Discrete Correlate Summation Clustering of a Dopamine-Angiotensin Network

Brian M. Westwood[1], Liliya M. Yamaleyeva[1] and Rong Chen[2]

[1]Department of Surgery/Hypertension and Vascular Research and [2]Department of Translational Neuroscience
Wake Forest University School of Medicine, Winston-Salem, North Carolina, United States

KEYWORDS - discrete correlate summation clustering, angiotensin, dopamine, framework, contingency test

ABSTRACT: We present a model of directed investigation of KEGG pathway analysis, beginning with discrete correlate summation (DCS) clustering of Swiss-Prot sets. Dopamine and angiotensin related proteins define the boundary for pathway analyses upon a network of proteins also associated with a compositely defined matrix of Swiss-Prot keywords to provide a framework for their biological indices (Figure 1).

Figure 1. Graphical abstract.

INTRODUCTION

This study is a framework. It is designed to visualize a system between two disparate topics, connected by common relationships.

Generally, we categorize genes with differential expression between treatments, to investigate enrichment of an effected gene set, and decide if there is evidence to indicate a concerted alteration of a pathway and not just stochastic changes.

This framework is constructed using the classification structure of the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa et al., 2000) molecular network database, evaluated with the ConsensusPathDB-human (CPDB) (Kamburov et al., 2011) over-representation analysis, using very specific starting sets (Swiss-Prot keyword searches) (UniProt Consortium, 2021).

METHODS

This is an investigation of 3309 proteins from the 20 Swiss-Prot keyword searches [oxidative&stress, glutamate, inflammation, memory, glutamatergic, blood&pressure, dopamine, hypertension, blood&brain&barrier, gaba, angiotensin, prefrontal&cortex, dopaminergic, gabaergic, cognition, midbrain, cerebral&vascular, striatum, renin, ageing].

Iterative 'original matrix total array average was calculated right to left' (RL-OriMaTArA) DCS clustering (Bronson et al., 2022) yields 3 clear count domains and 2 singleton searches for the 3309 proteins (Figure 2a, 2b). Briefly, upper cluster of a matrix was determined by sorting the totals of each variable column, then repeating on remaining extra-cluster submatrix.

RL-OriMaTArA produced 3 primary clusters [C, M, Y] and 4 combination sets [CM-blue, MY-red, CY-green, CMY-gray] to search in CPDB for pathway enrichment analysis (Figure 2c, 2d) of KEGG pathways.

The 7 CPDB searches of 3309 proteins yielded a primary output of 137 KEGG pathways with 1991 genes. Area under the curve (AUC) of receiver operating characteristic (ROC) curves and precision-recall F-measures (Fawcett, 2006) were calculated for neighborhoods 1-5 for 51 KEGG pathways and 1045 genes with 4 or more angiotensin or renin Swiss-Prot search proteins/genes and with 4 or more dopamine or dopaminergic Swiss-Prot search proteins/genes in each pathway (Table 1).

For each neighborhood, one through five, surrounding angiotensin/renin (yellow, RAS) nodes, pathways through edges were examined for dopamine/dopaminergic (magenta, DA) node proximity. If the path contained only DA associated (cyan, DA_{assoc}) and/or RAS associated (gray, RAS_{assoc}) nodes, the yellow node was marked with a blue cross (true positive, TP) for 'in' or a cyan cross (false negative, FN) for 'out' of the neighborhood being examined. If the contiguous path contained nodes other than RAS_{assoc} or DA_{assoc} (pale green), the yellow node was marked with a black cross (false positive, FP) for 'in' or a gray cross (true negative, TN) for 'out' of the neighborhood being examined (Figure 3).

All calculations were performed in Excel (Microsoft Corp., WA), ImageJ (Rasband, 1997-2018) was used for tabulation of supplemental gif images and Cytoscape (Shannon et al., 2003) was used for plotting the network in figure 4.

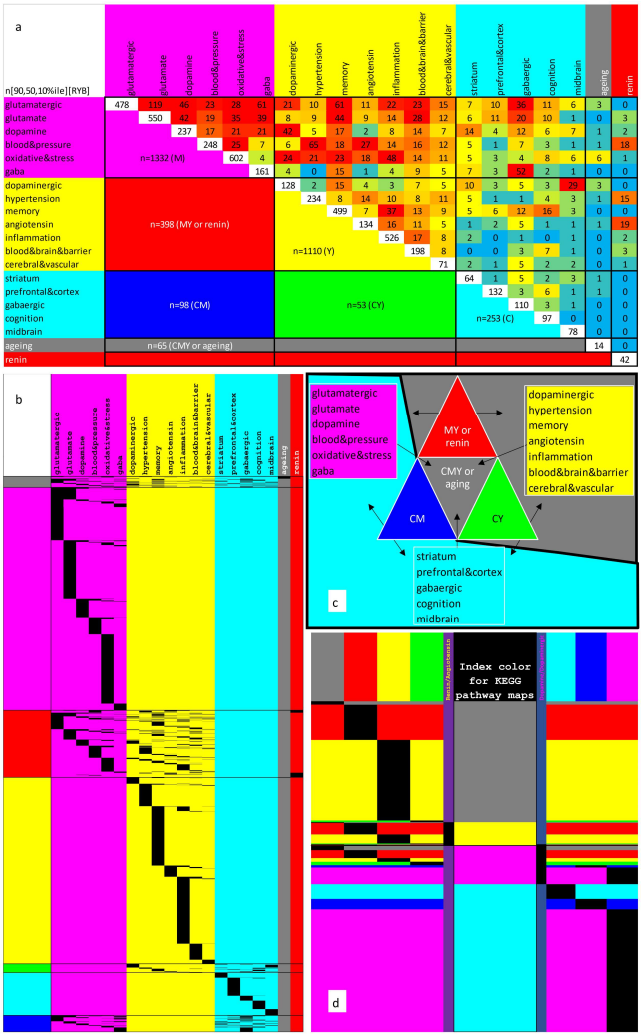


Figure 2. Binary count clustering for Swiss-Prot keyword searches (a); keyword group membership columns by cluster row groups [(b) rows are individual proteins, black cell = membership]; ternary diagram of cluster layout [(c) with underlay of node coloring for associated groups]; groups of clusters organized by index color for KEGG pathway map contingency tests [(d) RAS associated genes (gray) RAS genes (yellow) DA genes (magenta) DA associated genes (cyan)].

RESULTS

The 3309 proteins from the 20 Swiss-Prot keyword searches are in Supplementary Table 1. For each entry, membership in a keyword search is tabulated as a '1'. This data was used to calculate the binary count clustering for Swiss-Prot keyword searches and visualize the membership maps in Figure 2.

The contingency testing shown in the 51 supplemental gif images as referenced in Table 1 [1st, 2nd, 3rd, 4th, 5th neighborhoods, unmasked KEGG output] is a graphical representation of the exact F-AUC ROC analysis procedures.

The 51 KEGG pathways in Table 1 are color coded for all F-AUC-3 [specifically $(\text{F-AUC})_{0.7}$-3, maximum AUC or F-measure at least 0.7 and 3 TP nodes (Westwood et al., 2022)] blue for F-AUC of 1, orange for F-AUC at least 0.85 and gray for F-AUC at least 0.7. Of the 25 $(\text{F-AUC})_{0.85}$-5 pathways (bold), there were 160 proteins in at least 4 KEGG pathways (KP). There were 23 RAS (7.7±0.72 KP, mean+sem), 51 $\text{RAS}_{\text{assoc}}$ (7.5±0.59 KP), 22 DA (7.6±0.98 KP), 64 DA_{assoc} (7.3±0.44 KP) proteins (Chi-square, p=0.44, ns; Figure 4).

Table 1. KEGG pathway contingency test output, sorted by TP then maximum for AUC of ROC curve and F-measure.

KEGG	Pathway	AUC	F	TP
hsa04270	**Vascular smooth muscle contraction**		**0.96**	**12**
hsa04925	**Aldosterone synthesis and secretion**		**1.00**	**10**
hsa04022	**cGMP-PKG signaling pathway**	**0.80**	**0.90**	**9**
hsa04071	**Sphingolipid signaling pathway**		**1.00**	**8**
hsa04810	**Regulation of actin cytoskeleton**		**1.00**	**8**
hsa04927	**Cortisol synthesis and secretion**		**1.00**	**8**
hsa04750	**Inflammatory mediator regulation of TRP channels**		**0.94**	**8**
hsa04611	**Platelet activation**		**1.00**	**7**
hsa04024	**cAMP signaling pathway**	**0.96**	**0.93**	**7**
hsa04510	**Focal adhesion**		**0.93**	**7**
hsa04921	**Oxytocin signaling pathway**	**0.89**	**0.88**	**7**
hsa04062	**Chemokine signaling pathway**	**0.75**	**0.88**	**7**
hsa04020	**Calcium signaling pathway**	**1.00**	**1.00**	**6**
hsa04066	**HIF-1 signaling pathway**		**1.00**	**6**
hsa04151	**PI3K-Akt signaling pathway**		**1.00**	**6**
hsa04360	**Axon guidance**	**0.99**	**0.92**	**6**
hsa04072	**Phospholipase D signaling pathway**	**0.97**	**0.92**	**6**
hsa04015	**Rap1 signaling pathway**	**0.75**	**0.92**	**6**
hsa04924	**Renin secretion**		**0.92**	**6**
hsa04926	**Relaxin signaling pathway**	**0.86**	**0.92**	**6**
hsa04261	**Adrenergic signaling in cardiomyocytes**		**1.00**	**5**
hsa04728	**Dopaminergic synapse**		**1.00**	**5**
hsa04911	**Insulin secretion**		**1.00**	**5**
hsa04913	**Ovarian steroidogenesis**		**1.00**	**5**
hsa04010	**MAPK signaling pathway**	**0.87**	**0.83**	**5**
hsa04670	Leukocyte transendothelial migration	0.40	0.83	5
hsa04720	Long-term potentiation		1.00	4
hsa04725	Cholinergic synapse		1.00	4
hsa04310	Wnt signaling pathway	0.81	0.89	4
hsa04726	Serotonergic synapse	0.75	0.89	4
hsa04928	Parathyroid hormone synthesis_secretion and action	0.81	0.89	4
hsa04912	GnRH signaling pathway	0.00	0.80	4
hsa04722	Neurotrophin signaling pathway	0.50	0.73	4
hsa04144	Endocytosis		1.00	3
hsa04371	Apelin signaling pathway	1.00	1.00	3
hsa04723	Retrograde endocannabinoid signaling		1.00	3
hsa04727	GABAergic synapse		1.00	3
hsa04014	Ras signaling pathway	0.83	0.86	3
hsa04919	Thyroid hormone signaling pathway	0.75	0.86	3
hsa04613	Neutrophil extracellular trap formation	0.44	0.75	3
hsa04935	Growth hormone synthesis_secretion and action		0.75	3
hsa04724	Glutamatergic synapse		1.00	2
hsa04350	TGF-beta signaling pathway	0.88	0.80	2
hsa04150	mTOR signaling pathway		0.80	2
hsa04370	VEGF signaling pathway	0.75	0.80	2
hsa04657	IL-17 signaling pathway	0.80	0.80	2
hsa04916	Melanogenesis		0.80	2
hsa04929	GnRH secretion	0.55	0.57	2
hsa04080	Neuroactive ligand-receptor interaction		1.00	1
hsa04668	TNF signaling pathway	1.00	1.00	1
hsa04918	Thyroid hormone synthesis		1.00	1

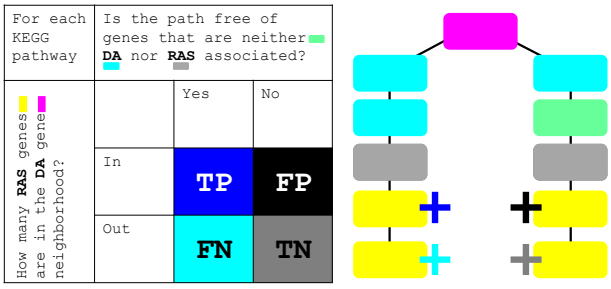


Figure 3. Contingency test and fourth neighborhood example.

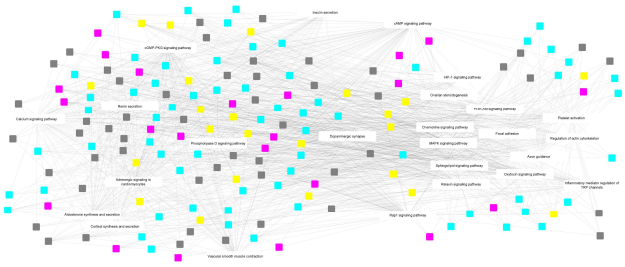


Figure 4. Naive unstructured network of 25 $(\text{F-AUC})_{0.85}$-5 pathways and 160 four KP proteins [RAS|$\text{RAS}_{\text{assoc}}$|DA|$\text{DA}_{\text{assoc}}$].

DISCUSSION

This framework ties the RAS and DA systems using targeted Swiss-Prot curated database search output membership in KEGG networks by implementing CPDB, resulting in an enriched dataset with a balanced representation from RAS, $\text{RAS}_{\text{assoc}}$, DA and DA_{assoc} proteins for further analysis. The utility if this method is in the simplicity of the contingency testing on display in the supplemental gif images. The ontological microenvironment defined by this framework array is broadly applicable to describing biological systems. This Swiss-Prot/KEGG/CPDB conduit can provide foundational information in the investigation of novel networks.

ACKNOWLEDGMENT

Supported by WFSM CVSC Pilot Award, CTSI NCATS UL1TR001420 and Odell Farley Foundation.

ABBREVIATIONS

DCS, discrete correlate summation; KEGG, Kyoto Encyclopedia of Genes and Genomes; CPDB, ConsensusPathDB-human; RL-OriMaTArA, original matrix total array average was calculated right to left; AUC, Area under the curve; ROC, receiver operating characteristic; RAS, angiotensin/renin; DA, dopamine/dopaminergic; DA_{assoc}, DA associated; $\text{RAS}_{\text{assoc}}$, RAS associated; TP, true positive; FN, false negative; FP, false positive; TN, true negative.

REFERENCES

Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000 Jan 1;28(1):27-30. doi: 10.1093/nar/28.1.27. PMID: 10592173; PMCID: PMC102409.

Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. ConsensusPathDB: toward a more complete picture of cell biology. Nucleic Acids Res. 2011 Jan;39(Database issue):D712-7. doi: 10.1093/nar/gkq1156. Epub 2010 Nov 11. PMID: 21071422; PMCID: PMC3013724.

The UniProt Consortium , UniProt: the universal protein knowledgebase in 2021, Nucleic Acids Research, Volume 49, Issue D1, 8 January 2021, Pages D480–D489, https://doi.org/10.1093/nar/gkaa1100

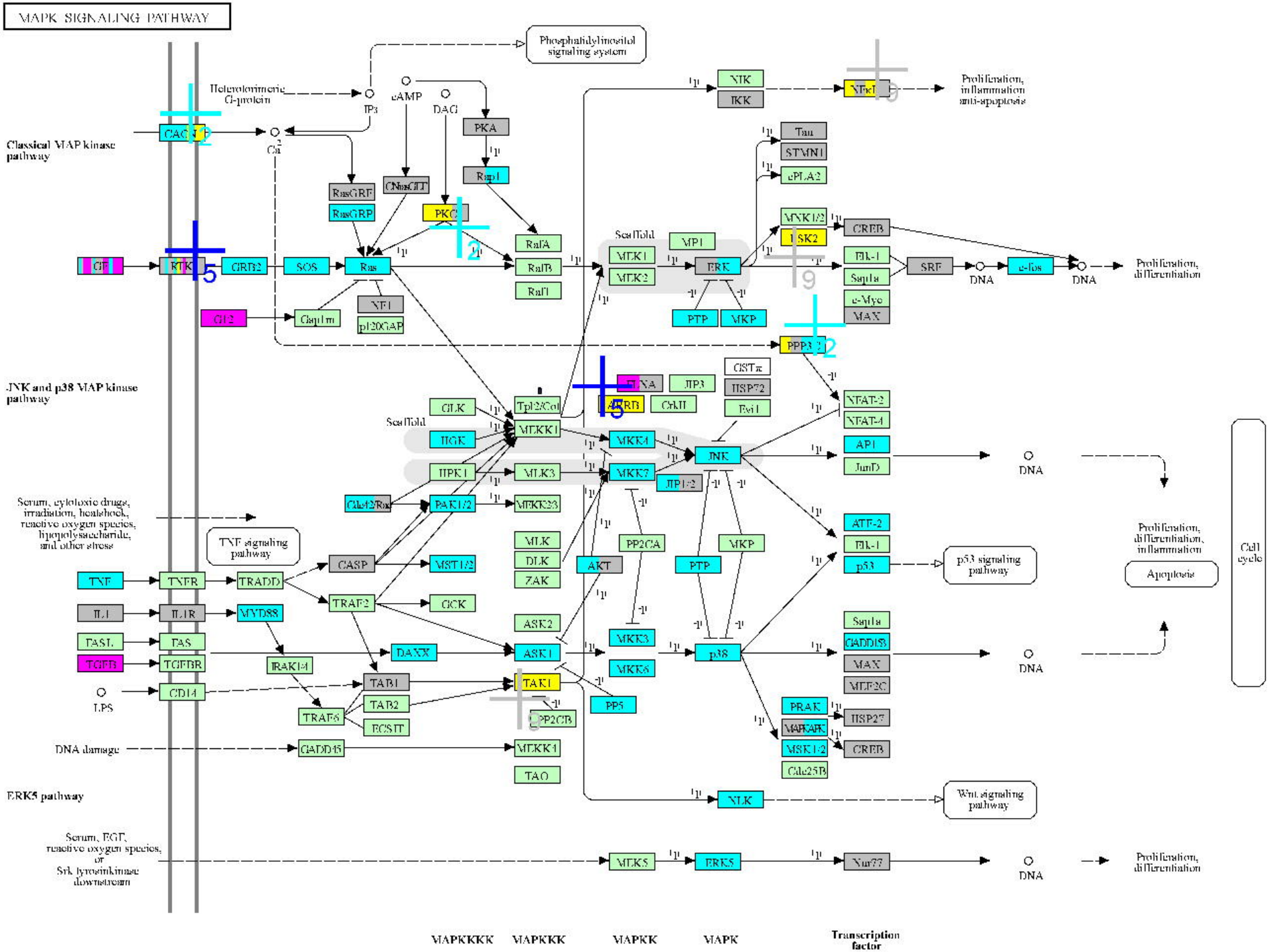
Bronson S, Westwood B, Cook K, Emenaker N, Chappell M, Roberts D, Soto-Pantoja D. Discrete Correlation Summation Clustering Reveals Differential Regulation of Liver Metabolism by Thrombospondin-1 in Low-Fat and High-Fat Diet-Fed Mice. Metabolites 2022, 12, 1036. https://doi.org/10.3390/metabo12111036

Fawcett T. An introduction to ROC analysis. Pattern Recogn Lett. 2006;27(8):861-74.

Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.net/ij/, 1997-2018.

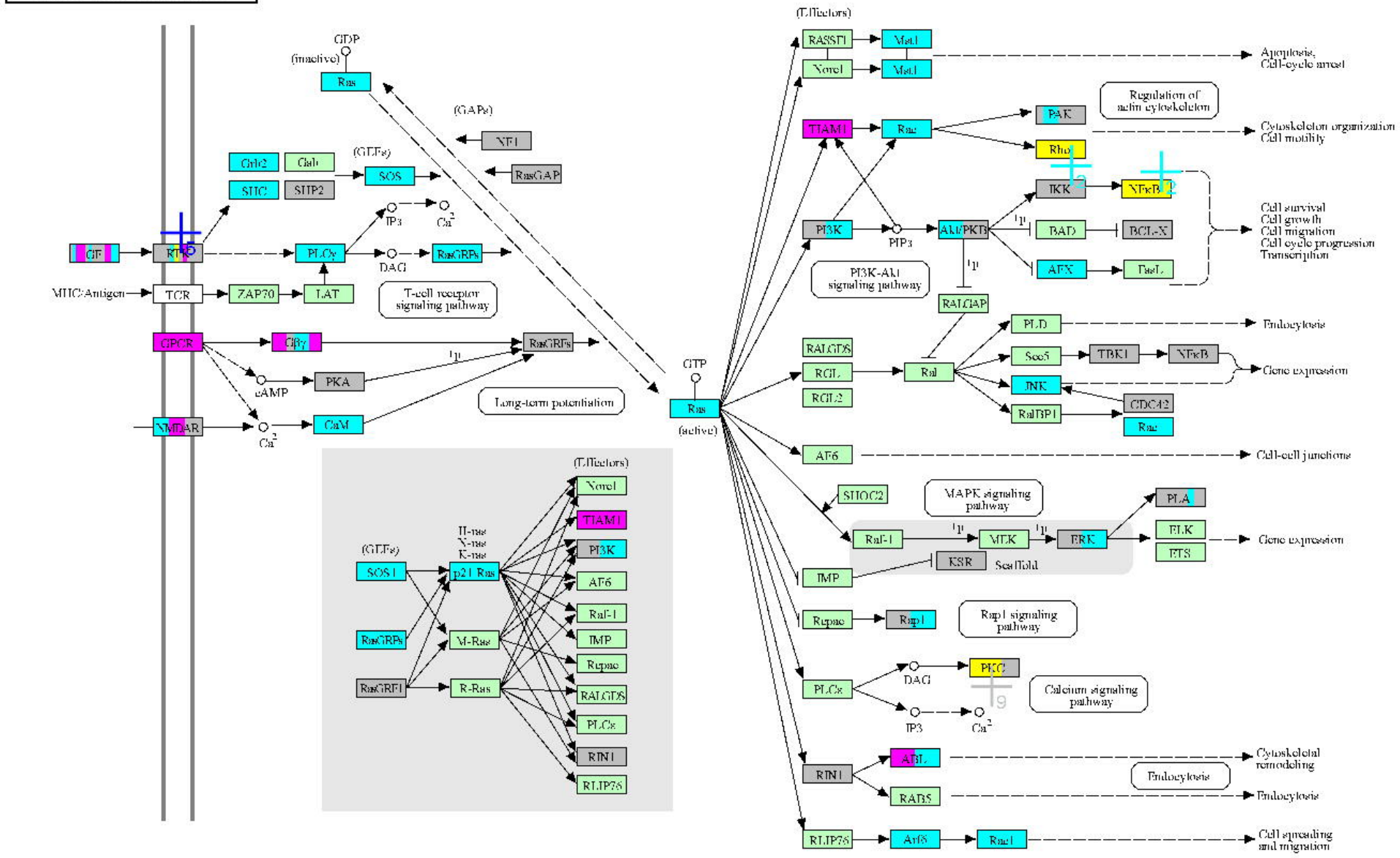
Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research 2003 Nov; 13(11):2498-504

Westwood B, Patil P, Tallant A, Gallagher P. Microbiome Metabolic Network Expansion Following Muscadine Grape Extract Intervention of Hypertensive Rats SB3C2022-266, 2022. https://archive.sb3c.org/sb3c2022/



