Pathway Analysis Report

This report contains the pathway analysis results for the submitted sample ''. Analysis was performed against Reactome version 76 on 05/06/2021. The web link to these results is:

https://reactome.org/PathwayBrowser/#/ANALYSIS=MjAyMTA2MDQxNjU0MzBfNzE2MjI%3D

Please keep in mind that analysis results are temporarily stored on our server. The storage period depends on usage of the service but is at least 7 days. As a result, please note that this URL is only valid for a limited time period and it might have expired.

Table of Contents

1. Introduction

Reactome is a curated database of pathways and reactions in human biology. Reactions can be considered as pathway 'steps'. Reactome defines a 'reaction' as any event in biology that changes the state of a biological molecule. Binding, activation, translocation, degradation and classical biochemical events involving a catalyst are all reactions. Information in the database is authored by expert biologists, entered and maintained by Reactome's team of curators and editorial staff. Reactome content frequently cross-references other resources e.g. NCBI, Ensembl, UniProt, KEGG (Gene and Compound), ChEBI, PubMed and GO. Orthologous reactions inferred from annotation for Homo sapiens are available for 17 non-human species including mouse, rat, chicken, puffer fish, worm, fly, yeast, rice, and Arabidopsis. Pathways are represented by simple diagrams following an SBGN-like format.

Reactome's annotated data describe reactions possible if all annotated proteins and small molecules were present and active simultaneously in a cell. By overlaying an experimental dataset on these annotations, a user can perform a pathway over-representation analysis. By overlaying quantitative expression data or time series, a user can visualize the extent of change in affected pathways and its progression. A binomial test is used to calculate the probability shown for each result, and the p-values are corrected for the multiple testing (Benjamini–Hochberg procedure) that arises from evaluating the submitted list of identifiers against every pathway.

To learn more about our Pathway Analysis, please have a look at our relevant publications:

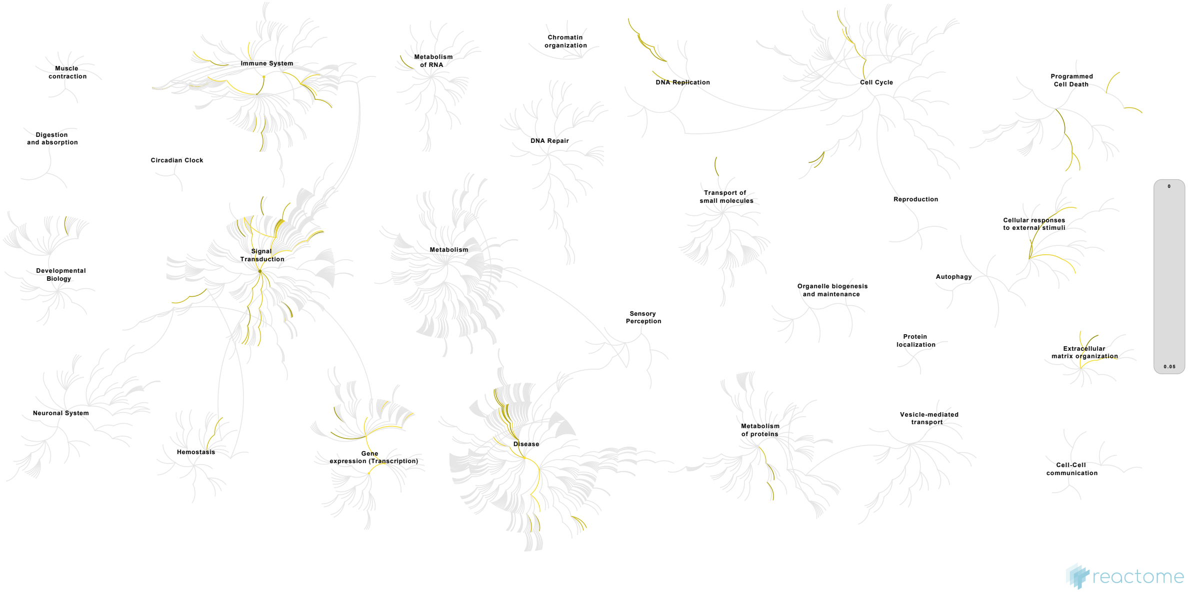
Fabregat A, Sidiropoulos K, Garapati P, Gillespie M, Hausmann K, Haw R, … D'Eustachio P (2016). The reactome pathway knowledgebase. Nucleic Acids Research, 44(D1), D481–D487. https://doi.org/10.1093/nar/gkv1351.

Fabregat A, Sidiropoulos K, Viteri G, Forner O, Marin-Garcia P, Arnau V, … Hermjakob H (2017). Reactome pathway analysis: a high-performance in-memory approach. BMC Bioinformatics, 18.

2. Properties

- This is an **overrepresentation** analysis: A statistical (hypergeometric distribution) test that determines whether certain Reactome pathways are over-represented (enriched) in the submitted data. It answers the question ‘Does my list contain more proteins for pathway X than would be expected by chance?’ This test produces a probability score, which is corrected for false discovery rate using the Benjamani-Hochberg method.
- 13 out of 13 identifiers in the sample were found in Reactome, where 256 pathways were hit by at least one of them.
- All non-human identifiers have been converted to their human equivalent.
- This report is filtered to show only results for species 'Homo sapiens' and resource 'all resources'.
- The unique ID for this analysis (token) is MjAyMTA2MDQxNjU0MzBfNzE2MjI%3D. This ID is valid for at least 7 days in Reactome’s server. Use it to access Reactome services with your data.

3. Genome-wide overview



This figure shows a genome-wide overview of the results of your pathway analysis. Reactome pathways are arranged in a hierarchy. The center of each of the circular "bursts" is the root of one top-level pathway, for example "DNA Repair". Each step away from the center represents the next level lower in the pathway hierarchy. The color code denotes over-representation of that pathway in your input dataset. Light grey signifies pathways which are not significantly over-represented.

4. Most significant pathways

The following table shows the 25 most relevant pathways sorted by p-value.

