of synapses



86×10^9 neurons (Suzana Herculano-Houzel, 2009)



~ 10^4 synapses in and out

Dependence of brain functions on genes

https://www.ncbi.nlm.nih.gov/books/NBK22197/

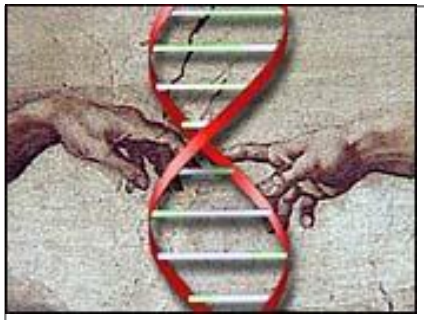
DISEASE	MUTATIONS OF GENES IDENTIFIED SO FAR	LOCATION OF GENES ON CHRO MOSOMES	BRAIN ABNORMALITY	SYMPTOMS	AGE OF ONSET
Alzheimer disease (AD)	PS2 (AD4) PS1 (AD3) unknown unknown	1 14 19 21	plaques made of fragmented brain cells surrounded by amyloid-family proteins, tangles of cytoskeleton filaments	progressive inability to remember facts and events and later to recognize friends and family	71 years
Fragile X syndrome	FMR1 (codes for FMRI protein with unknown function)	X	failure of the glutamate synapse formation and elimination	the most common inherited form of mental retardation	1 year
Huntington disease (HD)	HD gene (codes for the protein huntingtin that stimulates expression of BDNF)	4	dilatation of ventricles and atrophy of caudate nucleus and striatum	degenerative neurological disease that leads to dementia	between 30 and 50 years
Rett syndrome	MeCP2 (codes for a protein which controls gene expression in the cell)	X	generalized brain atrophy, decrease in neuronal cell size, increased cell packing density, reduction in cholinergic neurons	loss of purposeful use of hands and speech, wringing hand movements, seizures, mental retardation	6 to 18 months
Williams syndrome	LIM kinase and elastin coding	7	unknown	high competence in language, music and low IQ	At birth

Gene Variation Doubles Risk of Depression After Stressful Life Events

- Among people who suffered severe stressful life events over 5 years, **43%** with the "short," or stress-sensitive version of the *serotonin transporter gene* developed depression, compared to only 17% with the long version;

- Yet no matter how many stressful life events they endured, people with the "long" or protective version experienced no more depression than people who were totally spared from stressful life events.

- R. Poulton, A. Braithwaite, J.S. Mill (from University of Otago), *Science*, 2003.



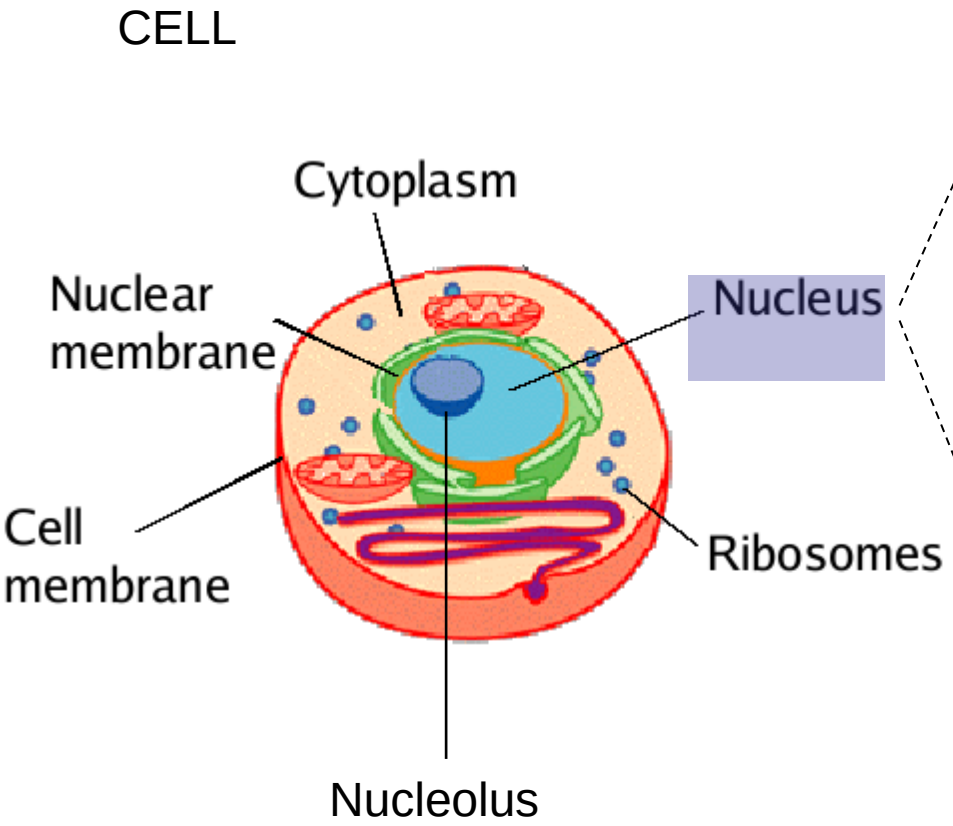
The "language gene" FOXP2

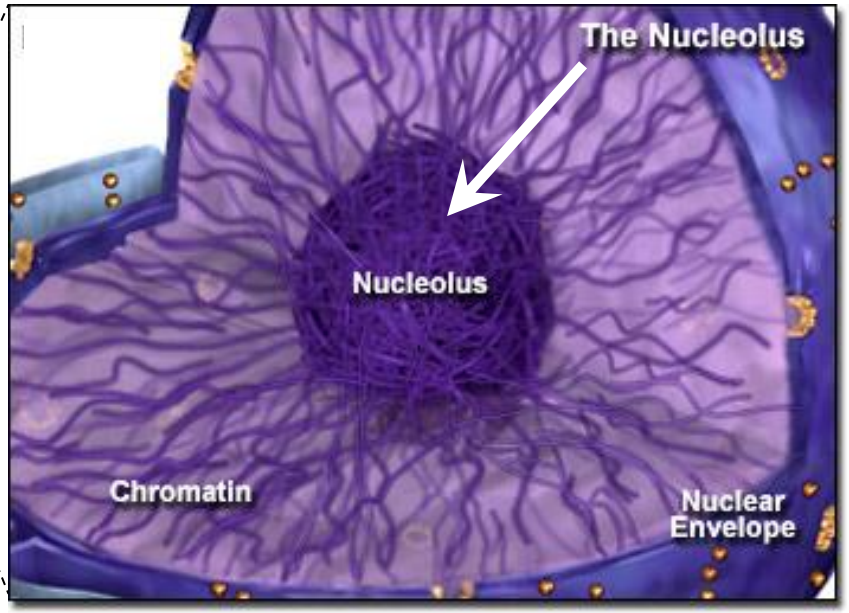
[Svante Pääbo et al, Nature, August 2002
Max Planck Institute of Evolutionary Anthropology]

- **Two functional copies** of FOXP2 (from mom and dad) are required for the development of the normal language in humans.
- Patients have delayed onset of speech, difficulty with articulation including, slurred speech, stuttering, and poor pronunciation, as well as dyspraxia.
- Mutations of FOXP2 causes decreased activation in the putamen and Broca's area in the fMRI studies.
- FOXP2 is fairly well conserved for about **200,000 years**.
- Modern humans share the same allele as Neanderthals.
- The FOXP2 gene is more active in girls than in boys genome.

How is it possible that genes affect brain functions ?

Cell and nucleus

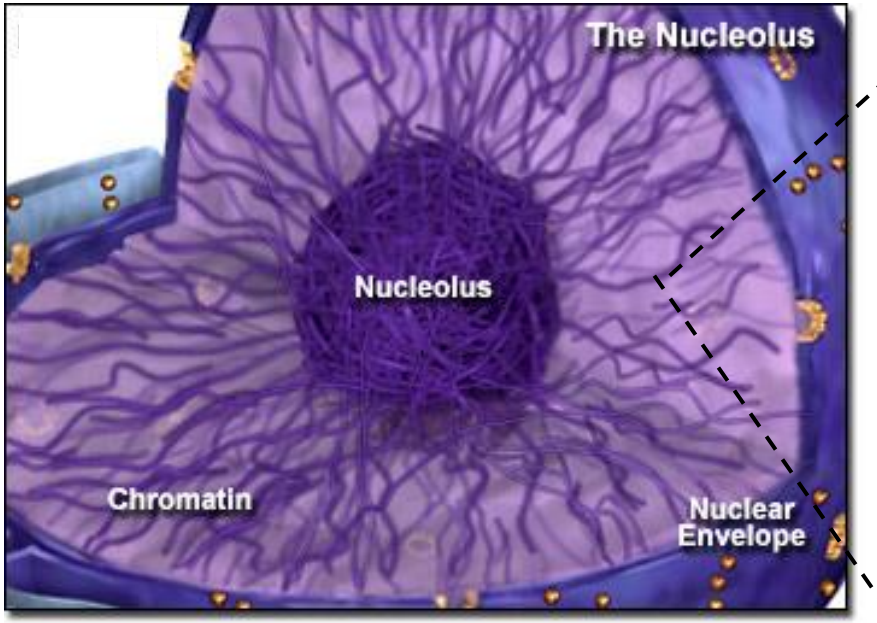




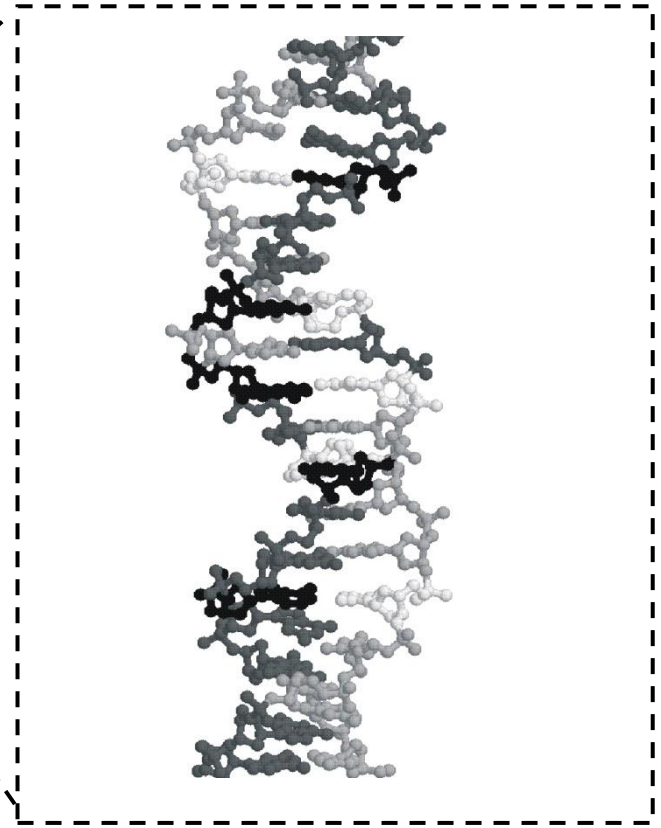
Note: A cell is the basic organizational unit of all living organisms. A cell is comprised of organelles, which are comprised of molecules, which are comprised of atoms.

Nucleus and DNA

Nucleus



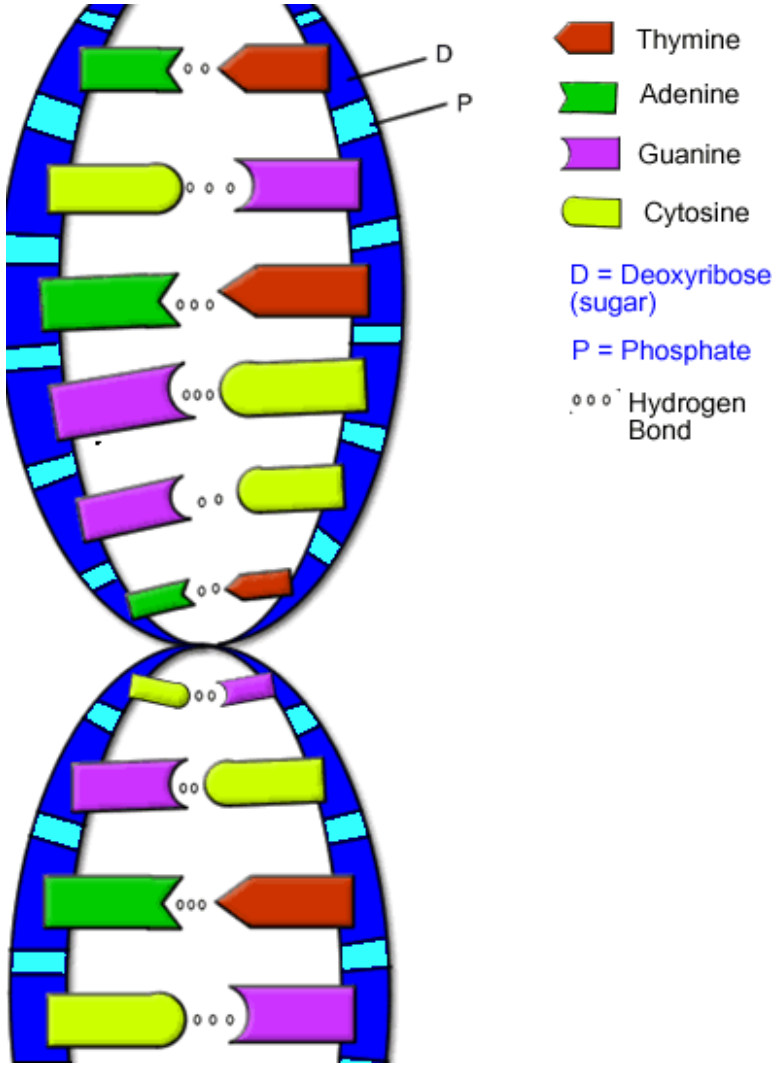
DNA



Note: Eukaryotes are organisms whose cells have a nucleus (animals, plants, fungi, protists). Bacteria and archaea are prokaryotes, i.e. are without a nucleus.

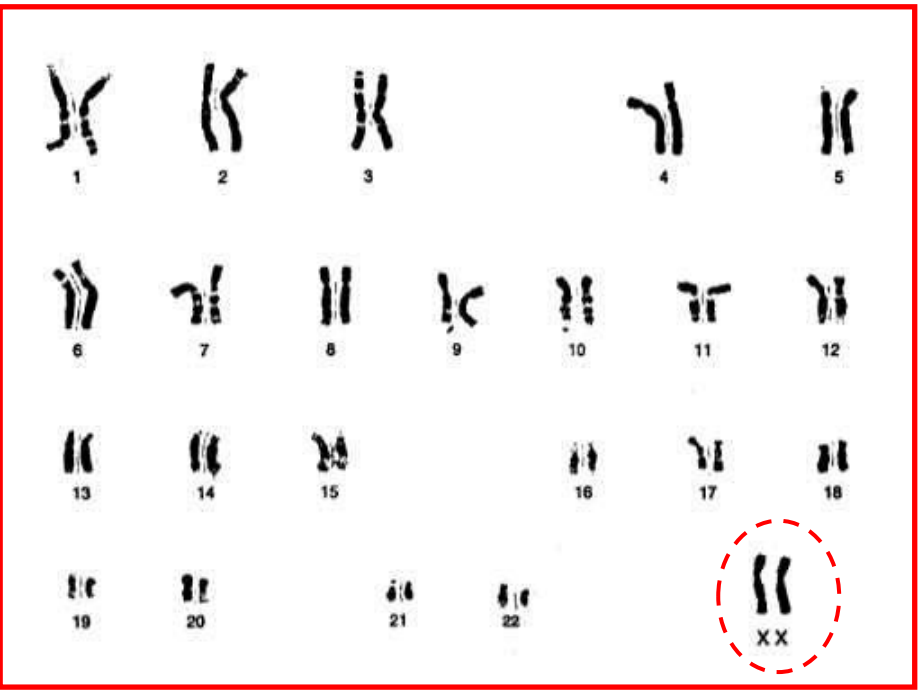
DNA - **D**eoxyribo**N**ucleic **A**cid

- Double-helix structure (discovered in 1953)
- 4 bases: Adenine (A), Thymine (T), Guanine (G), Cytosine (C)
- Bases pair in a base-pair or bp-rule: T-A and C-G
- The order (sequence) of the bases (**A, T, G, C**) determines the information similar to the way, in which letters of the alphabet form words and sentences.