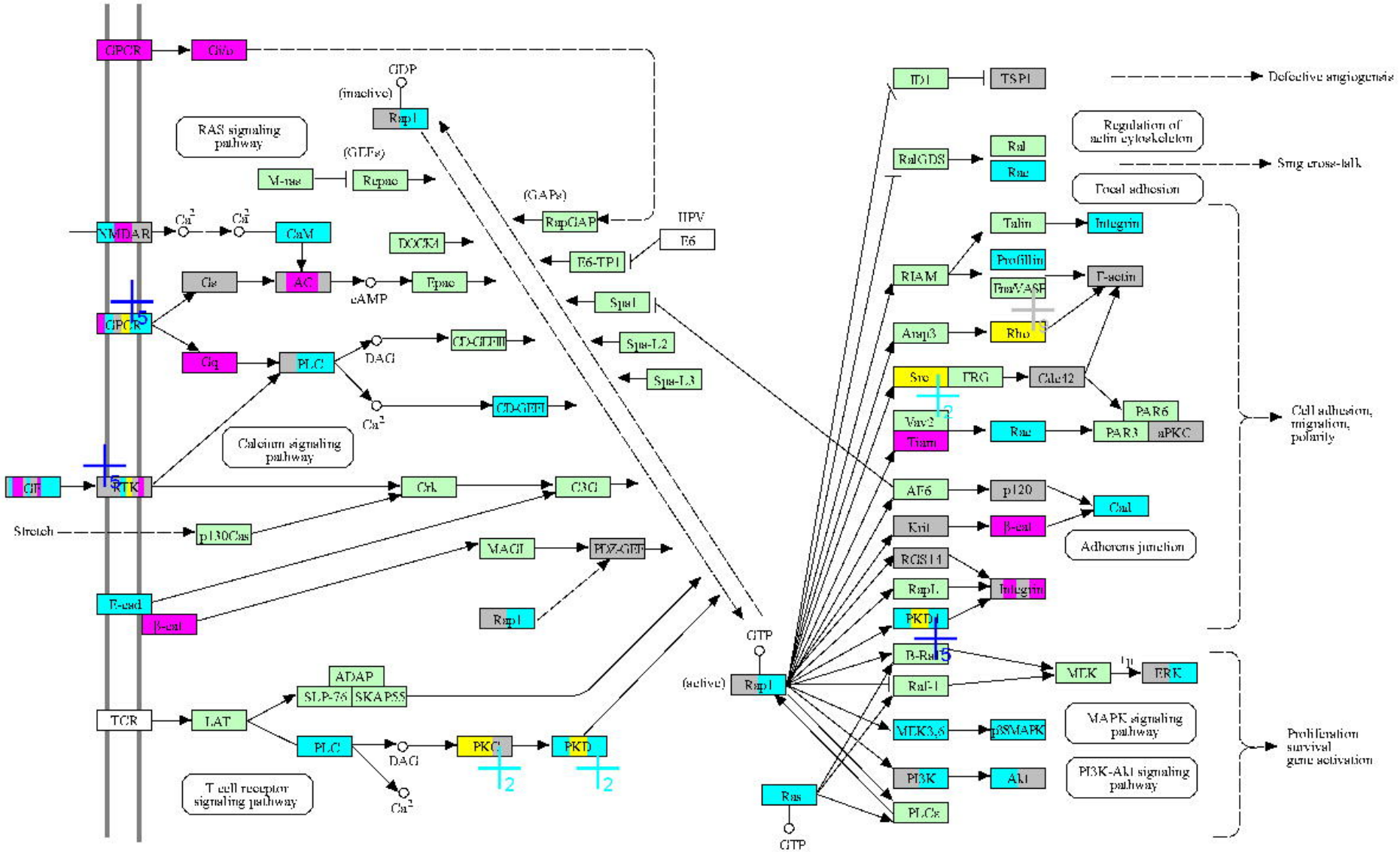
04010 4/11/23
(c) Kanehisa Laboratories

RAS SIGNALING PATHWAY



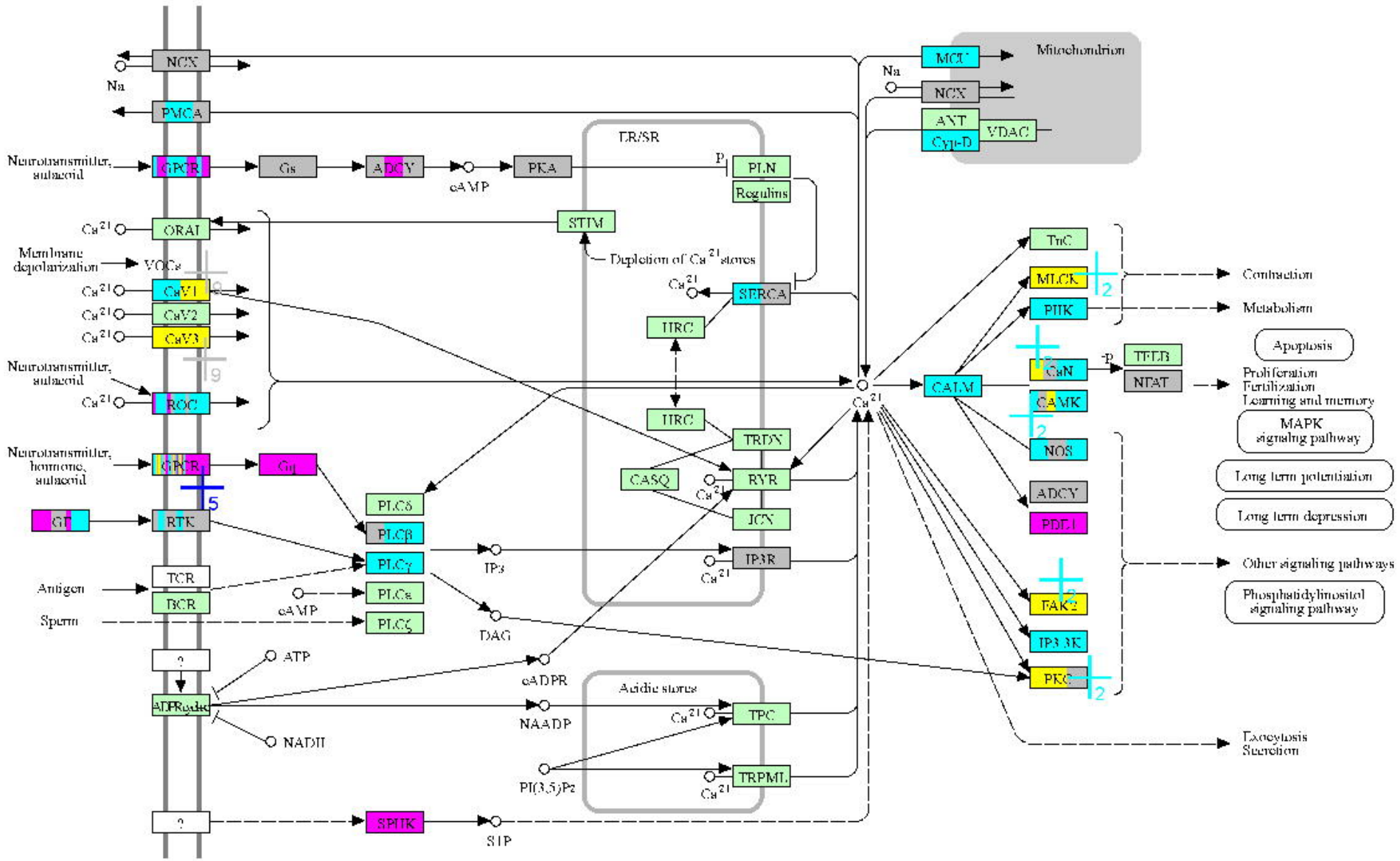
04014 4/27/22
(c) Kanehisa Laboratories

RAP1 SIGNALING PATHWAY

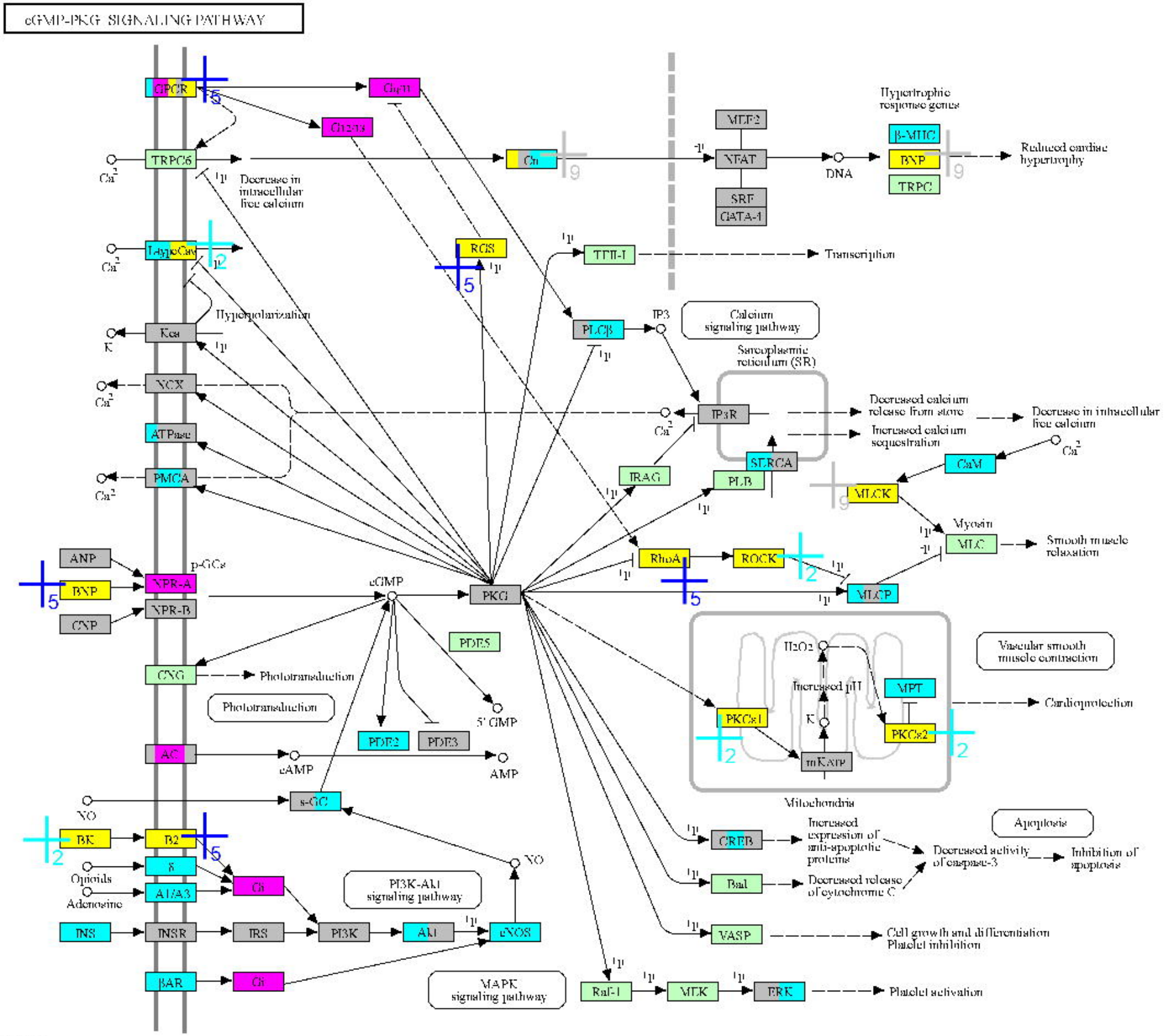


04015 6/25/19
(c) Kanehisa Laboratories

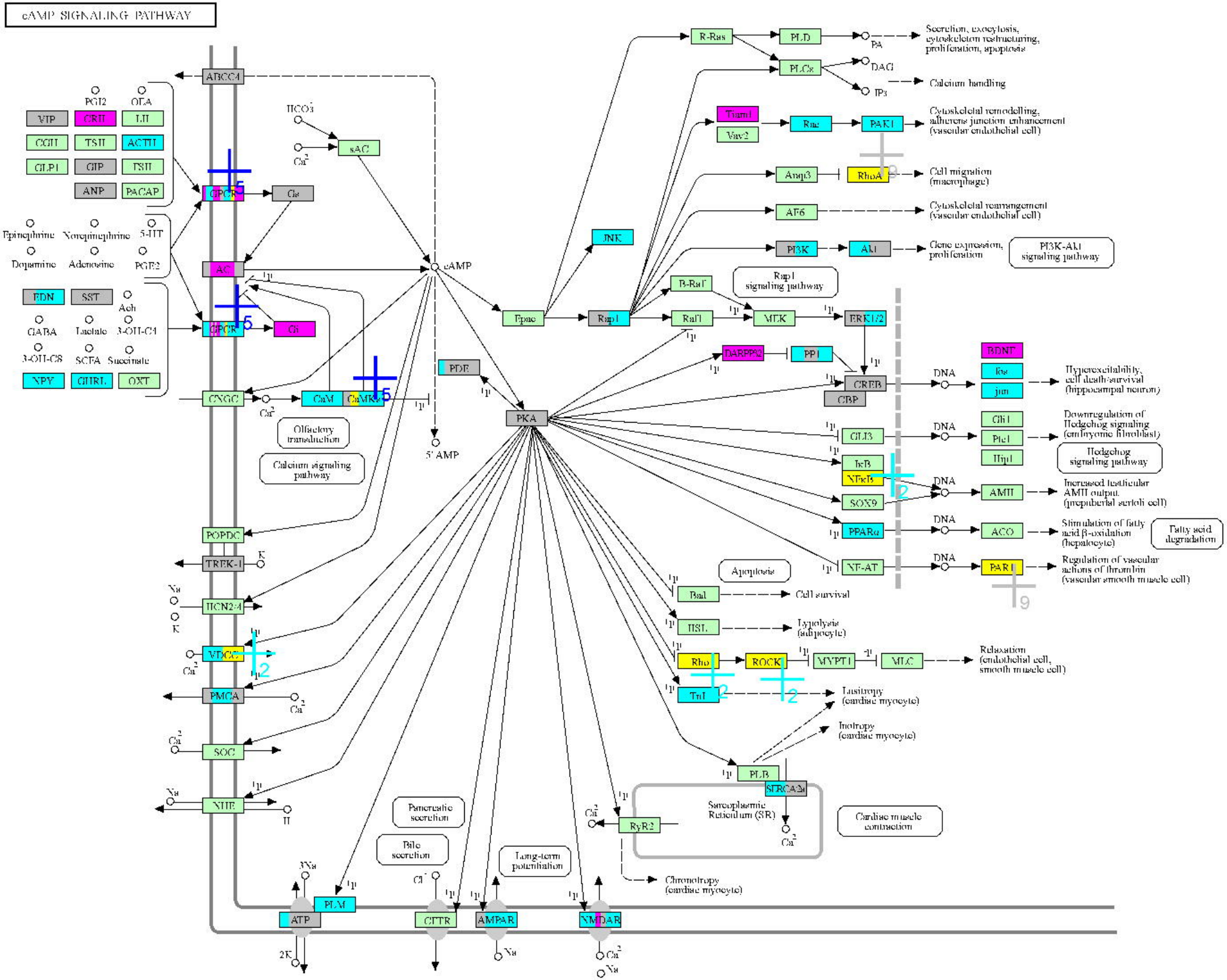
CALCIUM SIGNALING PATHWAY

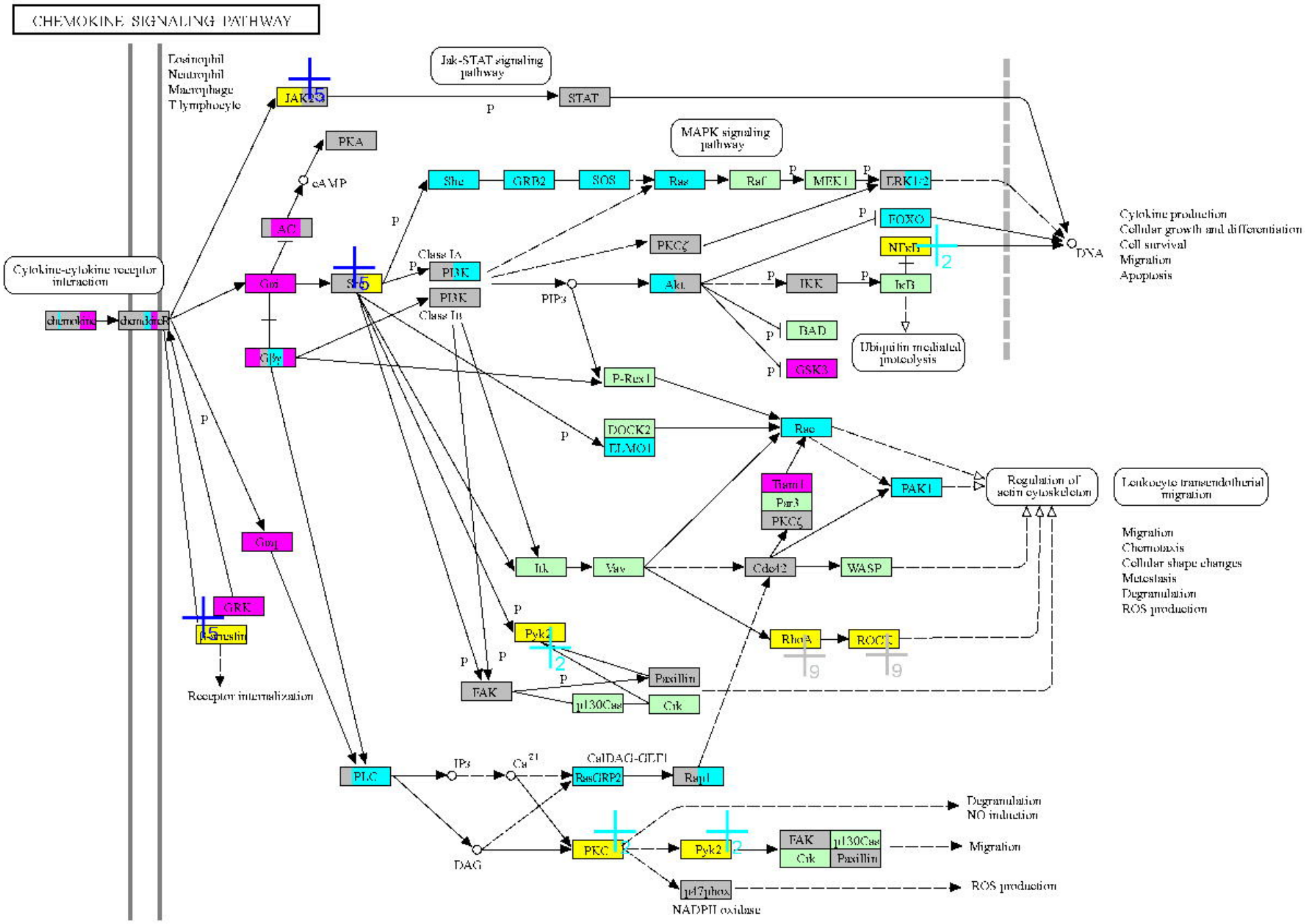


(c) Kanehisa Laboratories

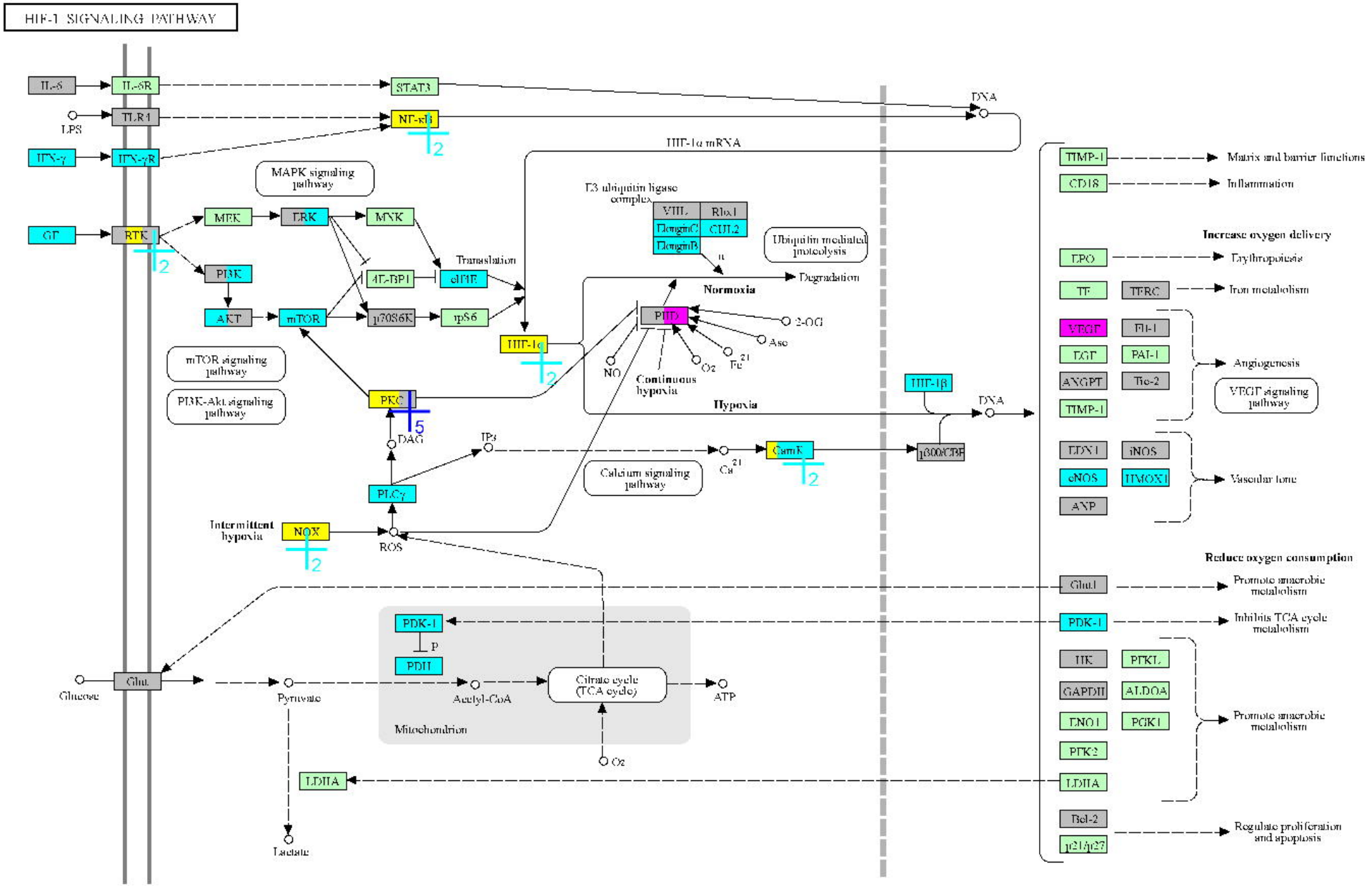


04022 6/4/19
(c) Kanehisa Laboratories

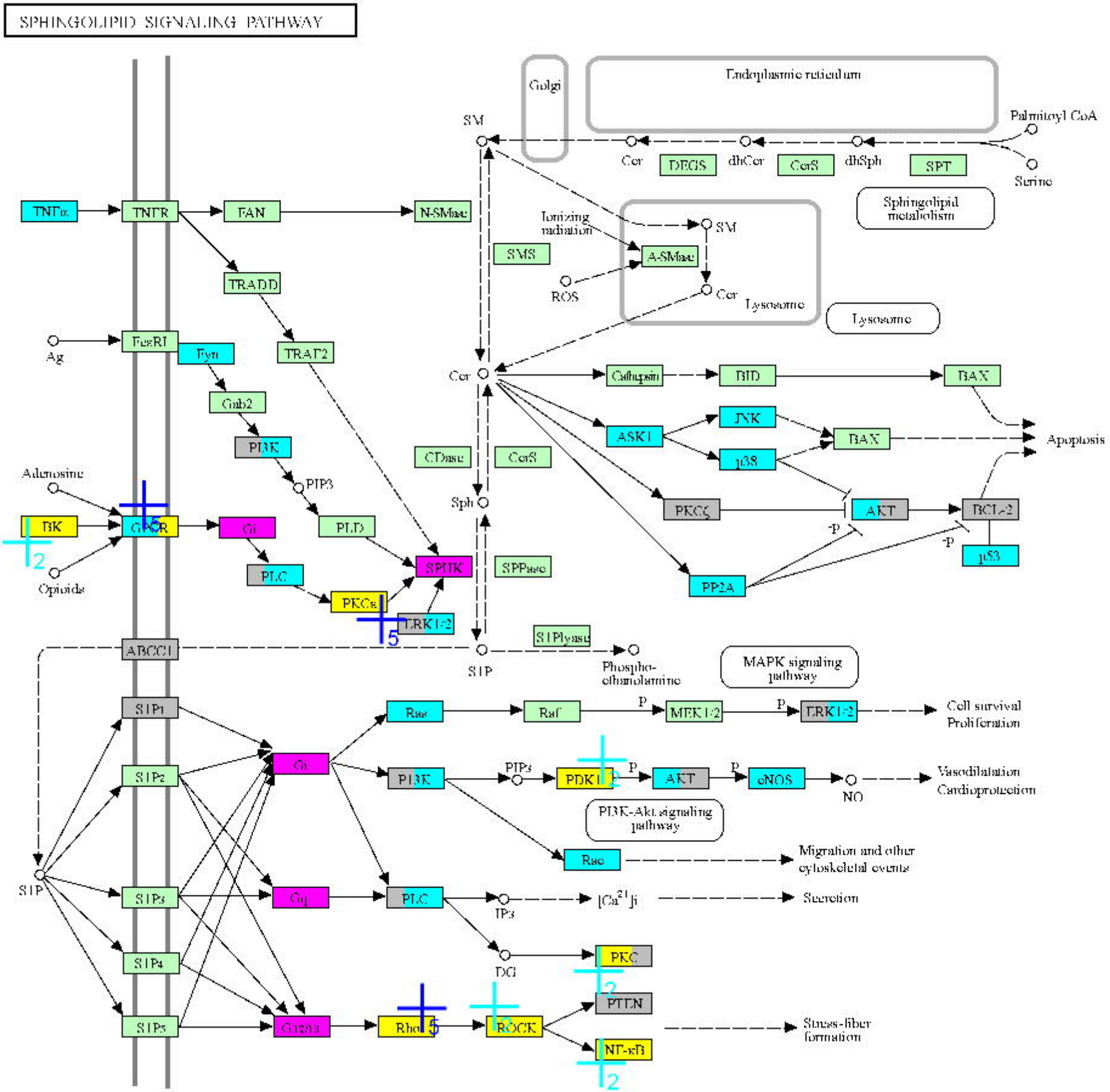

cAMP SIGNALING PATHWAY
ABCC4
PGI2
OLA
VIP
CRH
LH
CGH
TSH
ACTH
GLP1
GIP
TSH
ANP
PACAP
GPCR
Gs
Epinephrine
Norepinephrine
5-HT
Dopamine
Adenosine
PGE2
EDN
SST
Ach
GABA
Lactate
3-OH-C4
3-OH-C8
SCFA
Succinate
NPY
GHRL
OXT
AC
GPCR
Gi
CNGC
CaM
CaMKII
Olfactory transduction
Calcium signaling pathway
HCO3
sAC
cAMP
5' AMP
PDE
PKA
Epac
Rap1
JNK
R-Ras
PLD
PA
Secretion, exocytosis, cytoskeleton restructuring, proliferation, apoptosis
PLCε
DAG
IP3
Calcium handling
Tiam1
Vav2
Rac
PAK1
Cytoskeletal remodelling, adherens junction enhancement (vascular endothelial cell)
Arp2/3
RhoA
Cell migration (macrophage)
AF6
Cytoskeletal rearrangement (vascular endothelial cell)
PI3K
Akt
Gene expression, proliferation
PI3K-Akt signaling pathway
B-Raf
Rap1 signaling pathway
Raf1
MEK
ERK1/2
DARPP32
PP1
CREB
CBP
DNA
BDNF
fos
jun
Hyperexcitability, cell death/survival (hippocampal neuron)
GLI3
Gli1
Ptc1
Hip1
Downregulation of Hedgehog signaling (embryonic fibroblast)
Hedgehog signaling pathway
IκB
NFκB
SOX9
AMH
Increased testicular AMH output (prepubertal sertoli cell)
PPARα
ACO
Stimulation of fatty acid β-oxidation (hepatocyte)
Fatty acid degradation
NF-AT
PAR1
Regulation of vascular actions of thrombin (vascular smooth muscle cell)
Apoptosis
Bad
Cell survival
HSL
Lipolysis (adipocyte)
Rho
ROCK
MYPT1
MLC
Relaxation (endothelial cell, smooth muscle cell)
TnI
Lusitropy (cardiac myocyte)
Inotropy (cardiac myocyte)
PLB
SERCA2a
Sarcoplasmic Reticulum (SR)
RyR2
Cardiac muscle contraction
Chronotropy (cardiac myocyte)
POPDC
TREK-1
HCN2/4
VDCC
PMCA
SOC
NHE
Pancreatic secretion
Bile secretion
Long-term potentiation
PLM
ATP
CFTR
AMPAR
NMDAR
04024 12/9/22
(c) Kanehisa Laboratories