Pathway name	Entities				Reactions	
	found	ratio	p-value	FDR*	found	ratio
TFAP2 (AP-2) family regulates transcription of growth factors and their receptors	4 / 21	0.001	2.57e-08	6.88e-06	4 / 18	0.001
Potential therapeutics for SARS	5 / 84	0.006	1.22e-07	1.63e-05	6 / 32	0.002
SARS-CoV Infections	6 / 203	0.014	3.38e-07	3.01e-05	9 / 254	0.019
Transcriptional regulation by the AP-2 (TFAP2) family of transcription factors	4 / 52	0.004	9.37e-07	6.28e-05	4 / 44	0.003
Trafficking and processing of endosomal TLR	3 / 16	0.001	1.75e-06	9.29e-05	2 / 7	5.21e-04
Infectious disease	9 / 1,343	0.092	5.16e-05	0.002	16 / 750	0.056
VEGF ligand-receptor interactions	2 / 8	5.51e-04	6.33e-05	0.002	3 / 4	2.98e-04
VEGF binds to VEGFR leading to receptor dimerization	2 / 8	5.51e-04	6.33e-05	0.002	2 / 3	2.23e-04
RUNX1 regulates transcription of genes involved in differentiation of keratinocytes	2 / 11	7.57e-04	1.19e-04	0.003	1 / 8	5.96e-04
Assembly of collagen fibrils and other multimeric structures	3 / 67	0.005	1.23e-04	0.003	1 / 26	0.002
Toll-like Receptor Cascades	4 / 188	0.013	1.41e-04	0.003	14 / 185	0.014
Disease	11 / 2,360	0.163	1.48e-04	0.003	99 / 1,591	0.119
Generic Transcription Pathway	9 / 1,555	0.107	1.63e-04	0.003	17 / 824	0.061
Interleukin-4 and Interleukin-13 signaling	4 / 211	0.015	2.19e-04	0.004	2 / 47	0.004
Regulation of gene expression by Hypoxia-inducible Factor	2 / 15	0.001	2.21e-04	0.004	1 / 7	5.21e-04
RNA Polymerase II Transcription	9 / 1,694	0.117	3.14e-04	0.004	17 / 885	0.066
ERBB2 Activates PTK6 Signaling	2 / 18	0.001	3.18e-04	0.004	2 / 2	1.49e-04
ERBB2 Regulates Cell Motility	2 / 19	0.001	3.54e-04	0.005	2 / 2	1.49e-04
Collagen formation	3 / 104	0.007	4.44e-04	0.006	1 / 77	0.006
PI3K events in ERBB2 signaling	2 / 22	0.002	4.73e-04	0.006	5 / 7	5.21e-04
Gene expression (Transcription)	9 / 1,855	0.128	6.22e-04	0.007	23 / 1,000	0.074
Signaling by VEGF	3 / 137	0.009	9.83e-04	0.011	49 / 86	0.006
Signaling by ERBB2 KD Mutants	2 / 35	0.002	0.001	0.011	15 / 17	0.001
MHC class II antigen presentation	3 / 148	0.01	0.001	0.011	3 / 26	0.002

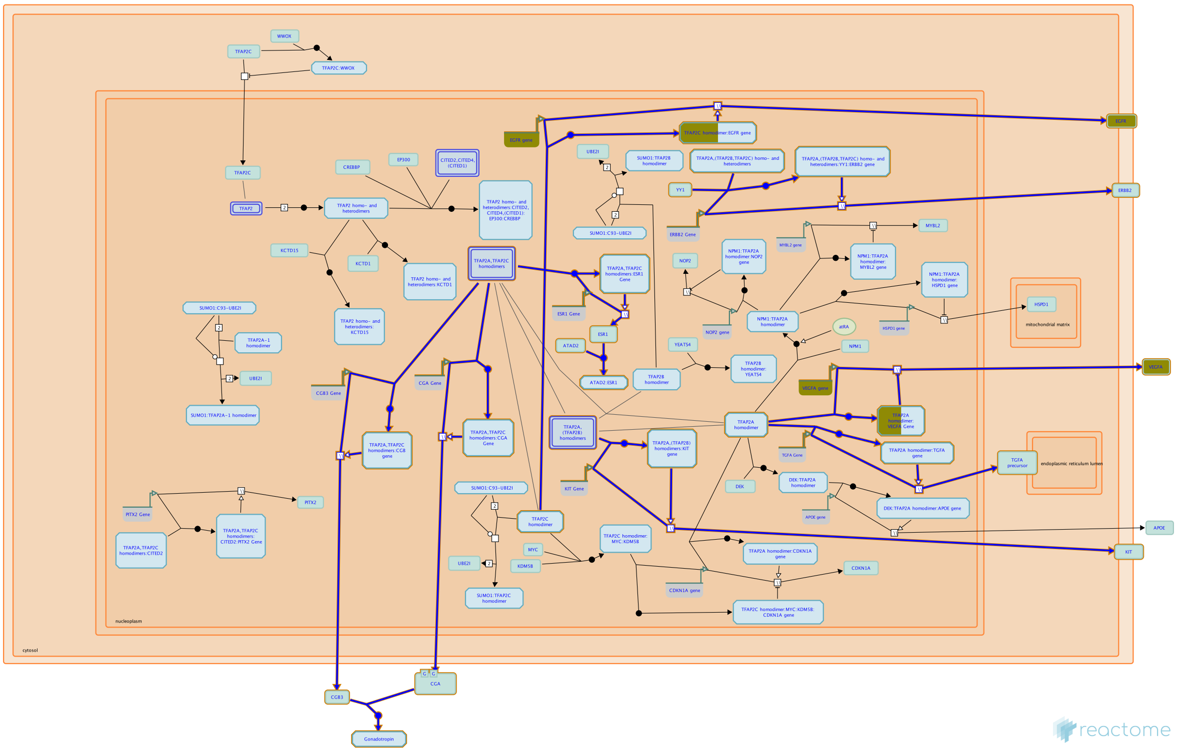
Pathway name	Entities				Reactions	
	found	ratio	p-value	FDR*	found	ratio
Degradation of the extracellular matrix	3 / 148	0.01	0.001	0.011	6 / 105	0.008

* False Discovery Rate

5. Pathways details

For every pathway of the most significant pathways, we present its diagram, as well as a short summary, its bibliography and the list of inputs found in it.