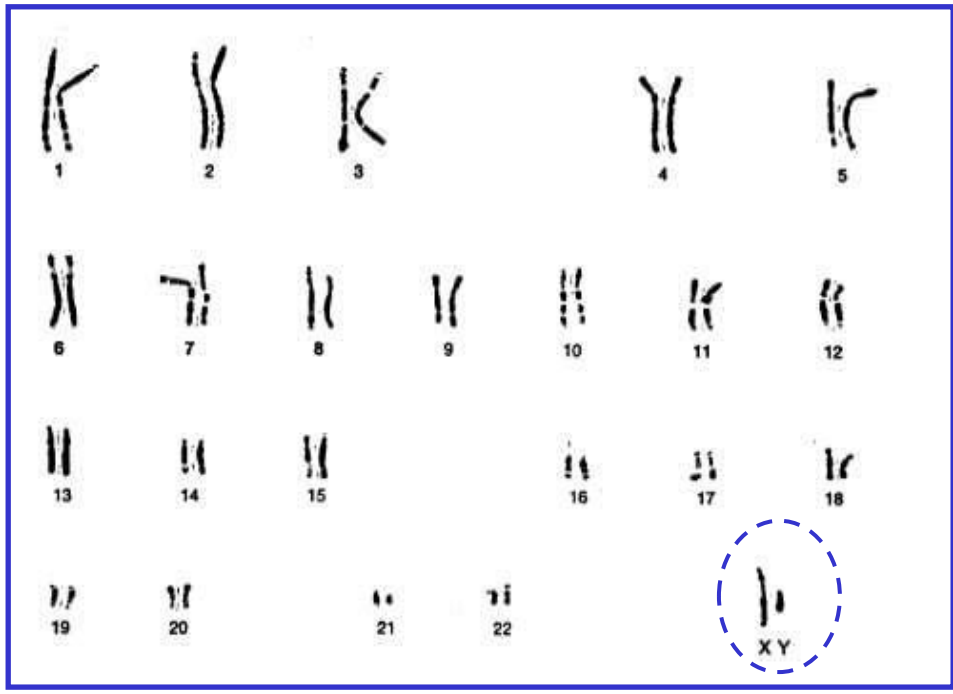


Genome

is the whole hereditary information encoded in the DNA



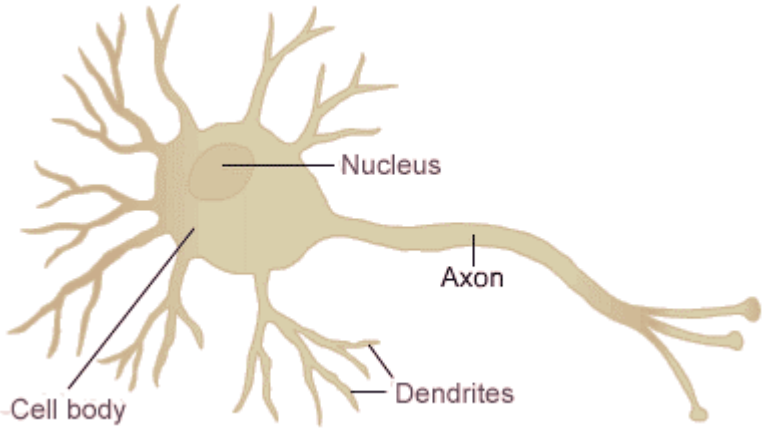
Female

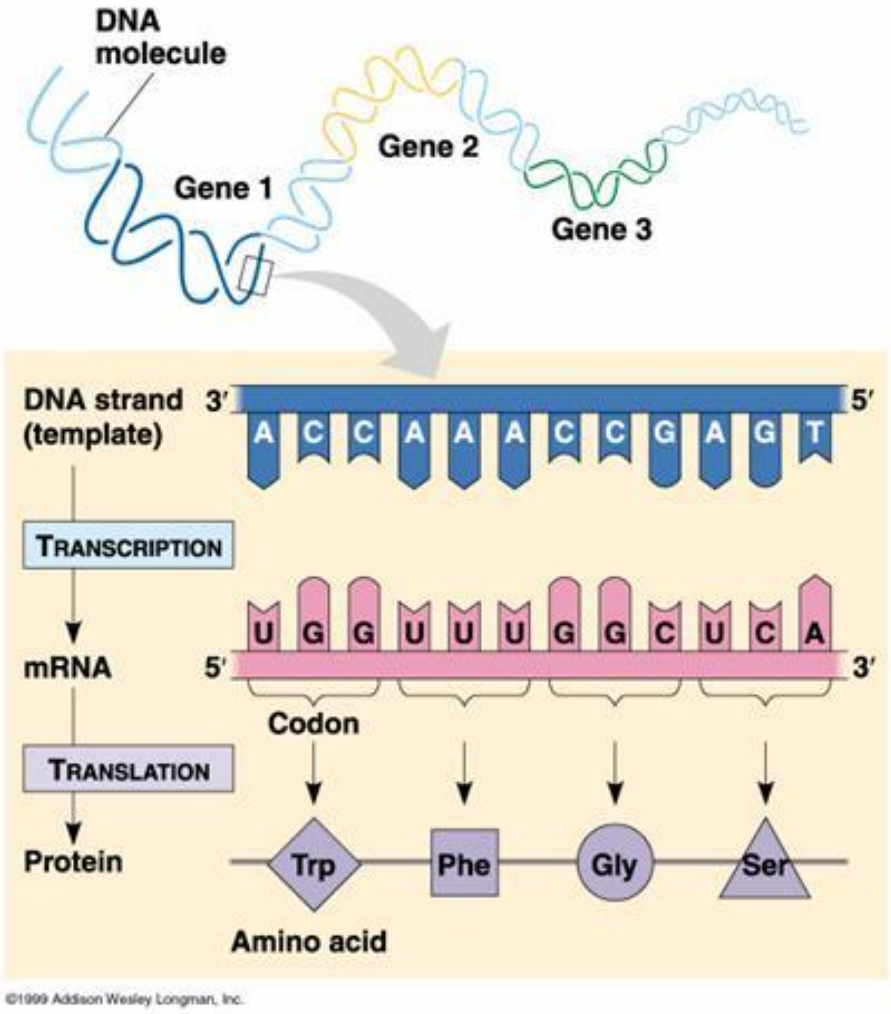


Male

In humans, there are 22 pairs of autosomes (non-sex chromosomes) and 2 sex chromosomes: XX in females and XY in males

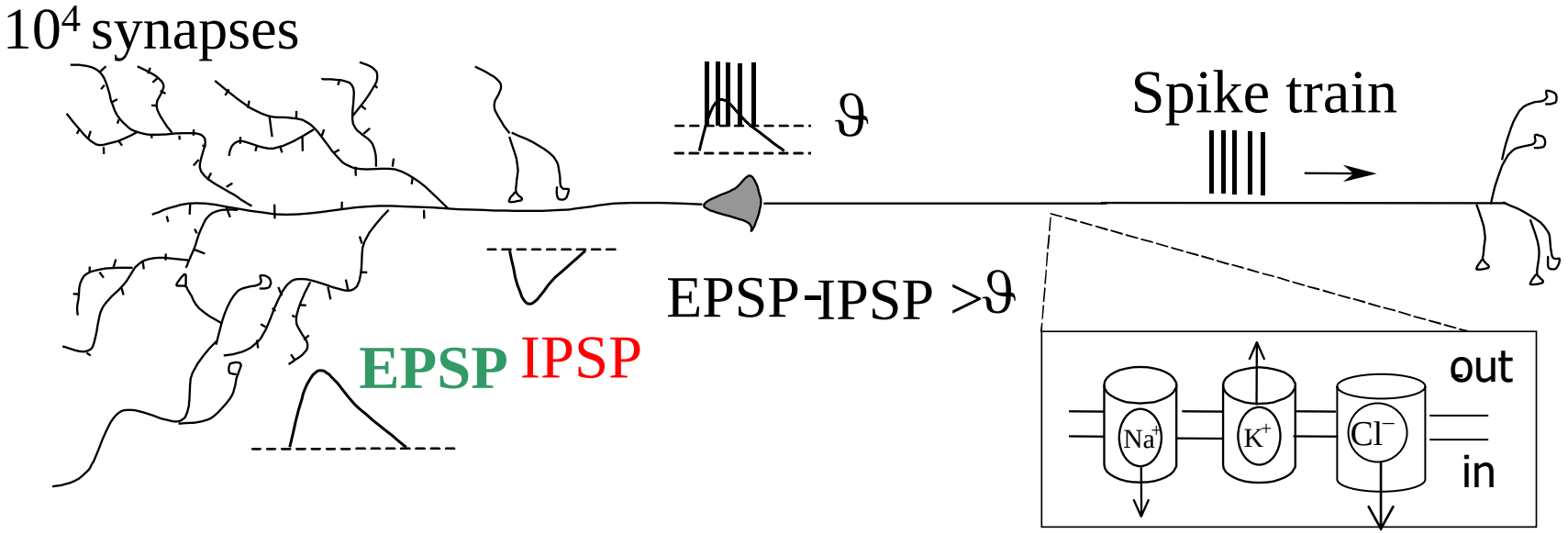
Neuron is a (nerve) cell with DNA, RNA, etc.





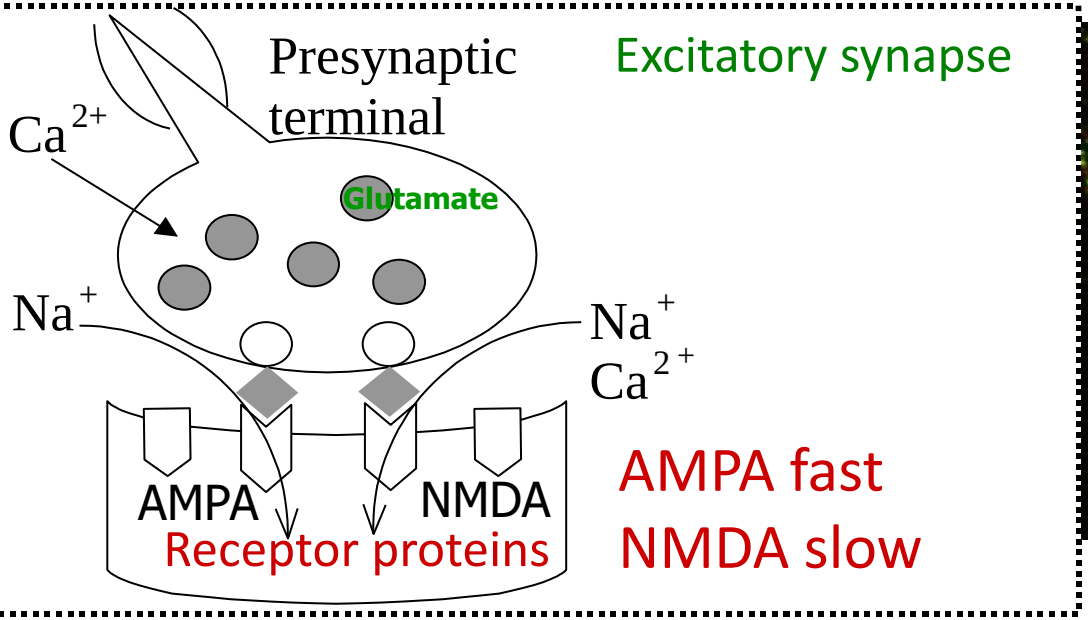
- Central dogma of molecular biology about gene expression:
- "DNA makes RNA, which makes proteins, which make us". (Crick 1958)

Where are proteins in neurons?



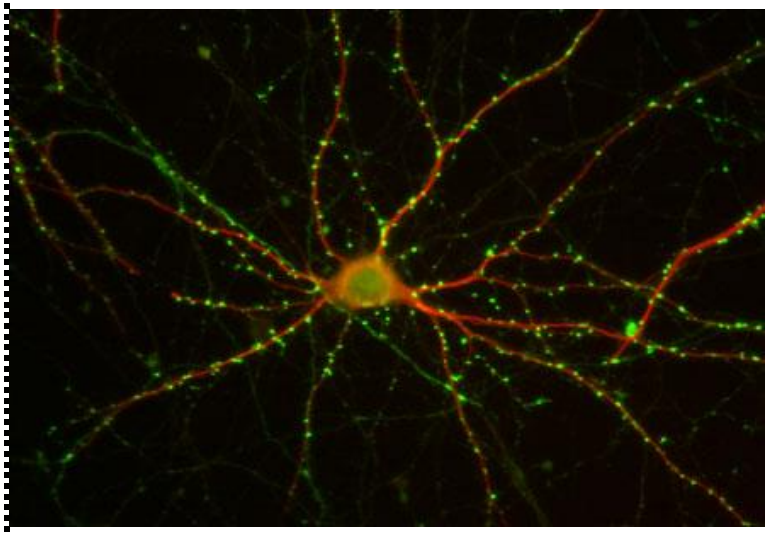
Ion channels for Na^+, K^+, Cl^-, neurotransmitter receptors (AMPA, NMDA, GABRA, GABRB), second messengers (enzymes) that mediate information processing in neurons are all **proteins.**

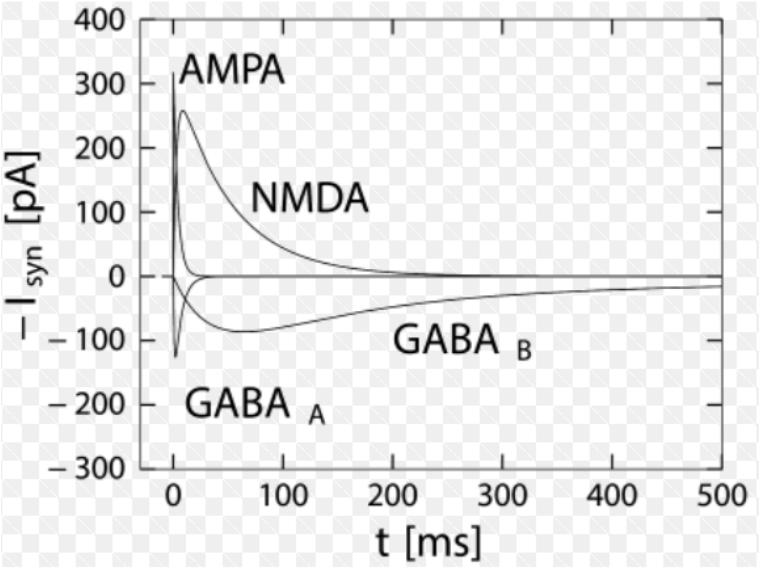
Where are proteins in neurons?