04062 8/24/20
(c) Kanehisa Laboratories

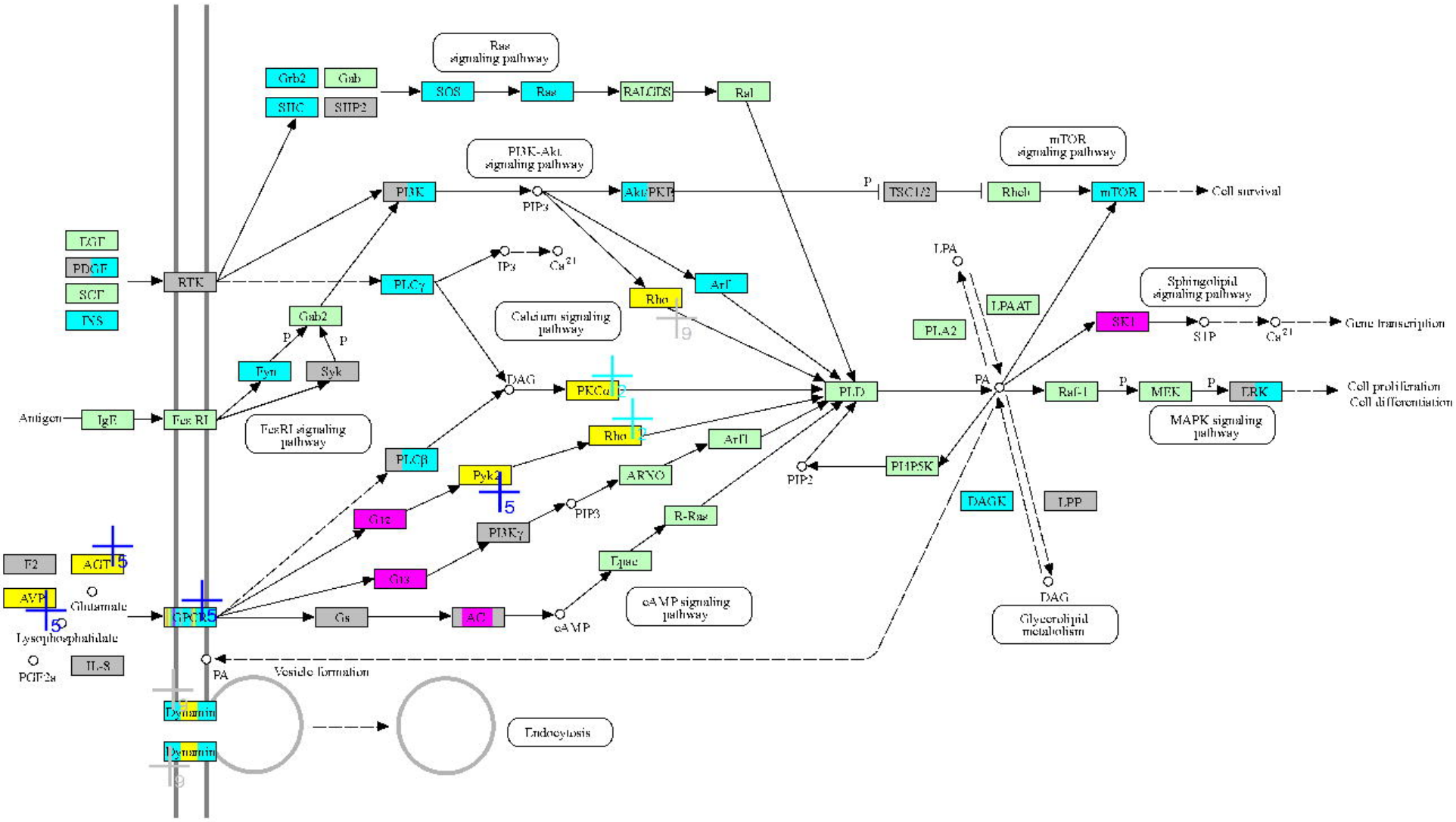


04066 8/5/21
(c) Kanehisa Laboratories



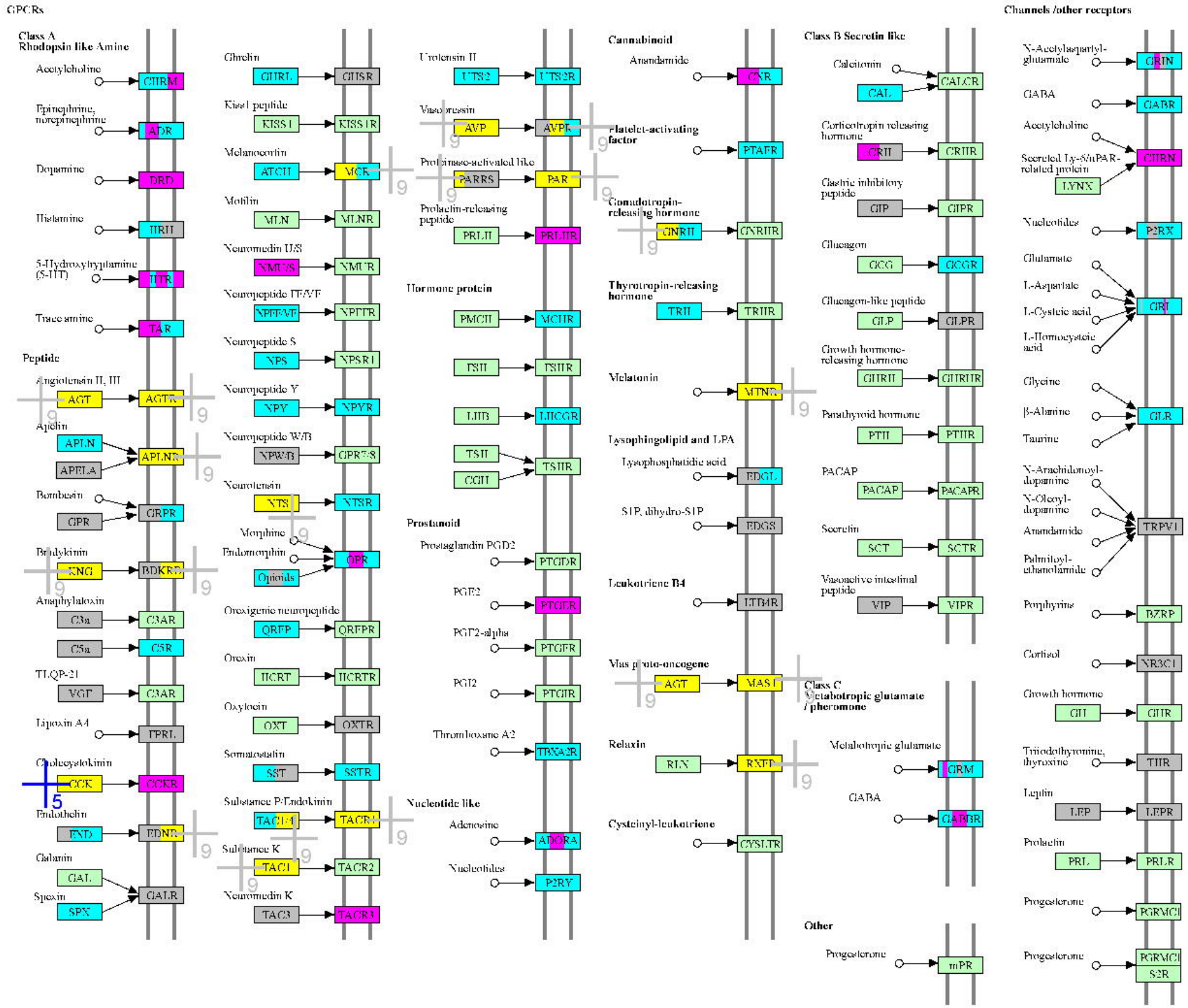
04071 2/1/19
(c) Kanehisa Laboratories

PHOSPHOLIPASE D SIGNALING PATHWAY

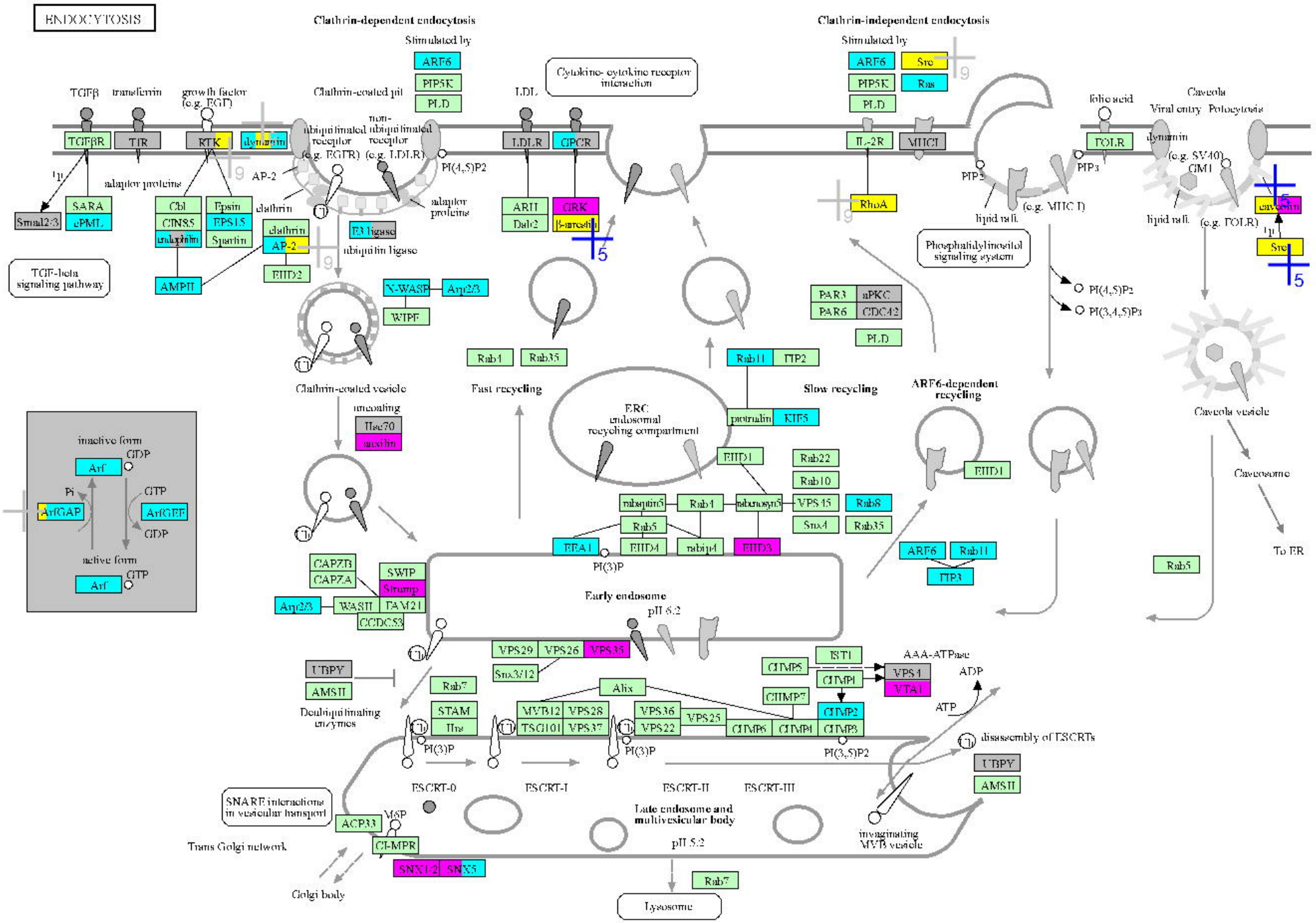


04072 8/29/19
(c) Kanehisa Laboratories

NEUROACTIVE LIGAND-RECEPTOR INTERACTION

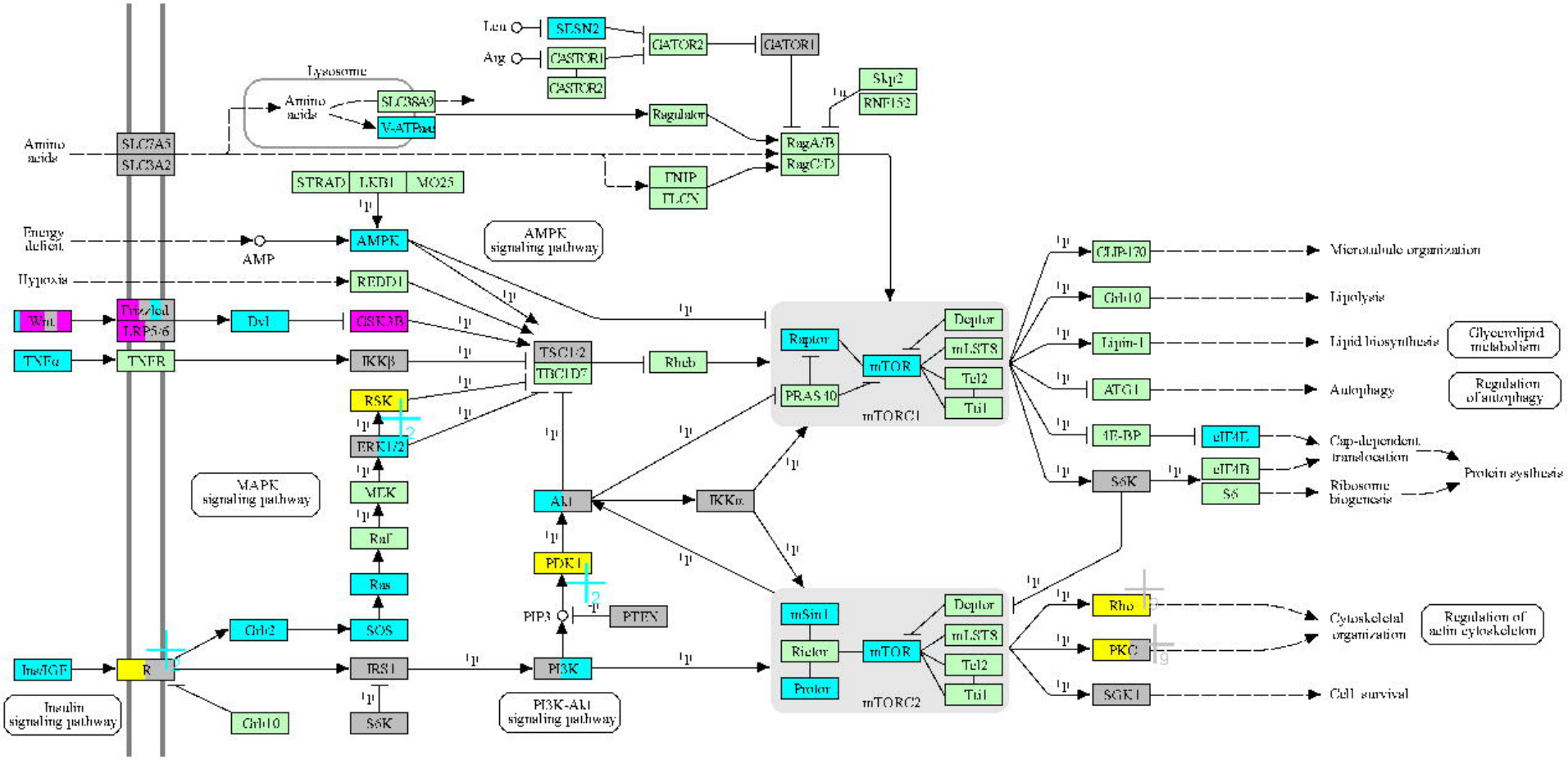


04080 11/30/22
(c) Kanehisa Laboratories

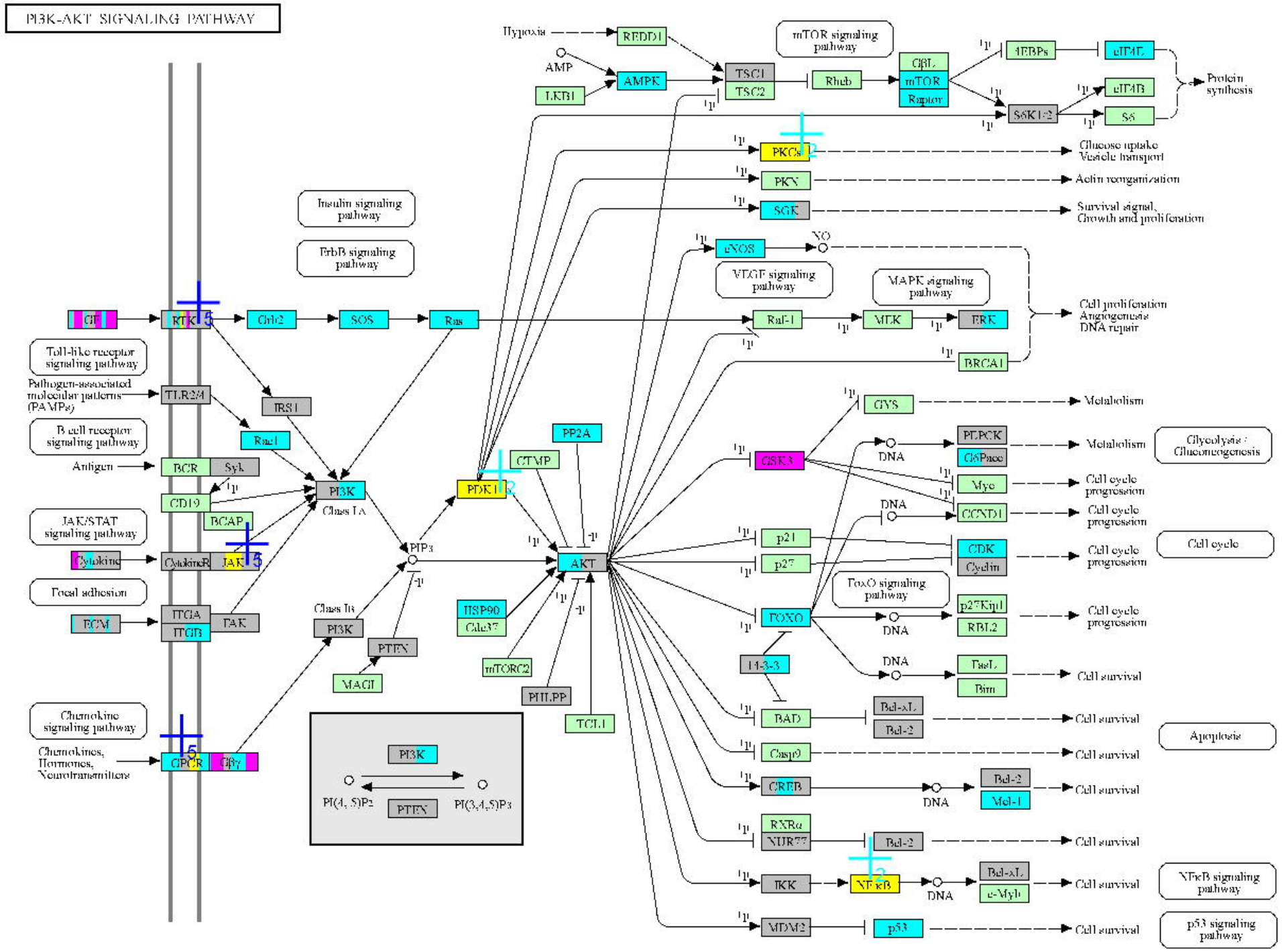


04144 3/4/21
(c) Kanehisa Laboratories

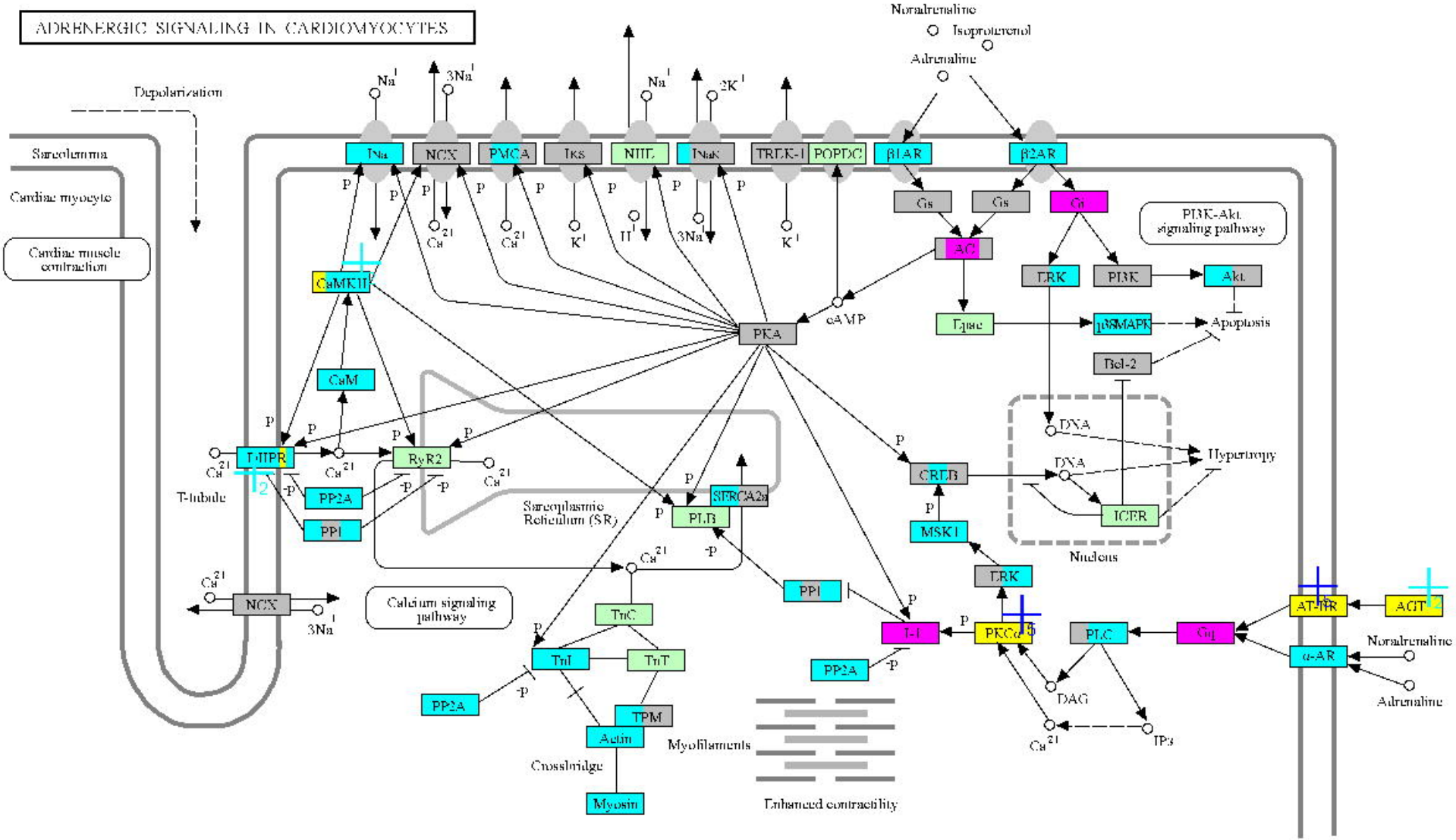
mTOR SIGNALING PATHWAY



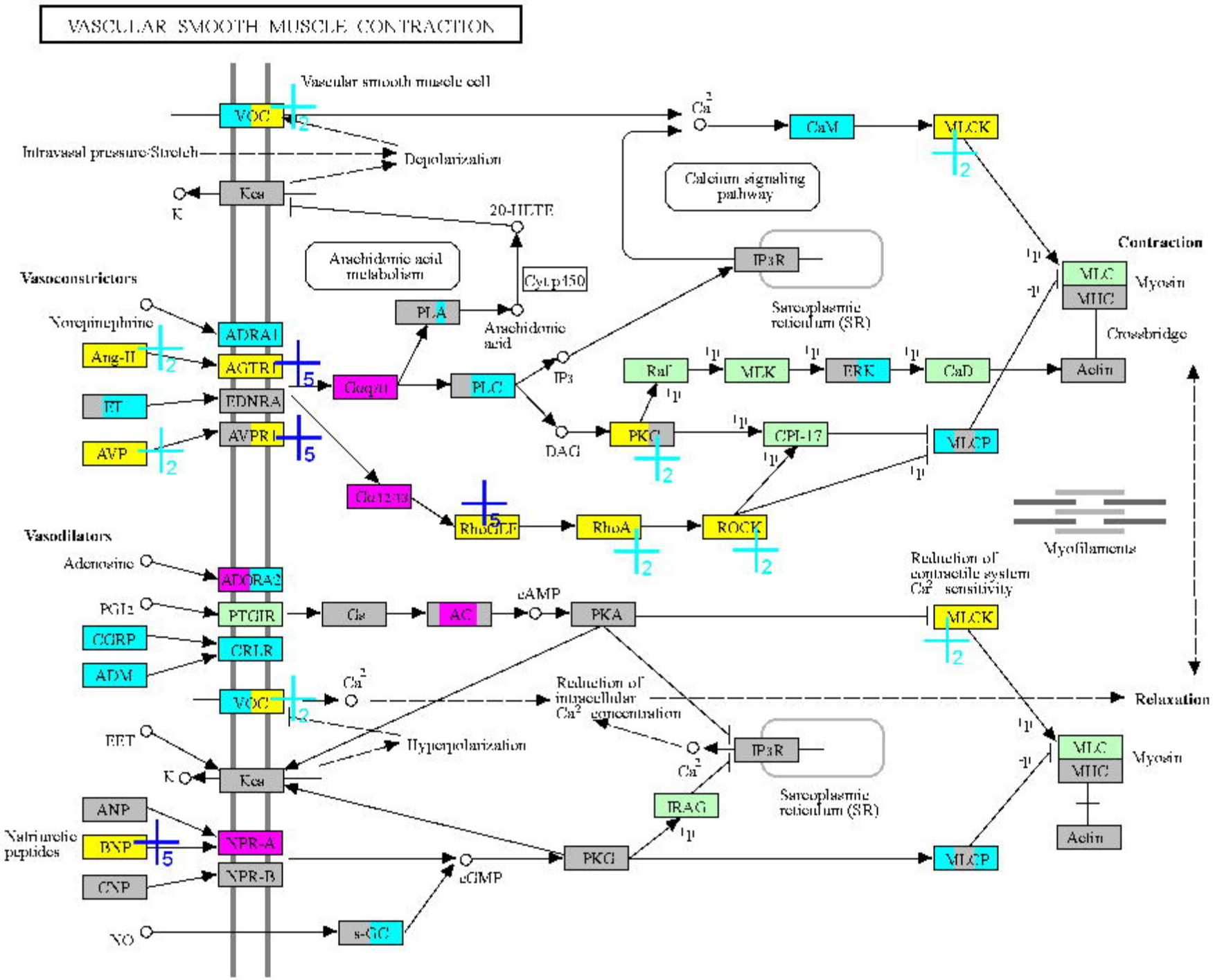
04150 7/7/21
(c) Kanehisa Laboratories



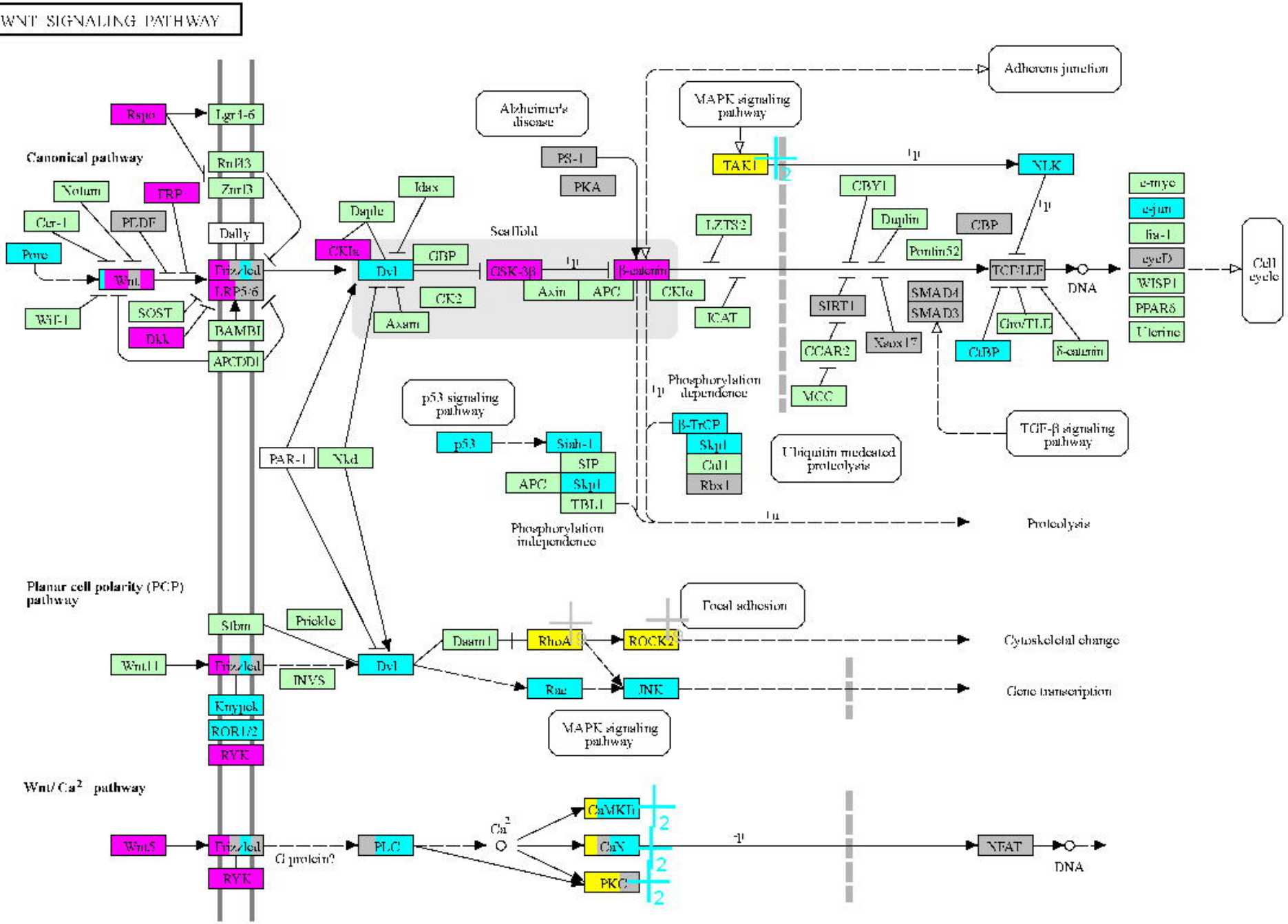
04151 7/7/21
(c) Kanehisa Laboratories



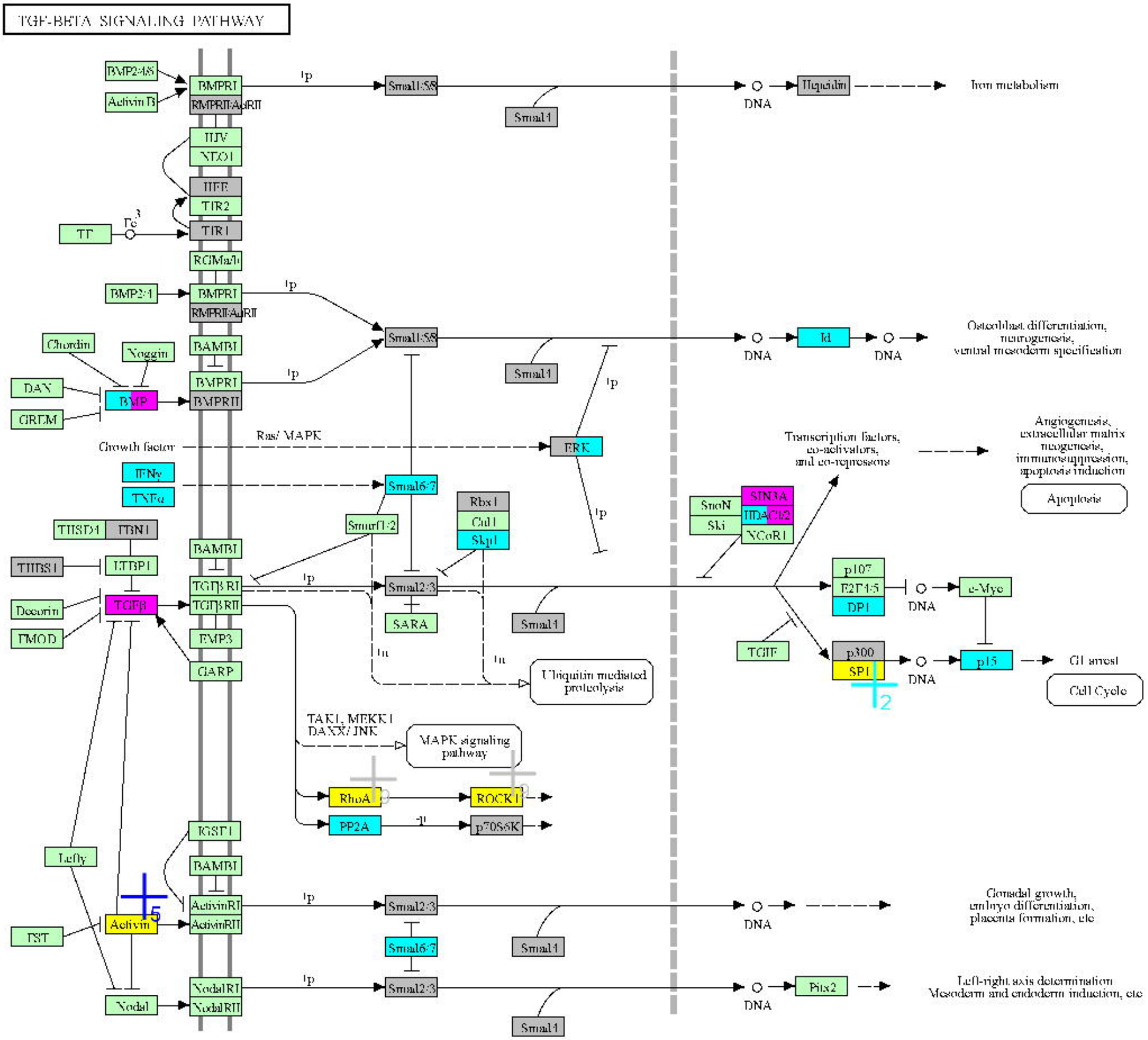
04261 12/9/22
(c) Kanehisa Laboratories



04270 6/11/20
(c) Kanehisa Laboratories