1. TFAP2 (AP-2) family regulates transcription of growth factors and their receptors (R-HSA-8866910)



TFAP2A and TFAP2C directly stimulate transcription of the estrogen receptor ESR1 gene (McPherson and Weigel 1999). TFAP2A expression correlates with ESR1 expression in breast cancer, and TFAP2C is frequently overexpressed in estrogen-positive breast cancer and endometrial cancer (deConinck et al. 1995, Turner et al. 1998). TFAP2A, TFAP2C, as well as TFAP2B can directly stimulate the expression of ERBB2, another important breast cancer gene (Bosher et al. 1996). Association of TFAP2A with the YY1 transcription factor significantly increases the ERBB2 transcription rate (Begon et al. 2005). In addition to ERBB2, the expression of another receptor tyrosine kinase, KIT, is also stimulated by TFAP2A and TFAP2B (Huang et al. 1998), while the expression of the VEGF receptor tyrosine kinase ligand VEGFA is repressed by TFAP2A (Ruiz et al. 2004, Li et al. 2012). TFAP2A stimulates transcription of the transforming growth factor alpha (TGFA) gene (Wang et al. 1997). TFAP2C regulates EGFR expression in luminal breast cancer (De Andrade et al. 2016). In placenta, TFAP2A and TFAP2C directly stimulate transcription of both subunits of the human chorionic gonadotropin, CGA and CGB (Johnson et al. 1997, LiCalsi et al. 2000).

References

McPherson LA & Weigel RJ (1999). AP2alpha and AP2gamma: a comparison of binding site specificity and trans-activation of the estrogen receptor promoter and single site promoter constructs. Nucleic Acids Res., 27, 4040-9.

deConinck EC, McPherson LA & Weigel RJ (1995). Transcriptional regulation of estrogen receptor in breast carcinomas. Mol. Cell. Biol., 15, 2191-6.

Turner BC, Zhang J, Gumbs AA, Maher MG, Kaplan L, Carter D, ... Williams T (1998). Expression of AP-2 transcription factors in human breast cancer correlates with the regulation of multiple growth factor signalling pathways. Cancer Res., 58, 5466-72.

Johnson W, Albanese C, Handwerger S, Williams T, Pestell RG & Jameson JL (1997). Regulation of the human chorionic gonadotropin alpha- and beta-subunit promoters by AP-2. J. Biol. Chem., 272, 15405-12.

LiCalsi C, Christophe S, Steger DJ, Buescher M, Fischer W & Mellon PL (2000). AP-2 family members regulate basal and cAMP-induced expression of human chorionic gonadotropin. Nucleic Acids Res., 28, 1036-43.

Edit history

Date	Action	Author
2016-03-14	Edited	Orlic-Milacic M
2016-03-14	Authored	Orlic-Milacic M
2016-04-04	Created	Orlic-Milacic M
2016-05-04	Reviewed	Dawid IB, Zarelli VE
2016-05-17	Reviewed	Bogachek MV, Weigel RJ
2021-03-05	Modified	Shorser S

Entities found in this pathway (2)

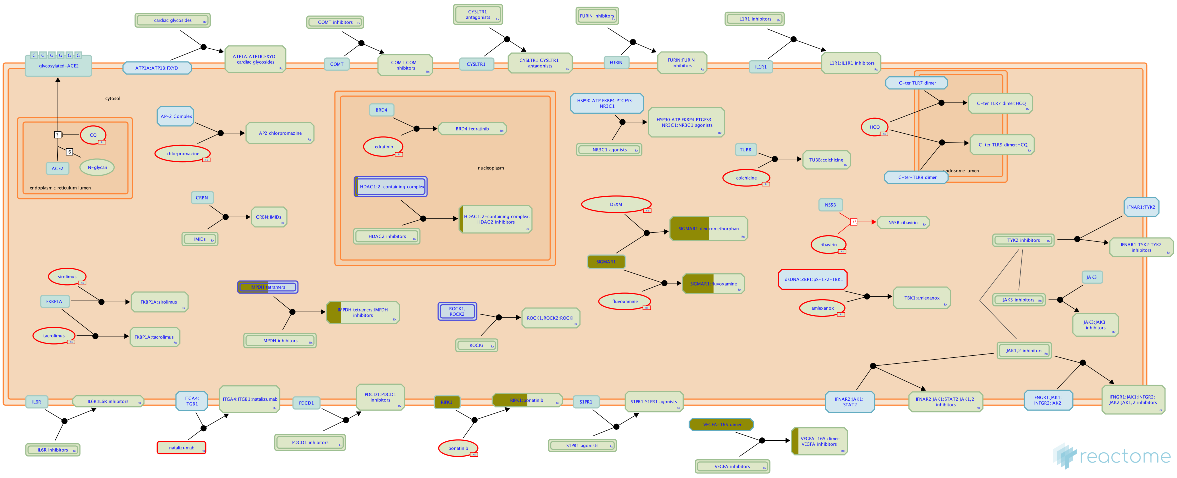
Input	UniProt Id
EGFR	P00533

Input	UniProt Id
VEGFA	P15692

Input	Ensembl Id
EGFR	ENSG00000146648

Input	Ensembl Id
VEGFA	ENSG00000112715

2. Potential therapeutics for SARS (R-HSA-9679191)



Diseases: viral infectious disease.

No drug has yet (April 2020) been shown in a randomized double-blind placebo-controlled trial to prevent or reduce the severity of human infection with SARS-CoV-1 or SARS-CoV-2. Nevertheless, a large number of promising drug candidates have been identified on the basis of their efficacy in treatment of human infections with other RNA viruses or in diminishing cytokine storms and other pathologies due to destructive host reactions to viruses similar to SARS-CoV-1 and SARS-CoV-2. The interactions of these candidate drugs with their known viral and human protein targets are annotated.

In addition, some drugs that inhibit Cytochrome P450 (CYP) oxidoreductases have been shown to be effective in prolonging the plasma half-lives of antiviral drugs with acceptable side effects, and CYP inhibition by these drugs is annotated.

Finally, effects of any of these drugs on unrelated essential human proteins, that might limit their use in vivo are annotated.

References

Rosa SGV & Santos WC (2020). Clinical trials on drug repositioning for COVID-19 treatment. Rev. Panam. Salud Publica, 44, e40.

Harrison C (2020). Coronavirus puts drug repurposing on the fast track. Nat. Biotechnol..