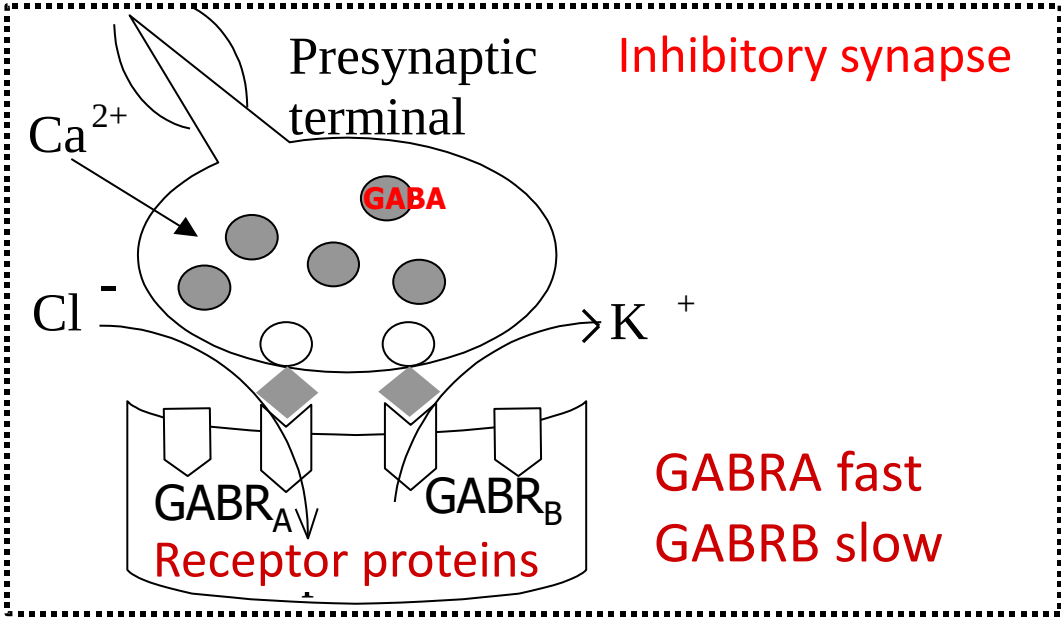


AMPA fast

NMDA slow



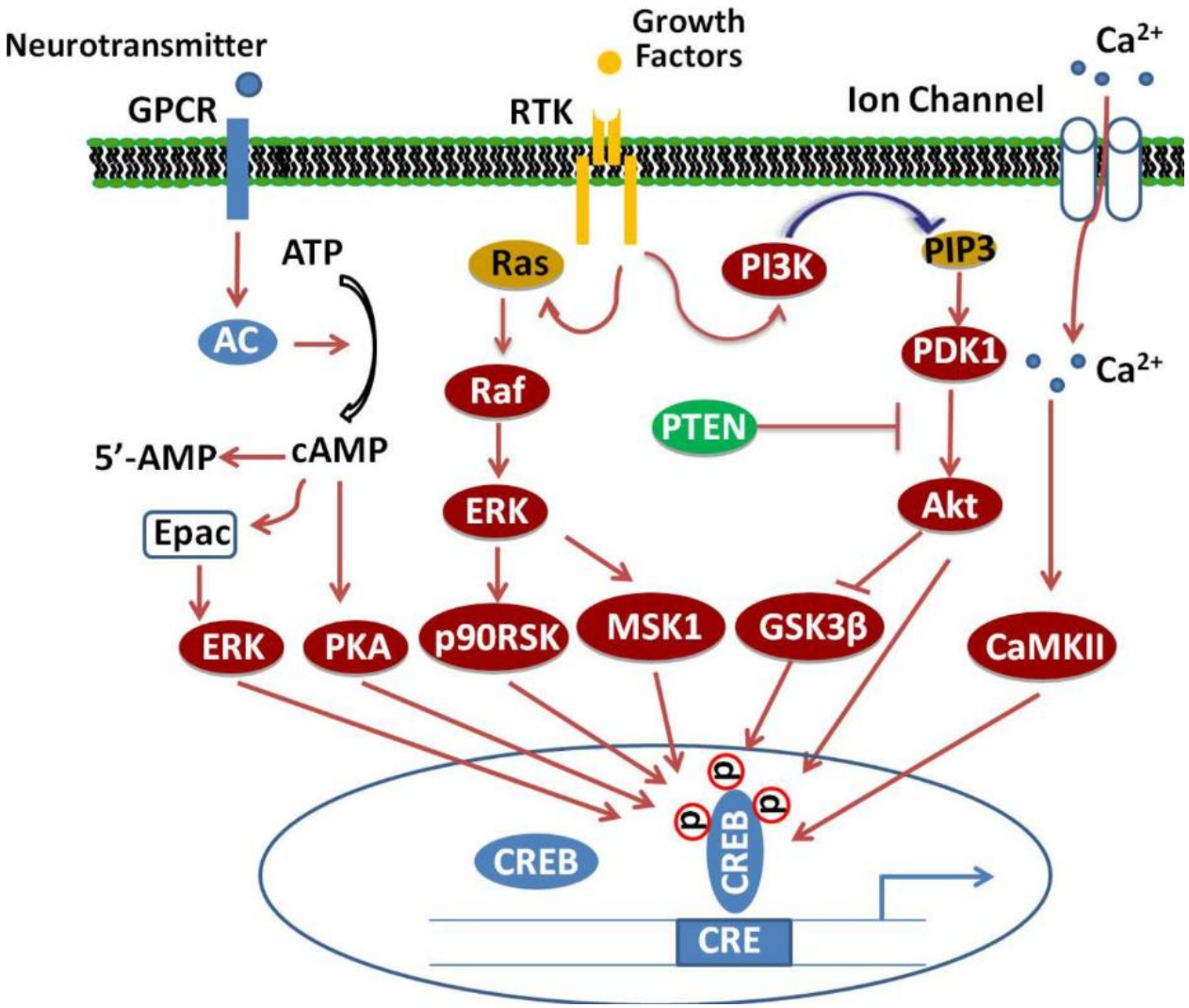




GABRA fast

GABRB slow

CREB activation pathways



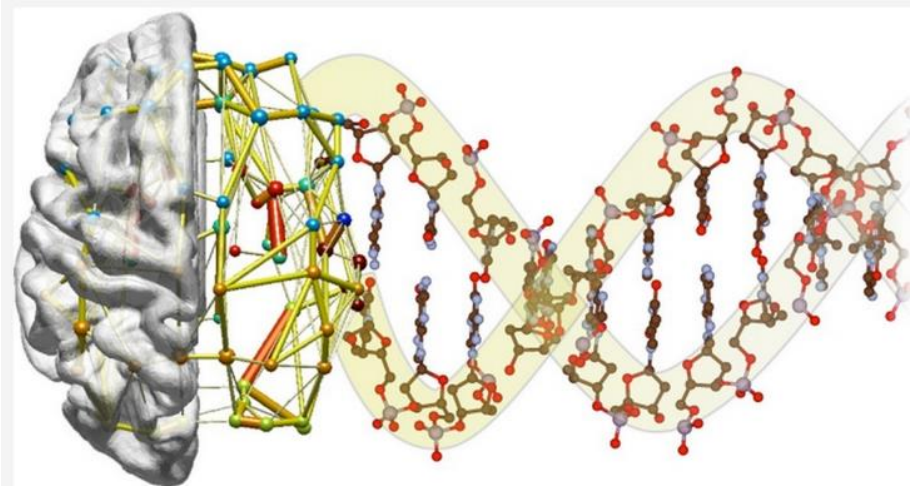
- Source: Wang H et al., Front. Mol. Neurosci., 2019.

Gene density and complexity of organisms

- Current estimates place the human genome DNA at just under 3 billion base pairs and about 20,000–25,000 genes.

- At least a third of the approximately 20,000 different genes that make up the human genome are active (expressed) primarily in the brain. This is the highest proportion of genes expressed in any organ.

- However, the density or number of genes are not a good measure of organism's complexity, because things are more complicated…



How many genes and what are they for?

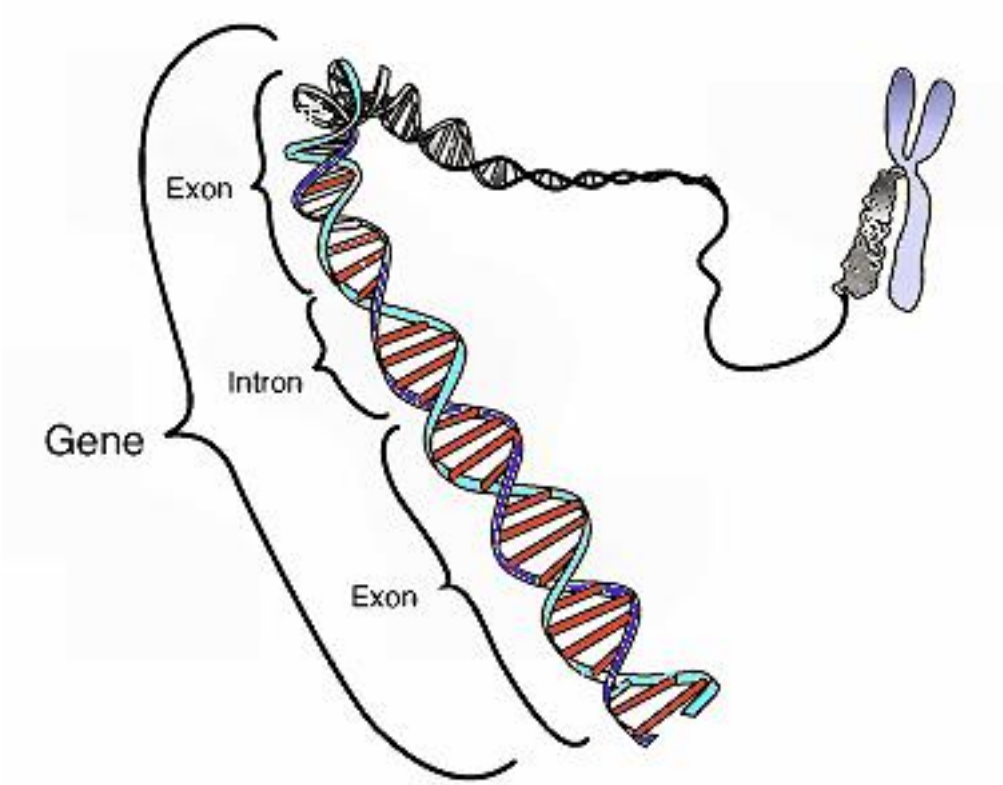
Species	Number of genes
Mycoplasma genitalium (bacterium)	500
Streptococcus pneumoniae (bacterium)	2,300
Escherichia coli (bacterium)	4,400
Saccharomyces cerevisiae (yeast)	5,800
Drosophila melanogaster (fruit fly)	13,700
Caenorhabditis elegans (roundworm)	19,000
Homo sapiens (human)	**20,500**
Sea urchin	23,300
Arabidopsis thaliana (plant)	25,500
Mus musculus (mouse)	29,000
Oryza sativa (rice)	50,000

Source: Watson JD, et al (2004). *Molecular Biology of the Gene*, 5th ed., Pearson Benjamin Cummings (Cold Spring Harbor Laboratory Press).

Exons and introns

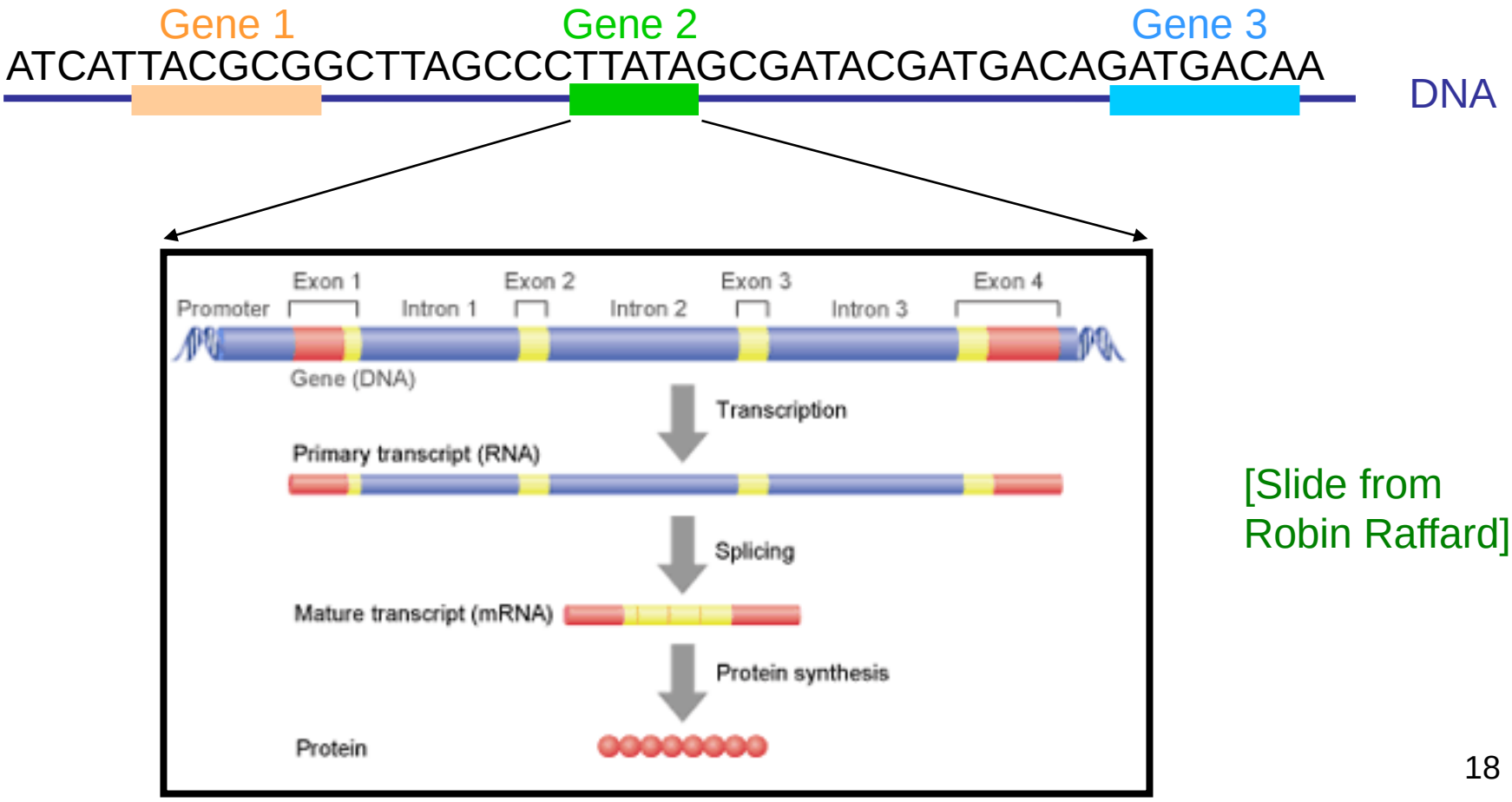
- In eukaryotic cells (including us) genes contain 2 types of regions
  - *Exons:* the regions *encoding* proteins
  - *Introns:* Regions that do not encode proteins.

- Introns are removed from the messenger RNA in a process called **splicing**.



Coding and non-coding DNA

DNA consists of genes (protein coding sequences) separated by non-coding sequences regions.

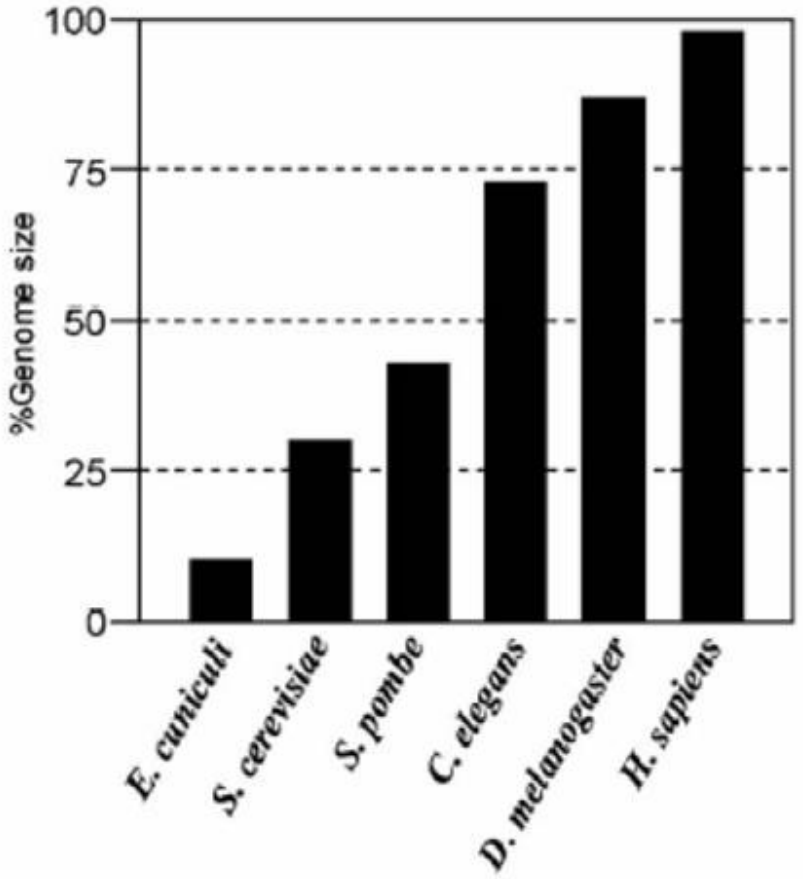


[Slide from Robin Raffard]

Proportion of "junk" or non-coding DNA

- Protein-coding DNA makes up barely 2% of the human genome!
- The rest 98% is the non-coding DNA.
- What is the role of non-coding DNA?

