Integrated tools to investigate the molecular basis of diseases: computational and experimental analysis of the impact of protein variants on protein stability and function.

Sapienza Università di Roma

Impact of non synonymous single nucleotide variants on protein fitness: experimental analysis for a comparative study

SAPIENZA
UNIVERSITÀ DI ROMA

Dipartimento di Scienze Biochimiche
"A. Rossi Fanelli"

Roberta Chiaraluce, Valerio Consalvi
Leonore Novak, Maria Petrosino

Sapienza University of Rome

ENEA
Italian National Agency for New Technologies, Energy and Sustainable Economic Development

Alessandra Pasquo

GOETHE UNIVERSITÄT FRANKFURT AM MAIN
SGC
Buchmann Institute for Molecular Life Sciences

Stefan Knapp

nsSNVs

Local structural changes
Modifications of the dynamic properties
Changes in stability

New and unknown interactions?

New pattern of interactions?

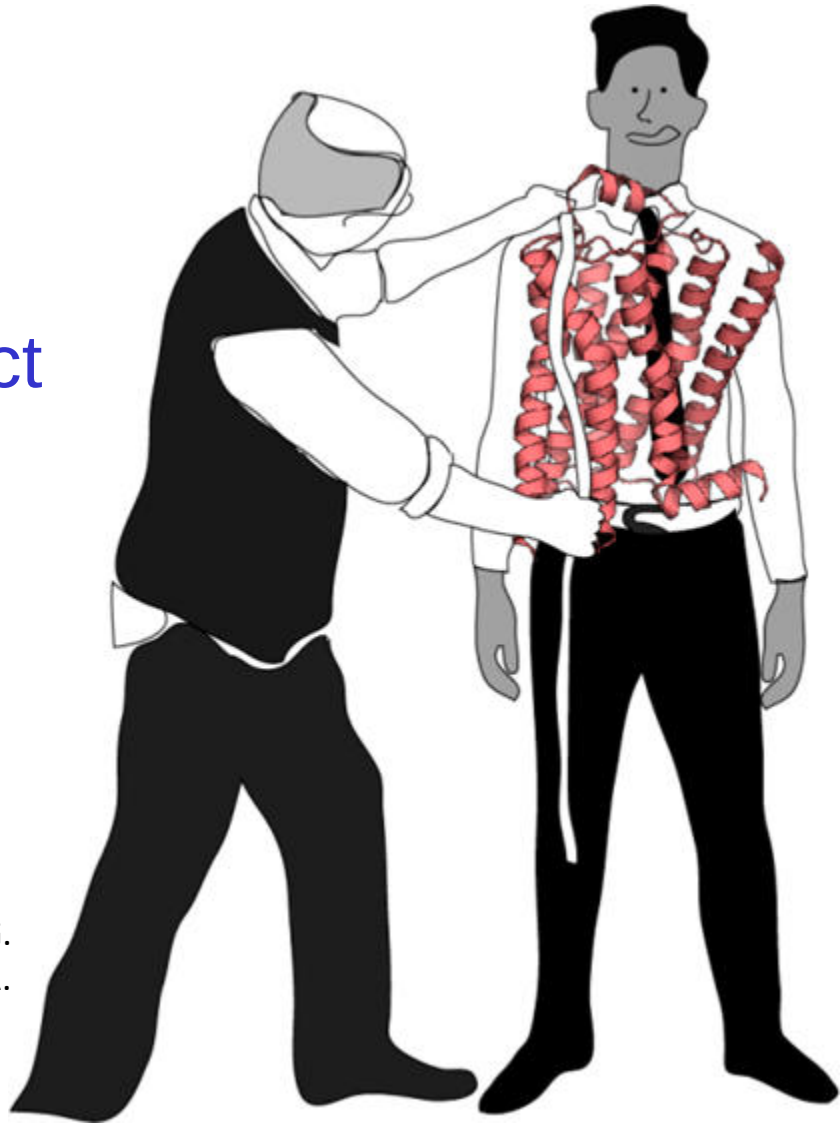
Alternative conformational states?

Novel and alternative network of protein-protein interconnection?

Novel and alternative functions?

nsSNVs

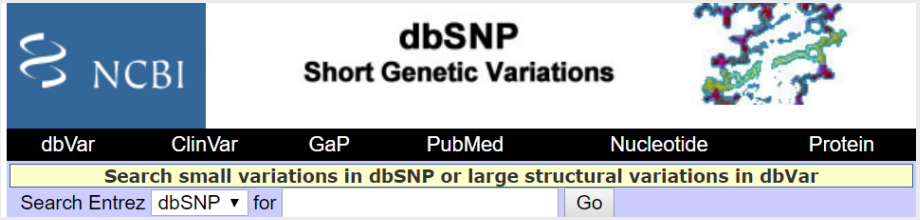
Do the natural variants affect the response to drugs?



Personalized Biochemistry and Biophysics Brett M. Kroncke, Carlos G. Vanoye, Jens Meiler, Alfred L. George, Jr. and Charles R. Sandershttp://dx.doi.org/10.1021/acs.biochem.5b00189

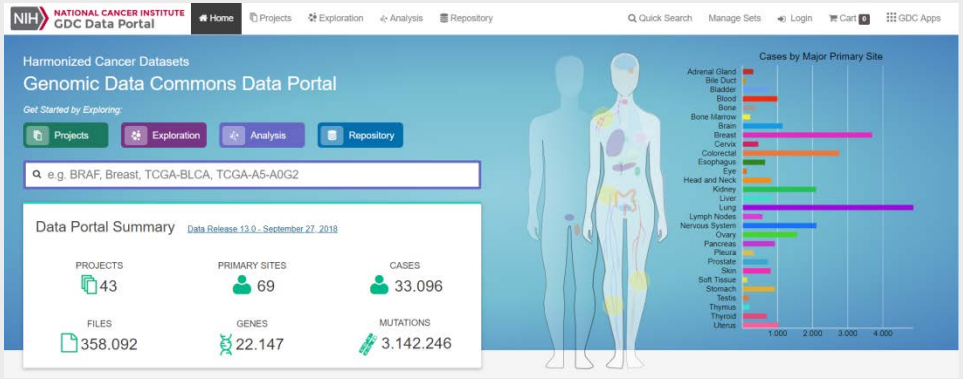
- Databases catalogue for human nsSNVs

dbSNP
(Single Nucleotide Polymorphism database)



COSMIC
(Catalogue of Somatic Mutations in Cancer)