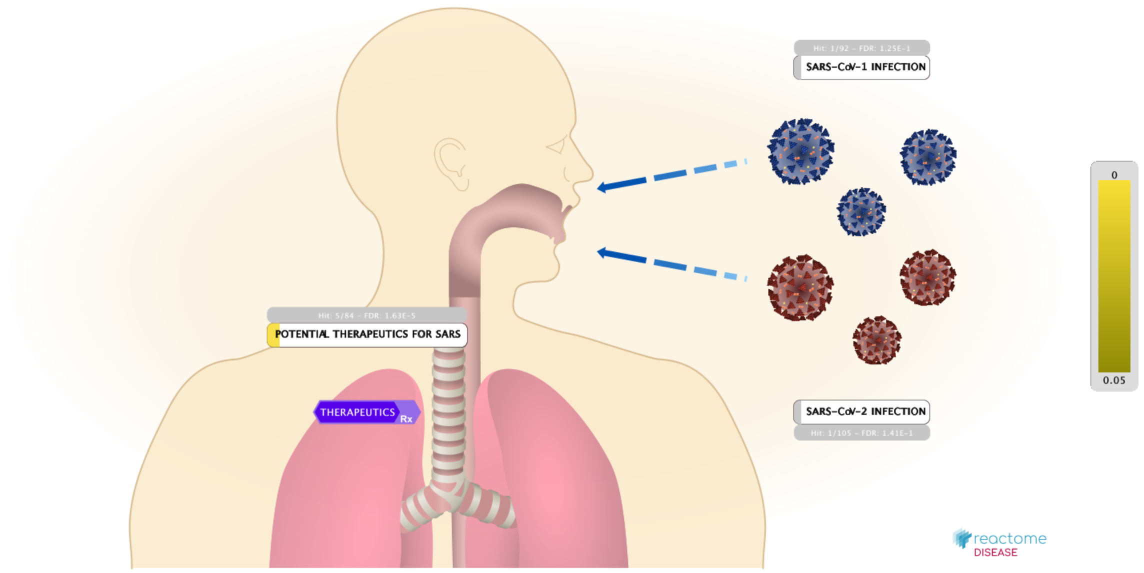
Edit history

Date	Action	Author
2020-03-23	Edited	Jassal B
2020-03-23	Authored	Jassal B
2020-03-23	Created	Jassal B
2020-05-14	Reviewed	Shoichet BK
2020-07-07	Modified	Jassal B

Entities found in this pathway (5)

Input	UniProt Id	Input	UniProt Id	Input	UniProt Id
HDAC2	Q92769	IMPDH2	P12268	RIPK1	Q13546
SIGMAR1	Q99720	VEGFA	P15692-4		

3. SARS-CoV Infections (R-HSA-9679506)



Diseases: Coronavirus infection.

Coronaviruses (CoVs) are large, enveloped, positive strand RNA viruses that can be classified into four genera: alpha, beta, delta, and gamma. Coronaviruses are ecologically diverse, infecting animals including camels, cattle, cats, and bats, with the greatest variety seen in bats, suggesting that bats are the reservoirs for many of these viruses. Rarely, A and B lineage beta coronaviruses of non-human origin can infect people and then spread directly between people. Four human coronaviruses (HCoVs), HCoV 229E, NL63, OC43, and HKU1, are endemic globally and account for 10% to 30% of upper respiratory tract infections in adults, typically presenting as common colds (van der Hoek 2007). However, in the 21st century, three highly pathogenic HCoVs - severe acute respiratory syndrome coronavirus (SARS-CoV-1) in 2003, Middle East Respiratory Syndrome coronavirus (MERS CoV) in 2012, and SARS-CoV-2 in 2019 - emerged from animal reservoirs to cause global epidemics with alarming morbidity and mortality (De Wit et al. 2016; Fung & Liu 2019; Marra et al. 2003; Paules et al. 2020).

During the 2003 outbreak of SARS-CoV-1, 8,098 people worldwide became sick. Of these, 774 died. In the United States, only eight people had laboratory evidence of SARS-CoV-1 infection. All of these people had traveled to other parts of the world where the disease was spreading. Community spread was not observed in the United States (De Wit et al. 2016; WHO - SARS). A second human coronavirus, MERS-CoV, first observed in 2012, has been identified in 2,494 patients with respiratory distress of whom 858 have died. Human-to-human transmission of the virus appears to be limited (De Wit et al. 2016; WHO – MERS).

In December 2019, yet another pathogenic HCoV, 2019 novel coronavirus (2019 nCoV), was recognized, initially in Wuhan, China. The World Health Organization has named the disease caused by the 2019 novel coronavirus COronaVIrus Disease 2019, or COVID-19. The disease has also been referred to as 2019 novel coronavirus or 2019 nCoV.

SARS-CoV-1 and SARS-CoV-2 viral infection pathways are annotated in this section, as are drugs that potentially modulate the infection processes. Many of the steps of SARS-CoV-1 infection have been characterized experimentally in the past 15 years (Fung & Liu 2019; Masters 2006), and this experimental work has allowed the annotation here of the infection process and interactions with host proteins in molecular detail. Less comparable data is yet available for SARS-CoV-2 infection, but the similarity in the genomes and predicted proteomes of the two viruses allows the inference of a detailed infection pathway for SARS-CoV-2 that has been manually annotated as well with experimental results published in the first half of 2020.

References

Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, ... Roper RL (2003). The Genome sequence of the SARS-associated coronavirus. Science, 300, 1399-404.