% of "junk" DNA in different species



What is the role of non-coding DNA?

- The 98% of DNA has many important functions like
  - *Retrotransposons* (42% of DNA), move in the genome by being transcribed to RNA and then back to DNA by reverse transcriptase; (can drive evolution by causing mutations)
  - Other *transposones*, mobile DNA pieces, "jumping genes"; (can cause mutations, too)
  - Non-functional remains of ancient genes, known as *pseudogenes*; (they were needed in the past and may be needed in the future)
  - DNA producing different types of RNAs that have important regulatory functions upon gene expression and which can regulate which protein will be produced by alternative splicing.
  - Reservoir of sequences from which potentially advantageous new genes can emerge?

Central dogma revisited

- Crick 1958: "DNA makes RNA makes protein, and proteins make us."

- Petsko 2000: "DNA makes RNA and RNA makes proteins, but sometimes RNA makes DNA and other times RNA makes RNA, which makes proteins different from what they would be if only DNA made the RNA, and once upon a time RNA made proteins, probably, but no-one knows for certain."

- Mattick 2006: "Things are even more complicated & regulatory RNAs represent the major output of the genomes of humans."

Regulation of gene transcription: basic concepts

- **Transcription factor** (TF) is a protein or RNA that binds to specific DNA domains and controls gene transcription into RNA.

- An **activator** is a TF that increases gene transcription, thus leading to the gene upregulation (activation, promotion). The activator binds to a DNA segment known as *enhancer*. Activator may increase transcription by itself, or may operate through one or more **co-activators**.

- A **repressor** is a TF that decreases gene transcription, thus leading to the gene downregulation (repression, suppression). Repressor proteins attach to a DNA segment known as the *operator*. If an inducer, a molecule that initiates gene expression, is present, then it can interact with the repressor protein and detach it from the operator. There exist also **co-repressor** TFs.

- Inducers function by disabling repressor proteins. Some inducers are modulated by activators.

Regulation of transcription: what is it good for?

- **Basal transcription regulation**. General transcription factors (GTFs) are necessary for any transcription to occur.
  - The most common GTFs are TFIIA(B, D, E, F, and H).

- **Cell cycle control.** Many TFs (especially oncogene and tumor suppressor proteins) help regulate cell differentiation, division and apoptosis (programmed cell death).
  - Examples of proto-oncogenes: RAS, WNT, MYC, ERK and TRK.

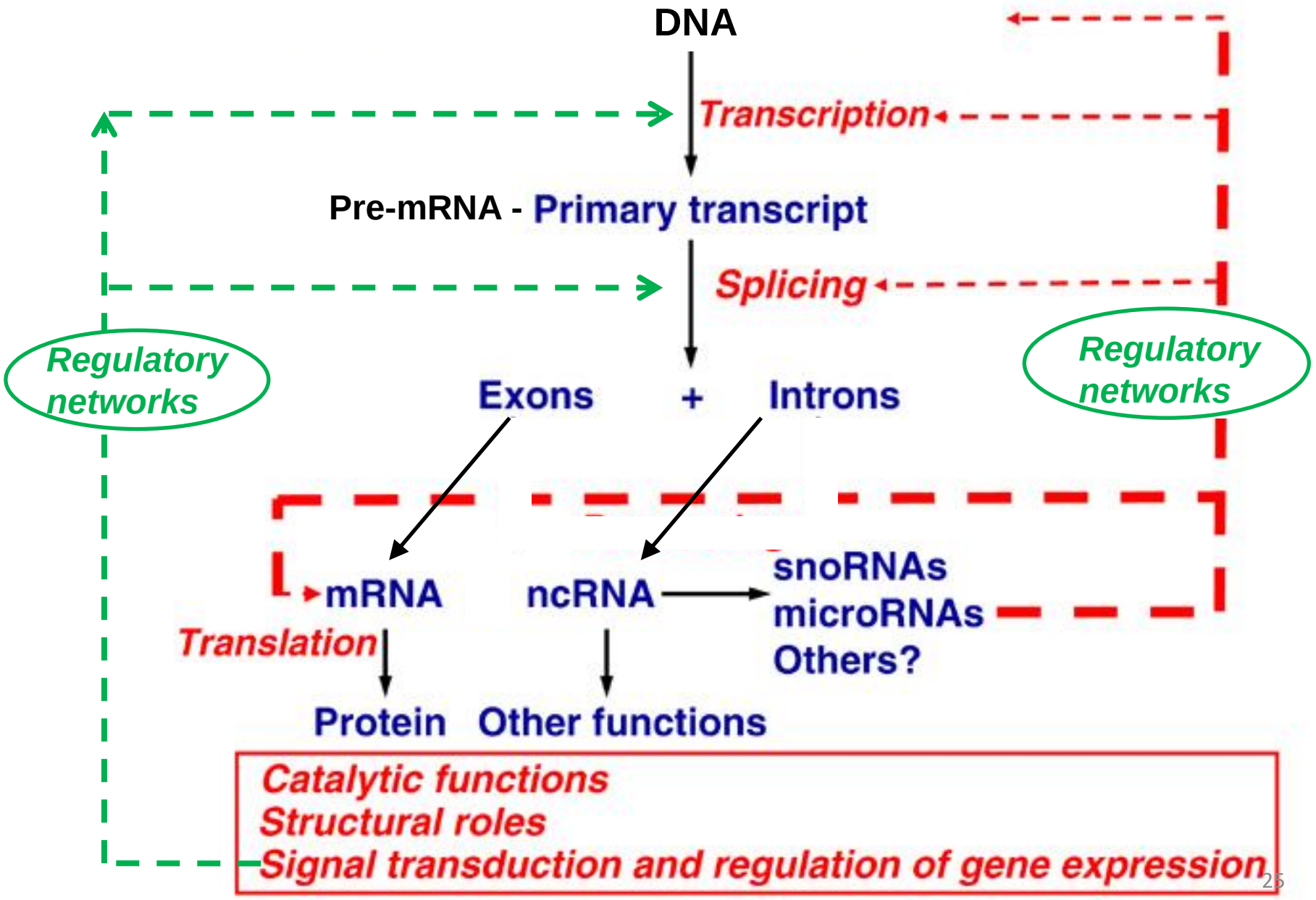
- **Development.** In response to intra- and extracellular stimuli, TFs turn on/off the transcription of the appropriate genes, which in turn allow for changes in cell morphology, differentiation and function.
  - Examples: the Hox TF family (body pattern formation).

Regulation of transcription: what is it good for?

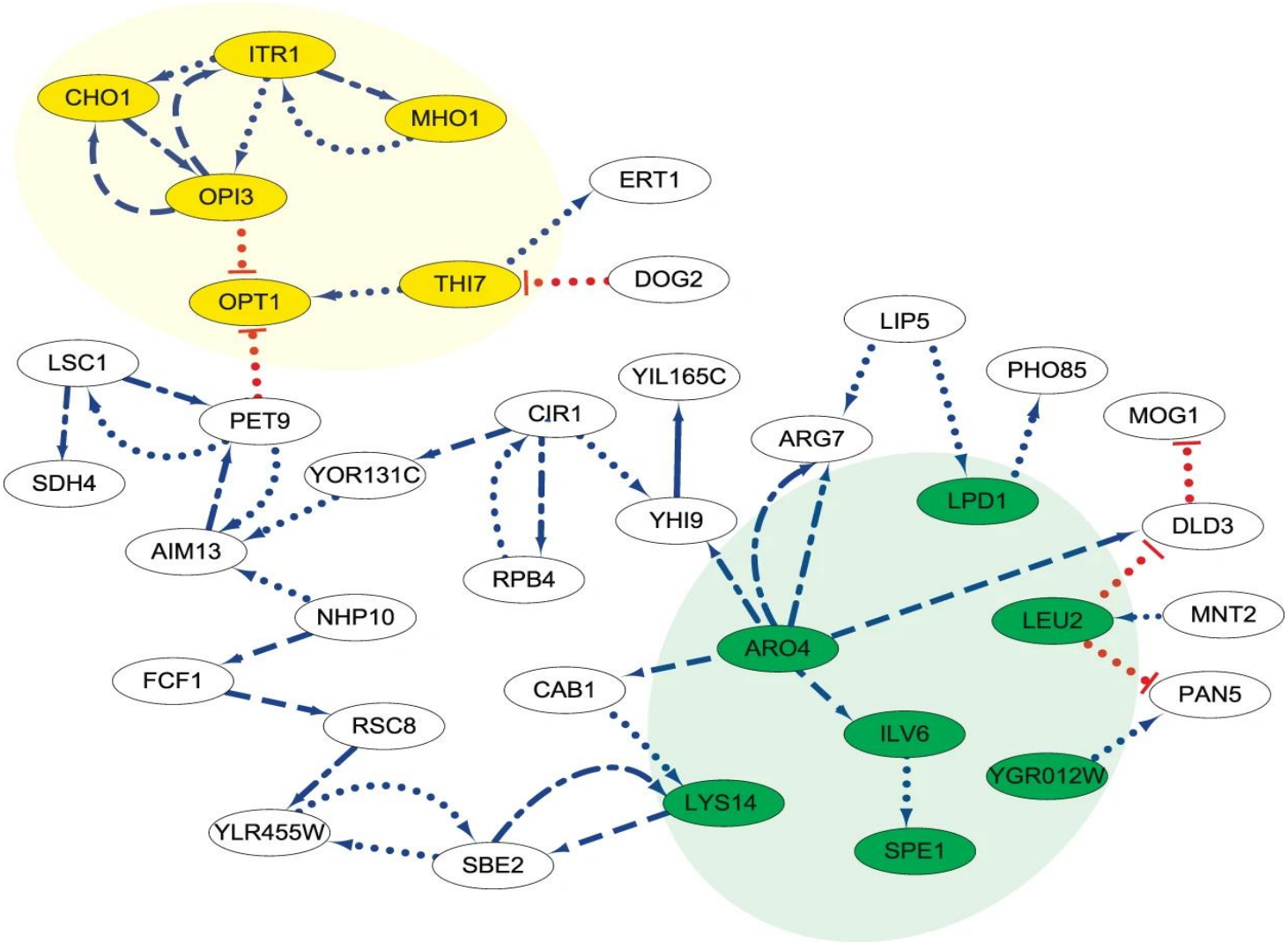
- **Response to environment**. Not only do transcription factors act downstream of signalling cascades related to biological stimuli, but they can also be downstream of signalling cascades involved in environmental stimuli.
  - Examples: heat shock factor (HSF) which upregulates genes necessary for survival at higher temperatures.

- **Response to inter-cellular signals**. Cells can communicate with each other by releasing molecules that produce signalling molecular chain cascades within other cells. Intracellular signalling pathways thus connect the cell surface to the nucleus, leading to changes in gene expression in response to extracellular stimuli.
  - Example: response to the release of neurotransmitter via CREB.

Regulation of gene expression by other genes

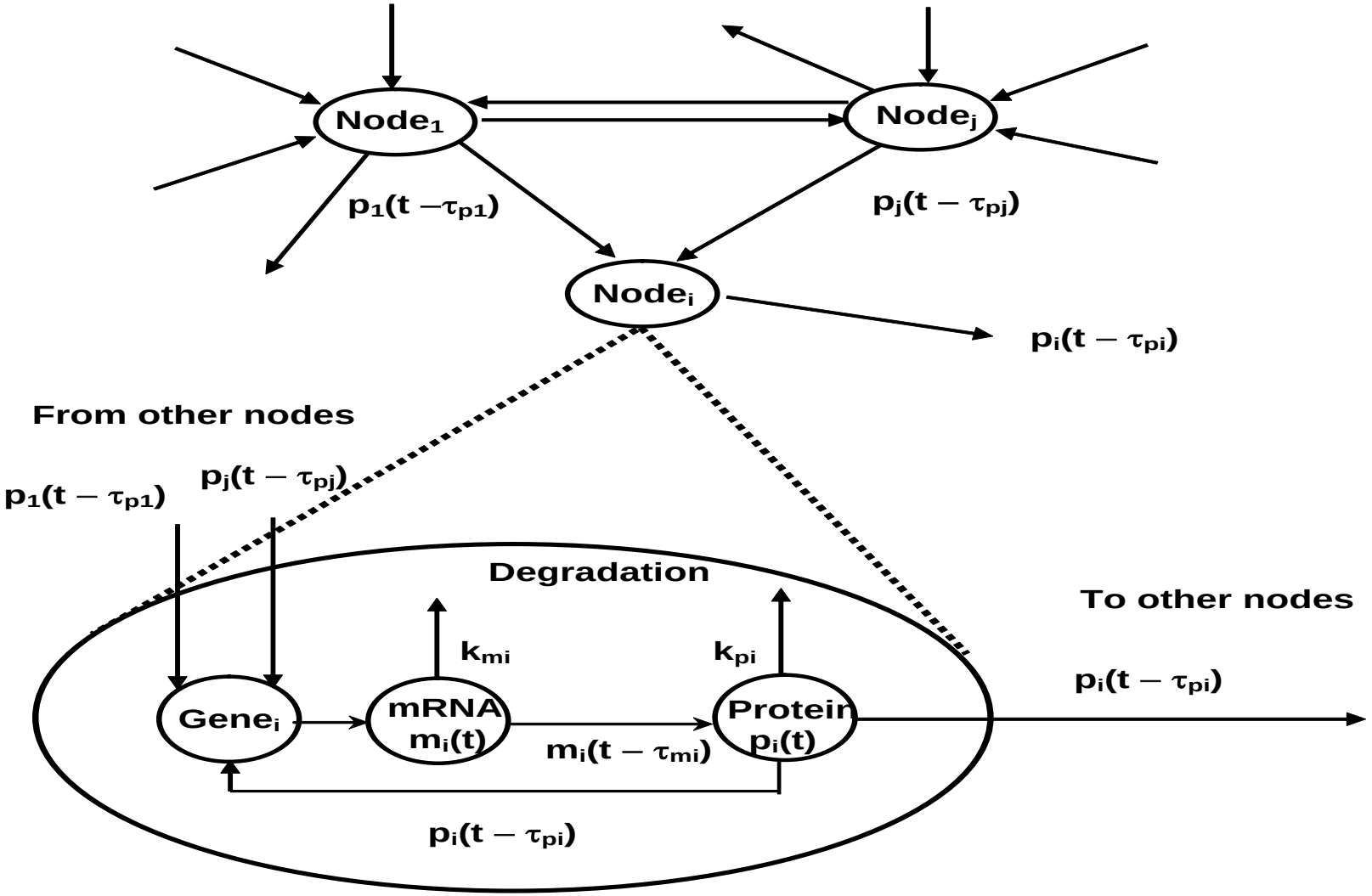


Yeast gene regulatory network (GRN)



- Source: Inferring GRN in yeast (Chen et al., Sci. Rep., 2019)