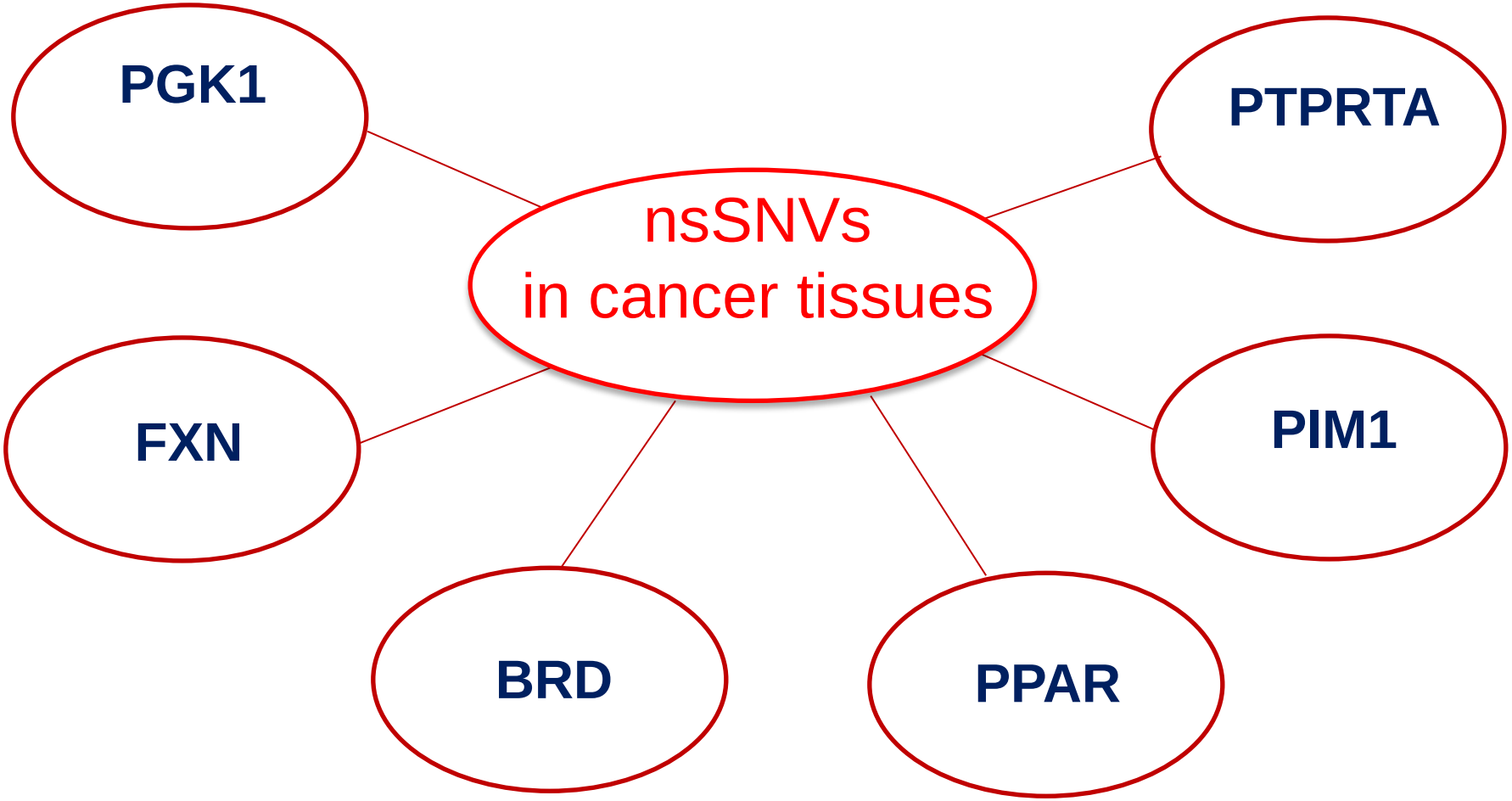
Pan - Cancer Atlas

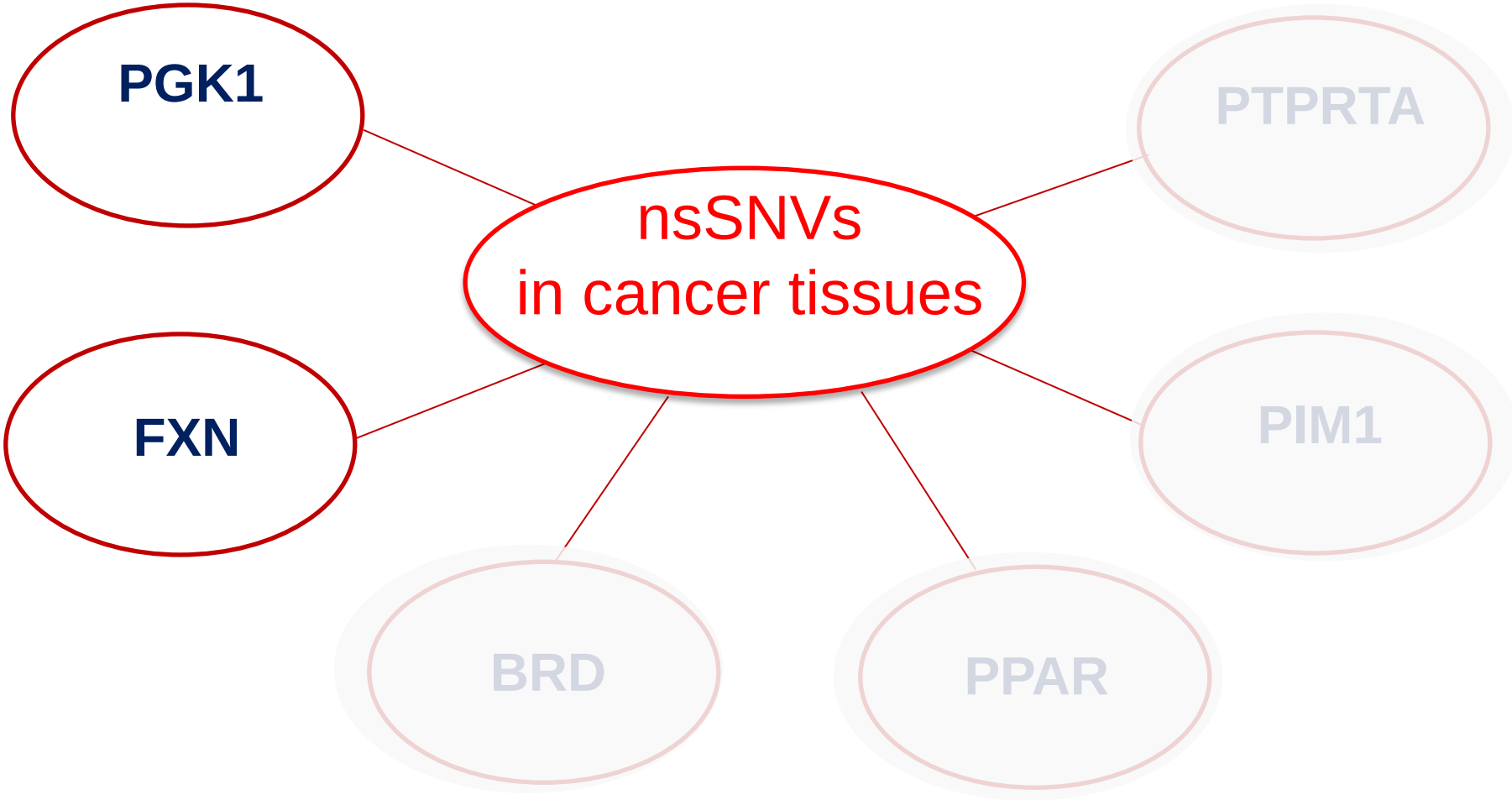


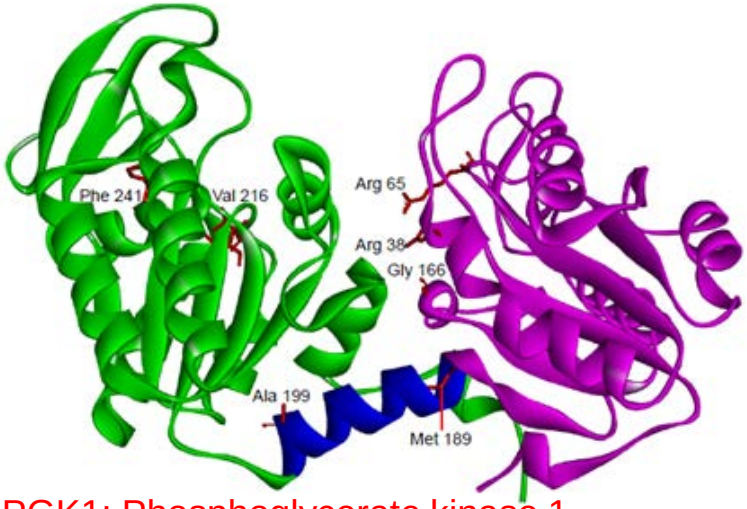
SOME CASE STUDIES…

nsSNVs related to pathological states in humans from available databases

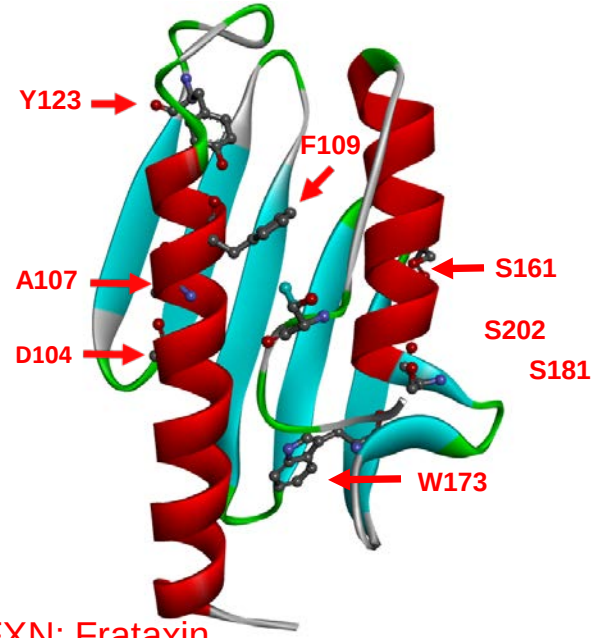
Proteins involved in disease with known crystal structure


PGK1
PTPRTA
nsSNVs
in cancer tissues
FXN
PIM1
BRD
PPAR

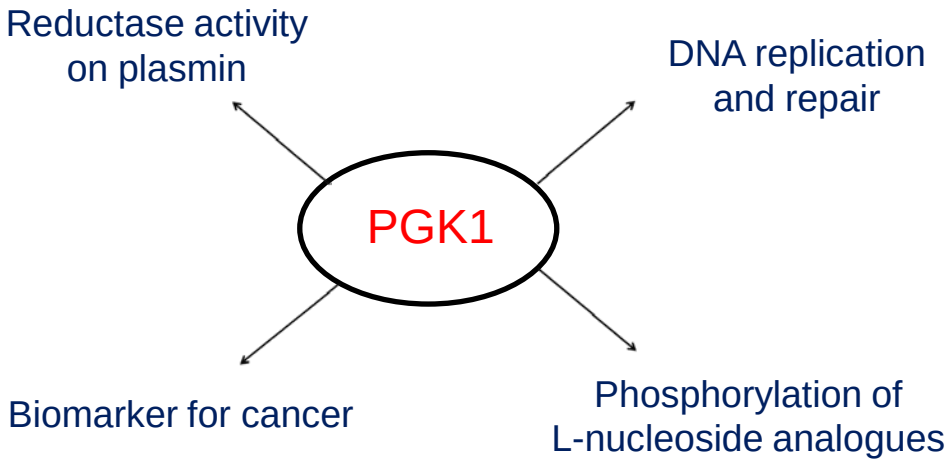

PGK1
FXN
nsSNVs
in cancer tissues
PTPRTA
PIM1
BRD
PPAR

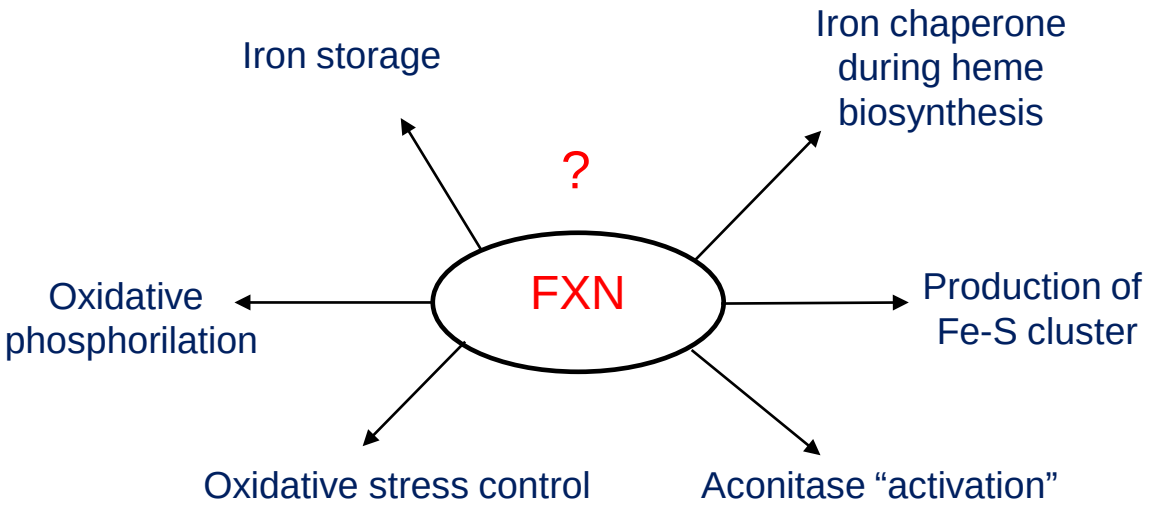


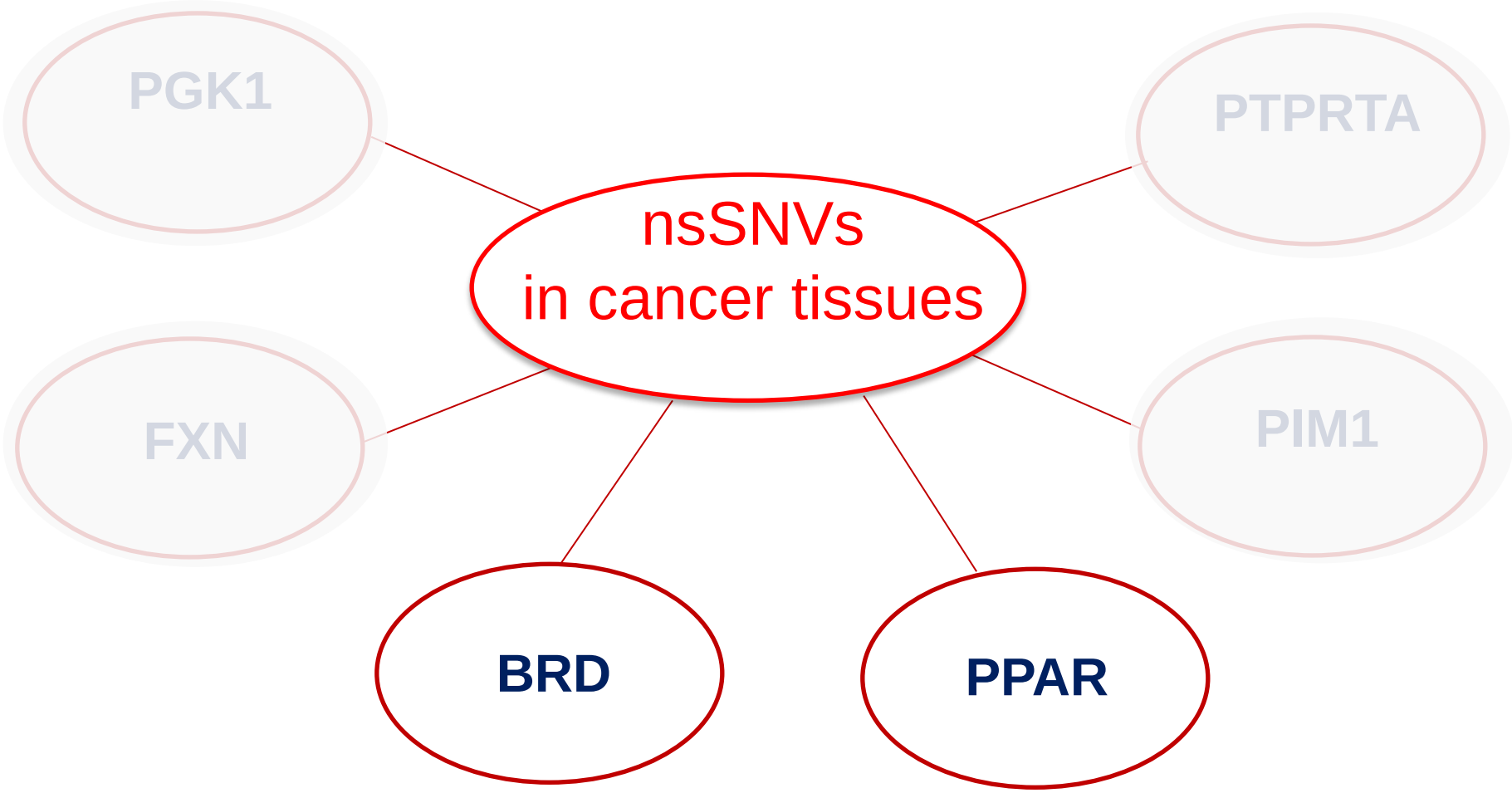
PGK1: Phosphoglycerate kinase 1
PDB code: 2XE7
Fiorillo A. et al., PlosOne 2018