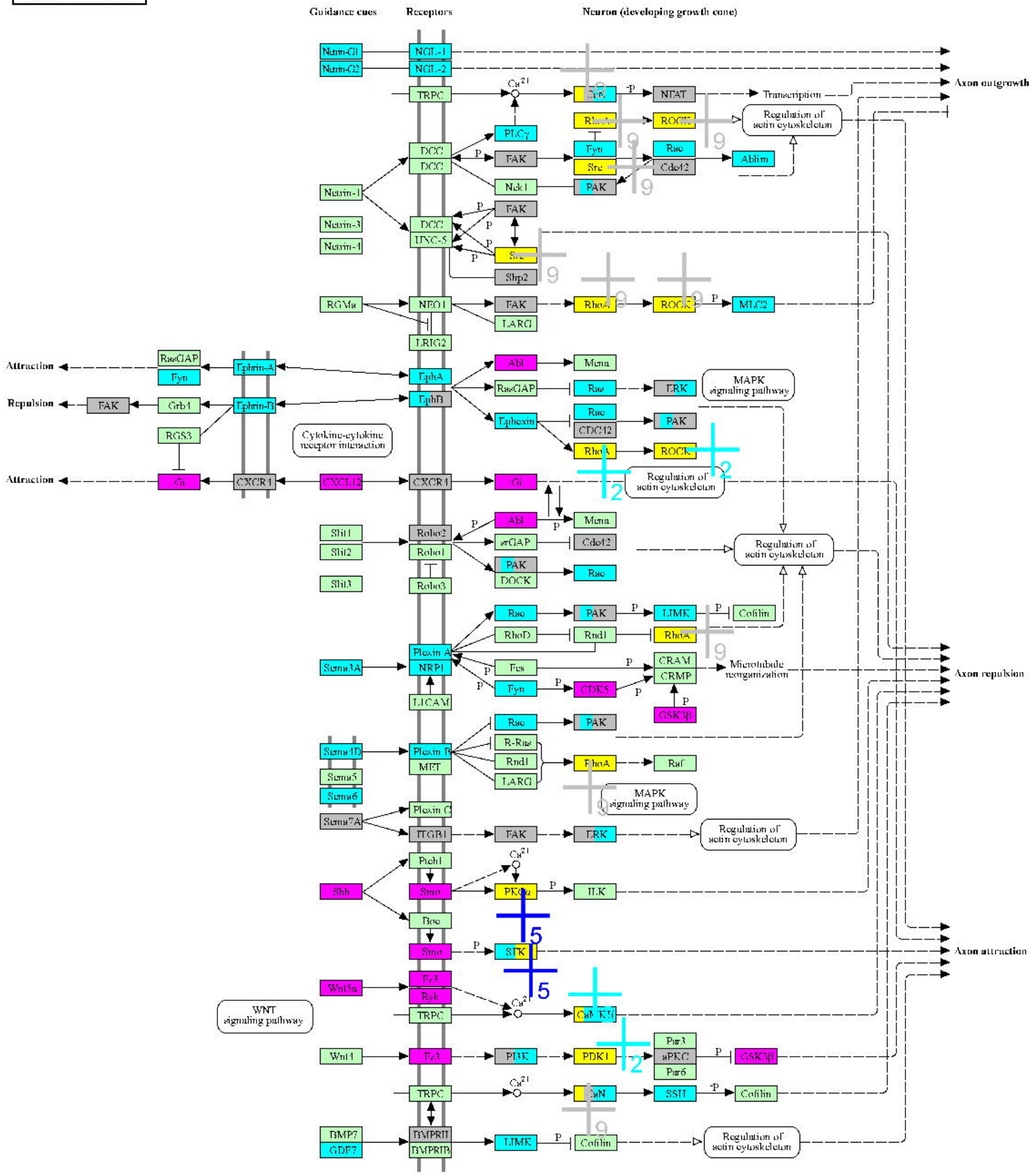


04310 8/9/23
(c) Kanehisa Laboratories



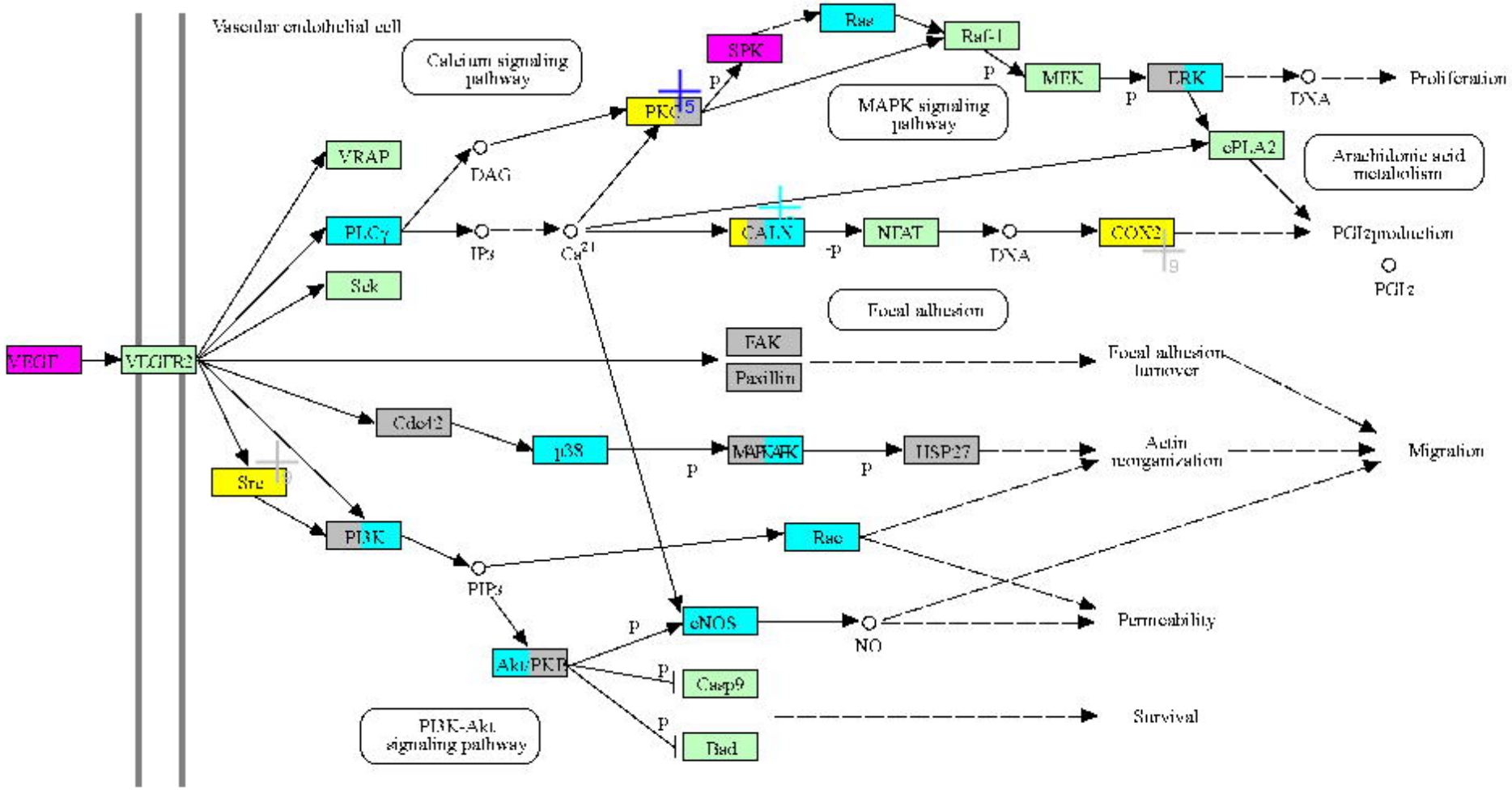
04350 5/25/23
(c) Kanehisa Laboratories

AXON GUIDANCE



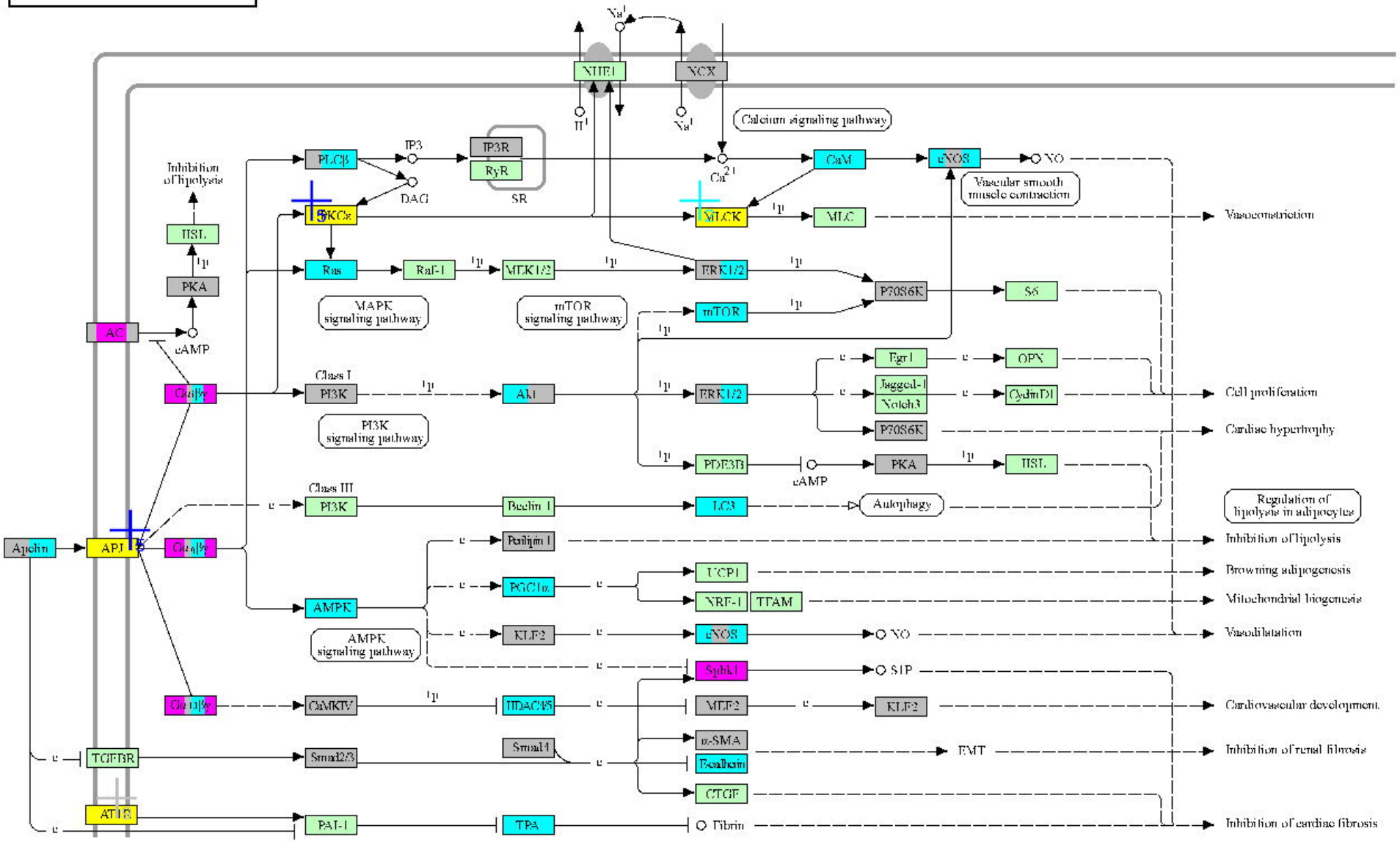
04360 3/19/21
(c) Kanehisa Laboratories

VEGF SIGNALING PATHWAY

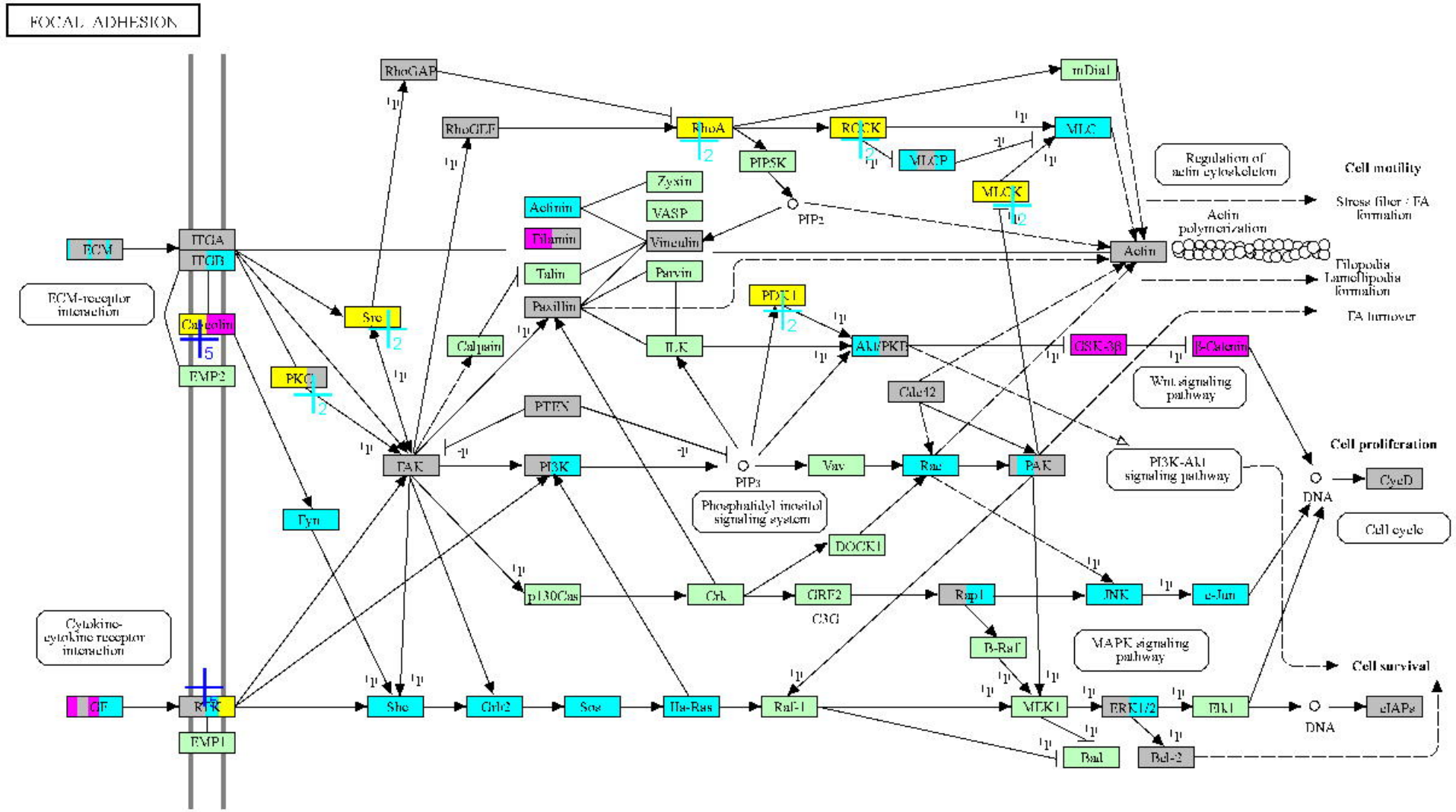


04370 8/28/17
(c) Kanehisa Laboratories

APELIN SIGNALING PATHWAY

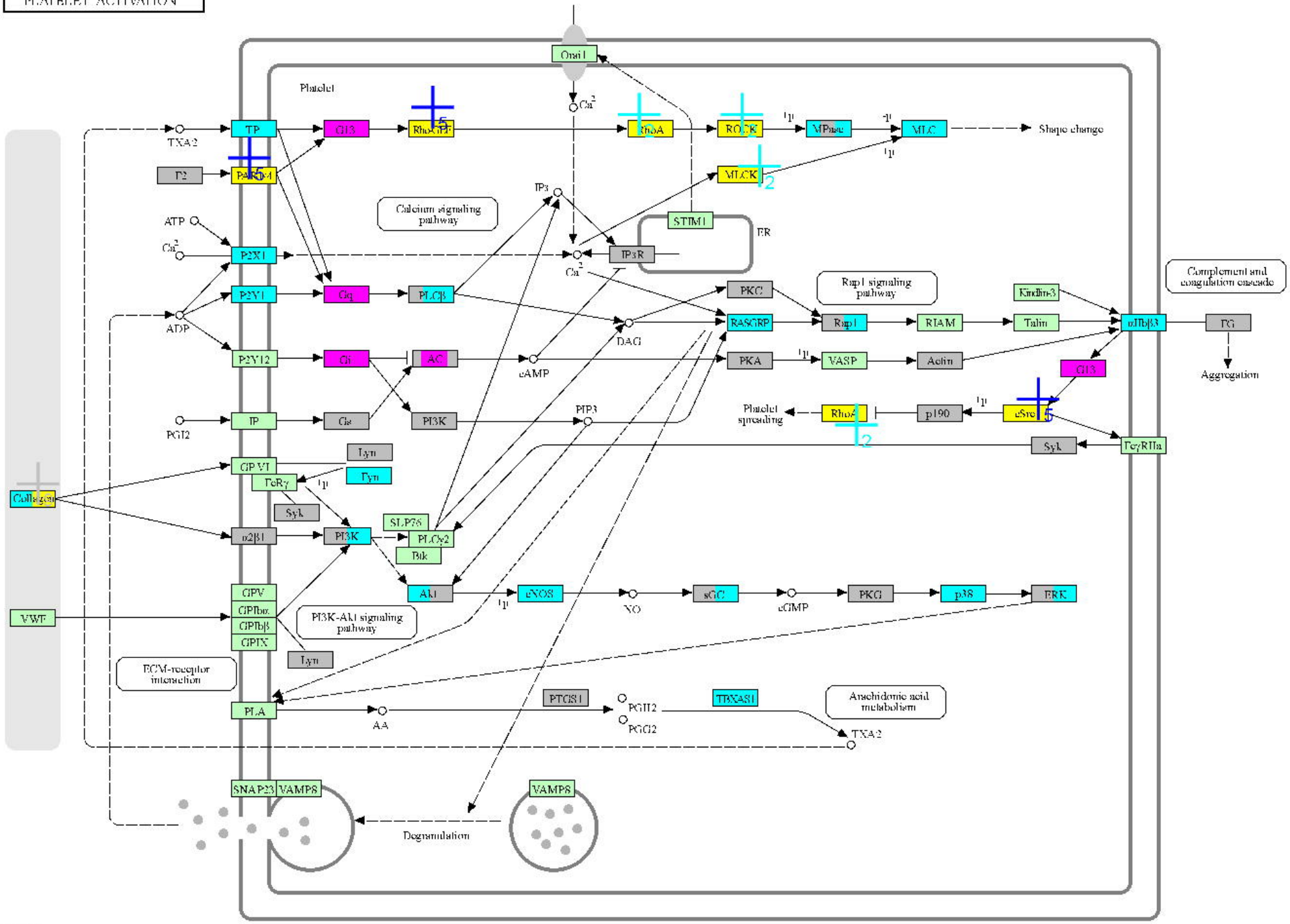


04371 6/21/21
(c) Kanehisa Laboratories



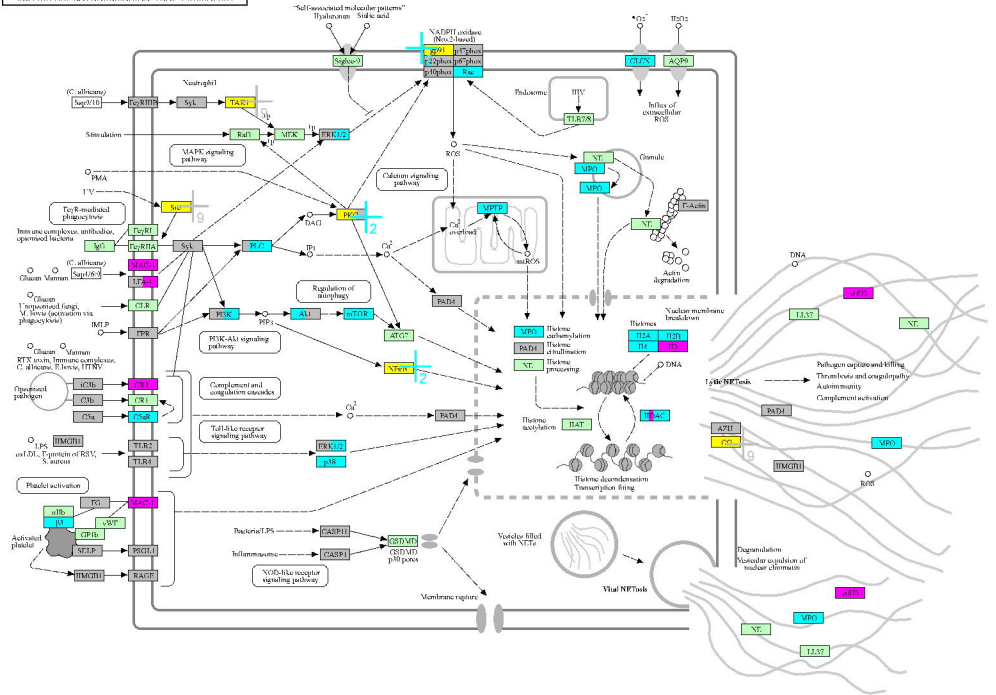
04510 1/25/23
(c) Kanehisa Laboratories

PLATELET ACTIVATION



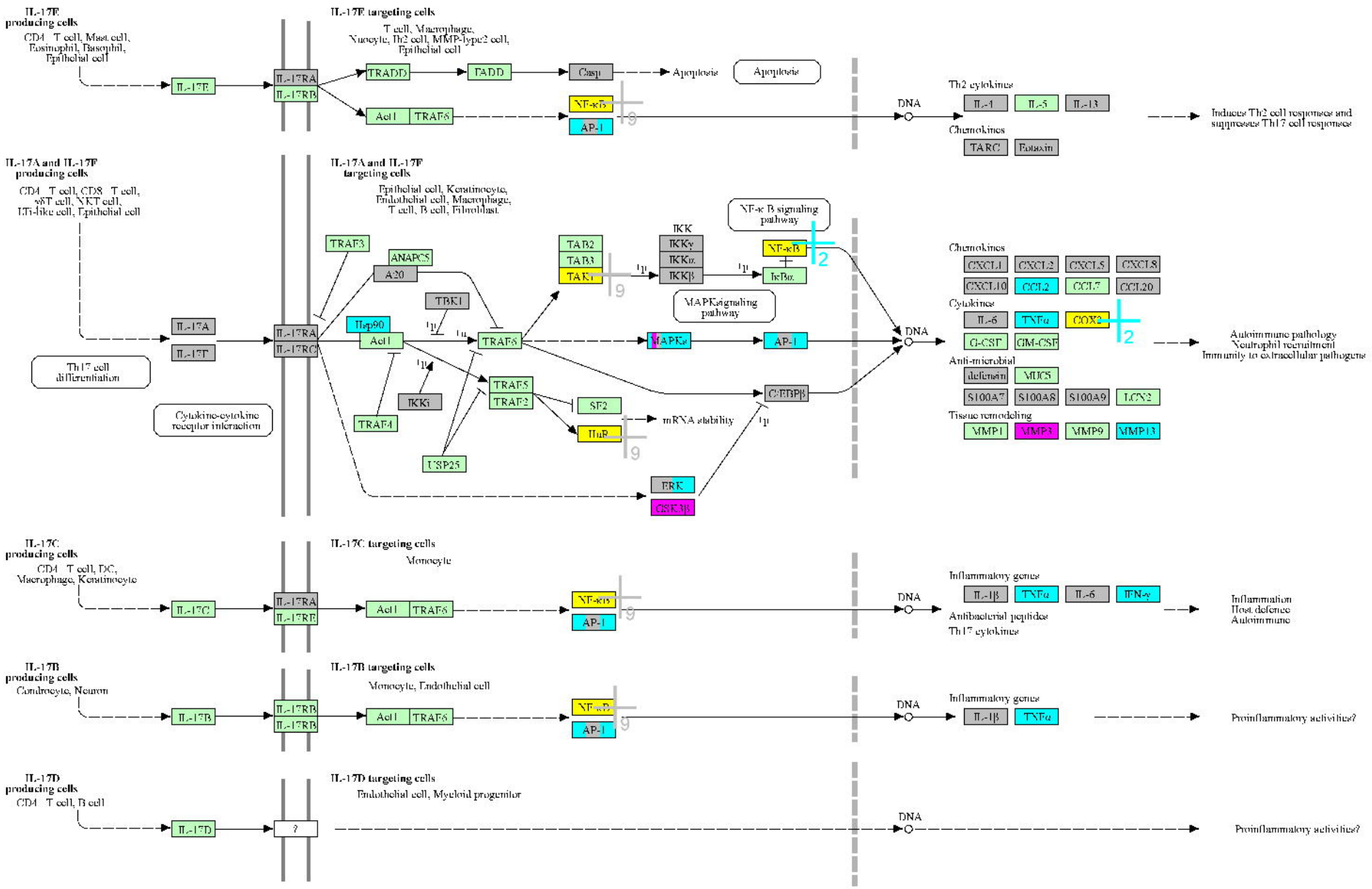
04611 2/1/19
(c) Kanehisa Laboratories

NEUTROPHIL EXTRACELLULAR TRAP FORMATION



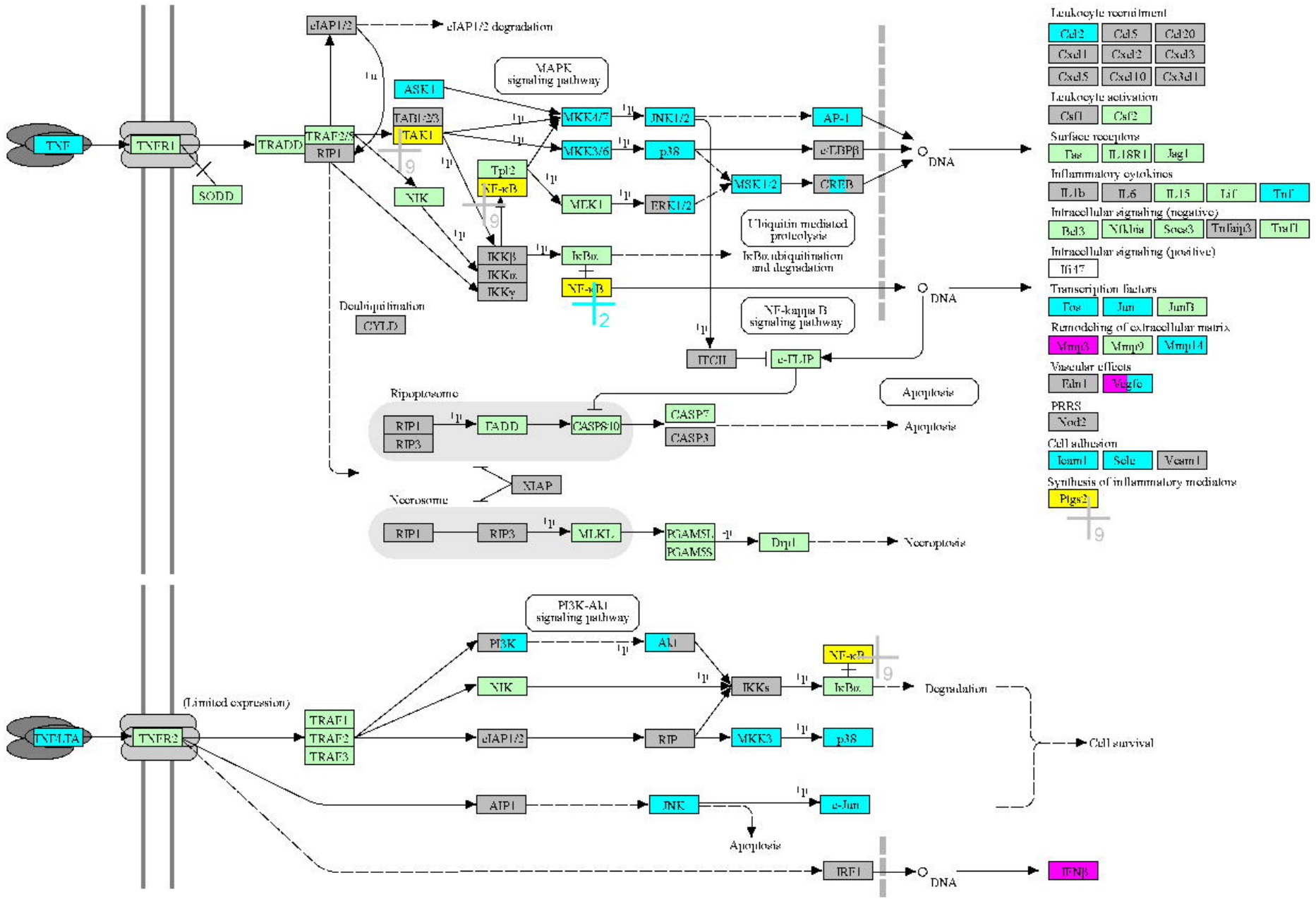
04613 3/29/31
(c) Kanehisa Laboratories

IL-17 SIGNALING PATHWAY

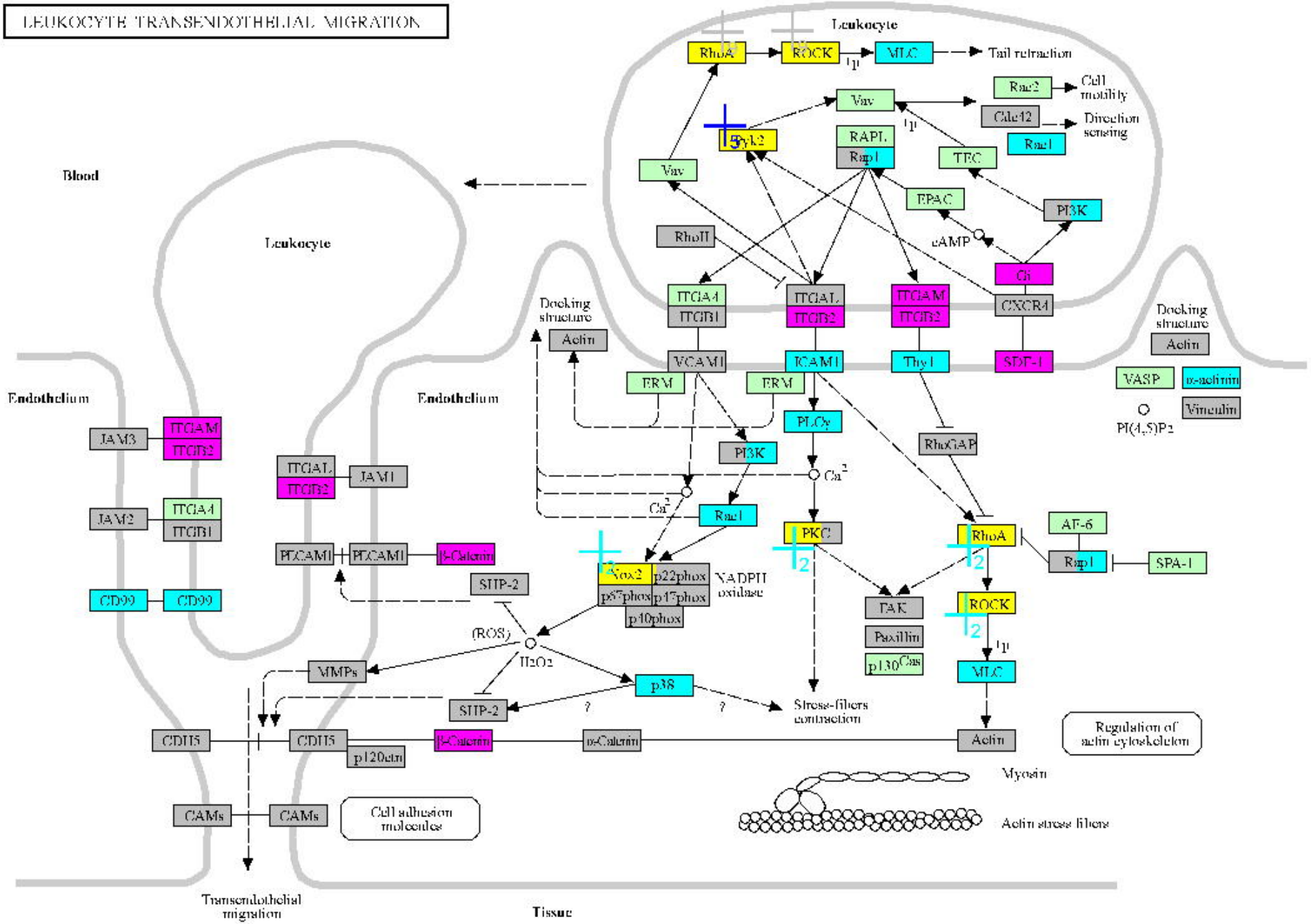


04657 3/15/23