GRN: towards mathematical model



System of differential equations

mRNA levels

w_{ij} coefficients of gene-protein interaction matrix

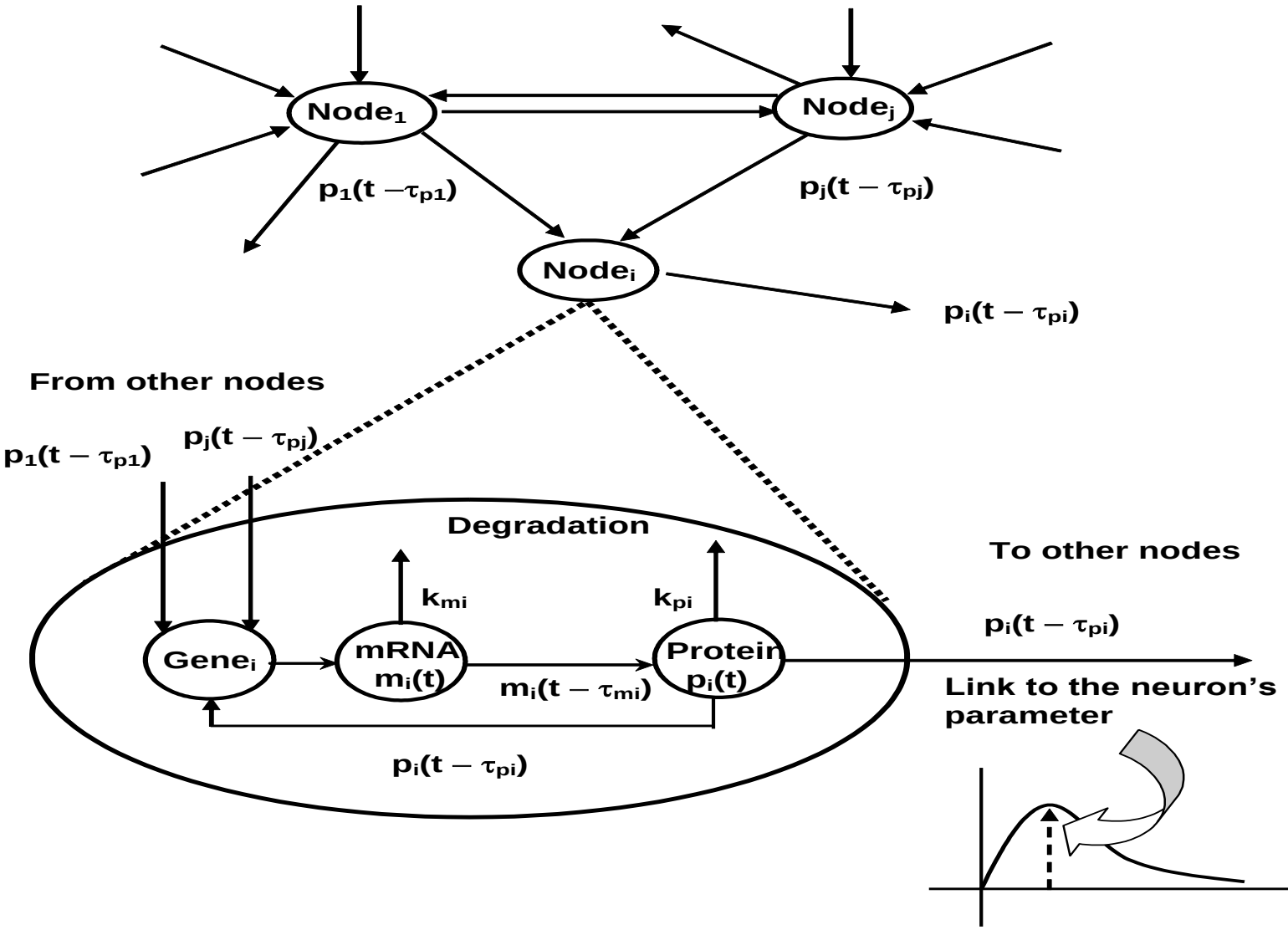
$$\frac{dm_i}{dt} = A_{m_i}\ \sigma_{m_i}\left(\sum_{j=1}^{n} w_{ij} p_j(t-\tau_{p_j}) + \sum_{k=1}^{K} v_{ik} x_k(t-\tau_{x_k}) + b_{m_i} \right) - \lambda_{m_i} m_i(t)$$

Protein levels

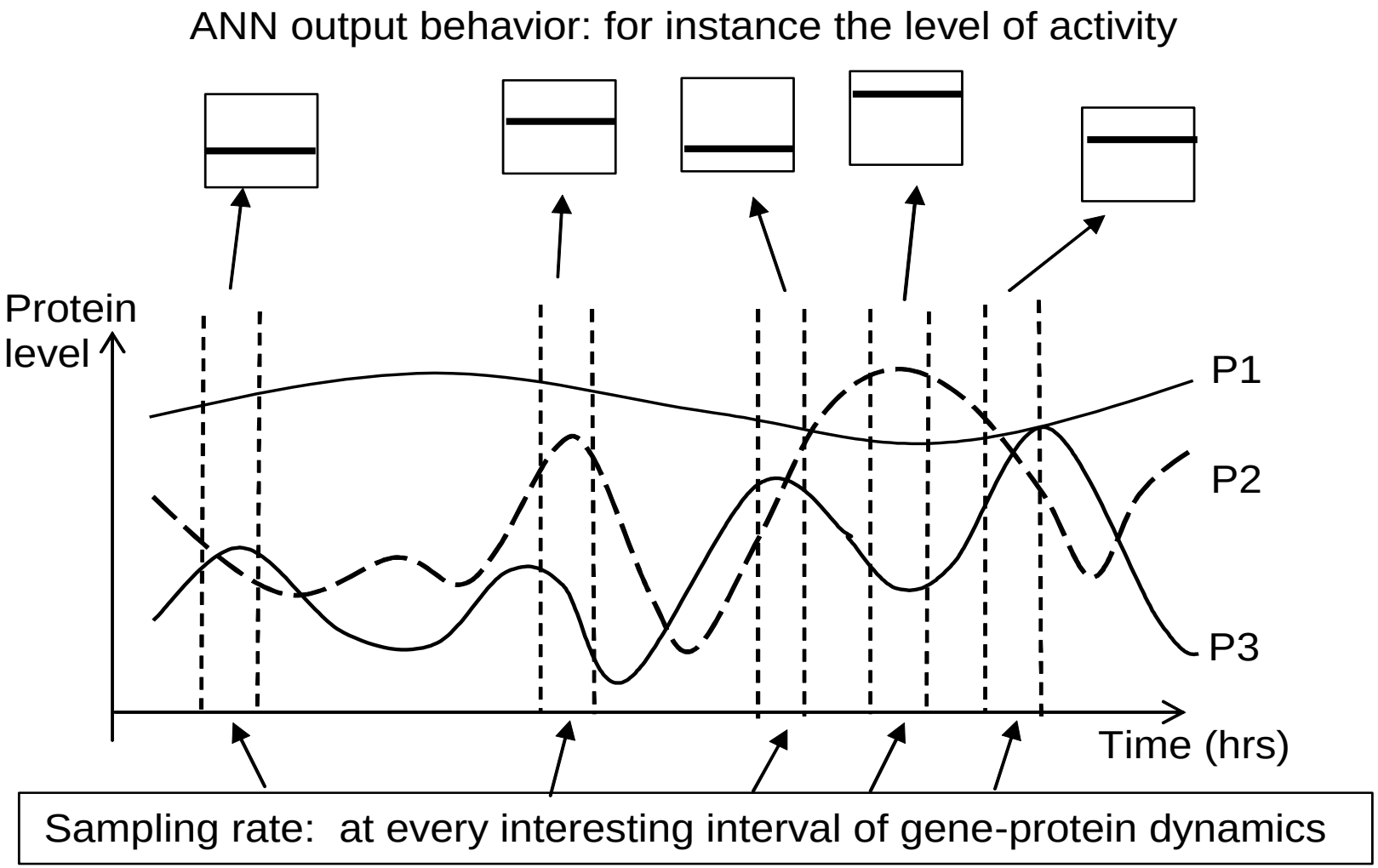
External factors (hormones, drugs,…)

$$\frac{dp_i}{dt} = A_{p_i}\ \sigma_{p_i}\left(m_i(t-\tau_{m_i}) + \sum_{k=1}^{K'} u_{ik} y_k(t-\tau_{y_k}) + b_{p_i} \right) - \lambda_{p_i} p_i(t)$$

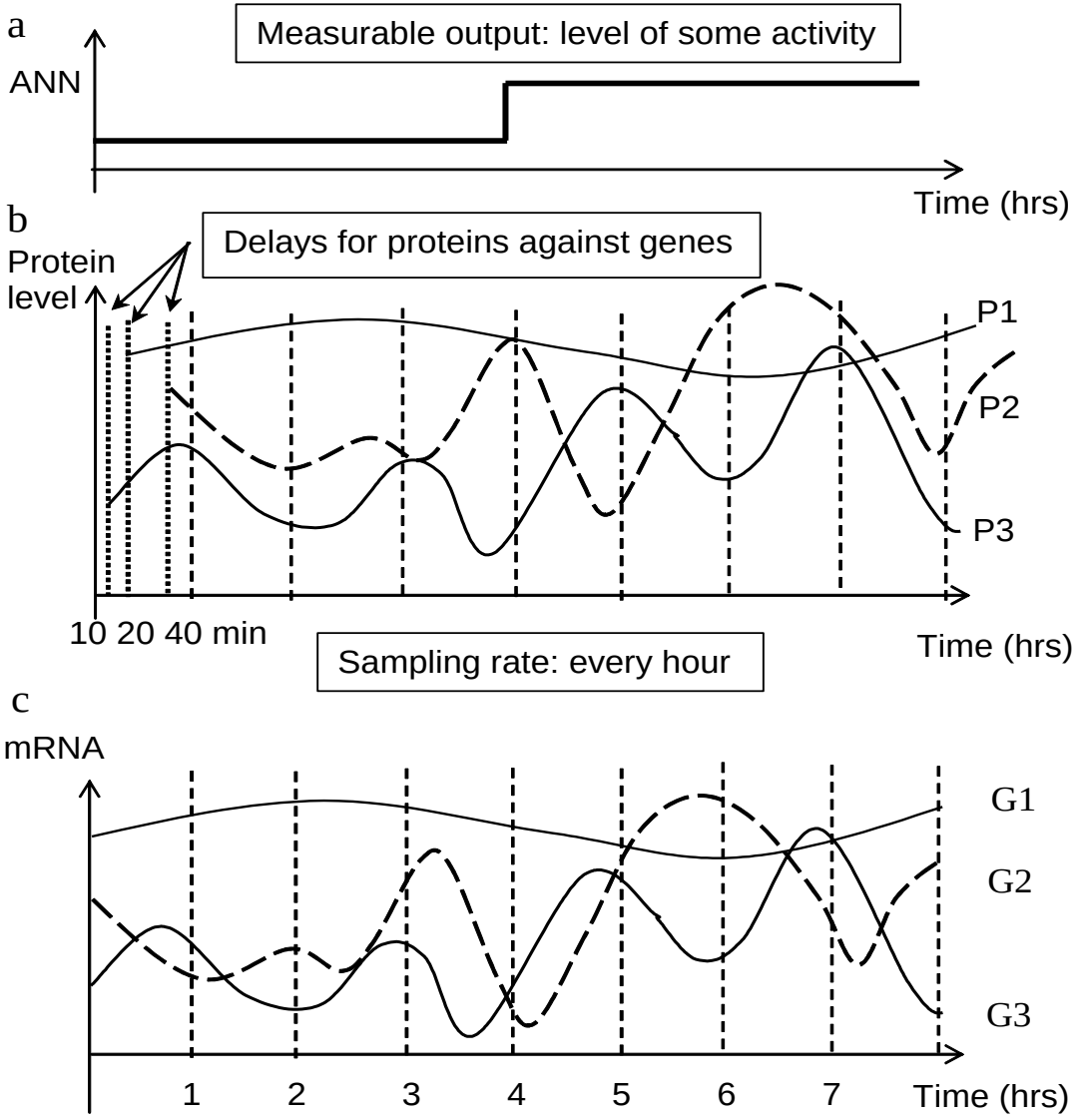
Principle of CNGM: neural model parameters are linked to protein levels $p_i(t)$



Output behaviour of model neurons depends on underlying gene/protein dynamics



CNGM: three embedded multi-scale dynamic systems



Examples of neurogenetic models

- Marnellos G, Mjolsness ED (2003) Gene network models and neural development. In: *Modeling Neural Development*, ed. by A. van Ooyen, MIT Press, Cambridge.

- Smolen P et al. (2004) Simulation of Drosophila circadian oscillations, mutations, and light responses by a model with VRI, PDP-1, and CLK. *Biophysical J.* 86, 2786–2802.

- Benuskova L, Kasabov N (2007) Modeling L-LTP based on changes in concentration of pCREB transcription factor. *Neurocomputing* 70, 2035-2040.

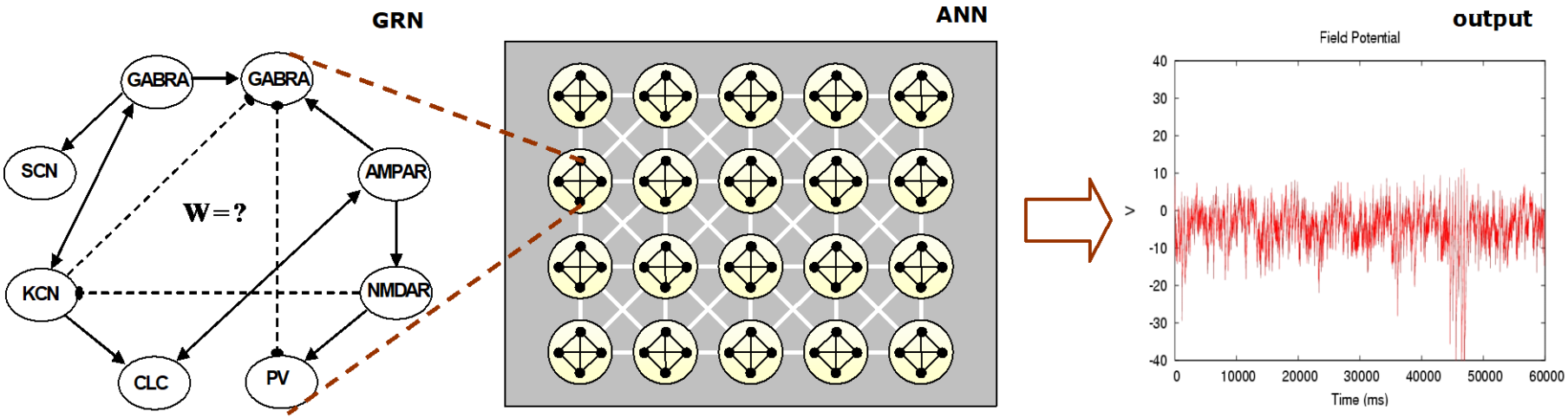
- Benuskova L, Kasabov N (2008) Modeling brain dynamics using computational neurogenetic approach. *Cognitive Neurodynamics* 2(4), 319-334.

- Benuskova L, Kasabov N (2014) Computational neurogenetic modeling: gene-dependent dynamics of cortex and idiopathic epilepsy. In: N. Kasabov (Ed), *Springer Handbook of Bio-/Neuroinformatics*, Springer-Verlag, Berlin/Heidelberg.

- Nido GS, Ryan MM, Benuskova L, Williams JM (2015) Dynamical properties of gene regulatory networks involved in LTP. *Front. Mol. Neuroscience* 8:42.

Example: electrical activity

Question: which gene interactions lead to the desired spectral characteristics of Local Field Potential?