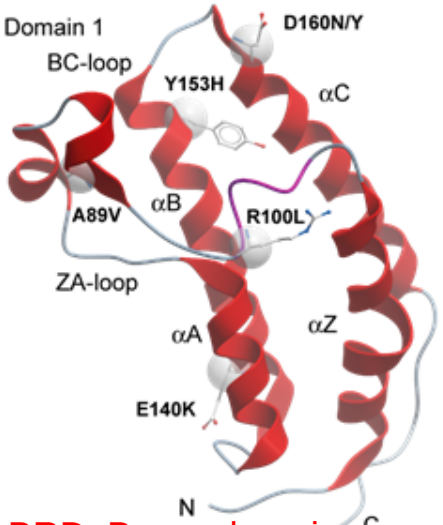


FXN: Frataxin
PDB code: 1EKG
Petrosino et al., Hum. Mut. 2019

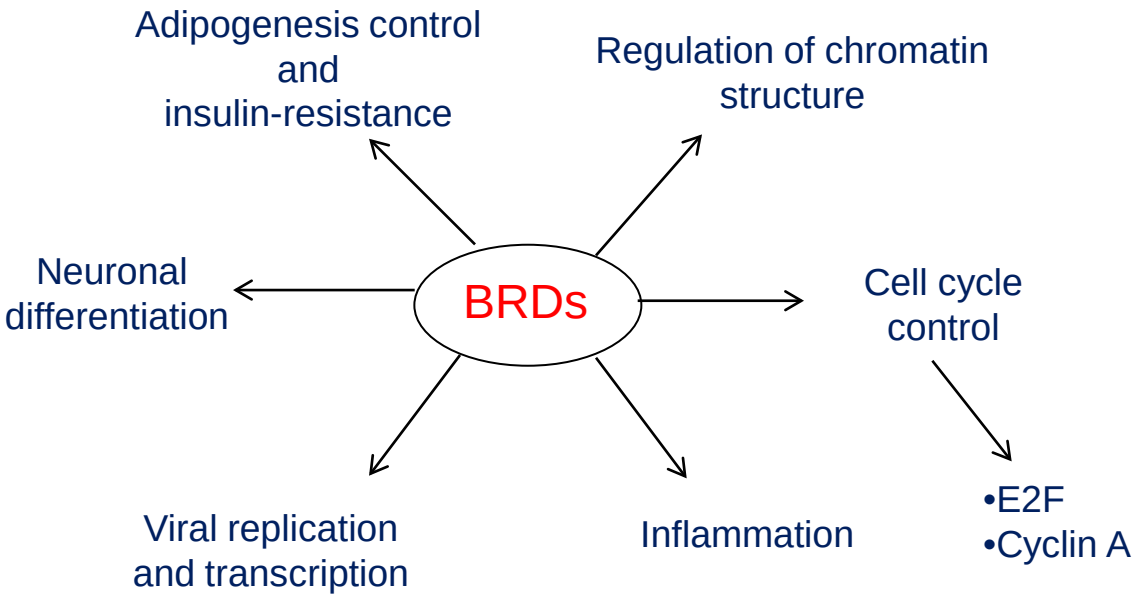


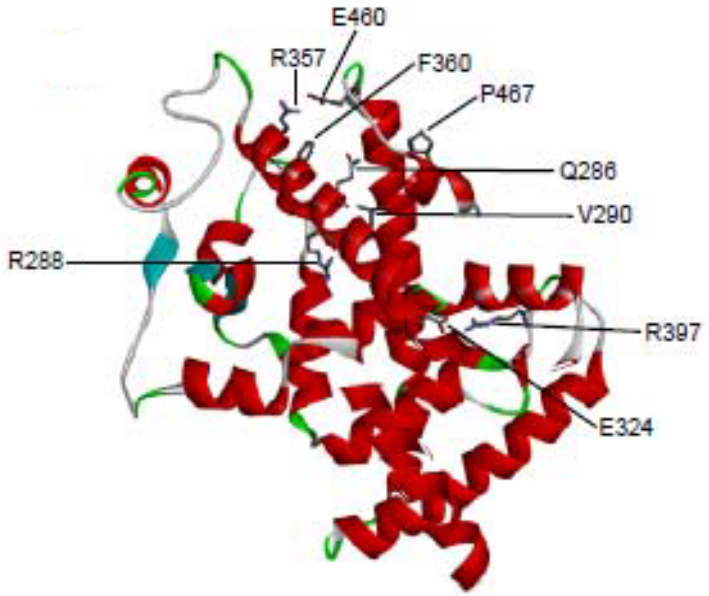



PGK1
PTPRTA
nsSNVs
in cancer tissues
FXN
PIM1
BRD
PPAR

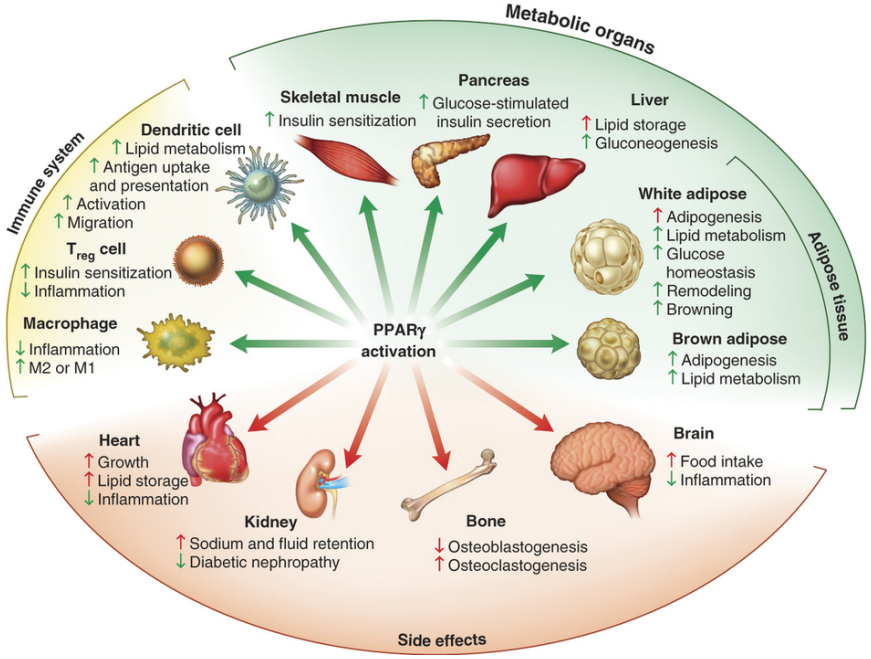


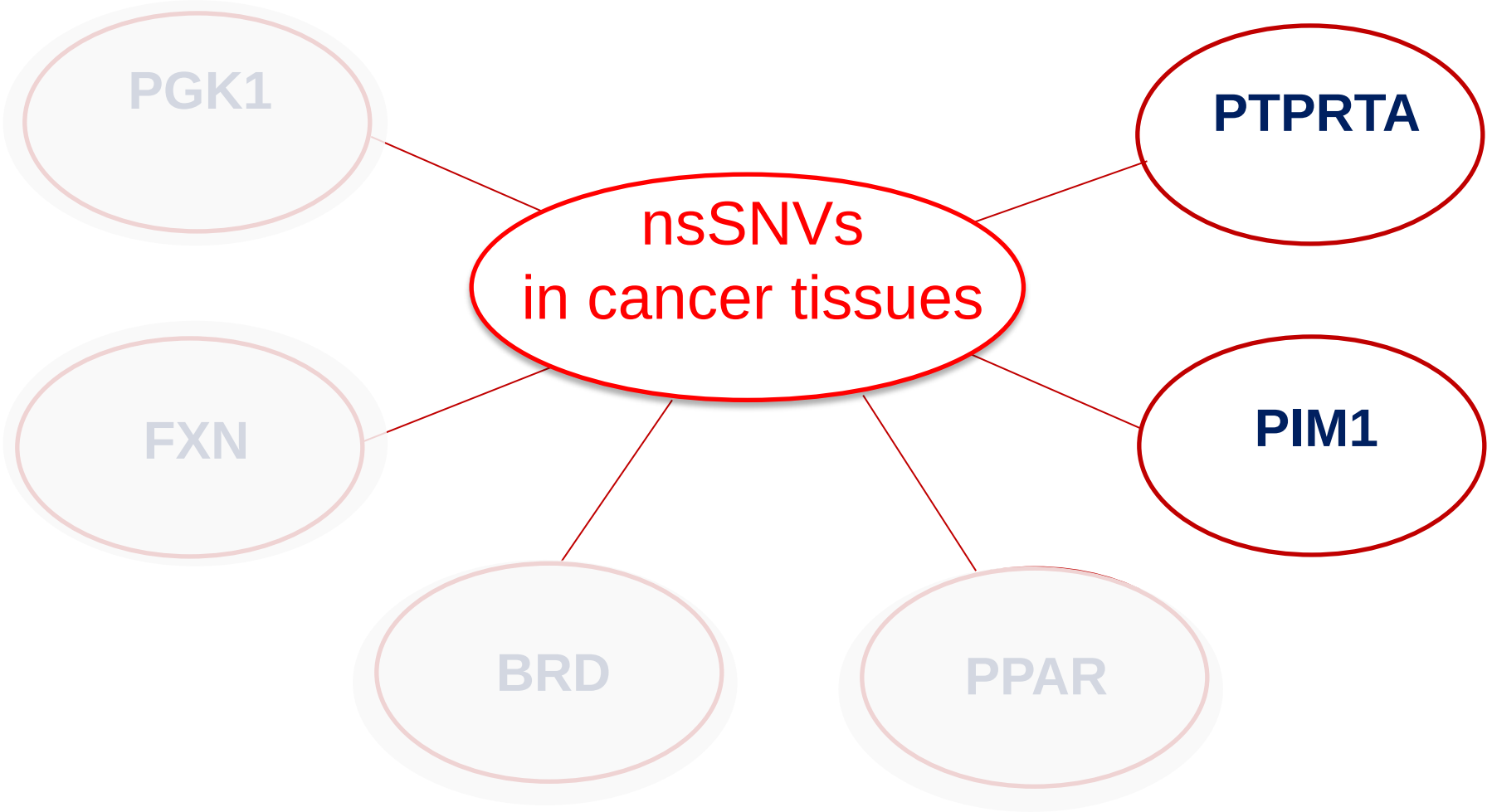
BRD: Bromodomains
PDB code: 3ONI
Lori L. et al, PlosOne, 2016