(c) Kanehisa Laboratories

TNF SIGNALING PATHWAY

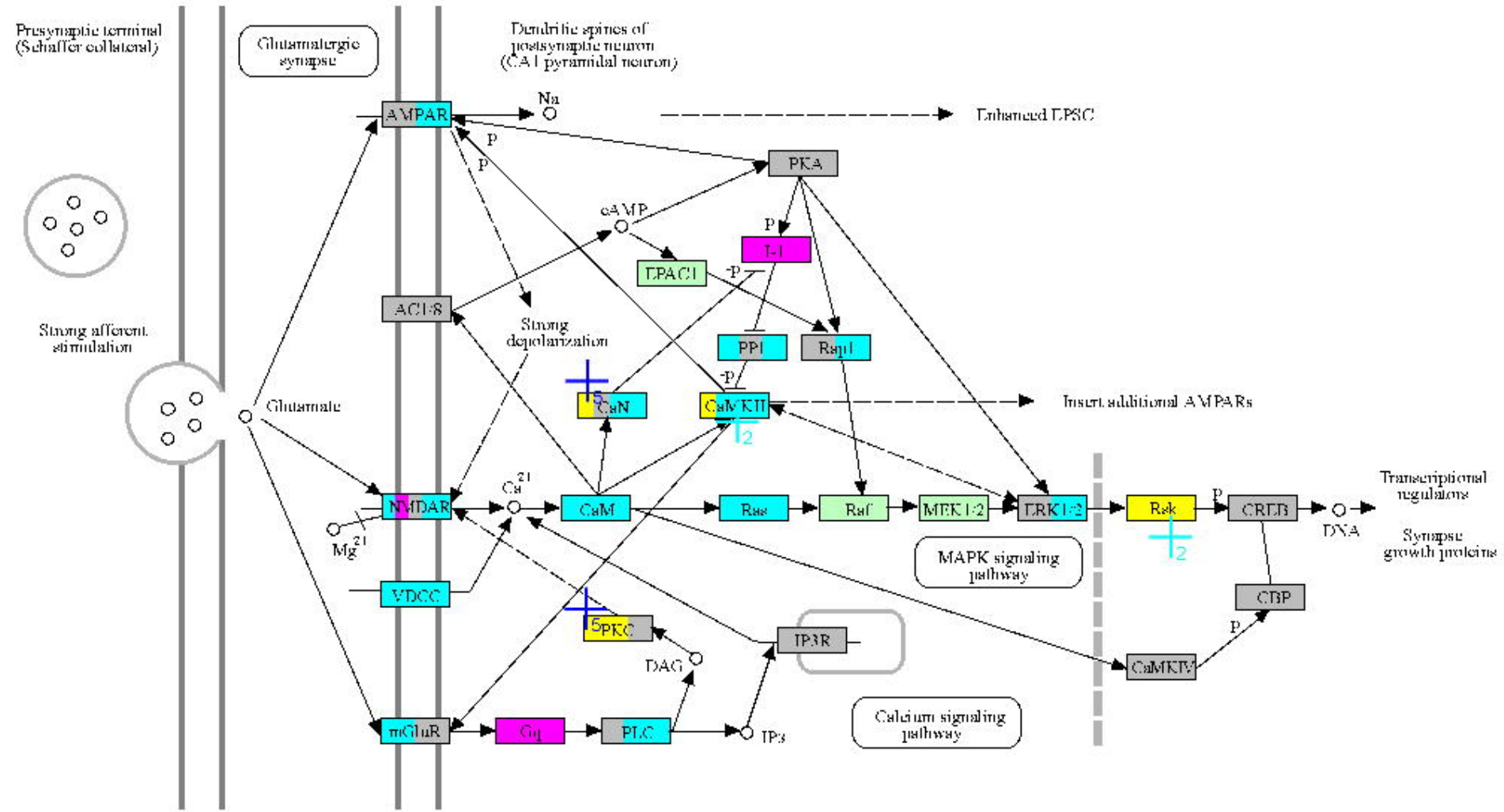


04668 12/23/22
(c) Kanehisa Laboratories

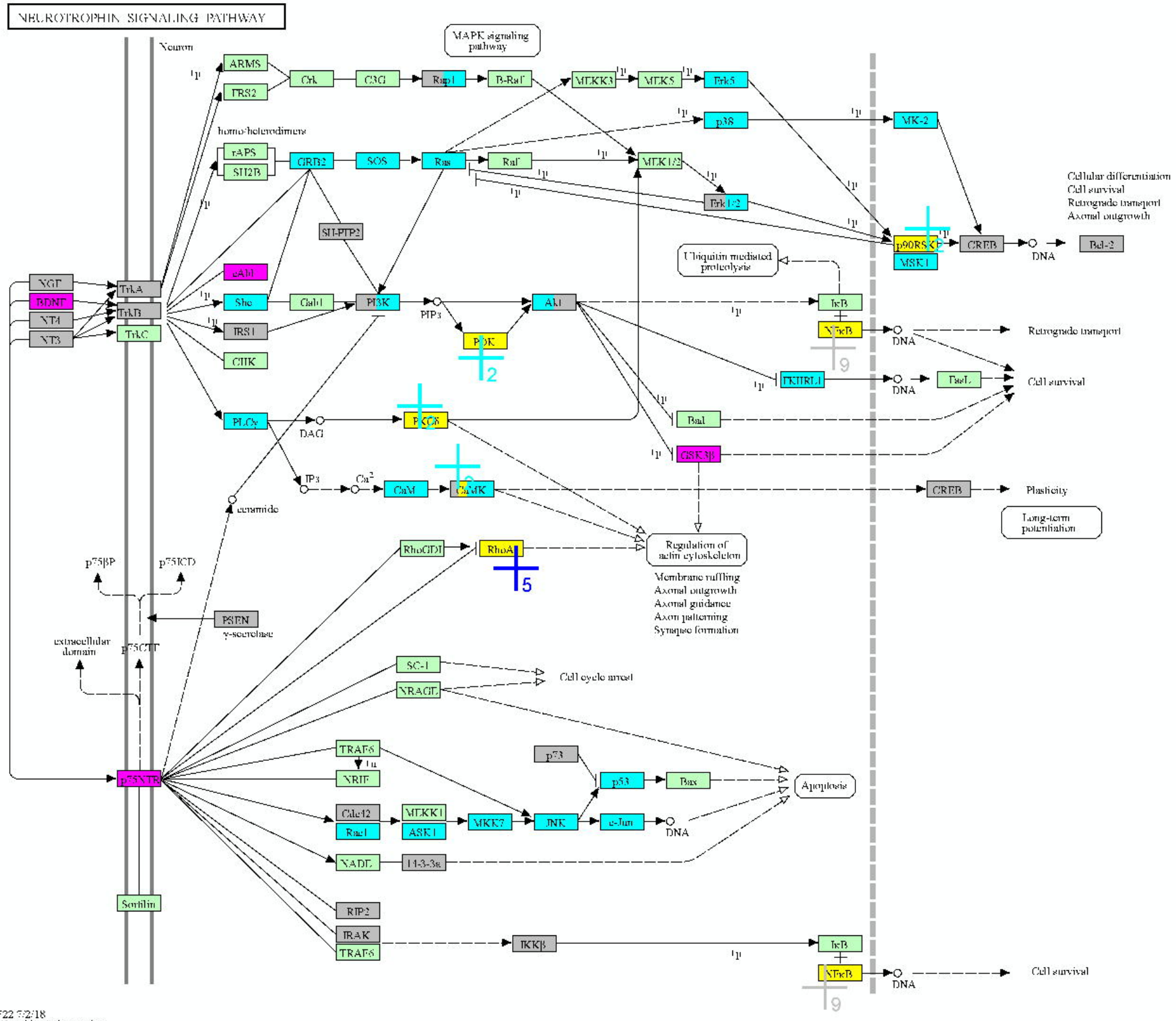


04670 3/24/17
(c) Kanehisa Laboratories

LONG-TERM POTENTIATION

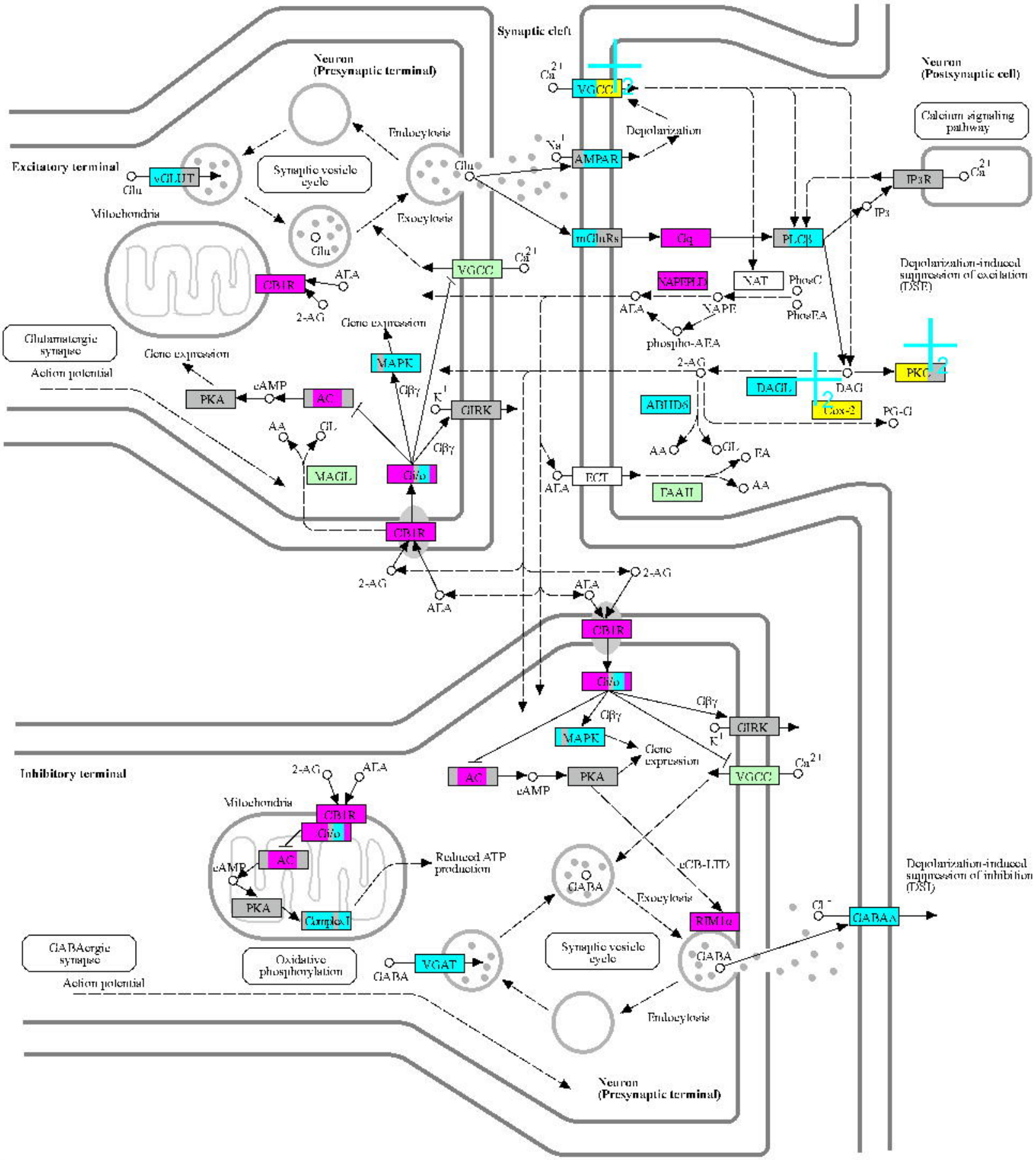


04720 10/23/15
(c) Kanehisa Laboratories



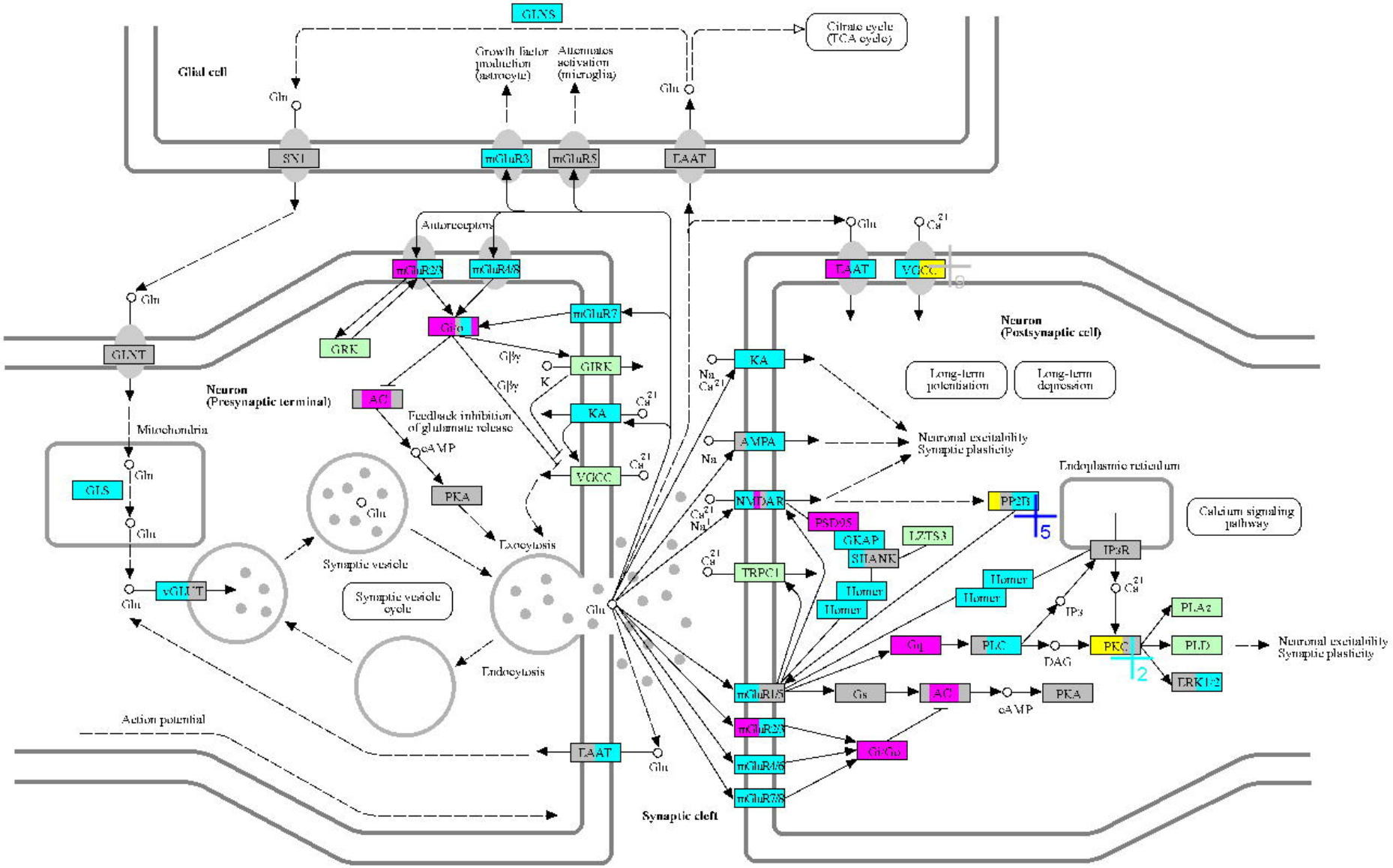
04722 7/2/18
(c) Kanehisa Laboratories

RETROGRADE ENDOCANNABINOID SIGNALING

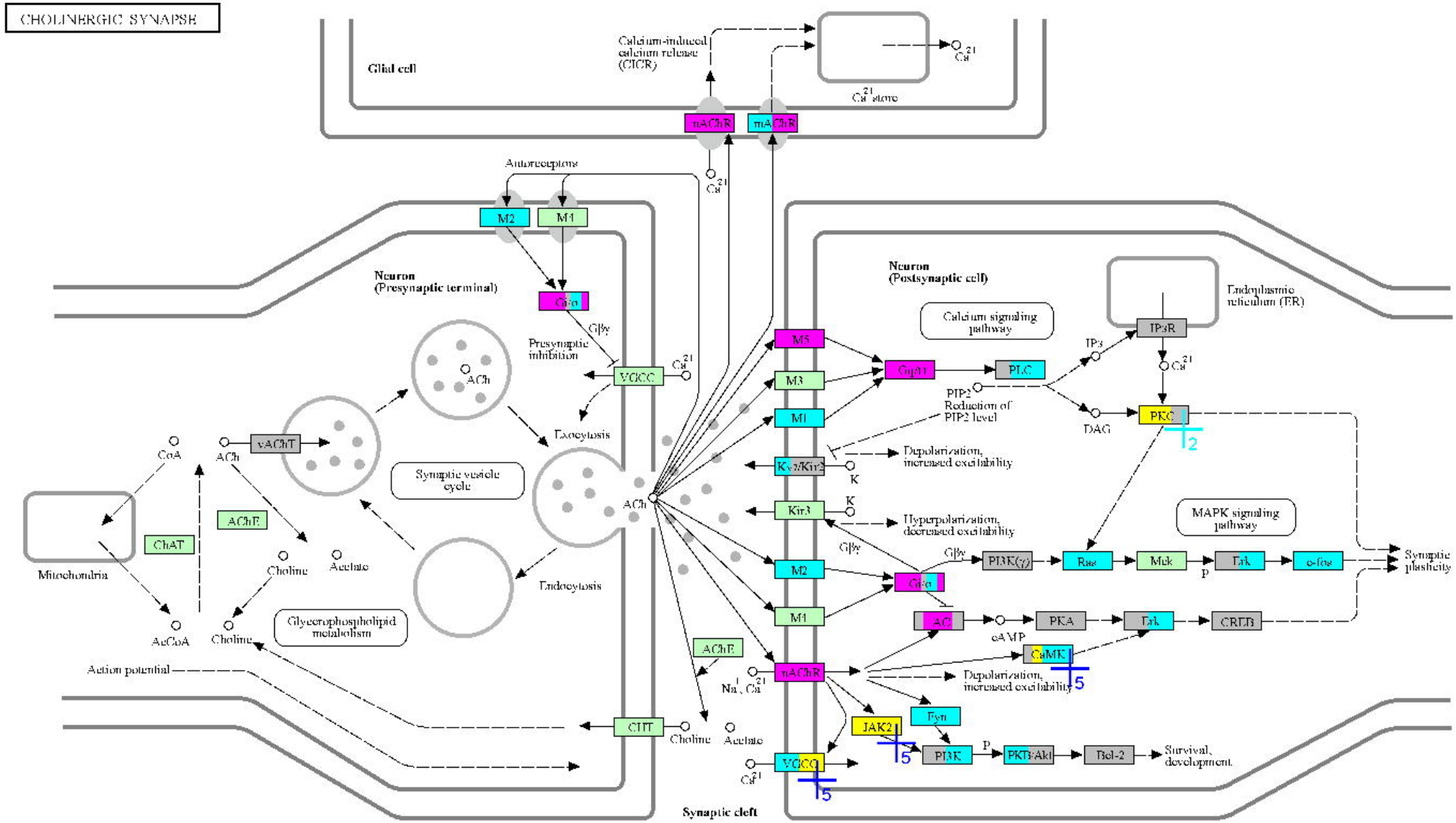


04723 4/17/17
(c) Kanehisa Laboratories

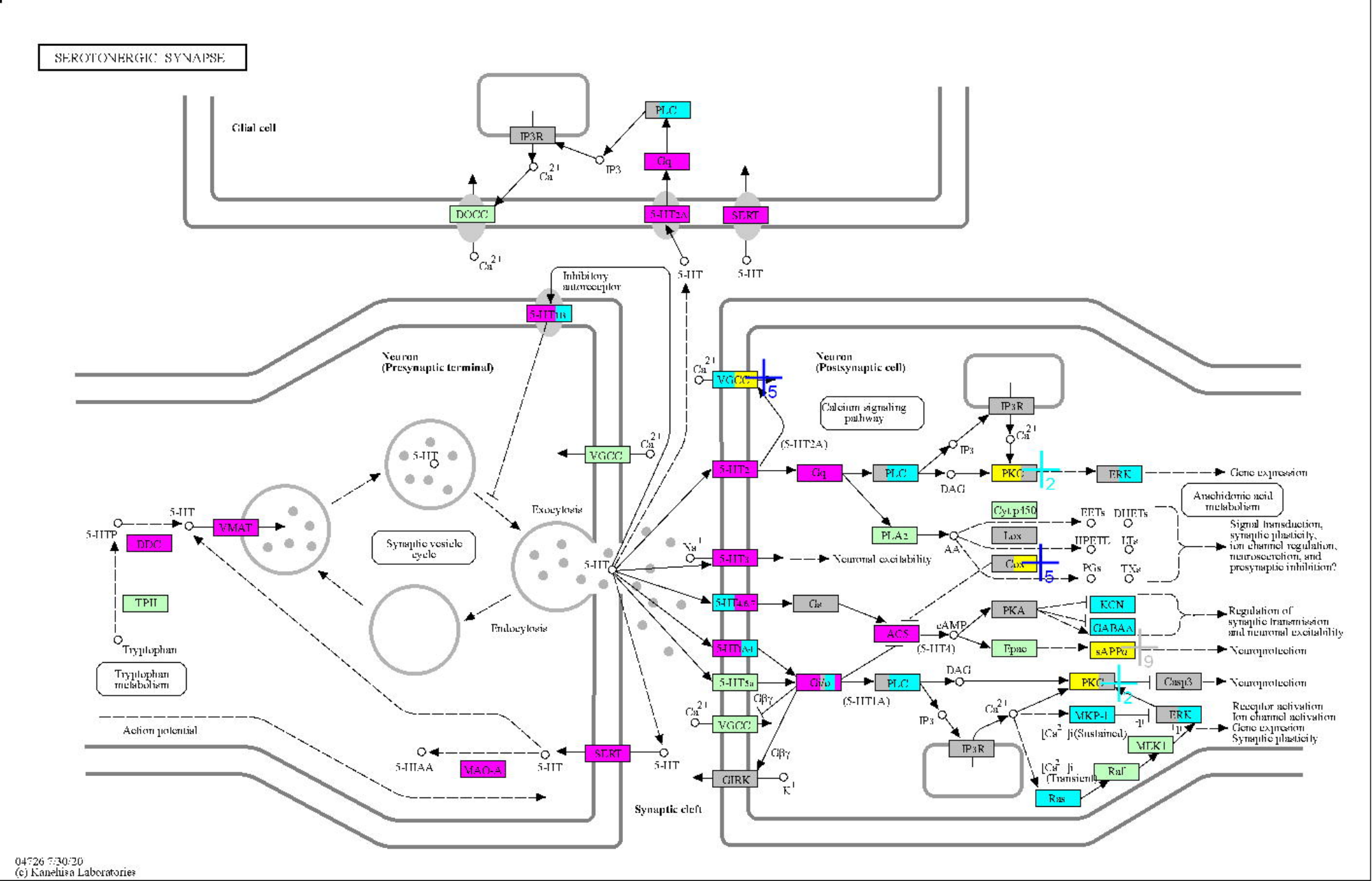
GLUTAMATERGIC SYNAPSE



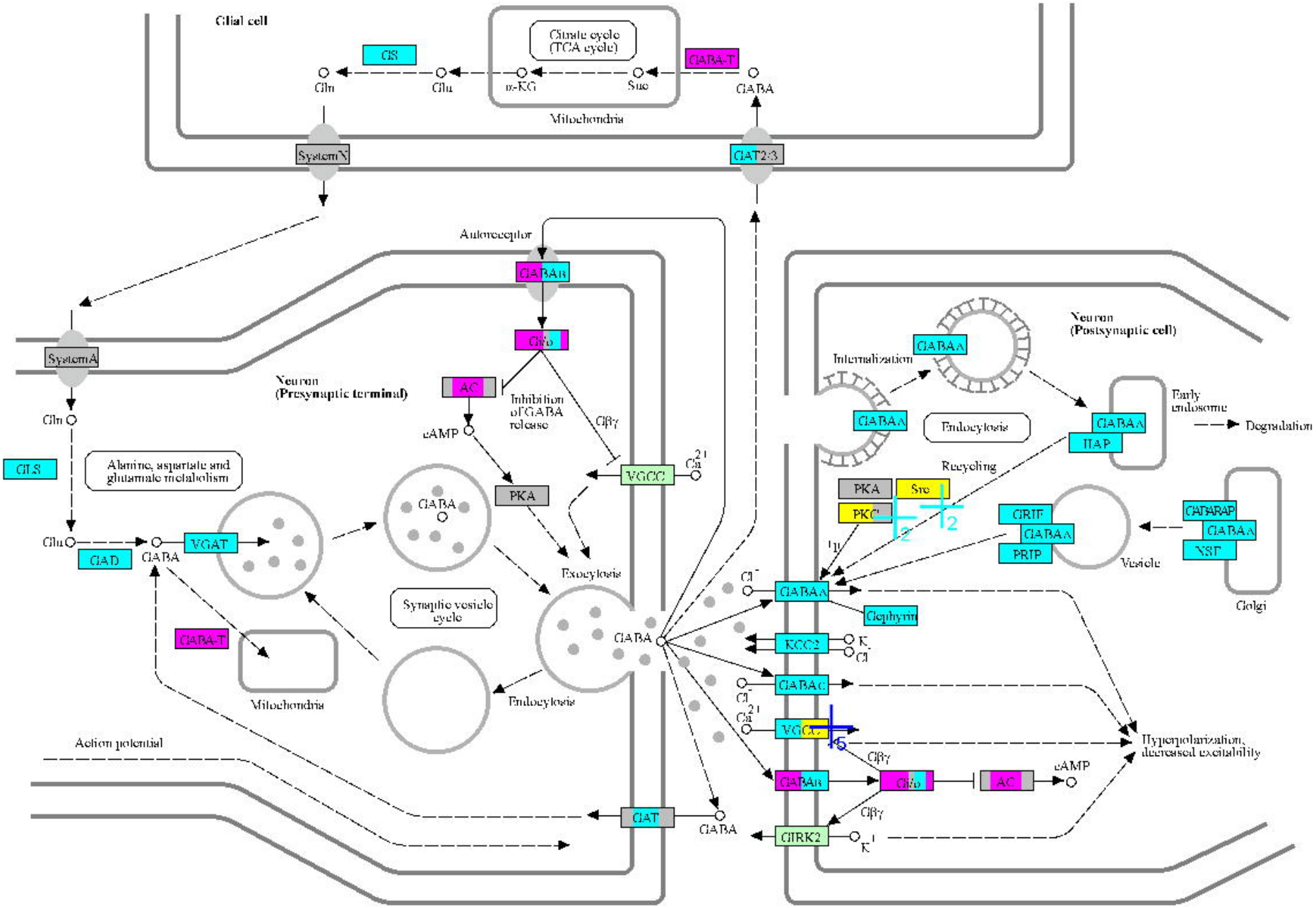
04724 4/20/23
(c) Kanehisa Laboratories



04725 2/15/17
(c) Kanehisa Laboratories

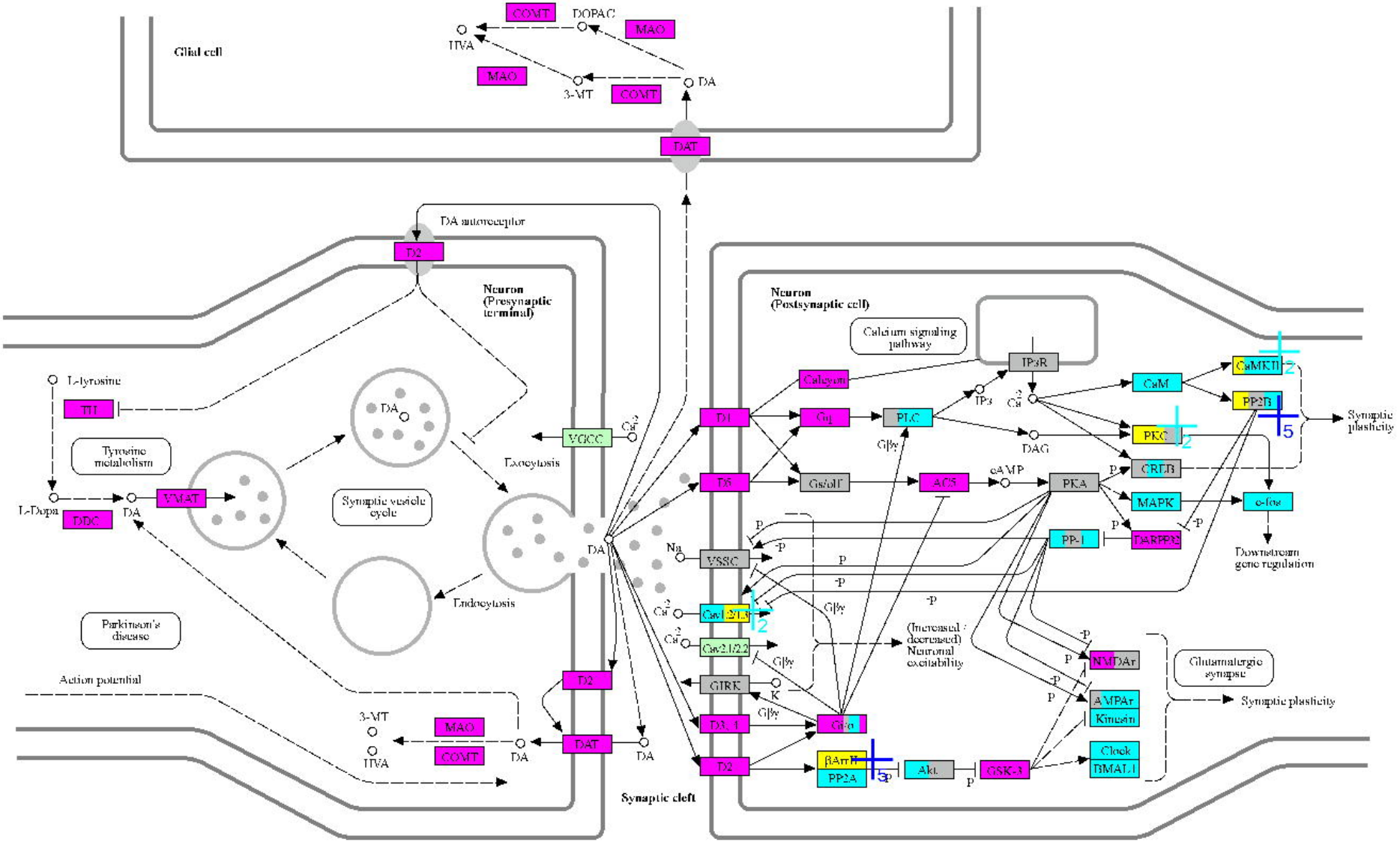

SEROTONERGIC SYNAPSE
Glial cell
IP3R
PLC
IP3
Gq
DOCC
5-HT2A
SERT
5-HT
Inhibitory autoreceptor
5-HT1B
Neuron (Presynaptic terminal)