Gene regulatory network

Abstract connections **W**

Model ANN

ANN output LFP

Gene-protein regulatory network model

mRNA levels

$$\frac{dm_i}{dt} = A_{m_i}\ \sigma_{m_i}\left(\sum_{j=1}^{n} w_{ij}\, p_j(t-\tau_{p_j}) + b_{m_i}\right) - \lambda_{m_i} m_i(t)$$

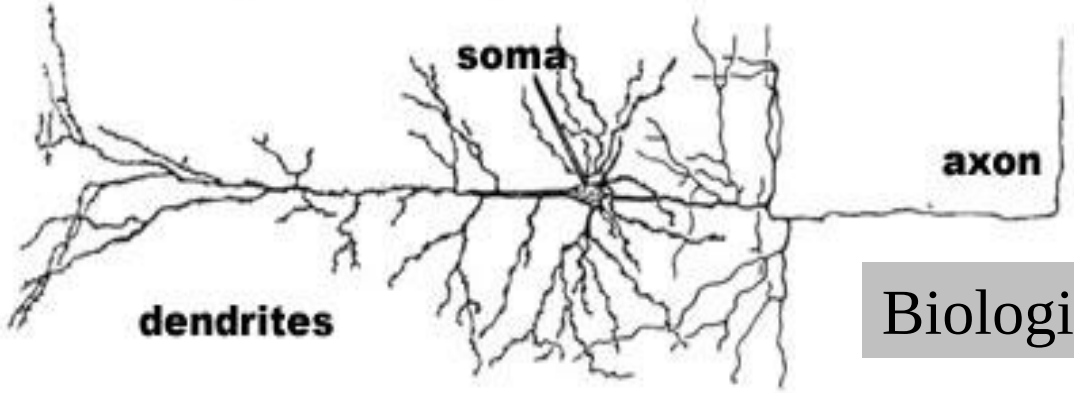
Protein levels

$$\frac{dp_i}{dt} = A_{p_i}\ \sigma_{p_i}\left(m_i(t-\tau_{m_i}) + b_{p_i}\right) - \lambda_{p_i} p_i(t)$$

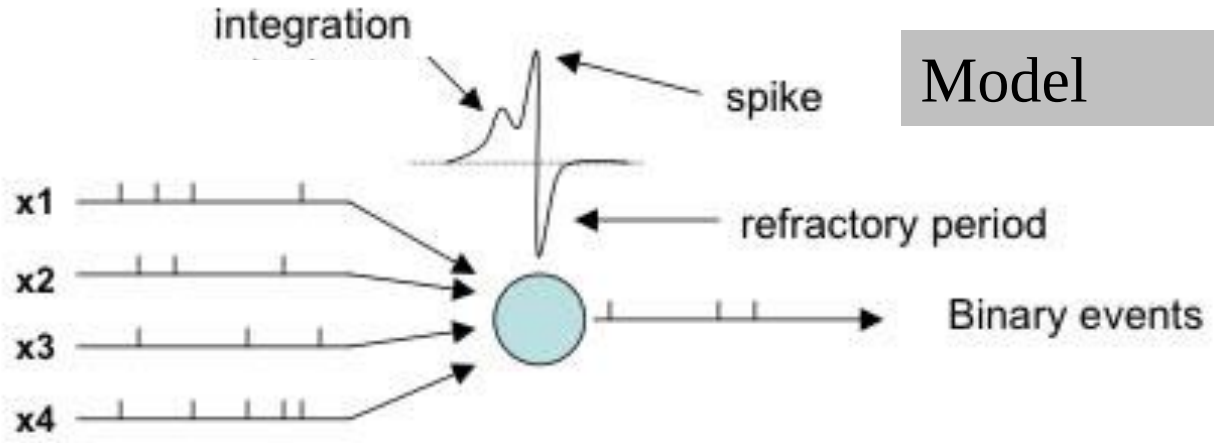
Values of neuronal parameters will depend on levels of proteins *p*

$$P_j(t) = P_j(0)\, p_j(t)$$

Neuron model



Biological neuron

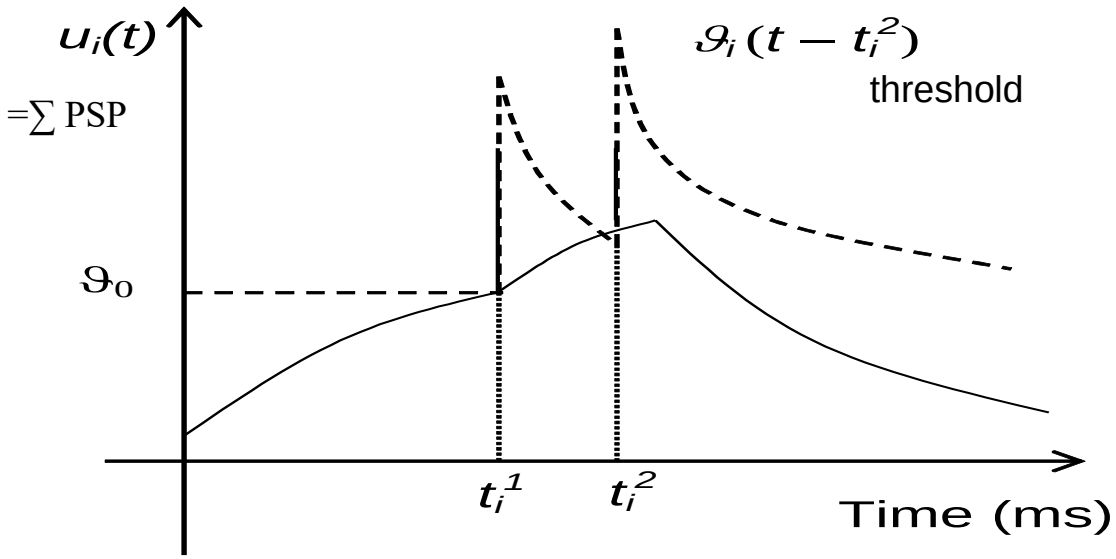


Model

Values of neuron model paramaters will be linked to the levels of proteins (like receptors and ion channels)

Spiking neuron model

Spike Response Model



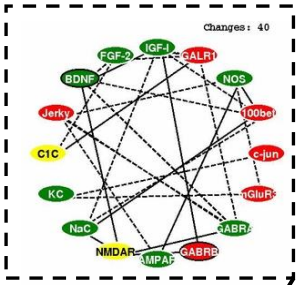
$$PSP_{ij}^{type}(t - t_j - \Delta_{ij}^{ax}) = A^{type}\left(\exp\left(-\frac{t - t_j - \Delta_{ij}^{ax}}{\tau_{decay}^{type}}\right) - \exp\left(-\frac{t - t_j - \Delta_{ij}^{ax}}{\tau_{rise}^{type}}\right)\right)$$

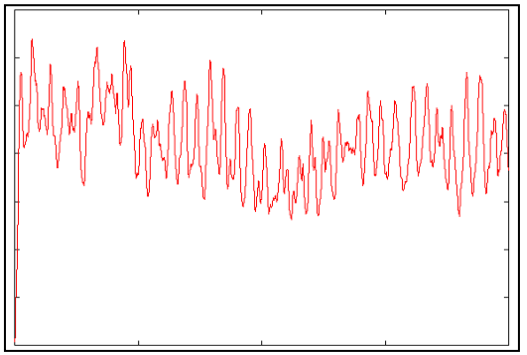
type = fast excitation, slow_excitation, fast_inhibition, slow_inhibition

Neural network model of cerebral cortex with input from thalamus

OUTPUT : Local Field Potential

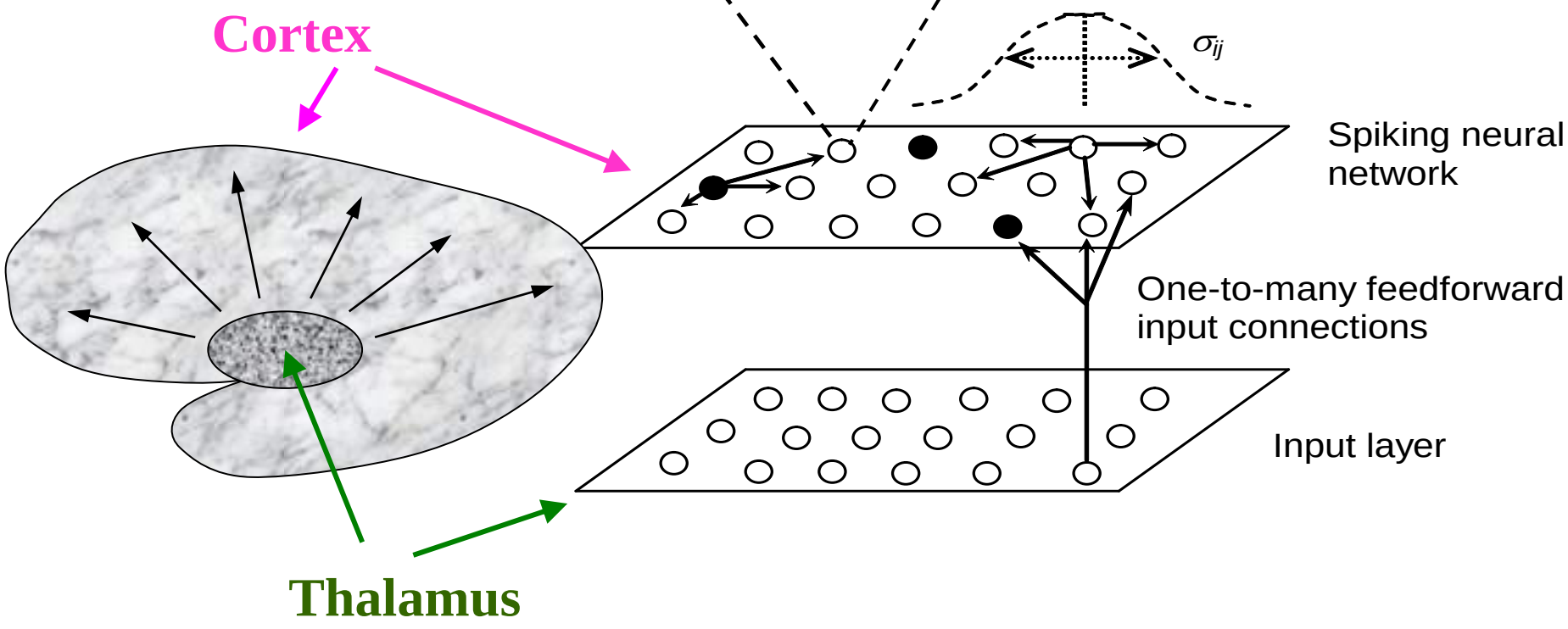
$$LFP(t) = \sum u_i(t) \approx EEG = \sum LFP(t)$$





J_{ij}

The same GRN in each neuron



Credits: Simei Gomes Wysoski

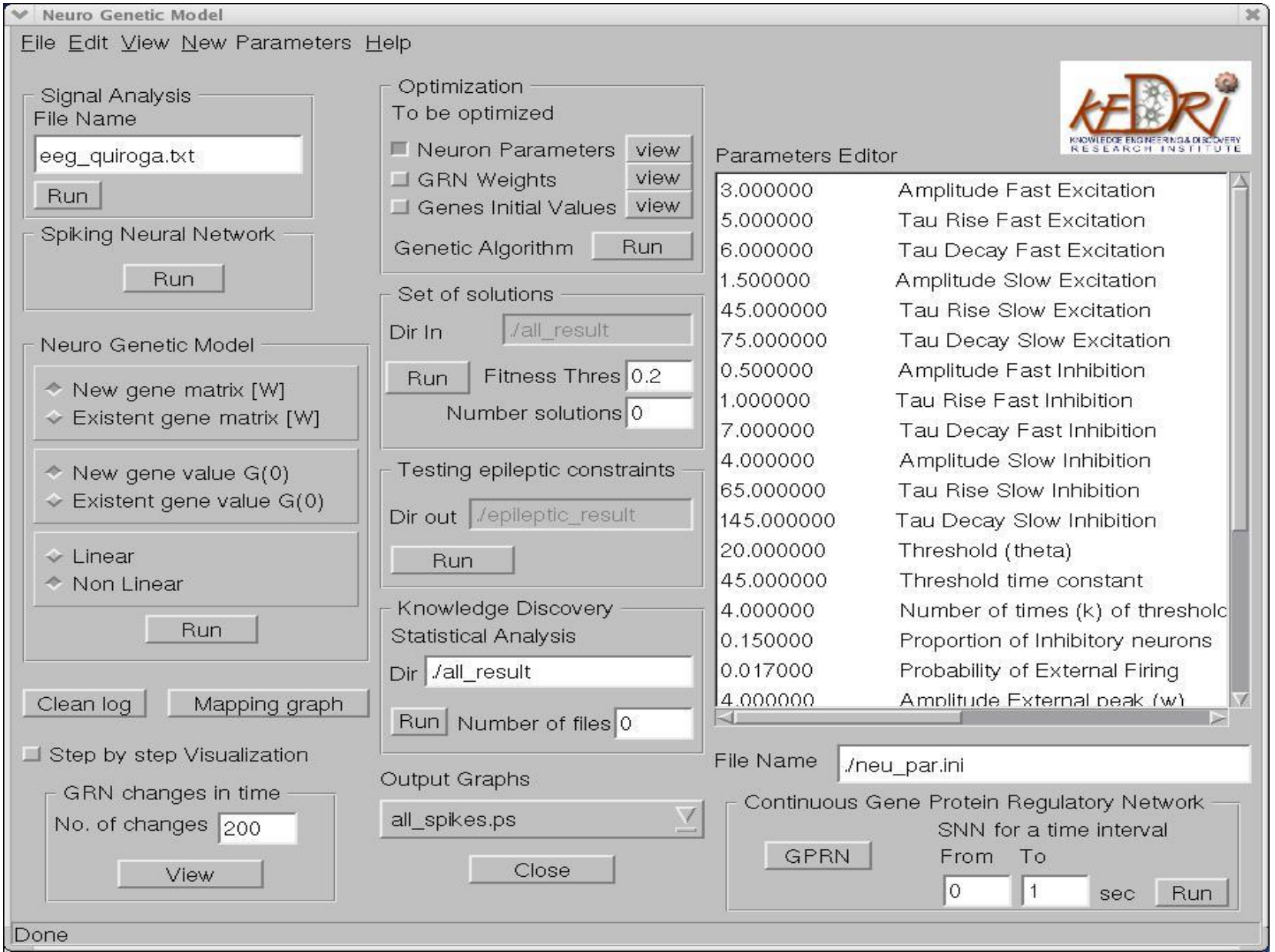


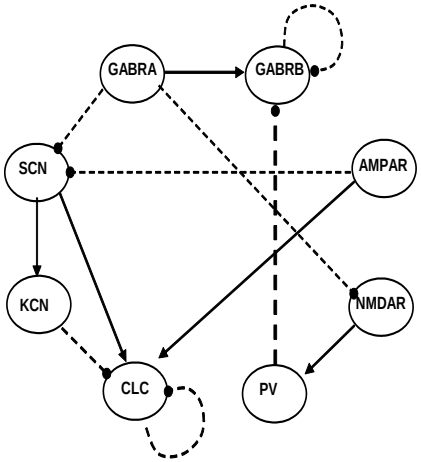
Table of neuron parameters

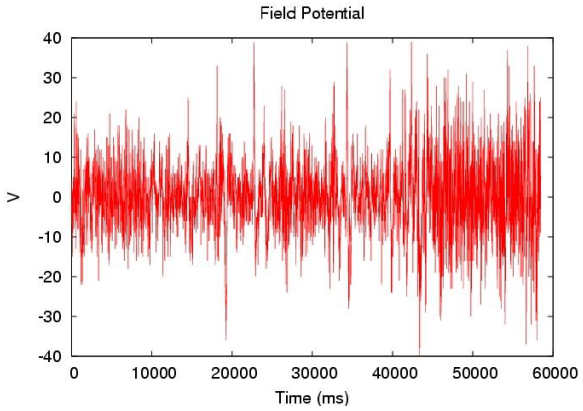
NEURON'S PARAMETERS	PROTEIN	RANGE of INTIAL VALUES
Fast excitation: Amplitude rise / decay time constants (ms)	AMPAR	0.5 – 3.0 1–5 / 5–10
Slow excitation: Amplitude rise / decay time constants (ms)	NMDAR	0.5 – 4.0 10–50 / 30–50
Fast inhibition: Amplitude rise / decay time constants (ms)	GABRA	4 – 8 5–10 / 20–30
Slow inhibition: Amplitude rise / decay time constants (ms)	GABRB	5 – 10 20–80 / 50–150
Resting firing threshold, decay time constant (ms)	SCN	17 – 25 5 – 50

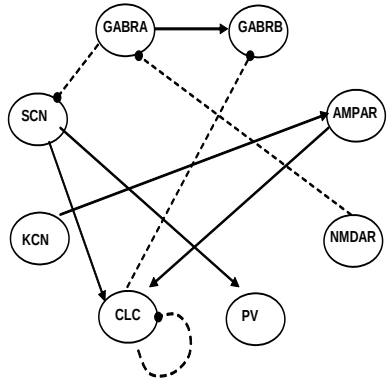
Table of network parameters

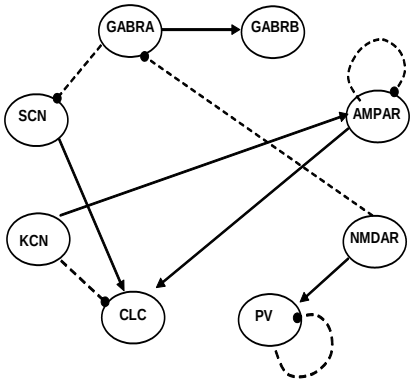
SNN PARAMETER	VALUE
Number of neurons	120
Proportion of inhibitory neurons	0.2
Probability of external (thalamic) fiber firing	0.015
Peak/sigma of external input (TC) weight	5 / 1
Peak/sigma of lateral excitatory weights	10 / 4
Peak/sigma of lateral inhibitory weights	40 / 6
Probability of connection	0.5
Unit delay in excitatory/inhibitory spike propagation	1 / 2 ms

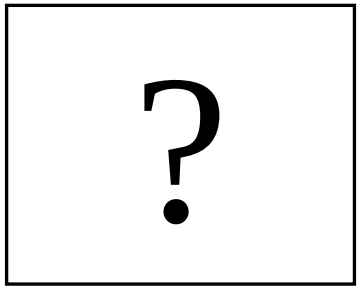
Question: which gene interactions lead to the desired spectral characteristics of LFP/EEG?