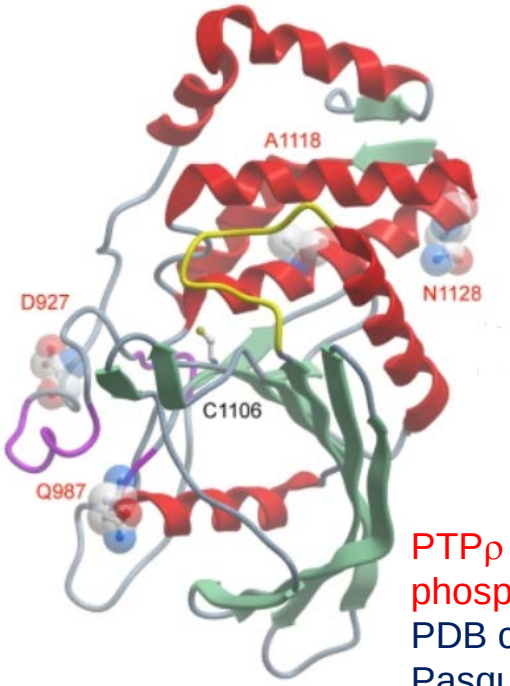




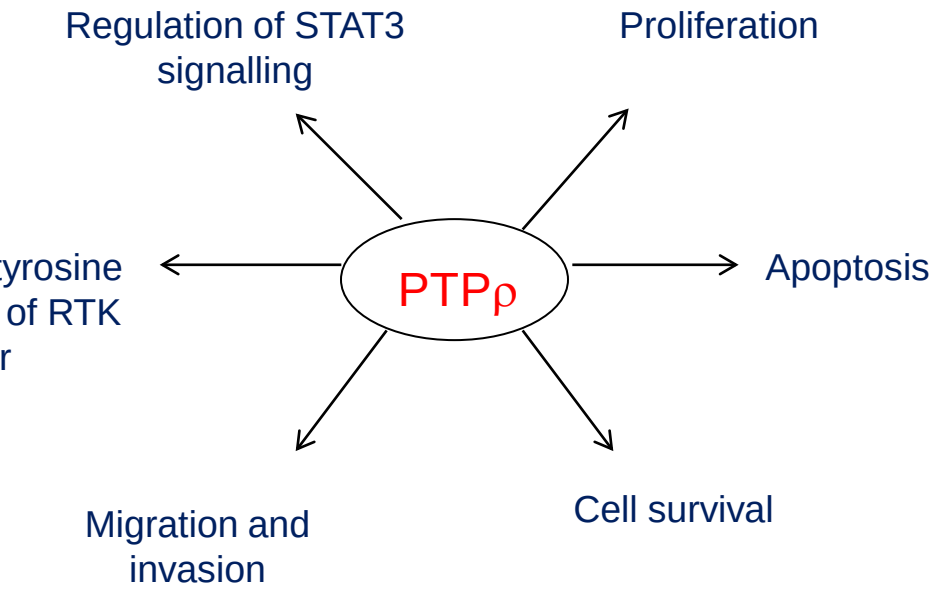
PPARγ Peroxisome Proliferator Receptor γ
PDB code: 1PRG
Petrosino M. et al, Int.Mol:Sci, 2017

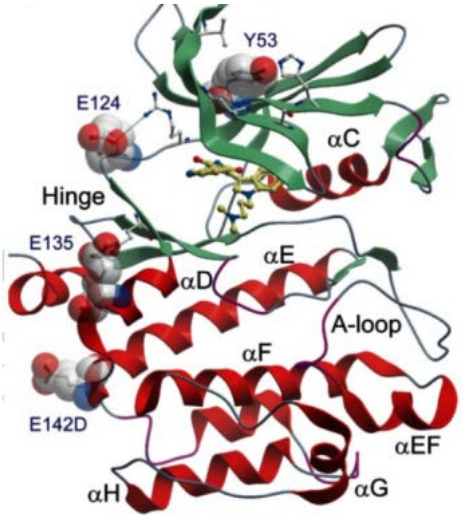



PGK1
FXN
BRD
PPAR
nsSNVs
in cancer tissues
PTPRTA
PIM1

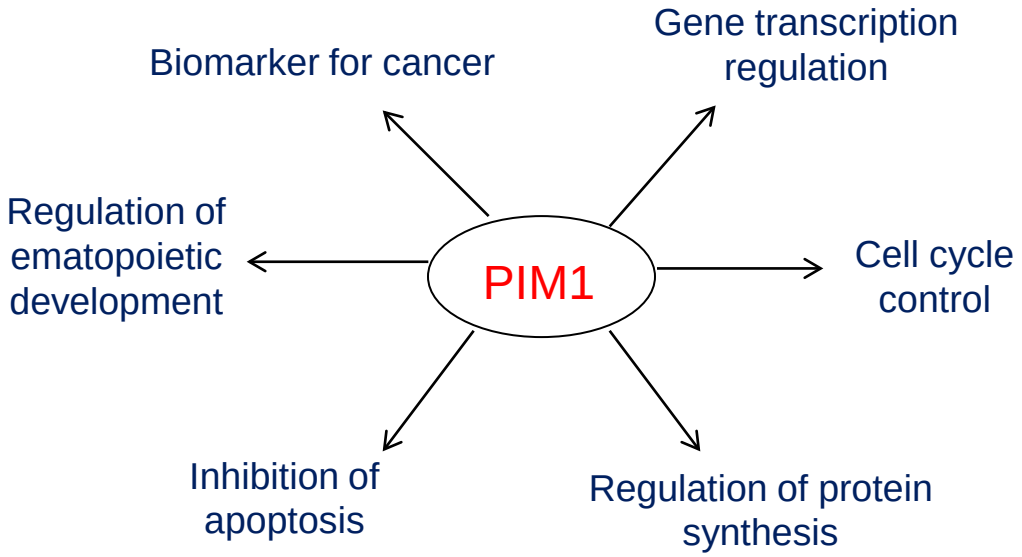


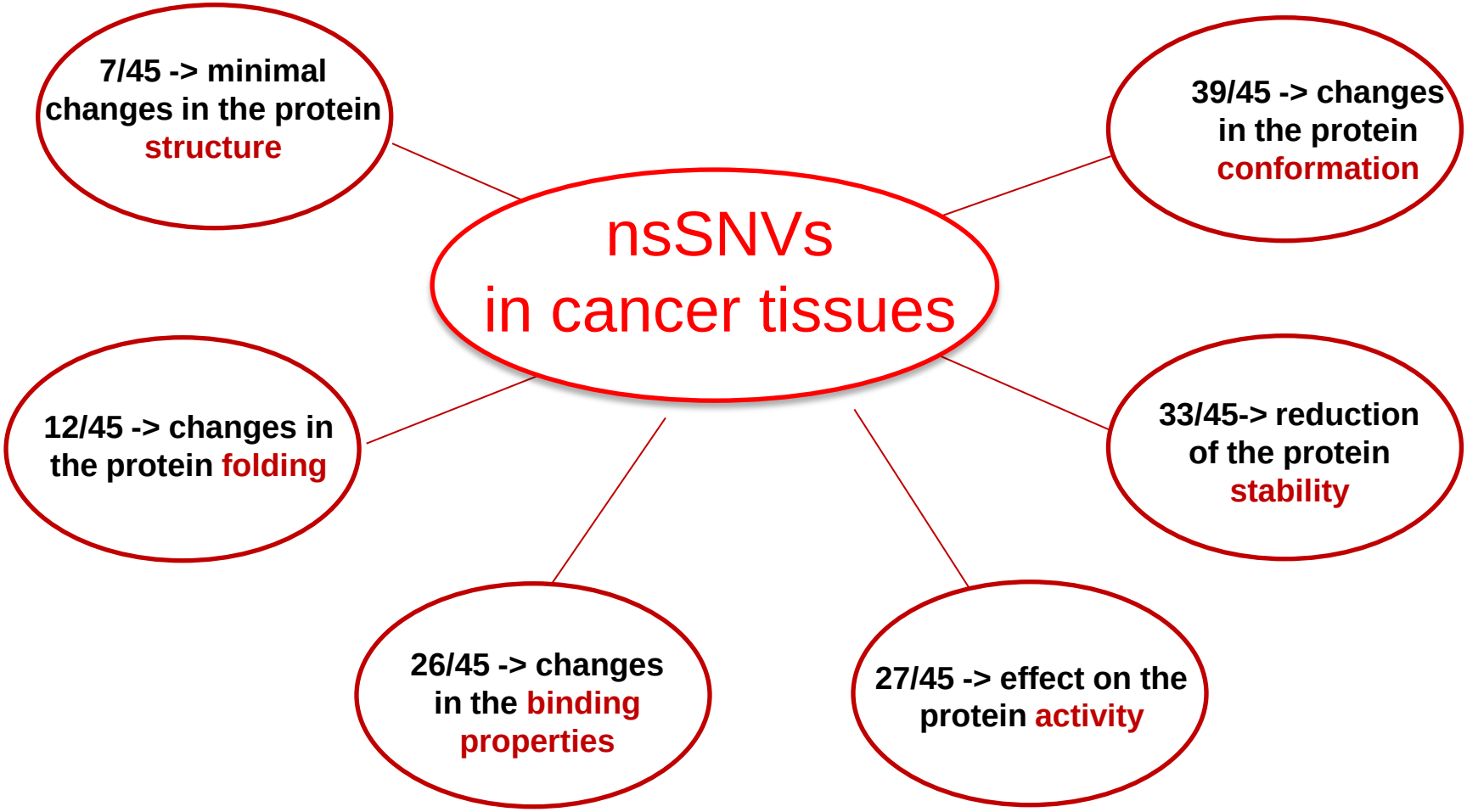
PTPρ Protein tyrosine phosphatase ρ
PDB code: 2OOQ
Pasquo A. et al, PlosOne,2012





Pim1 Kinase
PDB code: 1XWS
Lori C. et al, PlosOne, 2013




nsSNVs
in cancer tissues
7/45 -> minimal changes in the protein structure
39/45 -> changes in the protein conformation
12/45 -> changes in the protein folding
33/45-> reduction of the protein stability
26/45 -> changes in the binding properties
27/45 -> effect on the protein activity

Integrated tools to investigate the molecular basis of diseases: computational and experimental analysis of the impact of protein variants on protein stability and function.

The main aim of the project:
filling the gap between the collection of thermodynamic data and disease-related information on protein variants.

Sapienza Università di Roma

Computational

- collecting data from public resources on specific proteins and pathways (WP1).
- extracting experimental data from literature (WP1).
- development of a predictor of the impact of nsSNVs (WP2).