Neuron (Postsynaptic cell)
VGCC
Calcium signaling pathway
(5-HT2A)
5-HT2
PKC
ERK
Gene expression
DAG
Arachidonic acid metabolism
Cyt.p450
EETs
DHETs
Lox
HPETE
LTs
Cox
PGs
TXs
PLA2
AA
Signal transduction, synaptic plasticity, ion channel regulation, neurosecretion, and presynaptic inhibition?
Exocytosis
Synaptic vesicle cycle
Endocytosis
5-HTP
DDC
VMAT
TPH
Tryptophan
Tryptophan metabolism
5-HT3
Neuronal excitability
5-HT4,6,7
Gs
AC5
cAMP
(5-HT4)
PKA
KCN
GABAA
Regulation of synaptic transmission and neuronal excitability
Epac
sAPPα
Neuroprotection
5-HT1A-1
5-HT5a
Gi/o
(5-HT1A)
Casp3
Gβγ
MKP-1
Receptor activation
Ion channel activation
Gene expression
Synaptic plasticity
MEK1
Raf
Ras
Action potential
5-HIAA
MAO-A
GIRK
Synaptic cleft
04726 7/30/20
(c) Kanehisa Laboratories

GABAERGIC SYNAPSE



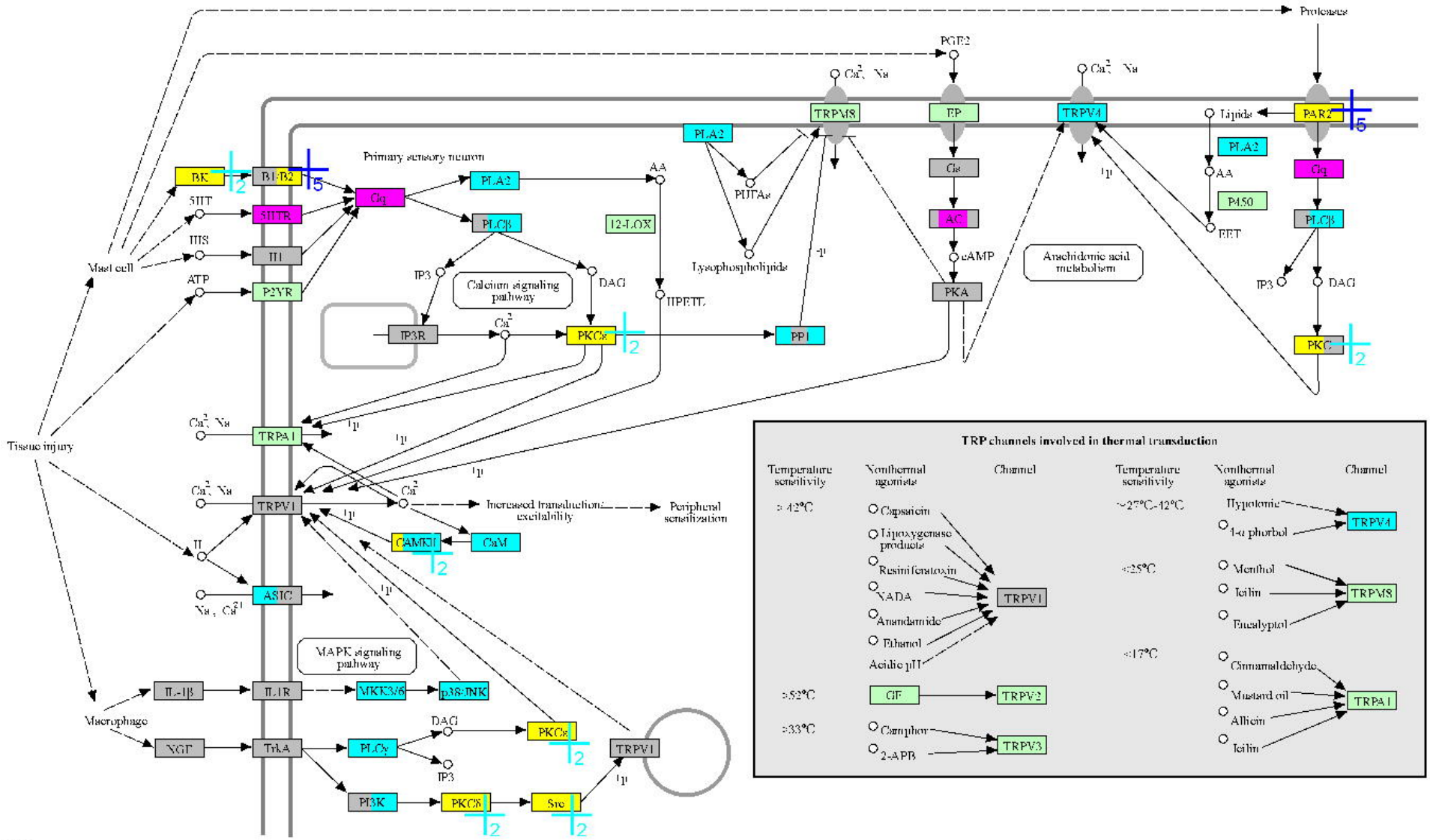
04/27/ 1:20:20
(c) Kanehisa Laboratories