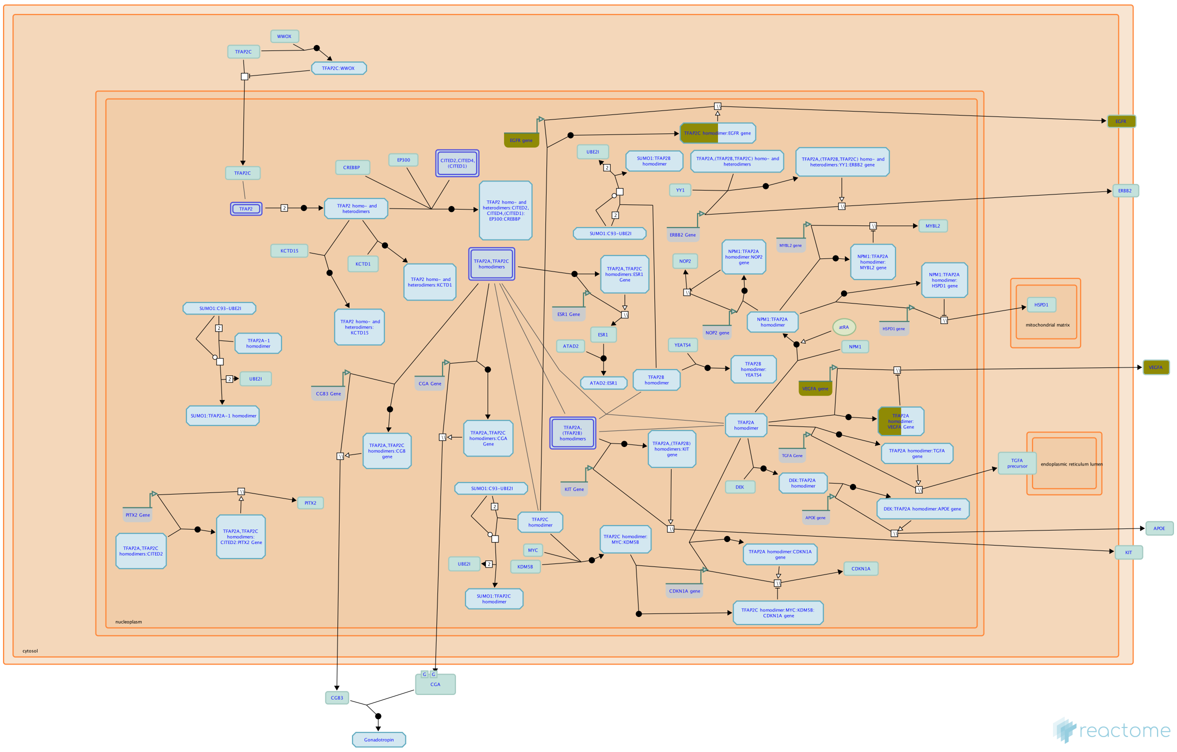
Edit history

Date	Action	Author
2020-03-24	Created	Gillespie ME
2020-05-26	Edited	Gillespie ME
2020-05-26	Authored	Gillespie ME
2020-10-28	Modified	Shamovsky V

Entities found in this pathway (6)

Input	UniProt Id	Input	UniProt Id	Input	UniProt Id
CTSL	P07711	HDAC2	Q92769	IMPDH2	P12268
RIPK1	Q13546	SIGMAR1	Q99720	VEGFA	P15692-4

4. Transcriptional regulation by the AP-2 (TFAP2) family of transcription factors (R-HSA-8864260)



The AP-2 (TFAP2) family of transcription factors includes five proteins in mammals: TFAP2A (AP-2 alpha), TFAP2B (AP-2 beta), TFAP2C (AP-2 gamma), TFAP2D (AP-2 delta) and TFAP2E (AP-2 epsilon). The AP-2 family transcription factors are evolutionarily conserved in metazoans and are characterized by a helix-span-helix motif at the C-terminus, a central basic region, and the transactivation domain at the N-terminus. The helix-span-helix motif and the basic region enable dimerization and DNA binding (Eckert et al. 2005).

AP-2 dimers bind palindromic GC-rich DNA response elements that match the consensus sequence 5'-GCCNNNGGC-3' (Williams and Tjian 1991a, Williams and Tjian 1991b). Transcriptional co-factors from the CITED family interact with the helix-span-helix (HSH) domain of TFAP2 (AP-2) family of transcription factors and recruit transcription co-activators EP300 (p300) and CREBBP (CBP) to TFAP2-bound DNA elements. CITED2 shows the highest affinity for TFAP2 proteins, followed by CITED4, while CITED1 interacts with TFAP2s with a very low affinity. Mouse embryos defective for CITED2 exhibit neural crest defects, cardiac malformations and adrenal agenesis, which can at least in part be attributed to a defective Tfap2 transactivation (Bamforth et al. 2001, Braganca et al. 2002, Braganca et al. 2003). Transcriptional activity of AP-2 dimers in inhibited by binding of KCTD1 or KCTD15 to the AP-2 transactivation domain (Ding et al. 2009, Zarelli and Dawid 2013). Transcriptional activity of TFAP2A, TFAP2B and TFAP2C is negatively regulated by SUMOylation mediated by UBE2I (UBC9) (Eloranta and Hurst 2002, Berlato et al. 2011, Impens et al. 2014, Bogachek et al. 2014).

During embryonic development, AP-2 transcription factors stimulate proliferation and suppress terminal differentiation in a cell-type specific manner (Eckert et al. 2005).

TFAP2A and TFAP2C directly stimulate transcription of the estrogen receptor ESR1 gene (McPherson and Weigel 1999). TFAP2A expression correlates with ESR1 expression in breast cancer, and TFAP2C is frequently overexpressed in estrogen-positive breast cancer and endometrial cancer (deConinck et al. 1995, Turner et al. 1998). TFAP2A, TFAP2C, as well as TFAP2B can directly stimulate the expression of ERBB2, another important breast cancer gene (Bosher et al. 1996). Association of TFAP2A with the YY1 transcription factor significantly increases the ERBB2 transcription rate (Begon et al. 2005). In addition to ERBB2, the expression of another receptor tyrosine kinase, KIT, is also stimulated by TFAP2A and TFAP2B (Huang et al. 1998), while the expression of the VEGF receptor tyrosine kinase ligand VEGFA is repressed by TFAP2A (Ruiz et al. 2004, Li et al. 2012). TFAP2A stimulates transcription of the transforming growth factor alpha (TGFA) gene (Wang et al. 1997). TFAP2C regulates EGFR in luminal breast cancer (De Andrade et al. 2016).

TFAP2C plays a critical role in maintaining the luminal phenotype in human breast cancer and in influencing the luminal cell phenotype during normal mammary development (Cyr et al. 2015).

In placenta, TFAP2A and TFAP2C directly stimulate transcription of both subunits of the human chorionic gonadotropin, CGA and CGB (Johnson et al. 1997, LiCalsi et al. 2000).

TFAP2A and/or TFAP2C, in complex with CITED2, stimulate transcription of the PITX2 gene, involved in left-right patterning and heart development (Bamforth et al. 2004, Li et al. 2012).

TFAP2A and TFAP2C play opposing roles in transcriptional regulation of the CDKN1A (p21) gene locus. While TFAP2A stimulates transcription of the CDKN1A cyclin-dependent kinase inhibitor (Zeng et al. 1997, Williams et al. 2009, Scibetta et al. 2010), TFAP2C represses CDKN1A transcription (Williams et al. 2009, Scibetta et al. 2010, Wong et al. 2012). Transcription of the TFAP2A gene may be inhibited by CREB and E2F1 (Melnikova et al. 2010).

For review of the AP-2 family of transcription factors, please refer to Eckert et al. 2005.

References

Eckert D, Buhl S, Weber S, Jäger R & Schorle H (2005). The AP-2 family of transcription factors. Genome Biol., 6, 246.

Williams T & Tjian R (1991). Analysis of the DNA-binding and activation properties of the human transcription factor AP-2. Genes Dev., 5, 670-82.

Williams T & Tjian R (1991). Characterization of a dimerization motif in AP-2 and its function in heterologous DNA-binding proteins. Science, 251, 1067-71.