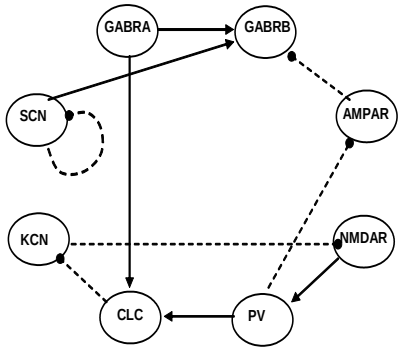








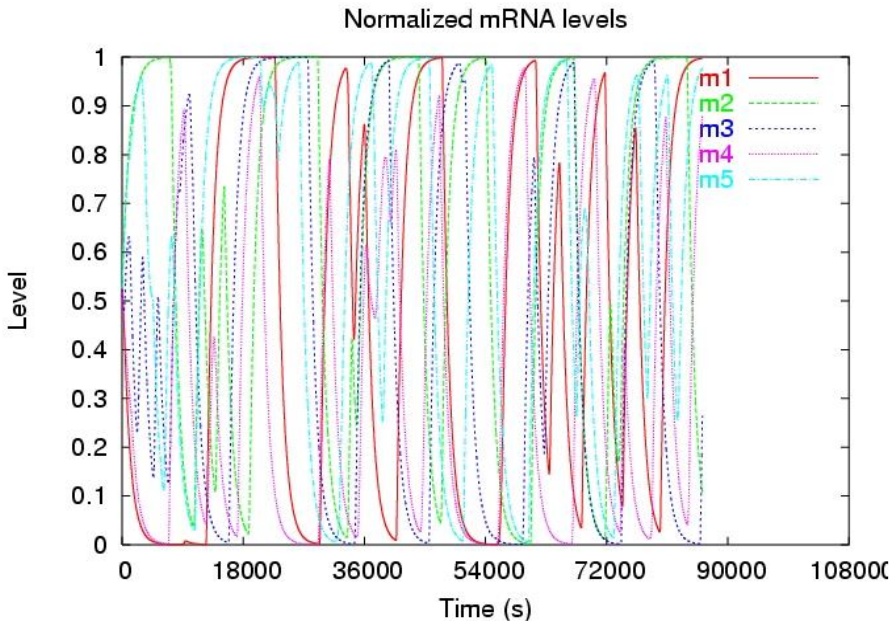


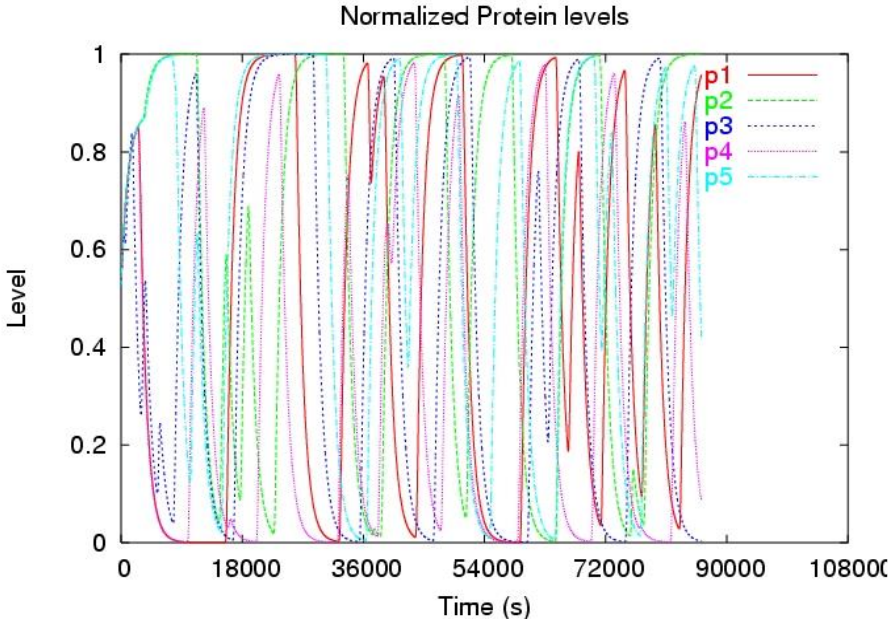


Toy example

- We generated an artificial gene interaction matrix W leading to a complex gene and protein dynamics over 24 hrs (only 5 genes/proteins)

- These were "genetic variables" for AMPAR (fast excitation), NMDAR (slow excitation, GABRA (fast inihibition), GABRB (slow inhibition), SCN (Firing threshold)

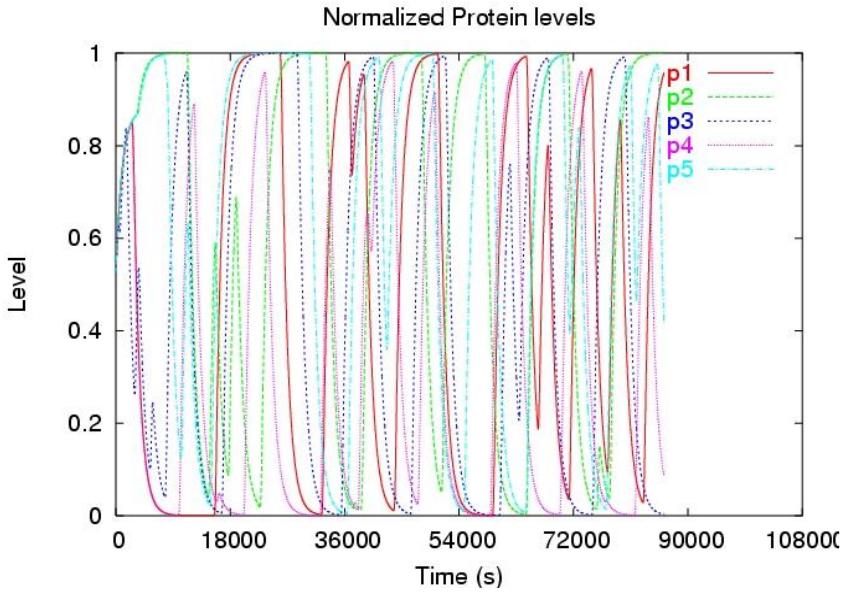


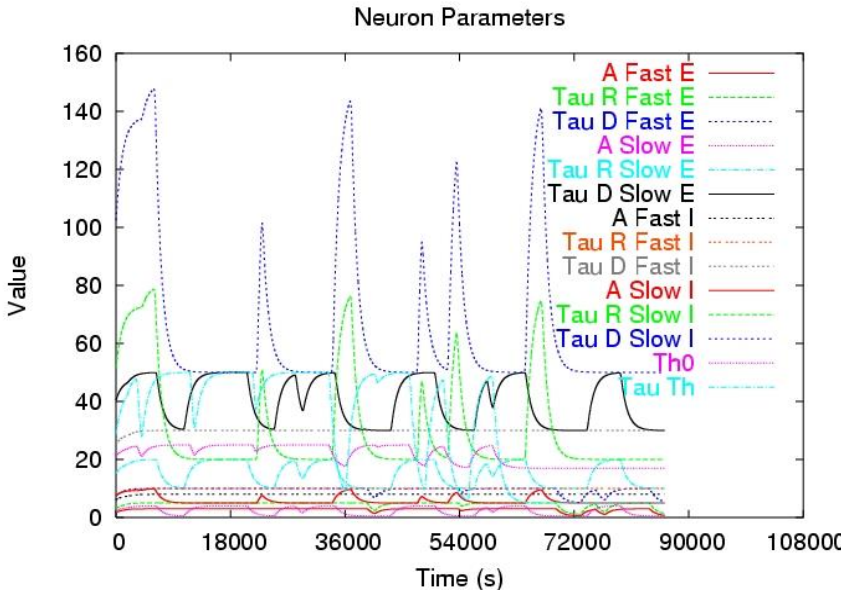


Toy example: complete genome

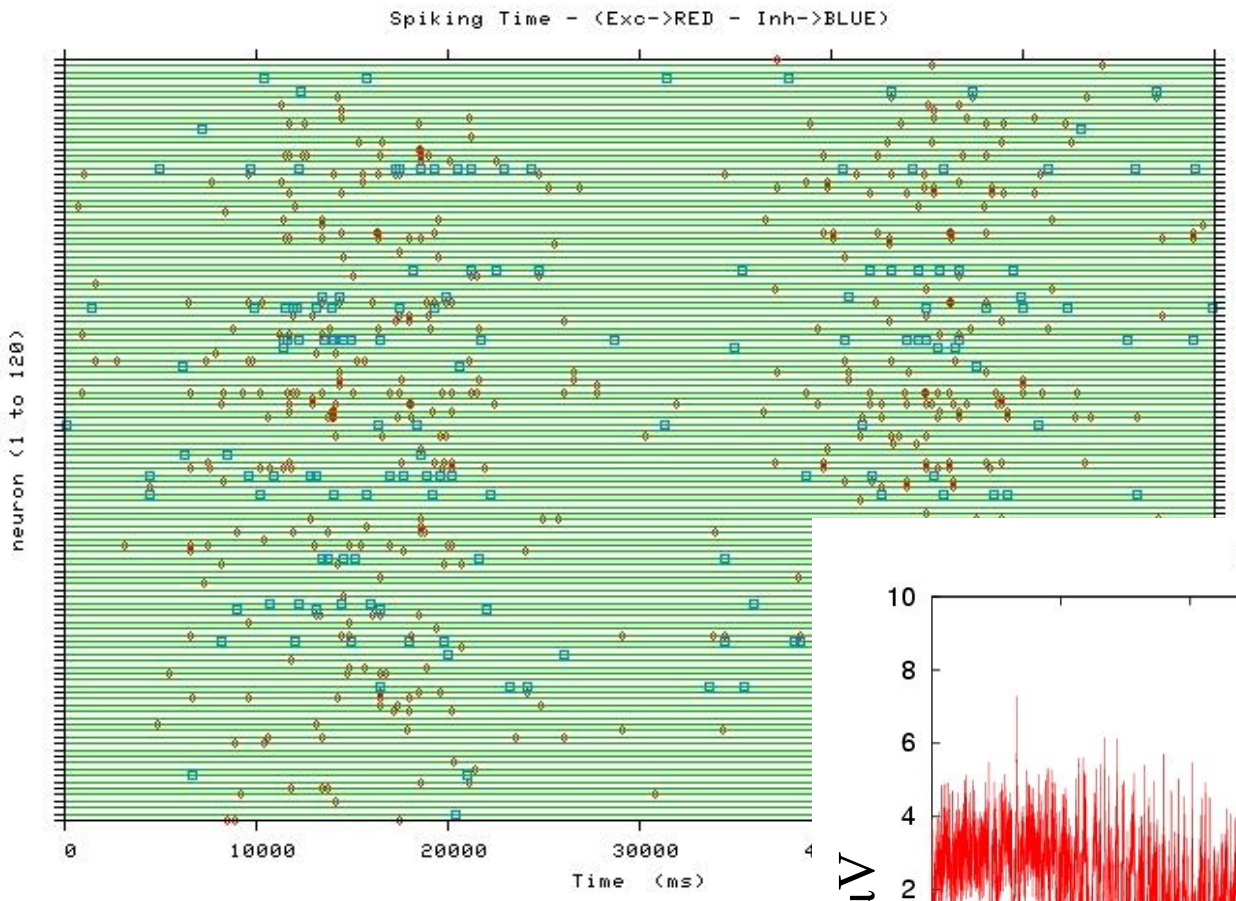
Value of parameter is proportional to the level of a protein

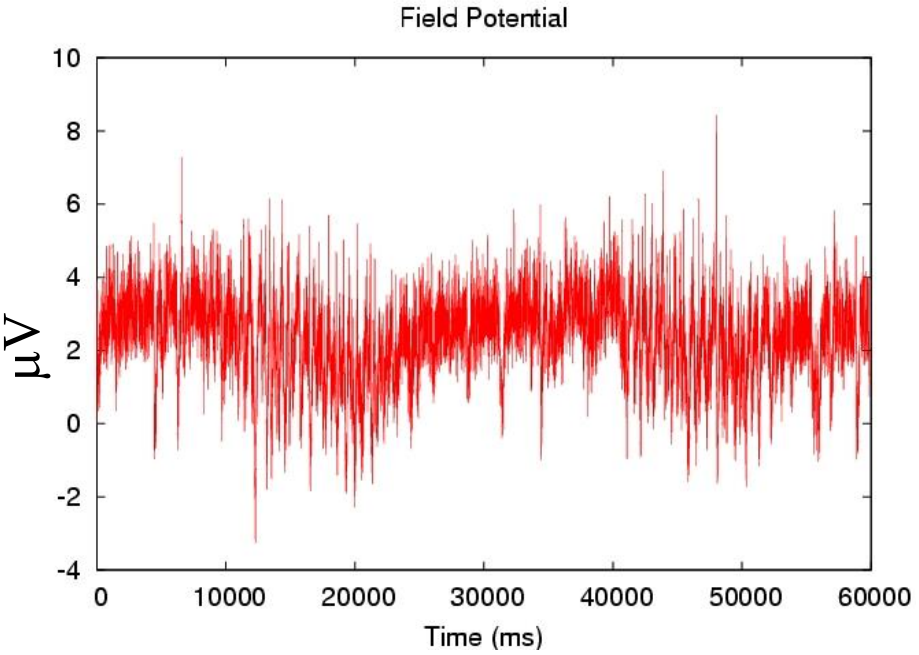
$$P_j(t) = P_j(0)\, p_j(t)$$





Asynchronous neural activity



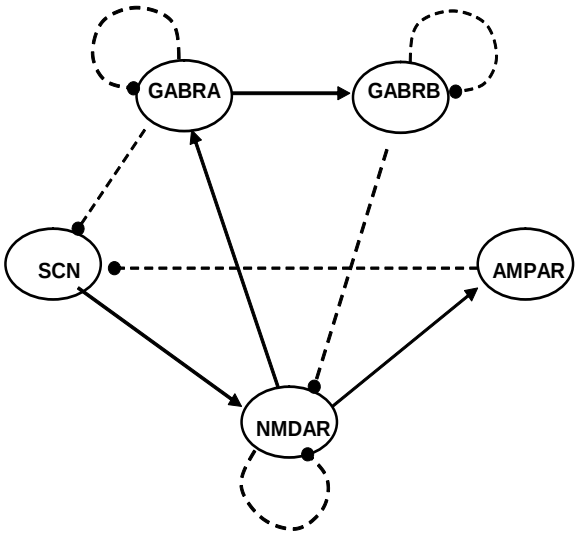


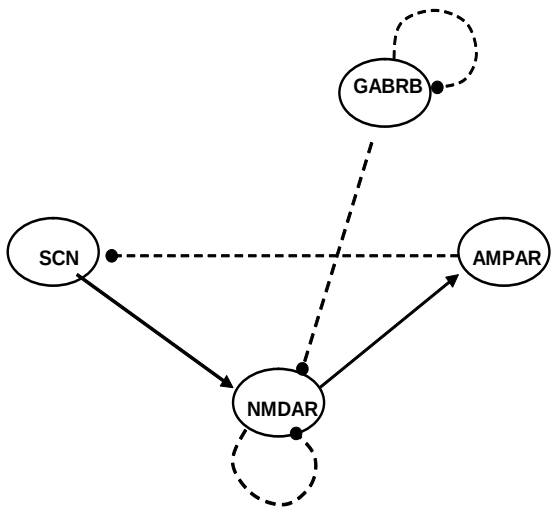
Some types of epilepsy are genetically caused