DOPAMINERGIC SYNAPSE

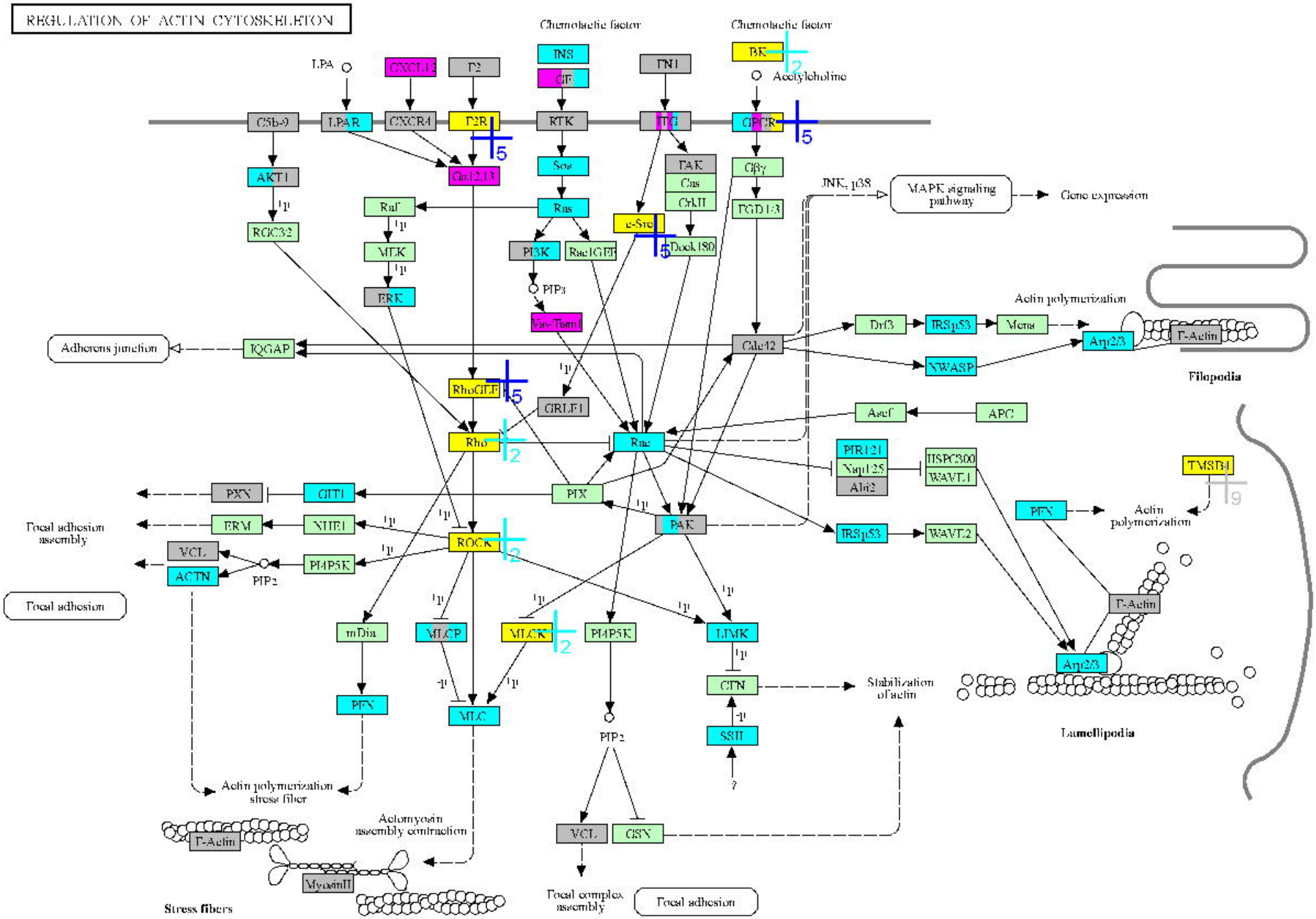


04728 1/21/20
(c) Kanehisa Laboratories

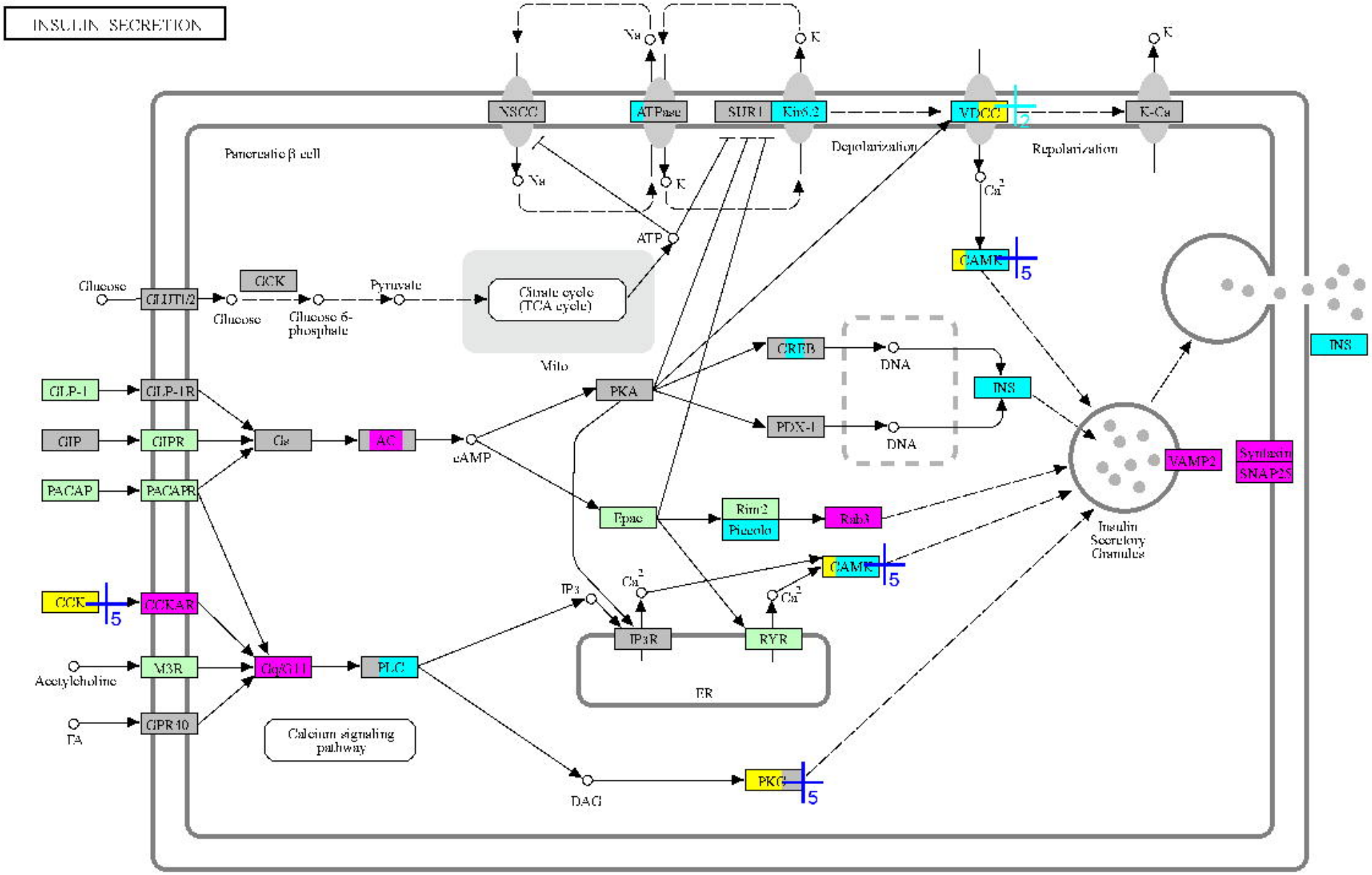
INFLAMMATORY MEDIATOR REGULATION OF TRP CHANNELS



04750 2/1/19
(c) Kanehisa Laboratories

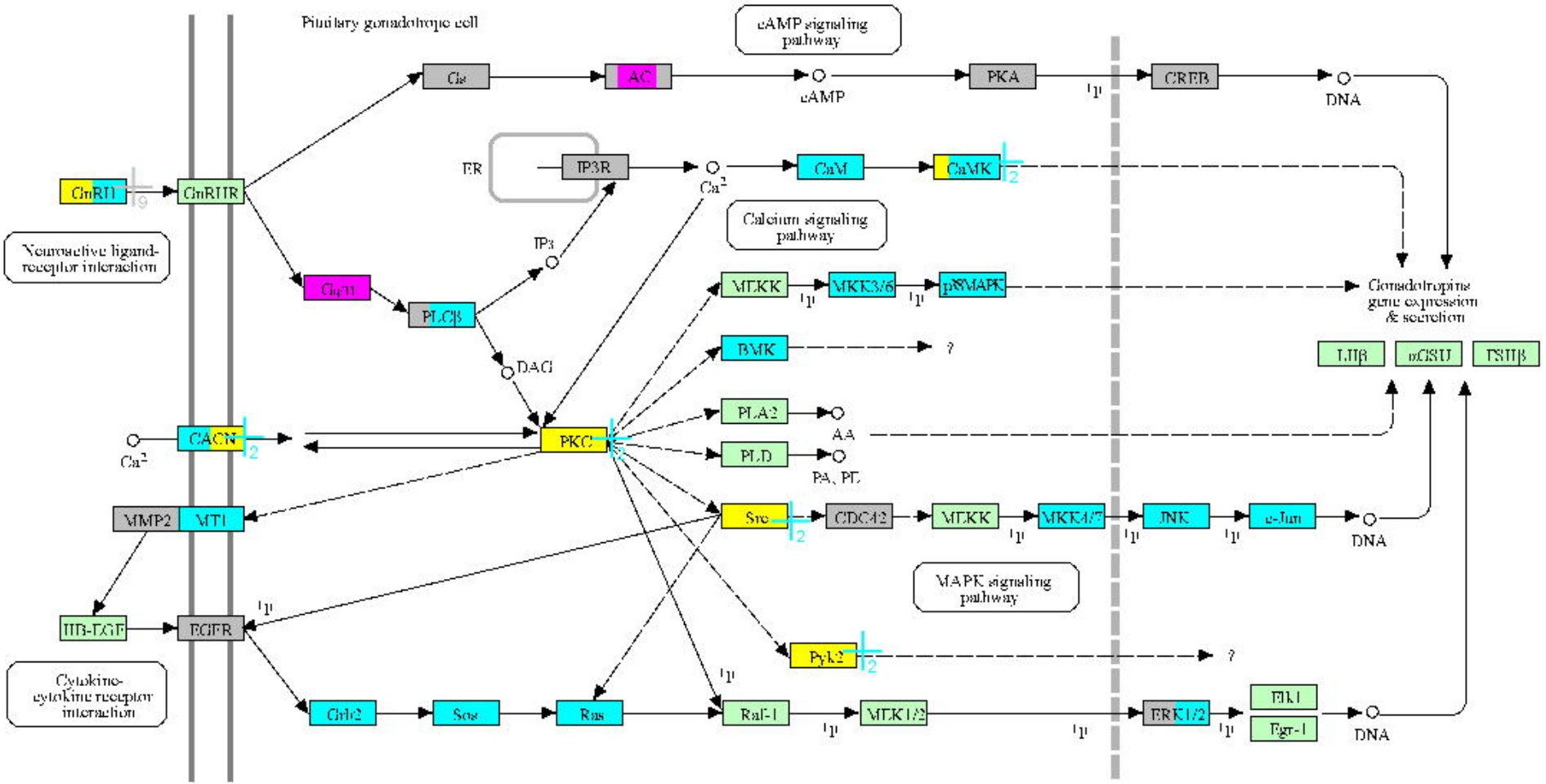


04810 10/18/22
(c) Kanehisa Laboratories



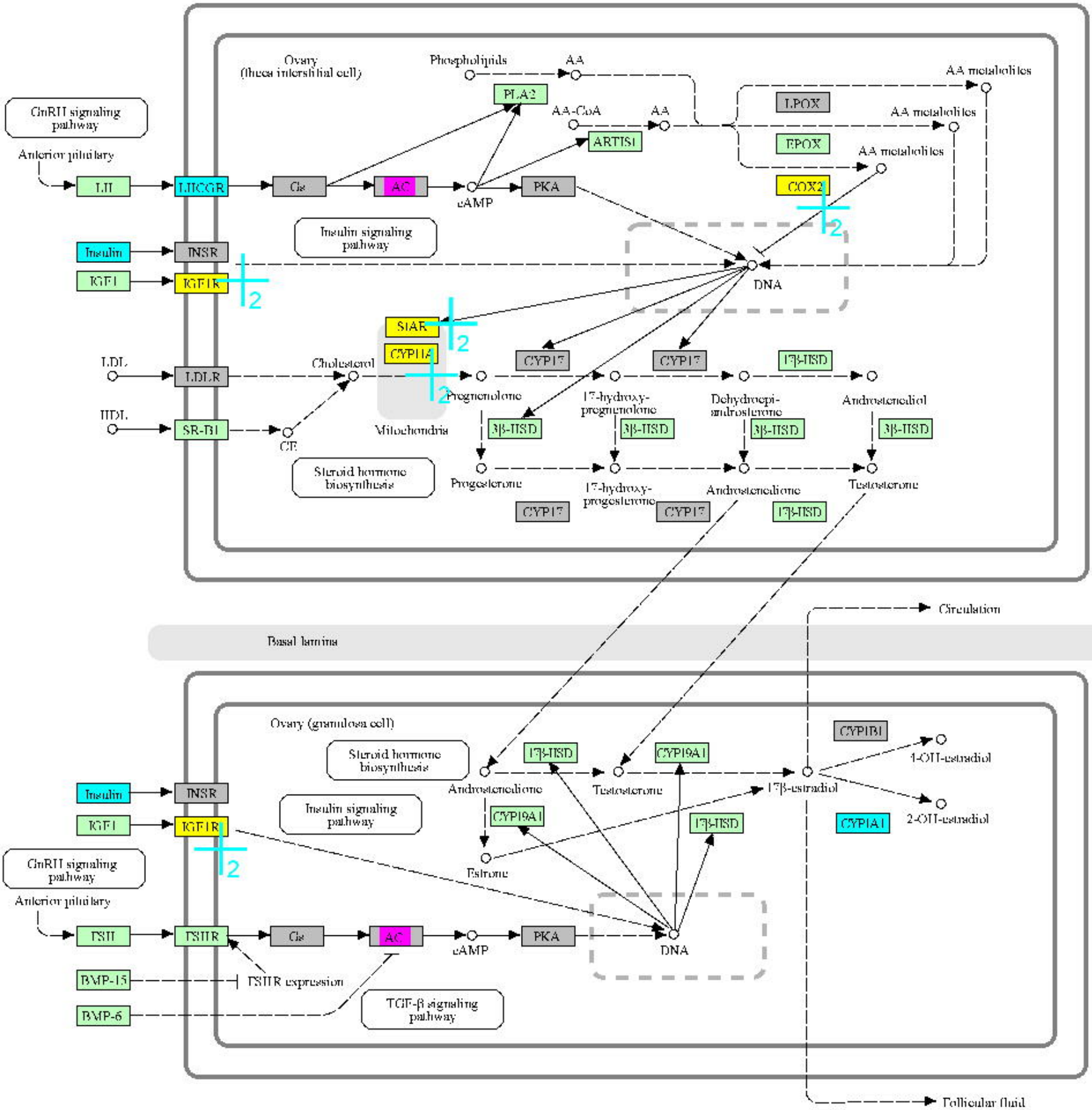
04911 8/5/21
(c) Kanehisa Laboratories

GnRH SIGNALING PATHWAY

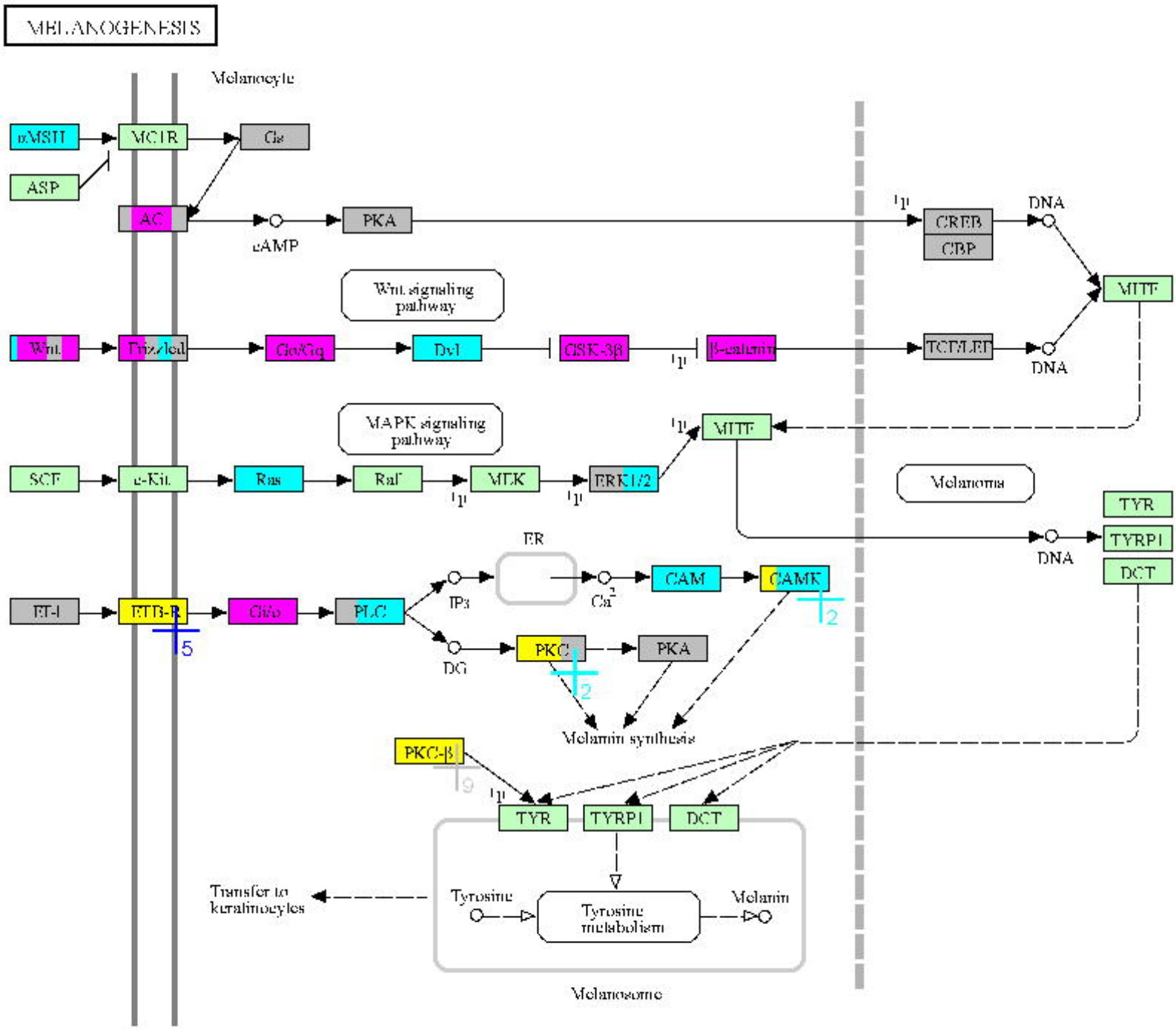


04912 4/6/18
(c) Kanehisa Laboratories

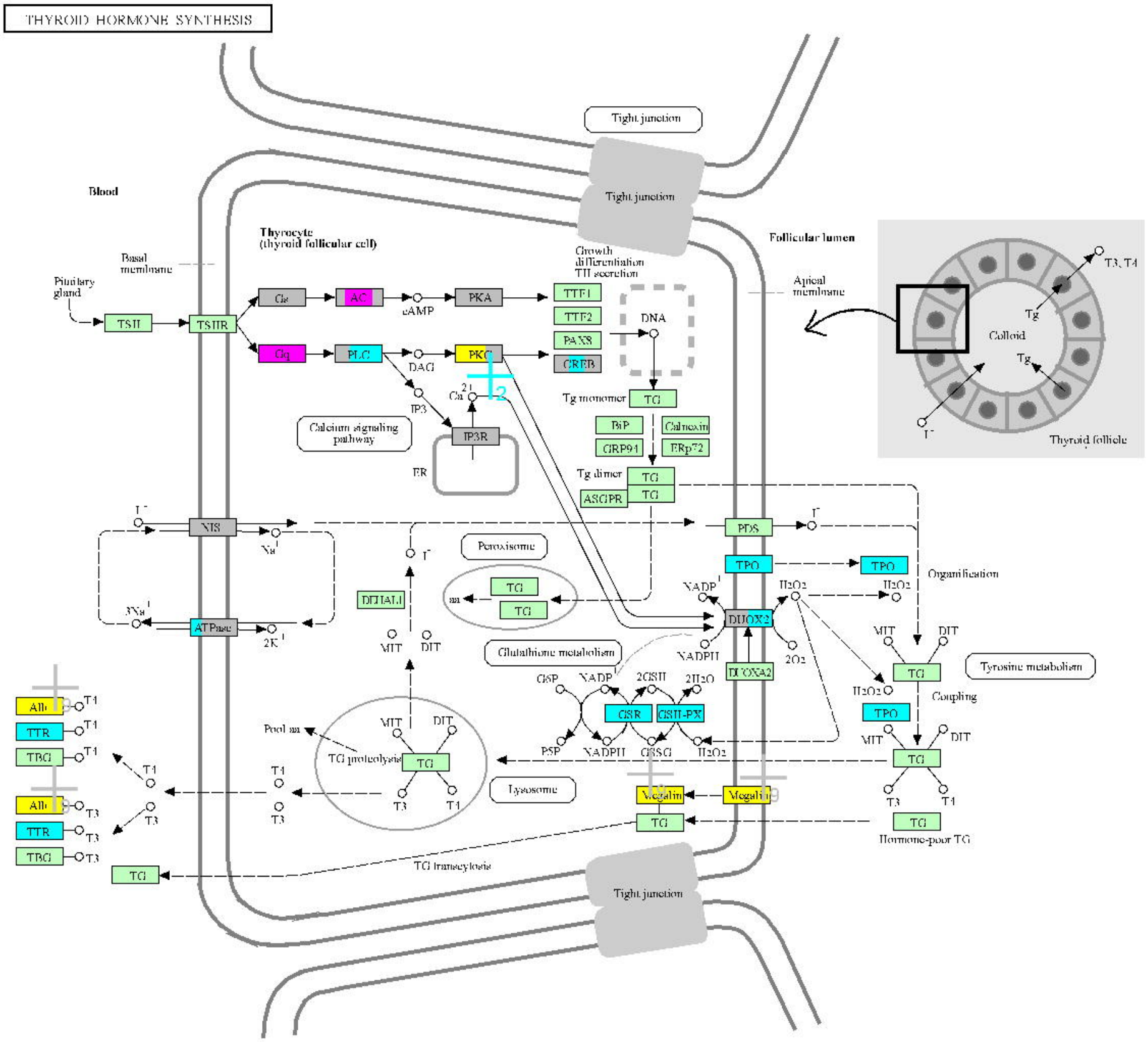
OVARIAN STEROIDOGENESIS



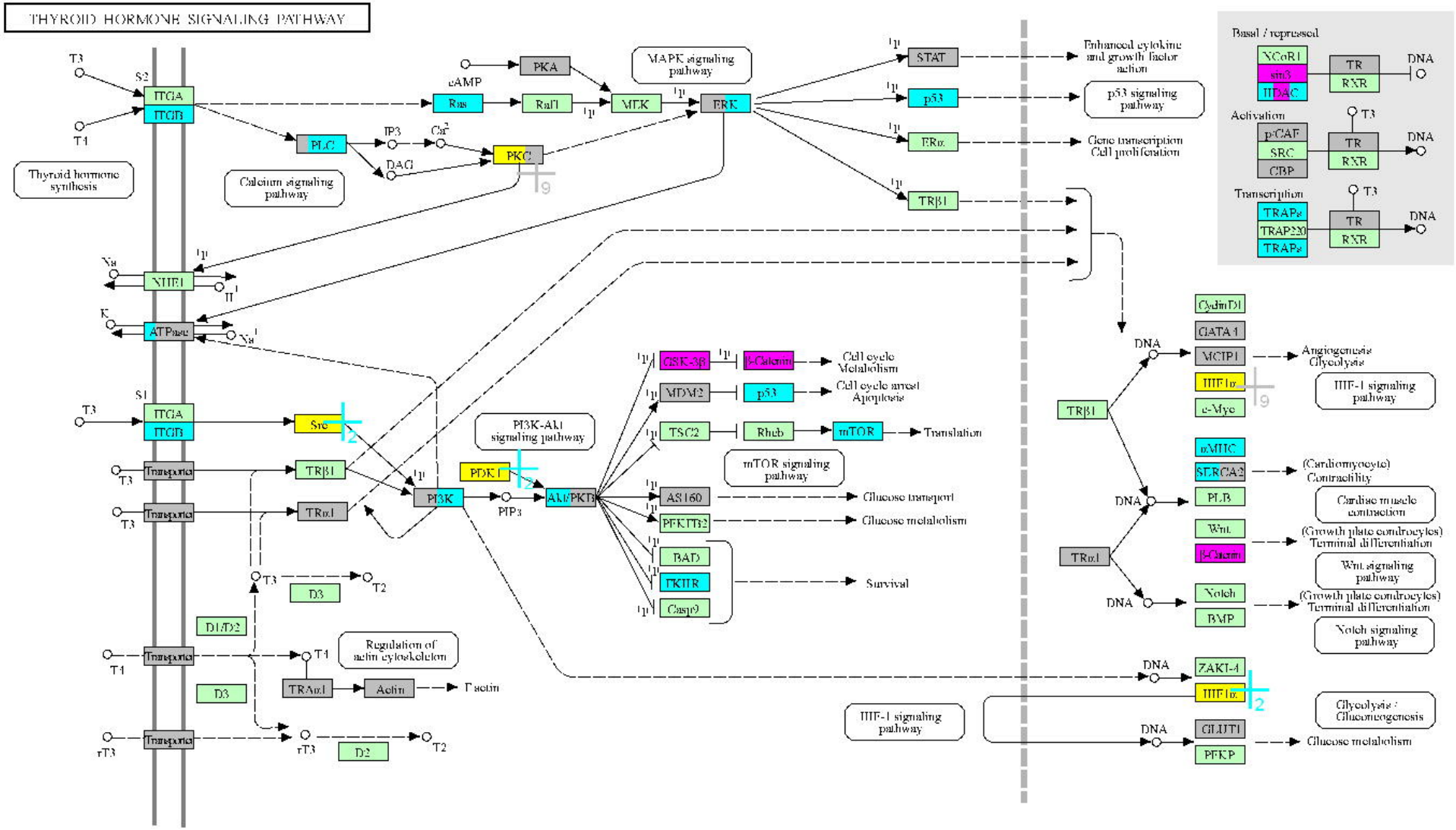
04913 1/23/20
(c) Kanehisa Laboratories



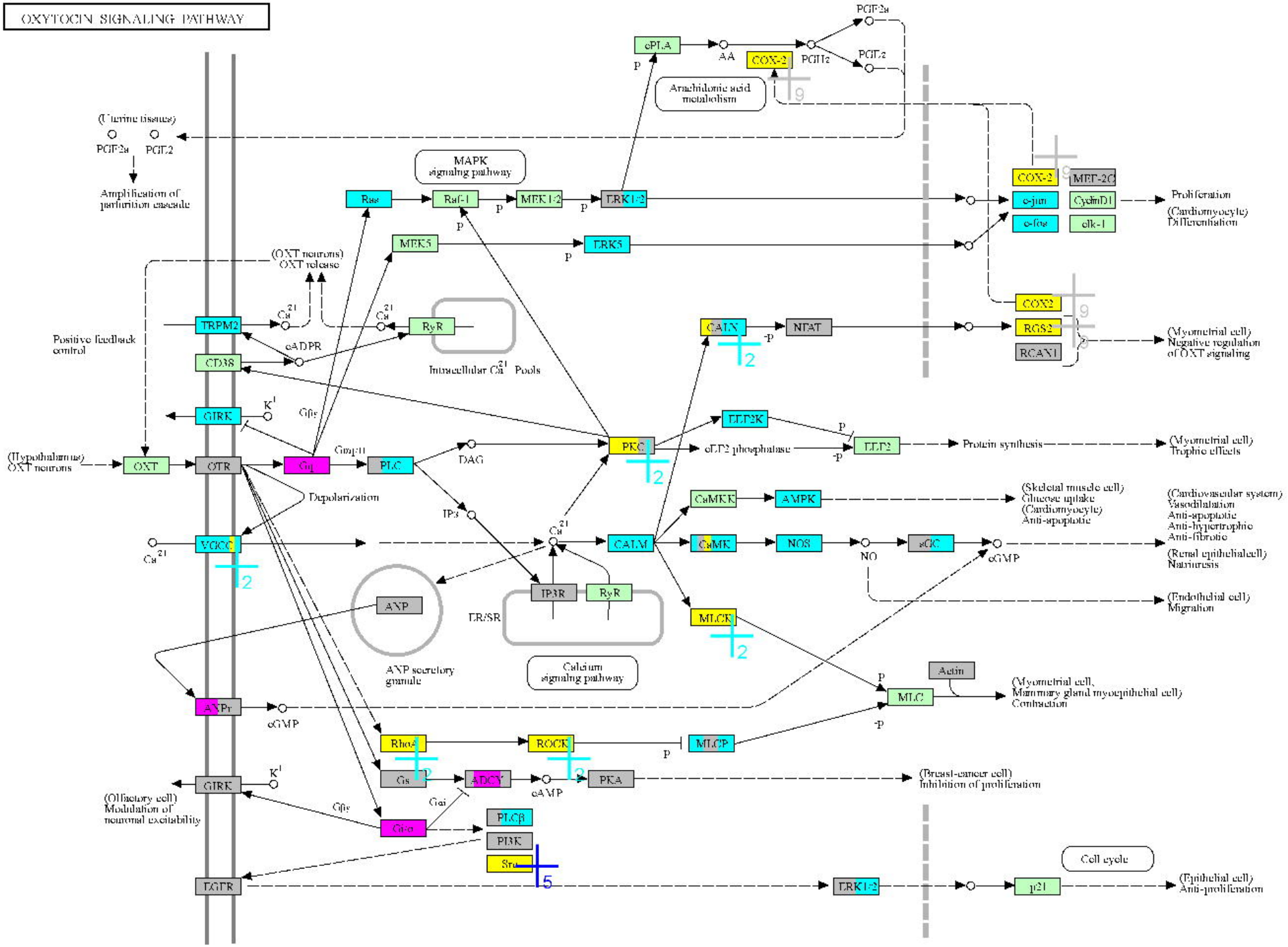
04916 7/7/21
(c) Kanehisa Laboratories



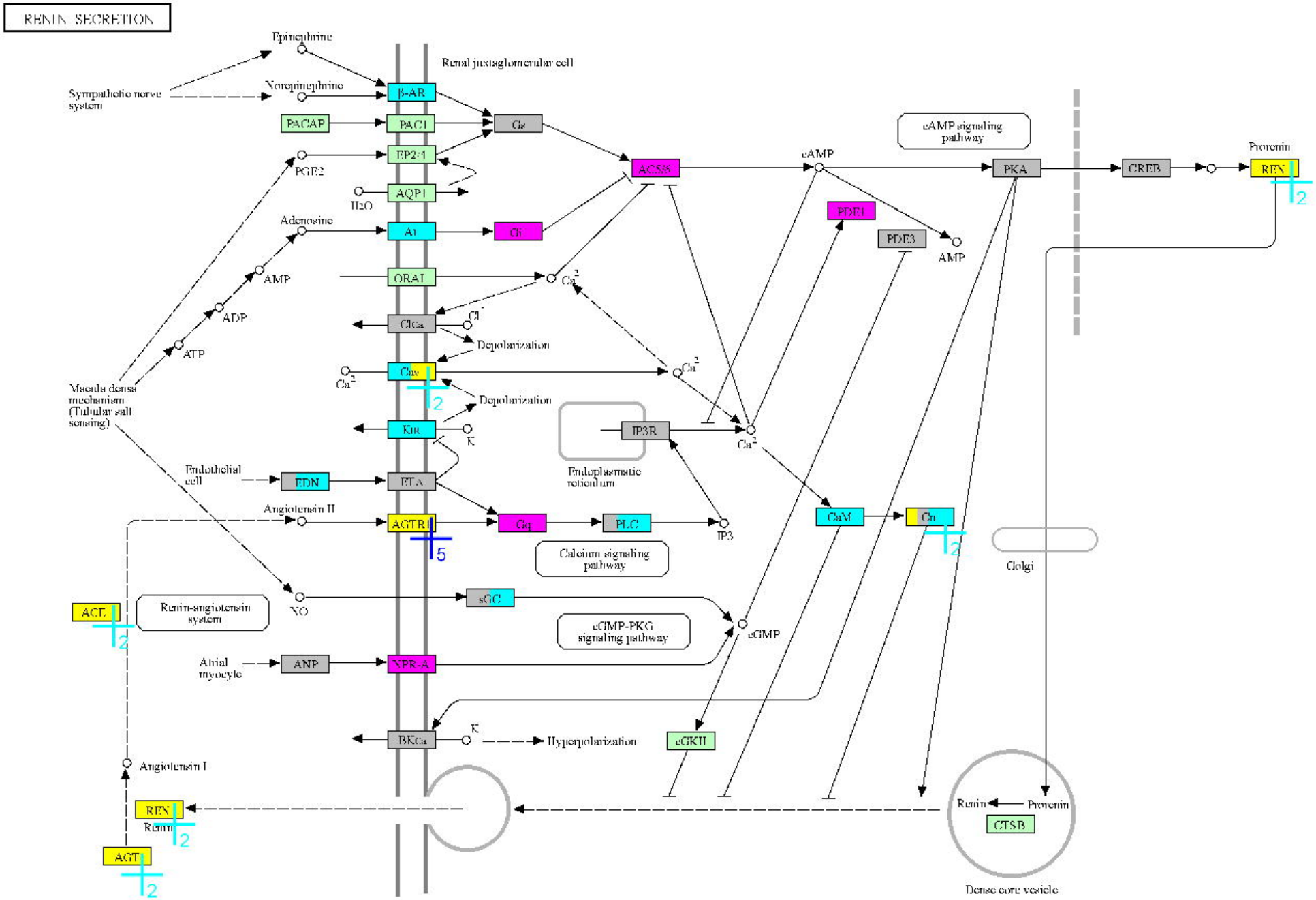
04918 5/18/20
(c) Kanehisa Laboratories



04919 8/5/21
(c) Kanehisa Laboratories

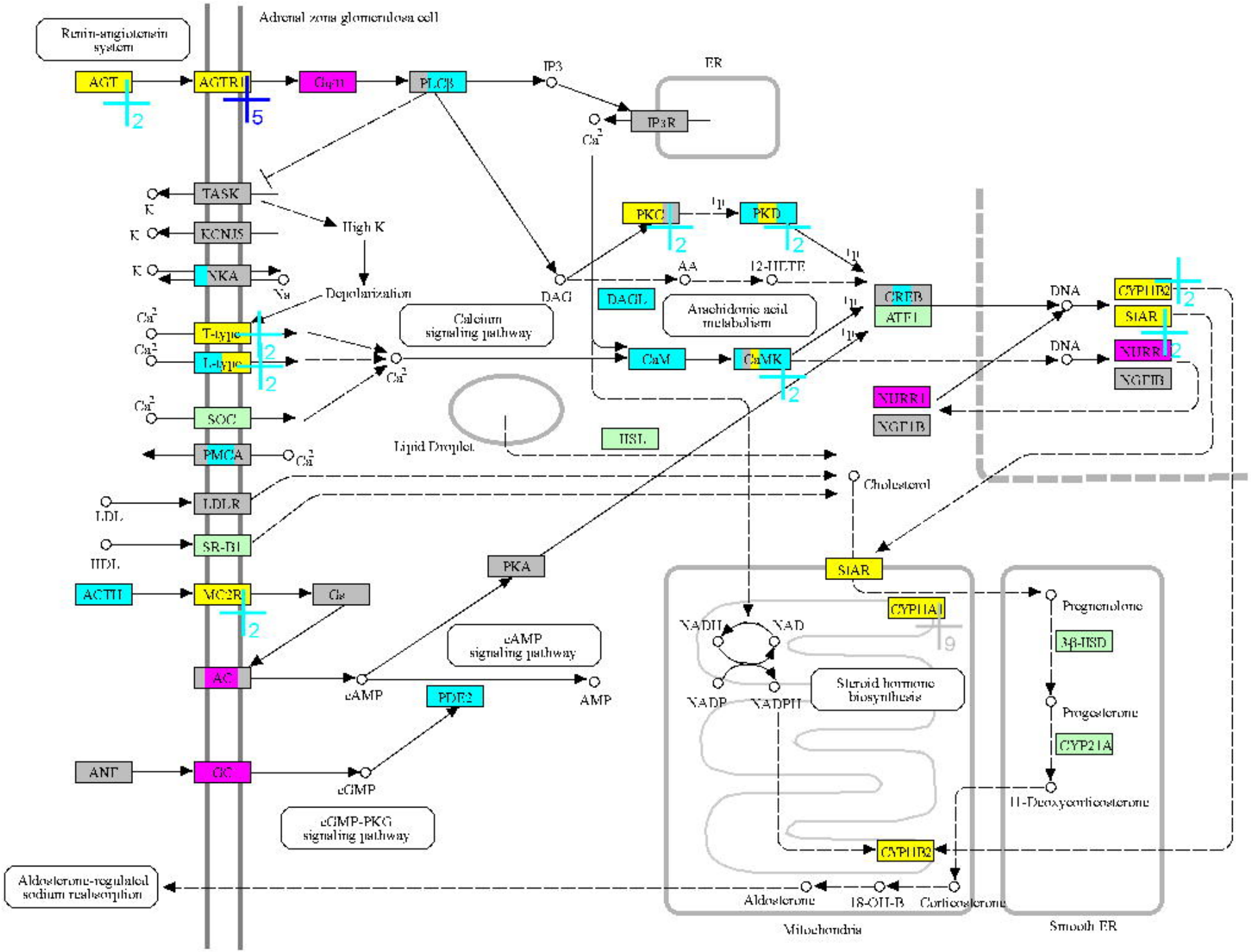


04921 2/5/19
(c) Kanehisa Laboratories



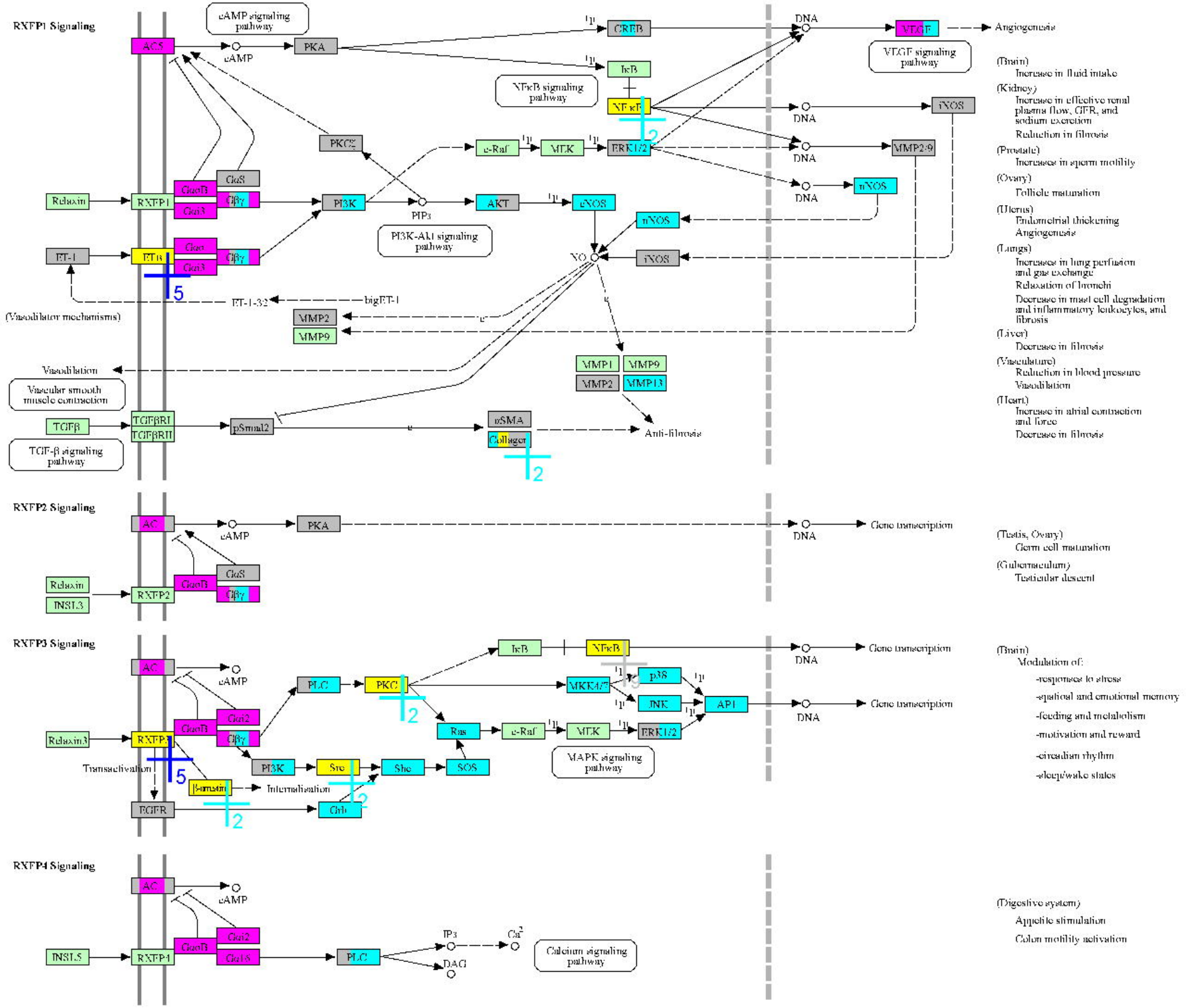
04924 9/26/23
(c) Kanehisa Laboratories

ALDOSTERONE SYNTHESIS AND SECRETION



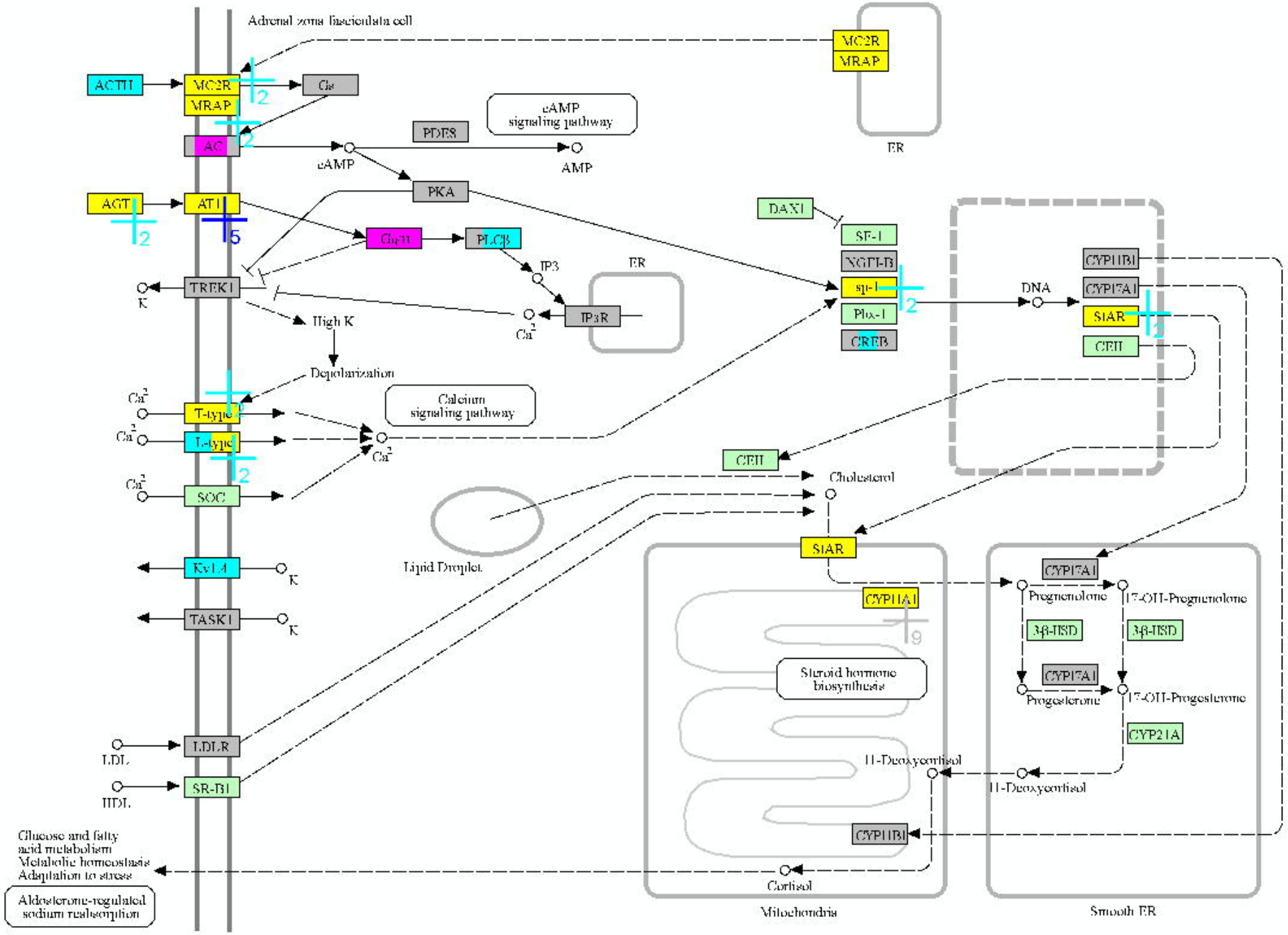
04925 9/26/23
(c) Kanehisa Laboratories