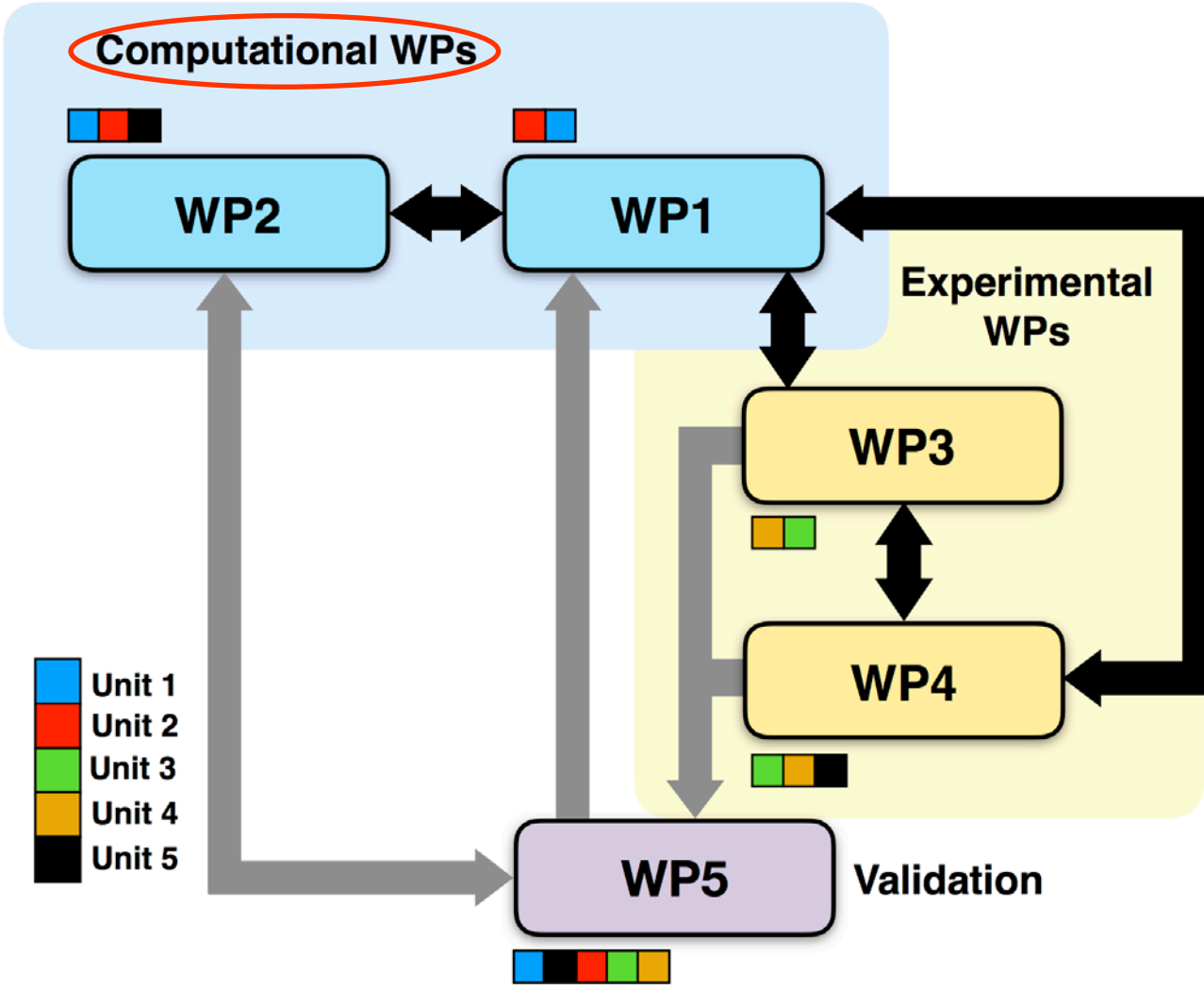
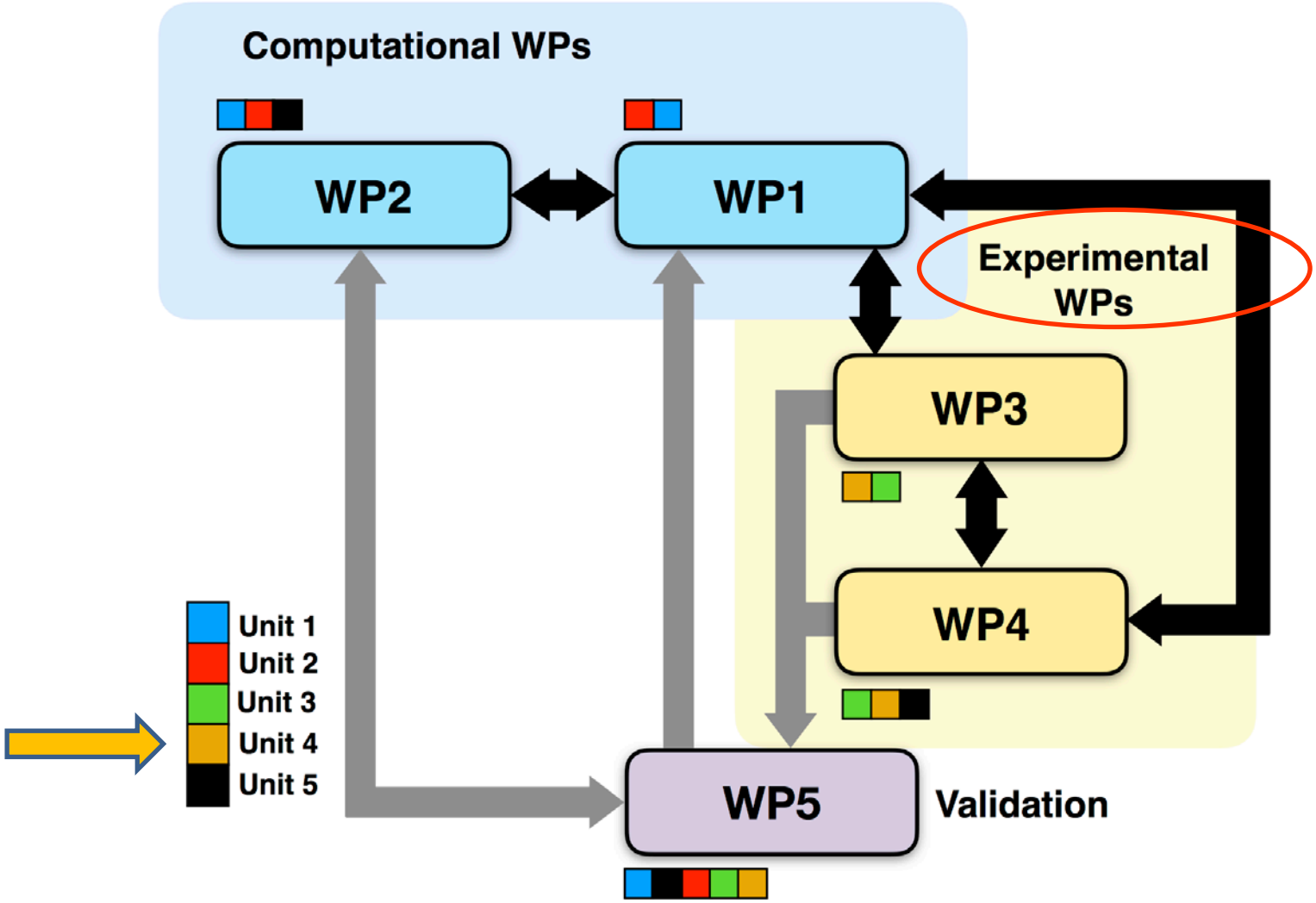


Fig. 3. Flowchart of the work packages. In black are indicated the connections taking place for 3 years. In gray are reported the interactions established during the final part of the project. Colored squares indicate the units that contribute to the realization of each work package. The first color represents the leader of each work package.



Experimental

Generation of new experimental data: structural, functional and stability

➢**WP3: Generation of new experimental data: structural, functional and stability (months 1-32)**
Leading unit: UNIT4
Participants: UNIT3 and UNIT5

➢Generate data for the development and testing of the new predictors.

➢Effect of somatic nsSNVs found in cancer tissues on the structural properties in solution of protein kinases MAPK and phosphatases PTPN and CaM variants

➢Determination of $\Delta\Delta G$ values (ΔGvariant- ΔGwt).

➢The results and data obtained in this WP will be inserted in the database WP1, and used for the development and testing of the predictors (WP2, WP5).

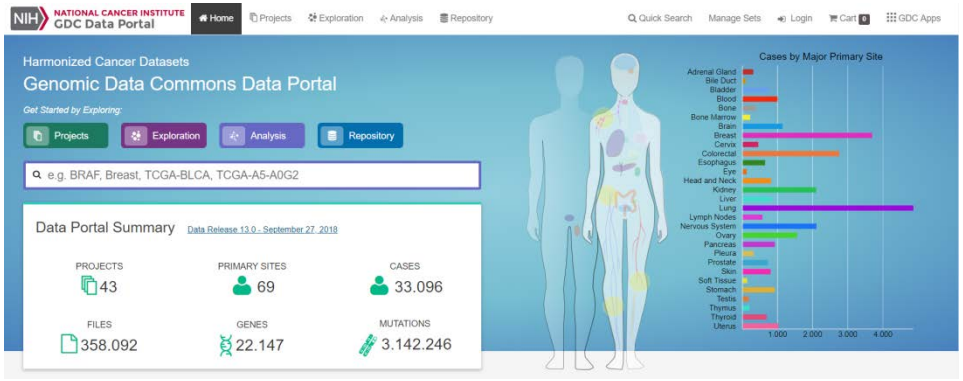
➢WP3: Generation of new experimental data: structural, functional and stability (months 1-32)
Leading unit: UNIT4
Participants: UNIT3 and UNIT5

Experimental characterization of two sets of protein variants

➢disease-related variants of the ubiquitous protein calmodulin (CaM)

➢protein kinases (MAPK1,3, 6, 8 and 11) and phosphatases (PTPN4, 11 and 14) detected in cancer cells

COSMIC
(Catalogue of Somatic Mutations in Cancer)



Genomic Data Commons Data Portal

Variazione dei livelli di espressione in vari tipi di cancro

(Davoli, T. et al., Cell 155, 948–962, 2013)

Frequentemente mutate in vari tipi di cancro

Potenziali target di farmaci antitumorali

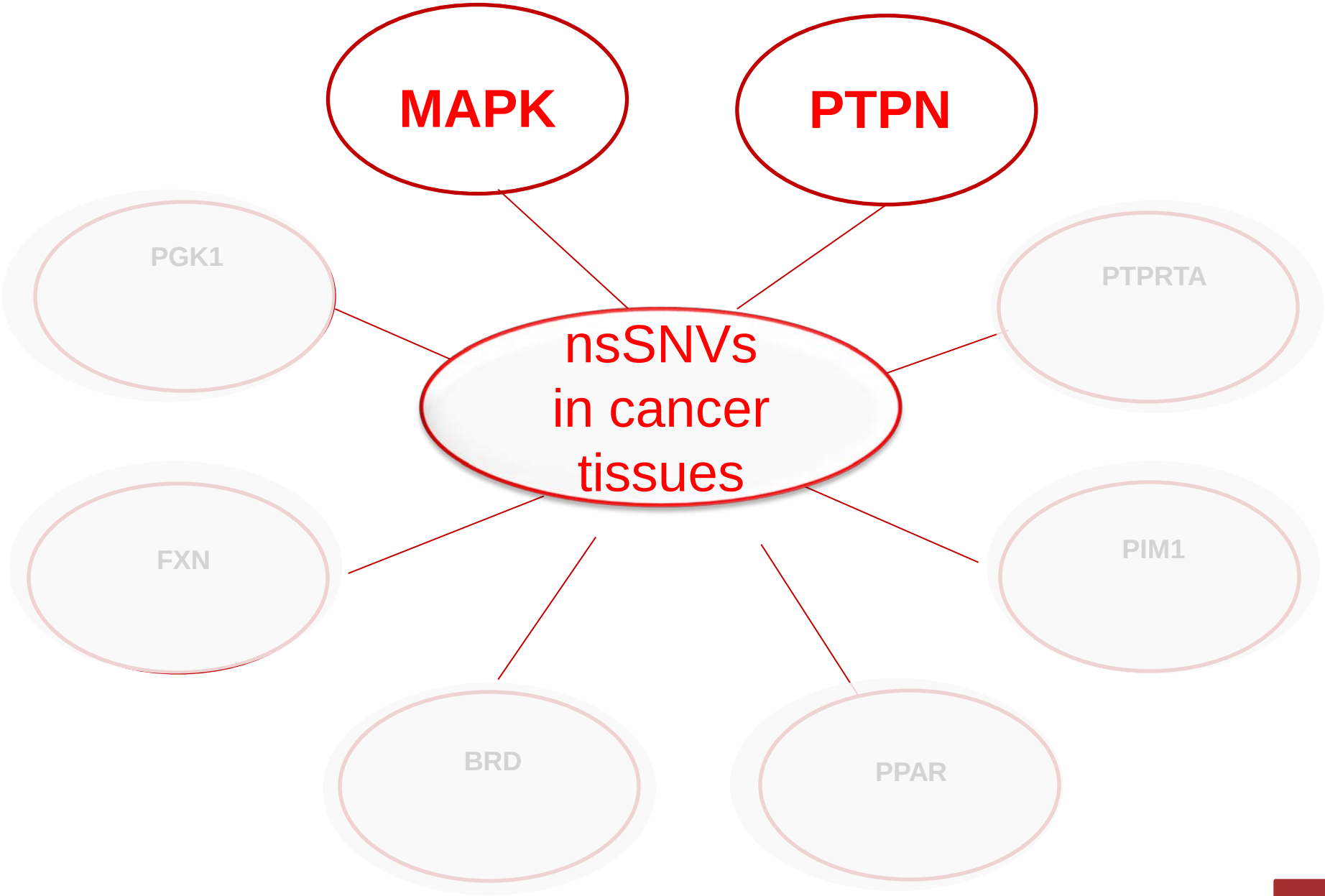
Tumor suppressor/oncogene ranking

Protein	TUSON ranking	
KINASE	TS	OG
MAPK1	10307	21
MAPK3	10314	1096
MAPK6	10316	3058
MAPK7	10317	6621
MAPK8	10318	163
MAPK11	10309	14169
PHOSPHATASE	TS	OG
PTPN3	1000	513
PTPN4	13602	3093
PTPN5	1420	14255
PTPN11	13595	52
PTPN14	1884	7199