Bamforth SD, Bragança J, Eloranta JJ, Murdoch JN, Marques FI, Kranc KR, ... Bhattacharya S (2001). Cardiac malformations, adrenal agenesis, neural crest defects and exencephaly in mice lacking Cited2, a new Tfap2 co-activator. Nat. Genet., 29, 469-74.

Bamforth SD, Bragança J, Farthing CR, Schneider JE, Broadbent C, Michell AC, ... Bhattacharya S (2004). Cited2 controls left-right patterning and heart development through a Nodal-Pitx2c pathway. Nat. Genet., 36, 1189-96.

Edit history

Date	Action	Author
2016-03-14	Edited	Orlic-Milacic M
2016-03-14	Authored	Orlic-Milacic M

Date	Action	Author
2016-03-14	Created	Orlic-Milacic M
2016-05-04	Reviewed	Dawid IB, Zarelli VE
2016-05-17	Reviewed	Bogachek MV, Weigel RJ
2021-03-06	Modified	Shorser S

Entities found in this pathway (2)

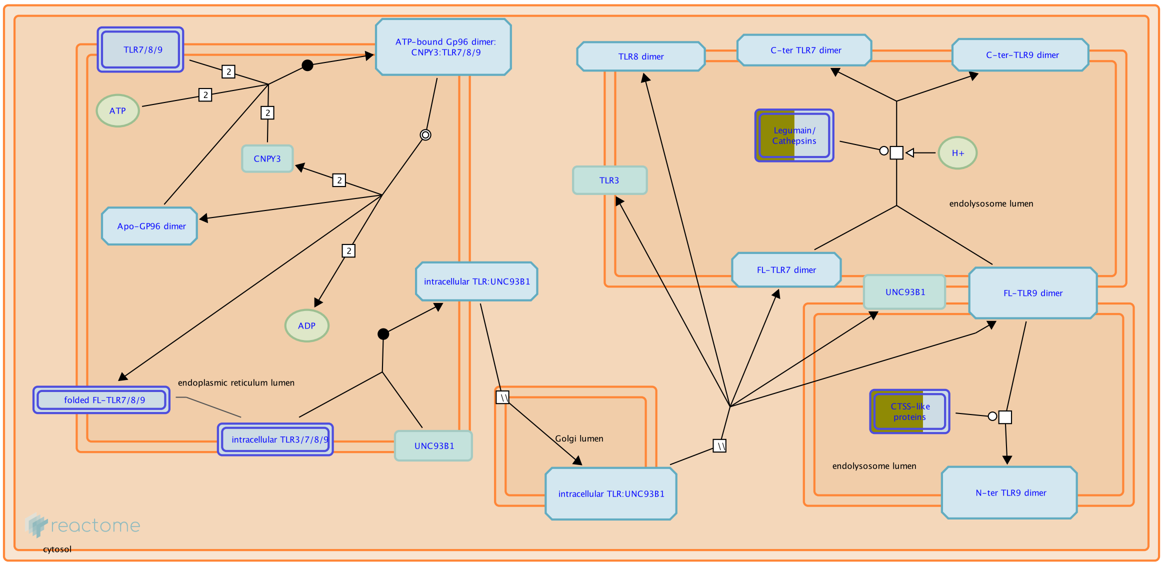
Input	UniProt Id
EGFR	P00533

Input	UniProt Id
VEGFA	P15692

Input	Ensembl Id
EGFR	ENSG00000146648

Input	Ensembl Id
VEGFA	ENSG00000112715

5. Trafficking and processing of endosomal TLR (R-HSA-1679131)



Mammalian TLR3, TLR7, TLR8, TLR9 are endosomal receptors that sense nucleic acids that have been released from endocytosed/phagocytosed bacteria, viruses or parasites. These TLRs have a ligand-recognition domain that faces the lumen of the endosome (which is topologically equivalent to the outside of the cell), a transmembrane domain, and a signaling domain that faces the cytosol.

Under normal conditions, self nucleic acids are not recognized by TLRs due to multiple levels of regulation including receptor compartmentalization, trafficking and proteolytic processing (Barton GM et al 2006, Ewald SE et al 2008). At steady state TLR3, TLR7, TLR8, TLR9 reside primarily in the endoplasmic reticulum (ER), however, their activation by specific ligands only occurs within acidified endolysosomal compartments (Hacker H et al 1998, Funami K et al 2004, Gibbard RJ et al 2006). Several chaperon proteins associate with TLRs in the ER to provide efficient translocation to endolysosome. Upon reaching endolysosomal compartments the ectodomains of TLR7 and TLR9 are proteolytically cleaved by cysteine endoproteases. Both full-length and cleaved C-terminus of TLR9 bind CpG-oligodeoxynucleotides, however it has been proposed that only the processed receptor is functional.

Although similar cleavage of TLR3 has been reported by Ewald et al 2011, other studies demonstrated that the N-terminal region of TLR3 ectodomain was implicated in ligand binding, thus TLR3 may function as a full-length receptor (Liu L et al 2008, Tokisue T et al 2008).

There are no data on TLR8 processing, although the cell biology of TLR8 is probably similar to TLR9 and TLR7 (Gibbard RJ et al 2006, Wei T et al 2009).

References

Hacker H, Mischak H, Miethke T, Liptay S, Schmid R, Sparwasser T, ... Wagner H (1998). CpG-DNA-specific activation of antigen-presenting cells requires stress kinase activity and is preceded by non-specific endocytosis and endosomal maturation. EMBO J, 17, 6230-40.

Funami K, Matsumoto M, Oshiumi H, Akazawa T, Yamamoto A & Seya T (2004). The cytoplasmic 'linker region' in Toll-like receptor 3 controls receptor localization and signaling. Int Immunol, 16, 1143-54.

Gibbard RJ, Morley PJ & Gay NJ (2006). Conserved features in the extracellular domain of human toll-like receptor 8 are essential for pH-dependent signaling. J Biol Chem, 281, 27503-11.

Barton GM, Kagan JC & Medzhitov R (2006). Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. Nat Immunol, 7, 49-56.

Ewald SE, Lee BL, Lau L, Wickliffe KE, Shi GP, Chapman HA & Barton GM (2008). The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor. Nature, 456, 658-62.