RELAXIN SIGNALING PATHWAY



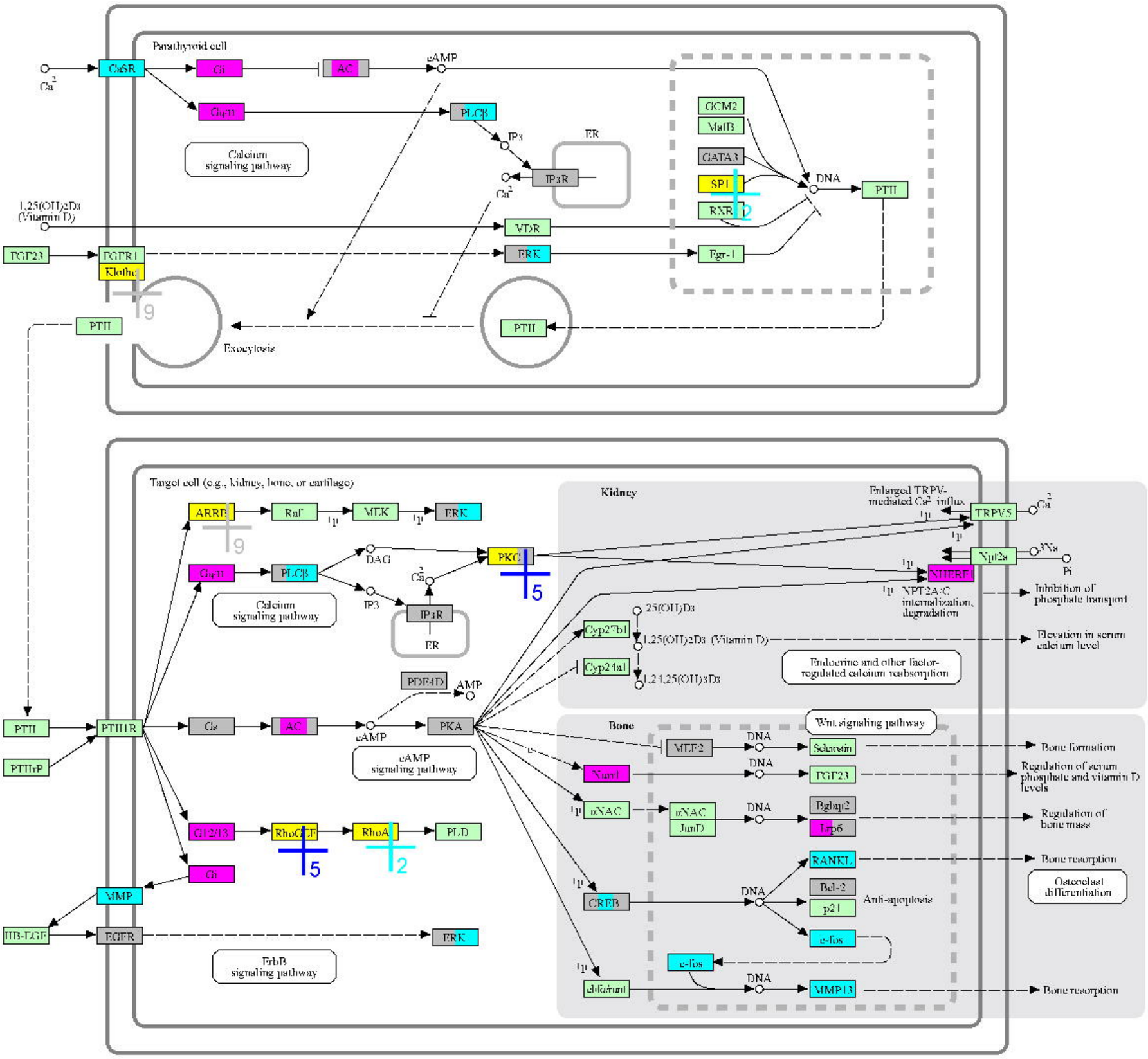
04926 7/7/21
(c) Kanehisa Laboratories

CORTISOL SYNTHESIS AND SECRETION

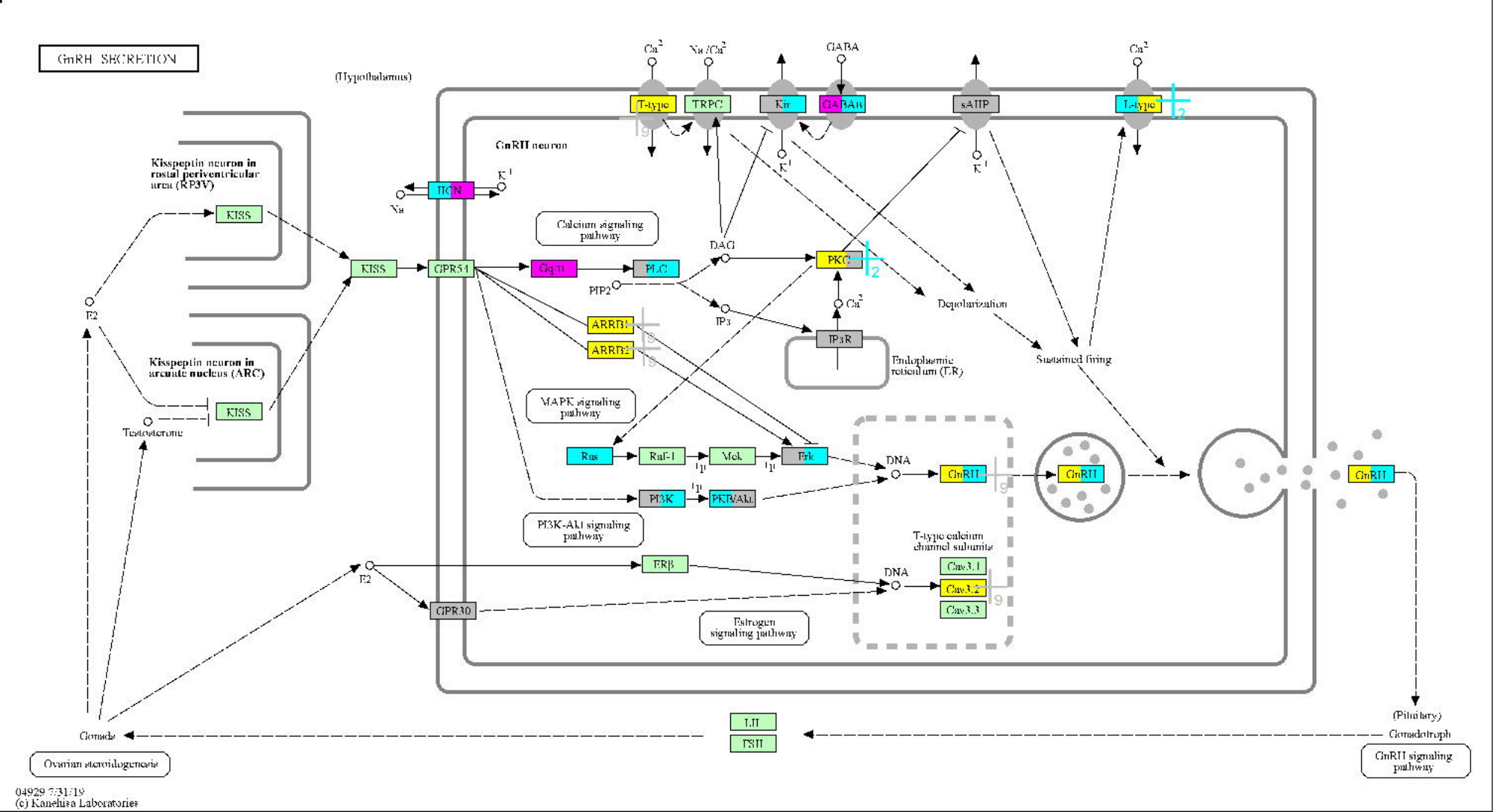


04927 11/25/20
(c) Kanehisa Laboratories

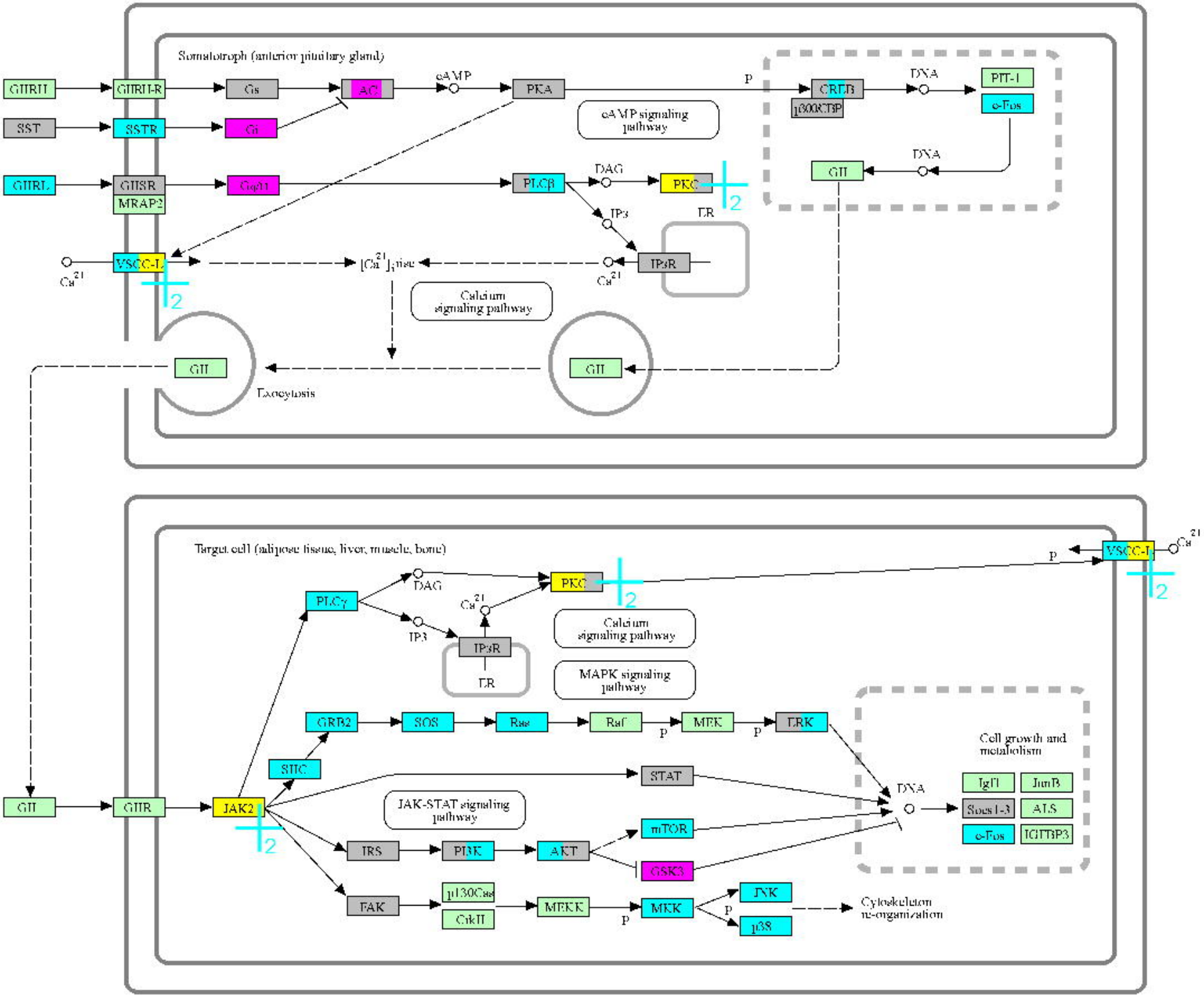
PARATHYROID HORMONE SYNTHESIS, SECRETION AND ACTION



04928 4/17/18
(c) Kanehisa Laboratories


GnRH SECRETION
(Hypothalamus)
Kisspeptin neuron in rostal periventricular area (RP3V)
KISS
Kisspeptin neuron in arcuate nucleus (ARC)
KISS
E2
Testosterone
HCN
Na
K
GnRH neuron
KISS
GPR54
Gq/11
PLC
PIP2
DAG
IP3
PKC
Ca2
IP3R
Endoplasmic reticulum (ER)
ARRB1
ARRB2
Calcium signaling pathway
MAPK signaling pathway
Ras
Raf-1
Mek
Erk
PI3K
PKB/Akt
PI3K-Akt signaling pathway
ERβ
GPR30
E2
Estrogen signaling pathway
Ca2
Na /Ca2
GABA
T-type
TRPC
Kir
GABAB
sAHP
L-type
Ca2
Depolarization
Sustained firing
DNA
GnRH
GnRH
GnRH
T-type calcium channel subunits
DNA
Cav3.1
Cav3.2
Cav3.3
LH
FSH
Gonads
Ovarian steroidogenesis
(Pituitary)
Gonadotroph
GnRH signaling pathway
04929 7/31/19
(c) Kanehisa Laboratories

GROWTH HORMONE SYNTHESIS, SECRETION AND ACTION



04935 8/4/22
(c) Kanehisa Laboratories