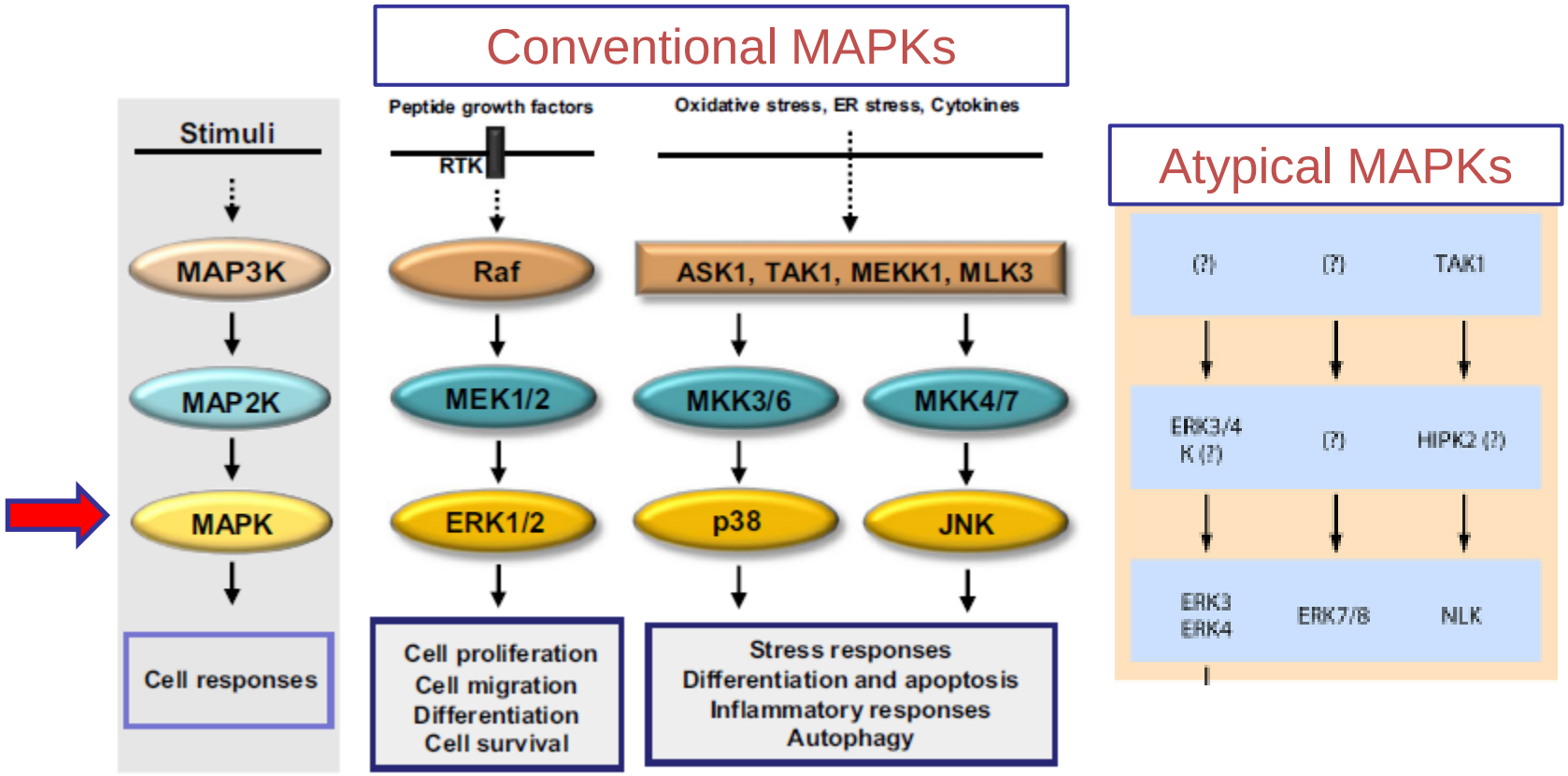
MAPK1,3,6,8,11
PTPN4,11,14

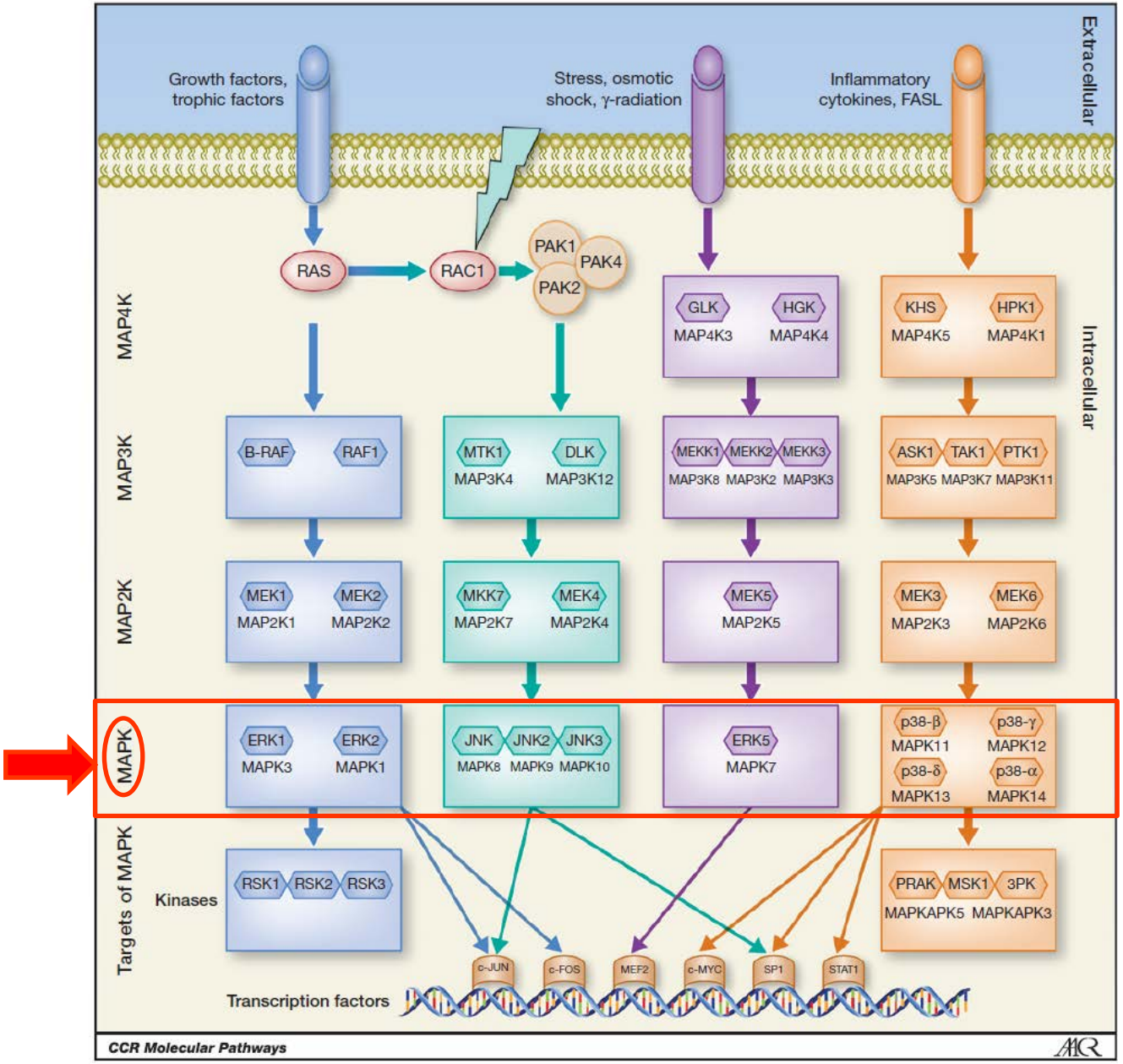
Struttura 3D disponibile

(Davoli, T. et al., Cell 155, 948–962, 2013)


MAPK
PTPN
PGK1
PTPRTA
nsSNVs
in cancer
tissues
FXN
PIM1
BRD
PPAR

MAPK- signaling cascade




Extracellular
Intracellular
Growth factors, trophic factors
Stress, osmotic shock, γ-radiation
Inflammatory cytokines, FASL
RAS
RAC1
PAK1
PAK2
PAK4
MAP4K
GLK
MAP4K3
HGK
MAP4K4
KHS
MAP4K5
HPK1
MAP4K1
MAP3K
B-RAF
RAF1
MTK1
MAP3K4
DLK
MAP3K12
MEKK1
MAP3K8
MEKK2
MAP3K2
MEKK3
MAP3K3
ASK1
MAP3K5
TAK1
MAP3K7
PTK1
MAP3K11
MAP2K
MEK1
MAP2K1
MEK2
MAP2K2
MKK7
MAP2K7
MEK4
MAP2K4
MEK5
MAP2K5
MEK3
MAP2K3
MEK6
MAP2K6
MAPK
ERK1
MAPK3
ERK2
MAPK1
JNK
MAPK8
JNK2
MAPK9
JNK3
MAPK10
ERK5
MAPK7
p38-β
MAPK11
p38-γ
MAPK12
p38-δ
MAPK13
p38-α
MAPK14
Targets of MAPK
Kinases
RSK1
RSK2
RSK3
PRAK
MAPKAPK5
MSK1
3PK
MAPKAPK3
c-JUN
c-FOS
MEF2
c-MYC
SP1
STAT1
Transcription factors
CCR Molecular Pathways
MCR