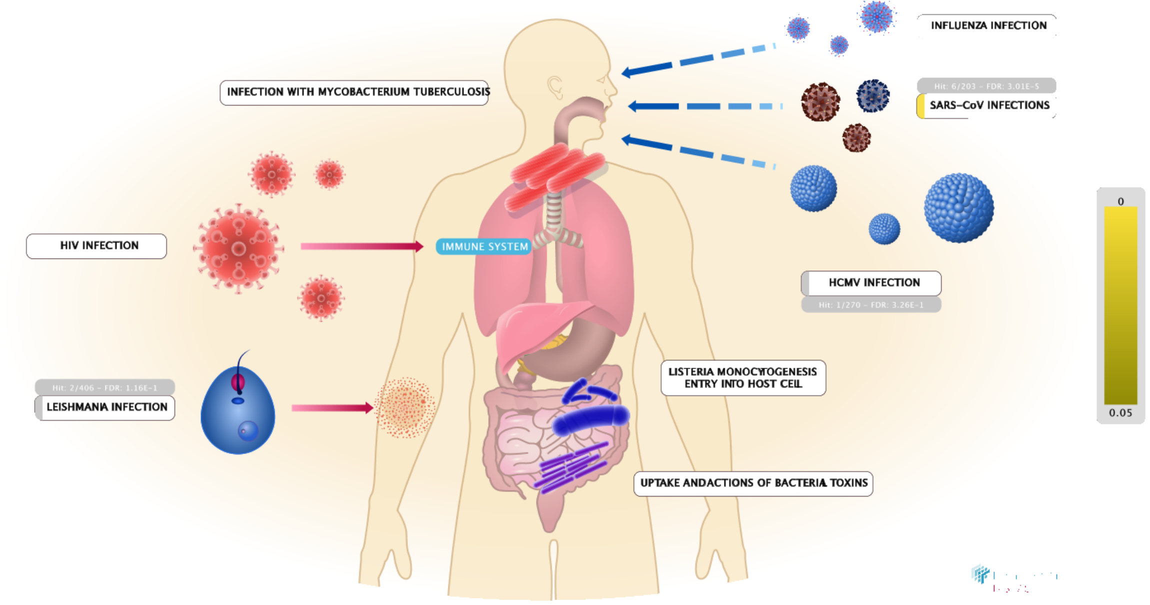
Edit history

Date	Action	Author
2011-10-19	Authored	Shamovsky V
2011-10-19	Created	Shamovsky V
2012-02-09	Reviewed	Gillespie ME
2012-02-19	Edited	Shamovsky V
2012-02-28	Reviewed	Rose II WA, Leifer CA
2021-03-05	Modified	Shorser S

Entities found in this pathway (2)

Input	UniProt Id	Input	UniProt Id
CTSB	P07858	CTSL	O60911, P07711

6. Infectious disease (R-HSA-5663205)



Diseases: disease by infectious agent.

Infectious diseases are ones due to the presence of pathogenic microbial agents in human host cells. Processes annotated in this category include the life cycles of SARS-CoV viruses, influenza virus and HIV (human immunodeficiency virus), some metabolic processes mediated by intracellular Mycobacterium tuberculosis, the actions of clostridial, anthrax, and diphtheria toxins, and the entry of Listeria monocytogenes into human cells.

References

Edit history

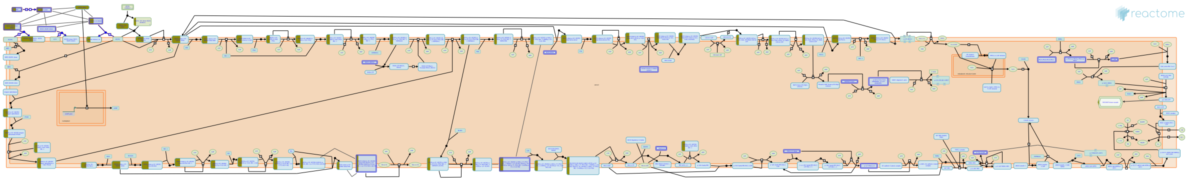
Date	Action	Author
2015-01-16	Created	D'Eustachio P
2021-03-02	Modified	Shorser S

Entities found in this pathway (8)

Input	UniProt Id	Input	UniProt Id	Input	UniProt Id
CTSL	P07711	EGFR	P00533	HDAC2	Q92769
IL6	P05231	IMPDH2	P12268	RIPK1	Q13546
SIGMAR1	Q99720	VEGFA	P15692-4		

Input	Ensembl Id
IL6	ENSG00000136244

7. VEGF ligand-receptor interactions (R-HSA-194313)



Cellular compartments: plasma membrane.

The VEGF family is encoded by seven genes (VEGF-A, B, C, D, E: PLGF (Placenta Growth Factor)-1, 2). Six isoforms of VEGF-A protein, containing 121, 145, 165, 183, 189, and 206 amino acid residues, and two isoforms of VEGF-B (167 and 186 residues) are specified by alternatively spliced mRNAs. The active form of each of these proteins is a homodimer.

The specificities of the three VEGF tyrosine kinase receptors, VEGFR-1, VEGFR-2 and VEGFR-3, for these ligands are shown in the figure (Hicklin and Ellis 2005). All VEGF-A isoforms bind both VEGFR-1 and VEGFR-2; PLGF-1 and -2, and VEGF-B isoforms bind only VEGFR-1; VEGF-E binds VEGFR-2; and VEGF-C and -D bind both VEGFR-2 and -3. VEGF-D undergoes a complex series of post-translational modifications that results in secreted forms with increased activity toward VEGFR-3 and VEGFR-2.

Two co-receptor proteins in the cell membrane, neuropilin (NRP)-1 and NRP-2, interact with VEGFR proteins to increase the affinity of the latter for their ligands (Neufeld et al.,2002). They differ from VEGFR proteins in not having intracellular signaling domains.

References

Olsson AK, Dimberg A, Kreuger J & Claesson-Welsh L (2006). VEGF receptor signalling - in control of vascular function. Nat Rev Mol Cell Biol, 7, 359-71.

Shibuya M & Claesson-Welsh L (2006). Signal transduction by VEGF receptors in regulation of angiogenesis and lymphangiogenesis. Exp Cell Res, 312, 549-60.

Matsumoto T & Mugishima H (2006). Signal transduction via vascular endothelial growth factor (VEGF) receptors and their roles in atherogenesis. J Atheroscler Thromb, 13, 130-5.

Cross MJ, Dixelius J, Matsumoto T & Claesson-Welsh L (2003). VEGF-receptor signal transduction. Trends Biochem Sci, 28, 488-94.