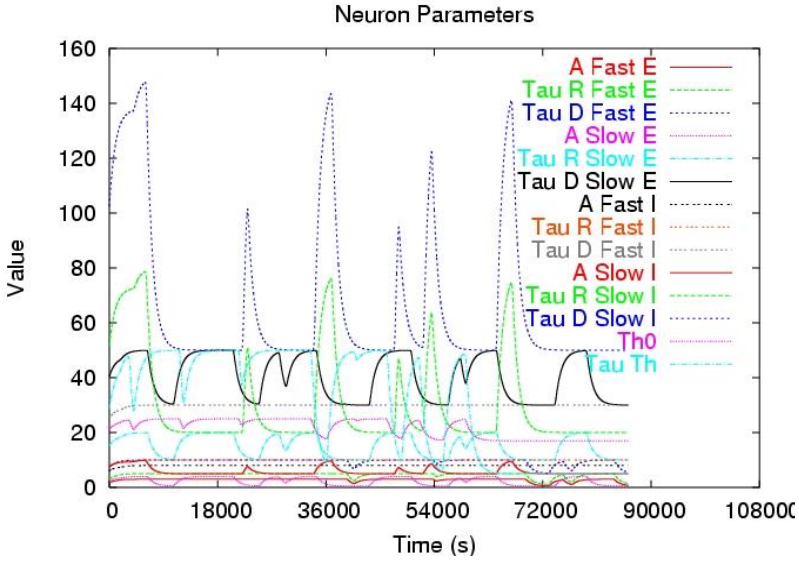
- Epilepsy is a group of neurological disorders characterized by recurrent epileptic seizures. An epileptic seizure is the clinical manifestation of an abnormal, excessive and synchronized electrical discharge of neurons.

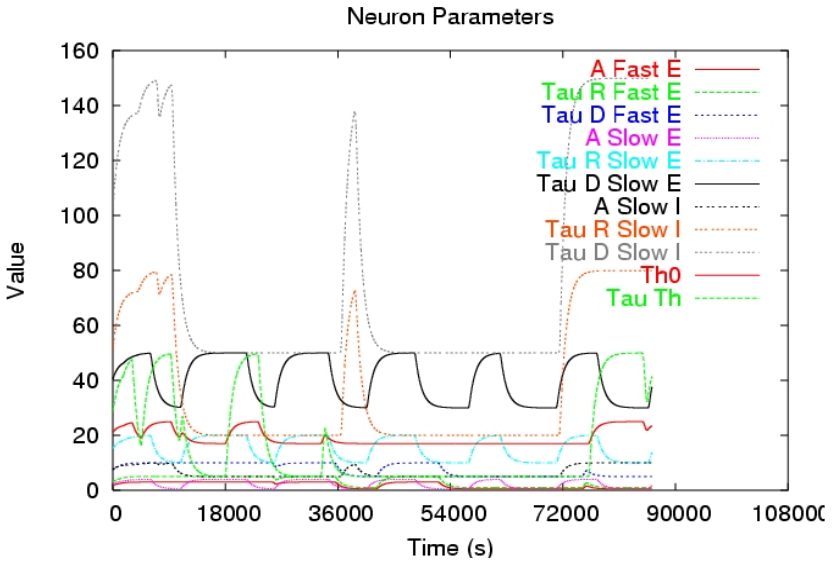
Childhood absence epilepsy (CAE)	γ_2 subunit gene for the $GABA_A$ receptor gene GABRG2/5q gene CLCN2/3q	Fast and part of slow GABAergic inhibition is reduced, voltage-gated Cl^- channel function is impaired	Absence seizures (consciousness impaired) up to 200 times a day, bilateral 2–4 Hz spike and slow-wave EEG
Generalized epilepsy and febrile seizures plus (GEFS+)	β_1 subunit of the Na^+ channel gene SCN1B/19q α_1 and α_2 subunits, gene SCN1A and gene SCN2A/2q GABRG2/5q	Normal inactivation kinetics of the Na^+ channel is reduced causing persistent Na^+ influx and hyperexcitability, reduced function of the $GABA_AR$	Childhood onset of febrile seizures, with febrile and afebrile generalized seizures continuing beyond 6 years of age
Intractable childhood epilepsy	α_1 subunit of the Na^+ channel, gene SCN1A/2q	Rapid recovery of the Na^+ channel from inactivation or very slow inactivation	Frequent intractable generalized tonic-clonic seizures
Juvenile absence epilepsy (JAE)	α_1/5q, α_5/15q, γ_2/5q subunit genes for the $GABA_A$ receptor gene (CLCN2)/3q	Fast and part of slow GABAergic inhibition is reduced, voltage-gated Cl channel function is impaired	Similar like CAE but the seizures start after year 10, seizures may be less frequent and last longer than few seconds

Toy example: GABRA deleted

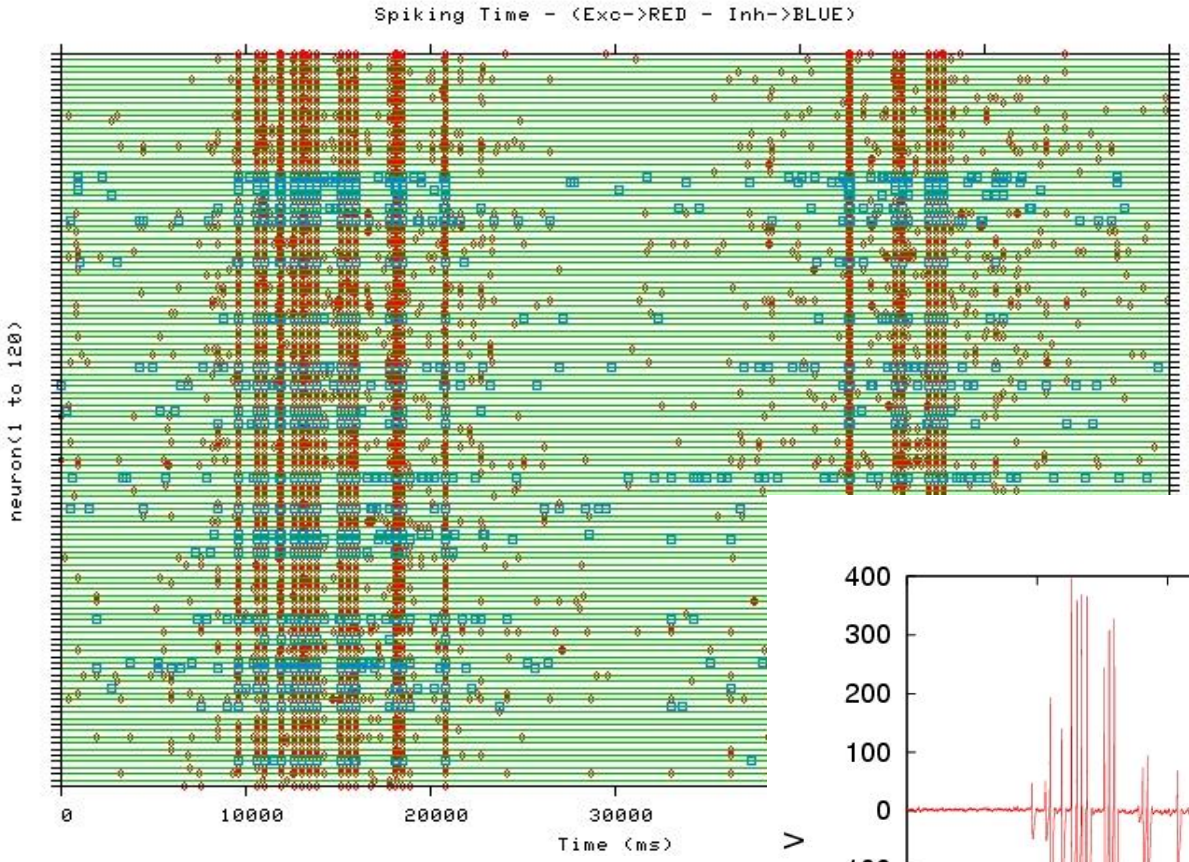


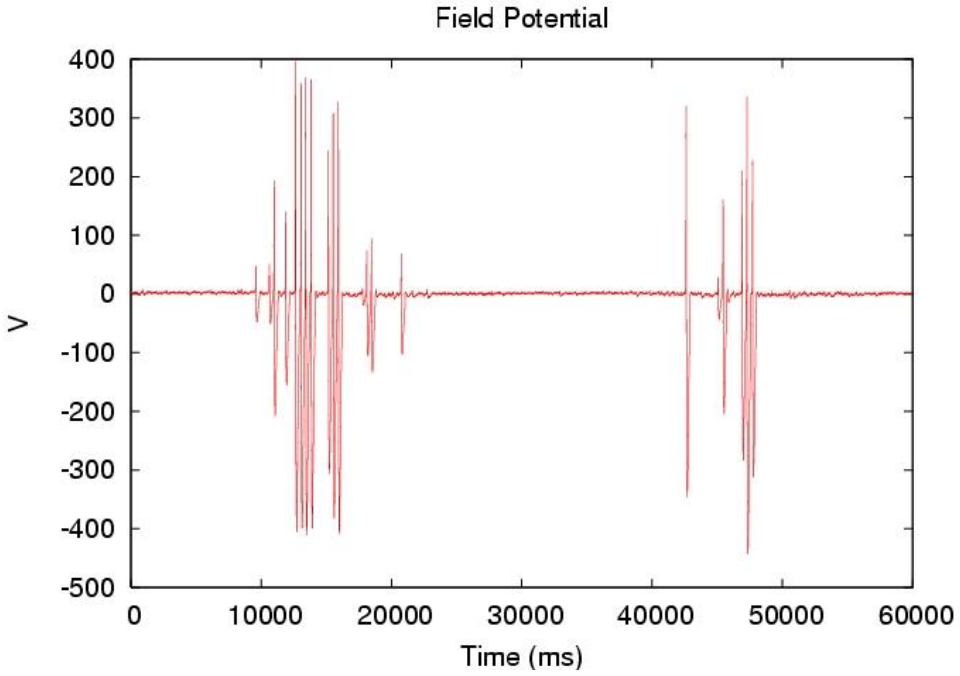




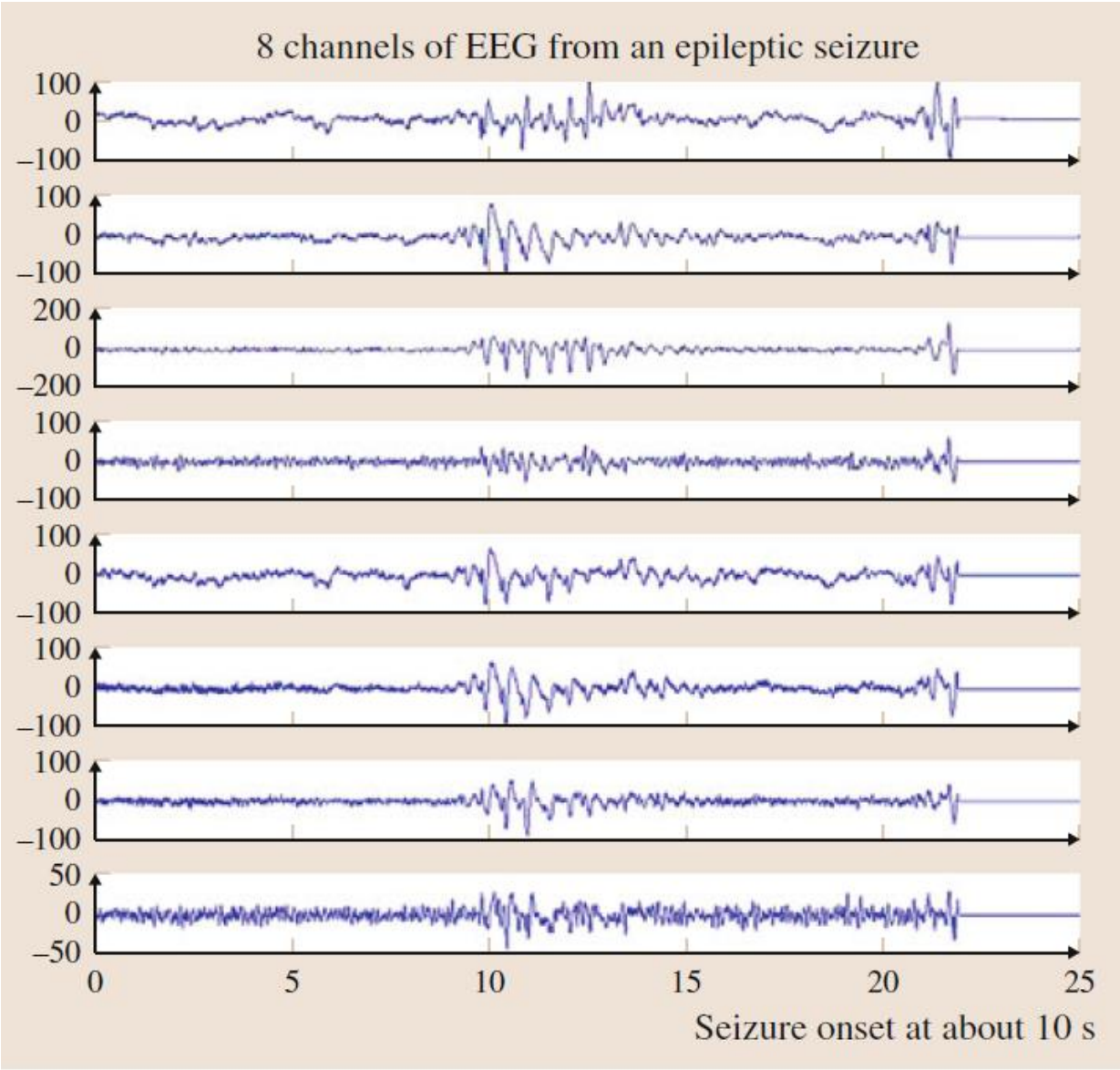


Frequent spontaneous global synchronisations





EEG in CAE



- Recording of eight channels of the normal and epileptic slow-wave discharge (SWD) in childhood absence epilepsy. SWD have large amplitudes and frequency of 2.5–4Hz.

Observations from the toy model

- Coefficients of gene-to-gene interaction were generated randomly. Many random gene-to-gene interactions yielded realistically looking LFP.

- We tested each interaction matrix by deleting the gene variable for GABRA. Some of interaction matrices yielded permanent synchronisation, some produced occasional synchronisations, some produced no synchronisations at al.

- This simple model shows gene mutations/deletions may have no effect on neural dynamics, everything depends on the rest of the GRN interactions and the function of the remaining genes/proteins.

- Or is this result only an artefact related to this simplified model?