MAPK

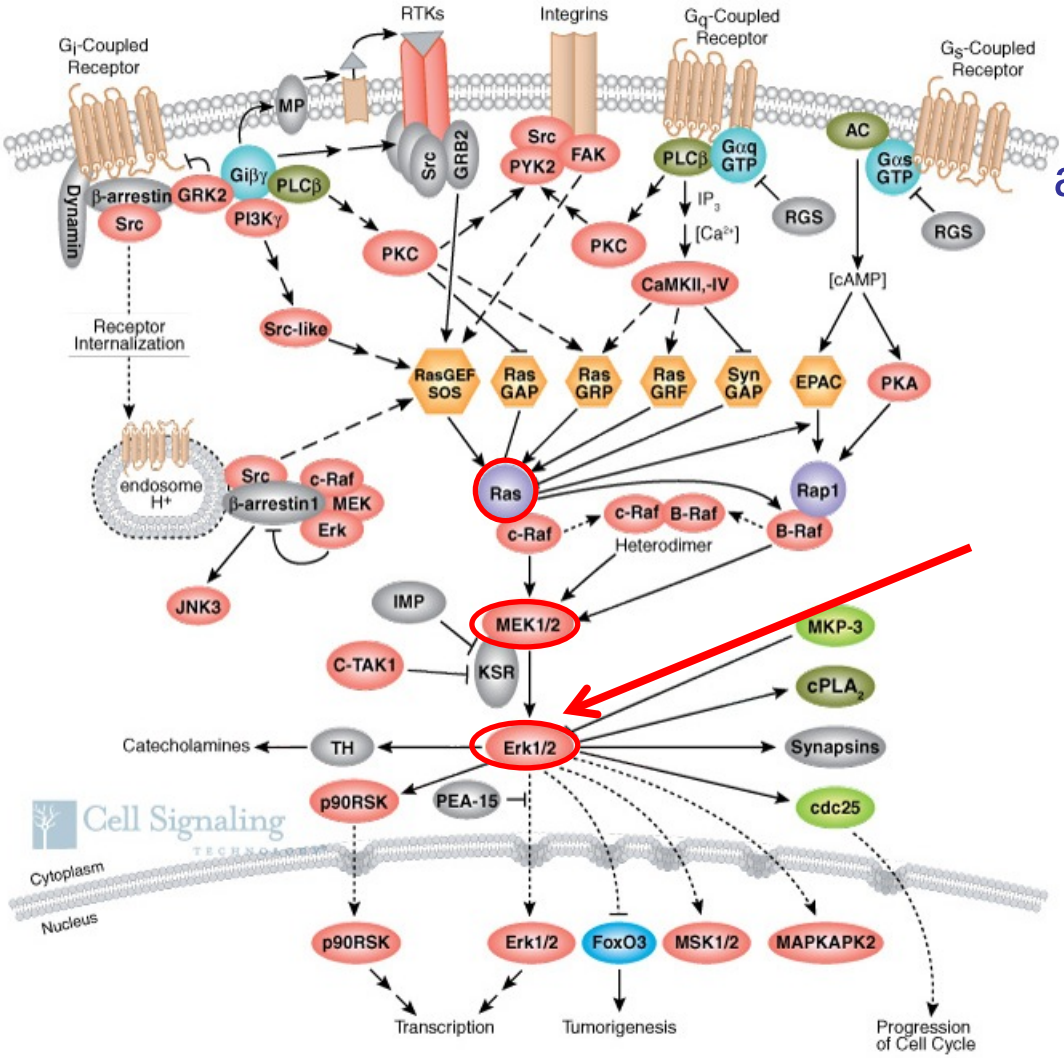
Ser/Thr chinasi

- ERK1 (MAPK3)
- ERK2 (MAPK1)
- ERK5 (MAPK7)
- p38α (MAPK8)
- p38β (MAPK11)
- p38γ (MAPK12)
- p38δ (MAPK13)
- JNK1 (MAPK8)
- JNK2 (MAPK9)
- JNK3 (MAPK10)

- ERK3 (MAPK6)
- ERK4 (MAPK4)
- ERK7/8 (MAPK15)

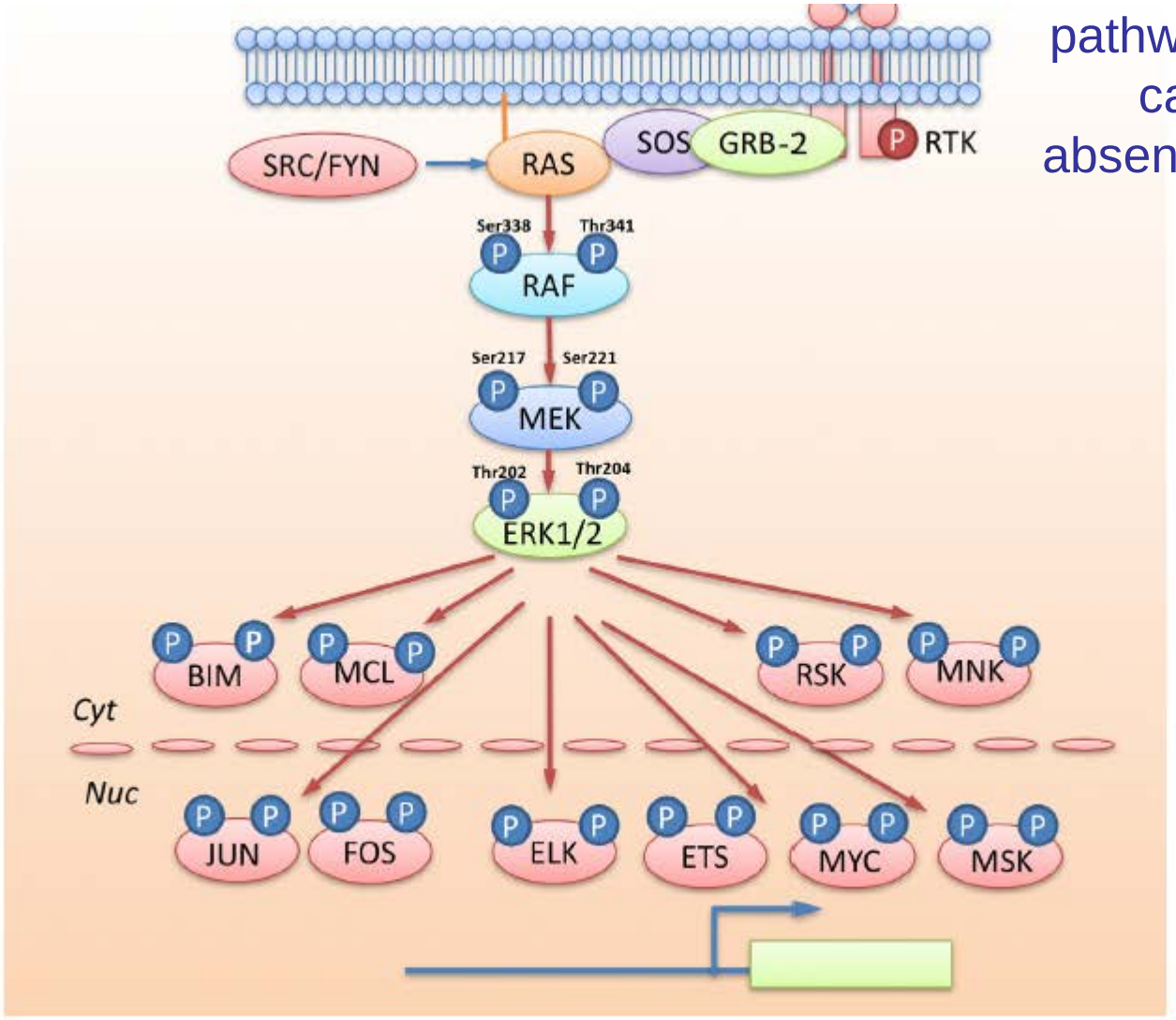
MAPK


Conventional MAPKs
ERK1/2
TEY
Kinase domain
p38 α/β (γ/δ)
TGY
Kinase domain
JNK1/2/3
TPY
Kinase domain
ERK5
TEY
Kinase domain
NLS
TAD
Atypical MAPKs
ERK3
SEG
Kinase domain
C34
ERK4
SEG
Kinase domain
C34
ERK7/8
TEY
Kinase domain
NLS
NLK
TQE
AHQr
Kinase domain



The Ras-Raf-MEK-ERK pathway is upregulated in many cancer types also in the absence of oncogenic mutations

The Ras-Raf-MEK-ERK pathway is upregulated in many cancer types also in the absence of oncogenic mutations

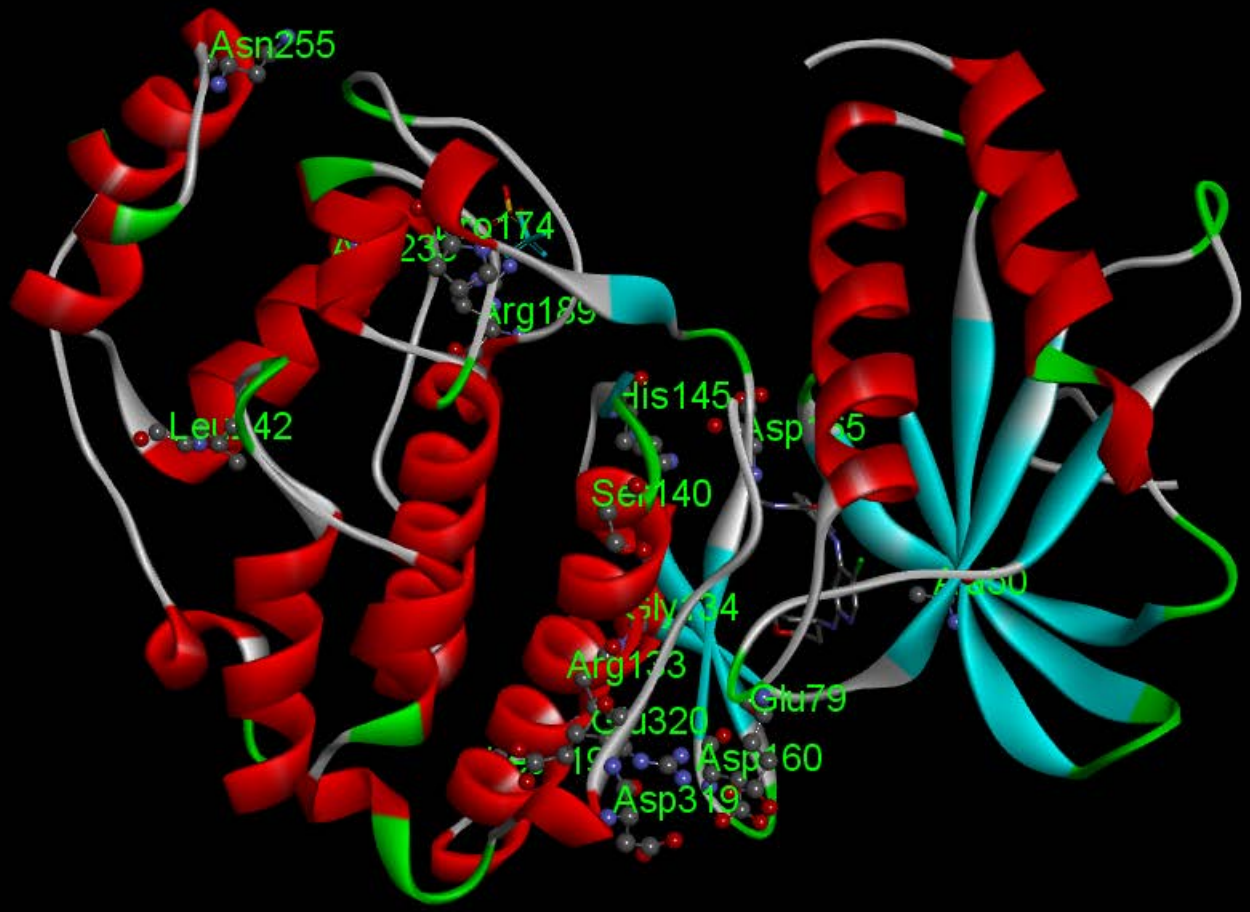


ERK1/2
Cytosolic and nuclear targets

Liu et al., 2017. Targeting ERK, an Achilles' Heel of the MAPK pathway, in cancer therapy

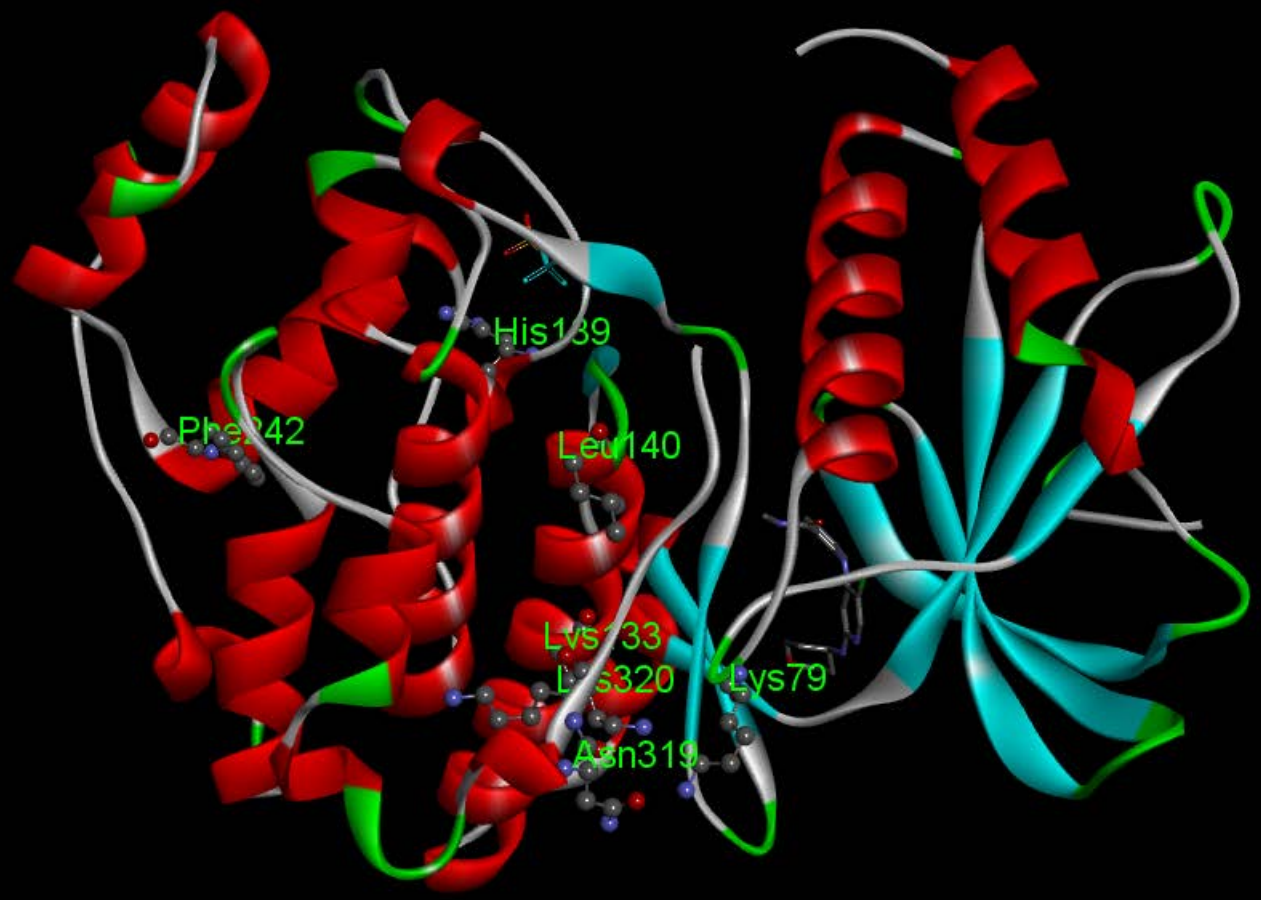
MAPK1 (ERK2)

PDB 4ZZN



MAPK1 (ERK2)

PDB 4ZZN



UNIT4

- Selection of protein variants from database
- Site-directed mutagenesis (PCR)
- Sequencing of the DNA containing the desired mutation
- Expression of recombinant protein variants in cell host (E.coli, eukarotic cells for post-translation modifications)
- Purification of protein variants
- Molecular mass and sequence of the purified protein (Mass spectrometry)

UNIT4

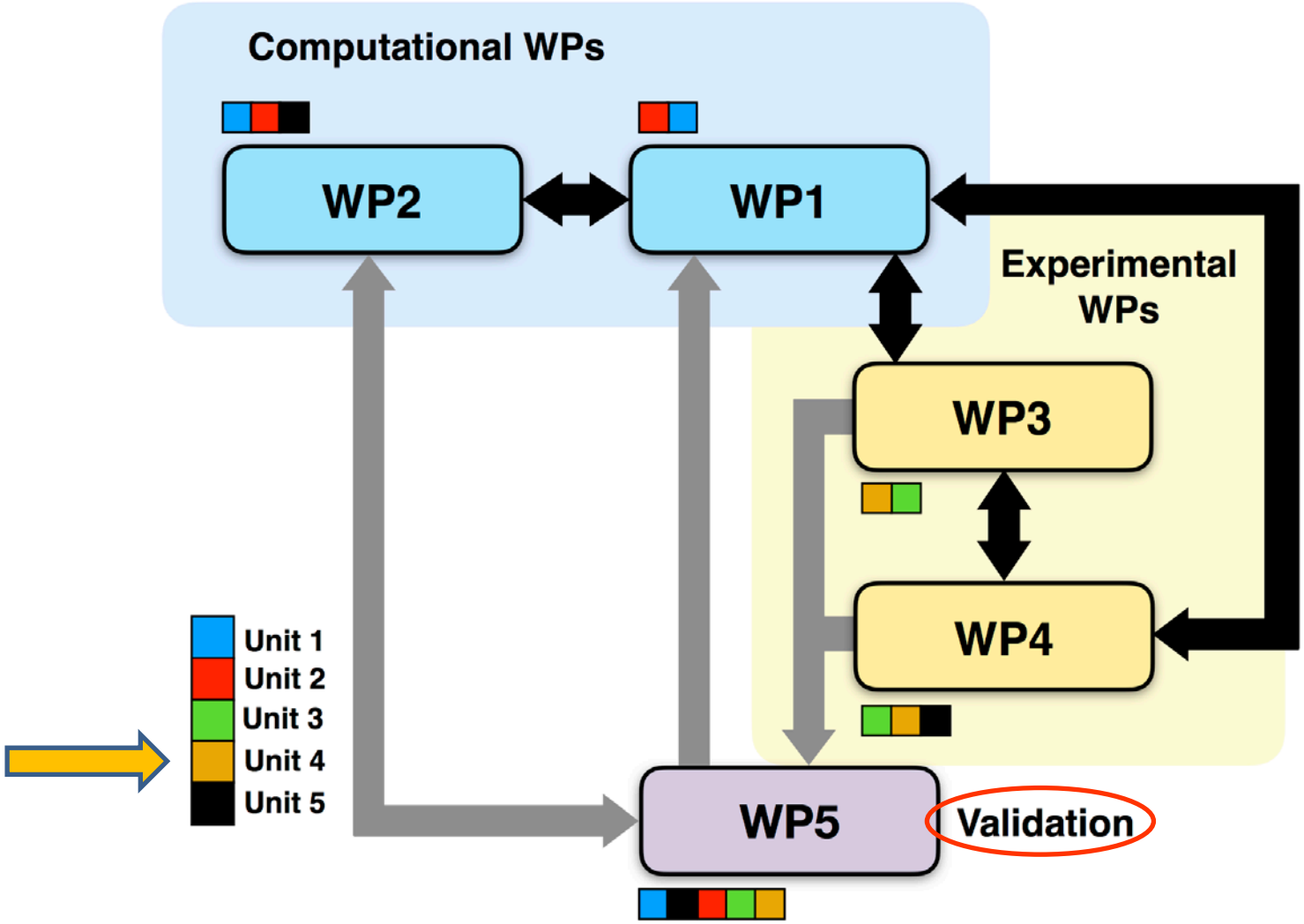
- Structural characterization of protein variants

- Functional activity in the presence of the appropriate substrates and/or ligands

- Determination of ΔG and $\Delta\Delta G$ (ΔGvariant- ΔGwt) associated to each variant.

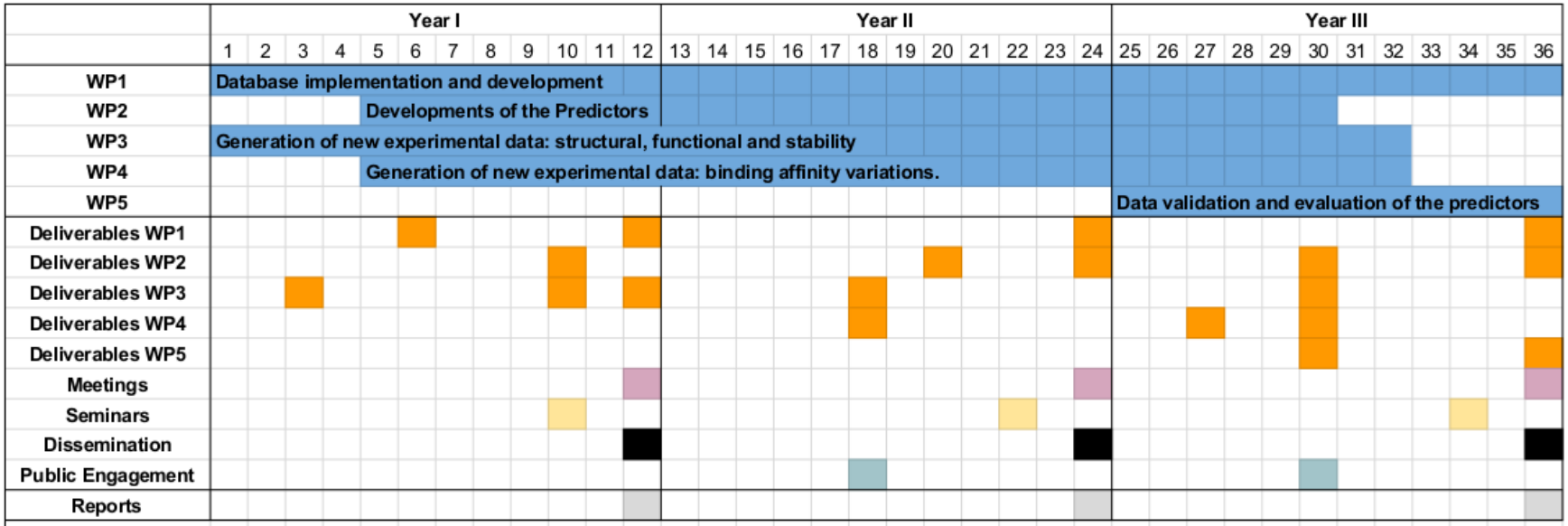
- Binding affinity with interacting partners and inhibitors (WP4)

UNIT4 will provide the selected variants to UNIT3 and to UNIT5 that will perform structural analysis by NMR and molecular dynamics



Validation

The validation of the developed predictor will be performed using new experimental data generated by UNITS 3, 4 and 5.



Dissemination	Month 12: Submission of a Poster to the ISMB Conference
	Month 24: Publication of a munuscript presenting the methodology for extracting data from literature
	Month 36: Punlication of a paper presenting the new database of protein variant data
Public Engagement	
Reports	Month 12: Report of the first year of activity
	Month 24: Report of the second year of activity
	Month 36: Final report of the project

Expected deliverables
Months 1-32

- **Month 3**: selection of protein variants from database
- **Month 10-12:** site-directed mutagenesis to produce the selected nsSNVs proteins; expression and purification of wild type and variant proteins
- **Month 30:** structural characterization of the purified selected variants in solution (CD, fluorescence, FTIR, NMR spectroscopy, XAS).
- **Month 36:** thermal (T_m value) and thermodynamic stability (ΔG and m) of nsSNVs and wild type proteins. Calculation of the difference in unfolding free energy ($\Delta\Delta G$) between the wild type and the selected nsSNVs (month 36)

WP4: Generation of new experimental data: binding affinity variations
Leading UNIT: UNIT3.
Participants: UNIT4, UNIT5

- Estimate the relative binding affinity $\Delta\Delta G$(ΔGvariant- ΔGwt) and provide useful experimental data for WPs 1, 2 and 5

- UNIT4 will provide the cancer-related, newly generated mutants to UNIT3 to study the impact of the single amino acid substitutions on the protein binding activity by surface plasmon resonance (SPR).

- UNIT4 will provide the results obtained from structural and functional studies on nsSNVs to UNIT5 that will study the impact of the single amino acid substitution on protein dynamics and interactions by non-equilibrium molecular dynamics.