Edit history

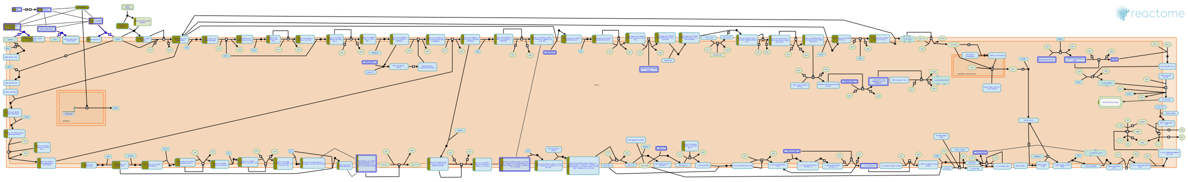
Date	Action	Author
2007-03-12	Created	Gopinathrao G
2008-02-28	Reviewed	Claesson-Welsh L
2013-08-30	Edited	Garapati P V
2013-08-30	Authored	Garapati P V
2021-03-05	Modified	Shorser S

Entities found in this pathway (1)

Input	UniProt Id
VEGFA	P15692, P49765

Input	UniProt Id

8. VEGF binds to VEGFR leading to receptor dimerization (R-HSA-195399)



Cellular compartments: plasma membrane.

The binding of VEGF ligands to VEGFR receptors in the cell membrane induces dimerization and activation of the latter, initiating intracellular signaling cascades that result in proliferation, survival, migration and increased permeability of vascular endothelial cells (Matsumoto and Mugishima, 2006). The receptors predominantly form homodimers but heterodimers between VEGFR-1 and -2 have been observed. Although both VEGFR-1 and -2 are expressed in the vascular endothelium, the angiogenic activities of VEGFs are transduced mainly through VEGFR-2 in vivo.

References

Matsumoto T & Mugishima H (2006). Signal transduction via vascular endothelial growth factor (VEGF) receptors and their roles in atherogenesis. J Atheroscler Thromb, 13, 130-5.

Olsson AK, Dimberg A, Kreuger J & Claesson-Welsh L (2006). VEGF receptor signalling - in control of vascular function. Nat Rev Mol Cell Biol, 7, 359-71.

Cross MJ, Dixelius J, Matsumoto T & Claesson-Welsh L (2003). VEGF-receptor signal transduction. Trends Biochem Sci, 28, 488-94.

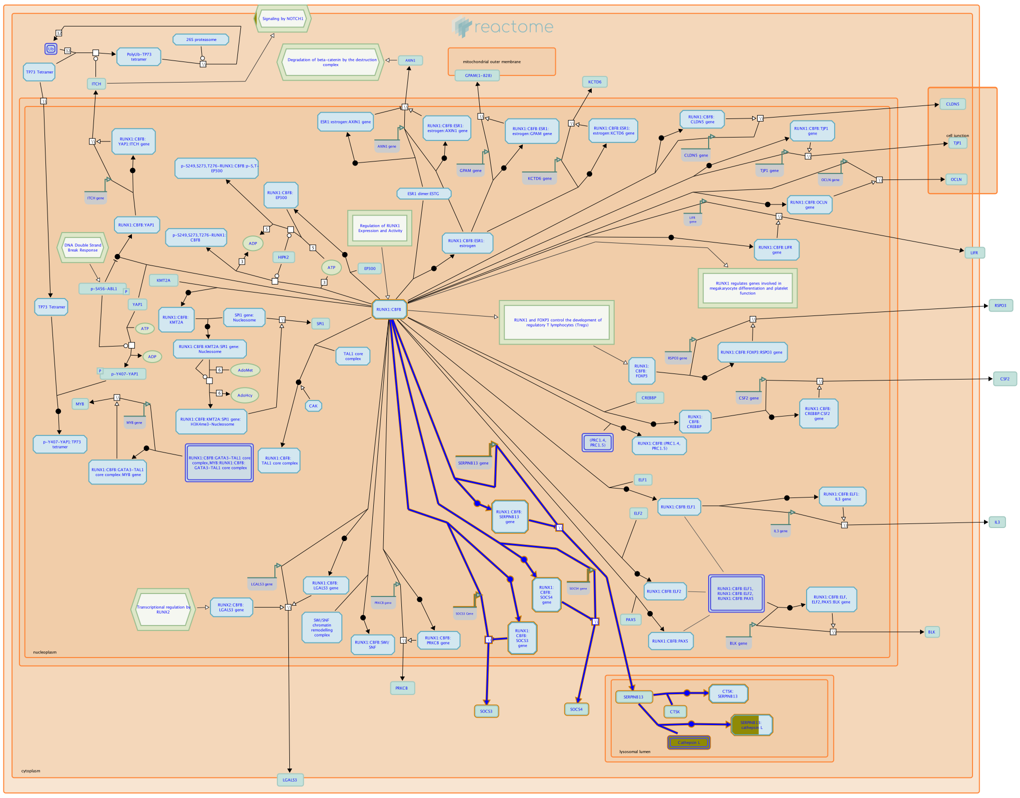
Edit history

Date	Action	Author
2007-04-06	Created	Gopinathrao G
2008-02-28	Reviewed	Claesson-Welsh L
2013-08-30	Edited	Garapati P V
2013-08-30	Authored	Garapati P V
2021-03-05	Modified	Shorser S

Entities found in this pathway (1)

Input	UniProt Id
VEGFA	P15692, P49765

9. RUNX1 regulates transcription of genes involved in differentiation of keratinocytes (R-HSA-8939242)



The RUNX1:CBFB complex directly inhibits transcription of the SERPINB13 gene (Nomura et al. 2005), a gene involved in keratinocyte differentiation that is frequently down-regulated in head and neck cancers (Boyapati et al. 2011). RUNX1 also inhibits transcription of STAT3 inhibitors SOCS3 and SOCS4, resulting in elevated STAT3 activity. RUNX1-mediated increase in STAT3 activity, first discovered in keratinocytes, is thought to be involved in the maintenance of epithelial stem cells and contributes to development of epithelial cancers, including squamous cell carcinoma (SCC) of the skin (Scheitz et al. 2012).

References

Boyapati A, Ren B & Zhang DE (2011). SERPINB13 is a novel RUNX1 target gene. Biochem. Biophys. Res. Commun., 411, 115-20.

Nomura T & Katunuma N (2005). Involvement of cathepsins in the invasion, metastasis and proliferation of cancer cells. J. Med. Invest., 52, 1-9.

Scheitz CJ, Lee TS, McDermitt DJ & Tumbar T (2012). Defining a tissue stem cell-driven Runx1/Stat3 signalling axis in epithelial cancer. EMBO J., 31, 4124-39.

Edit history

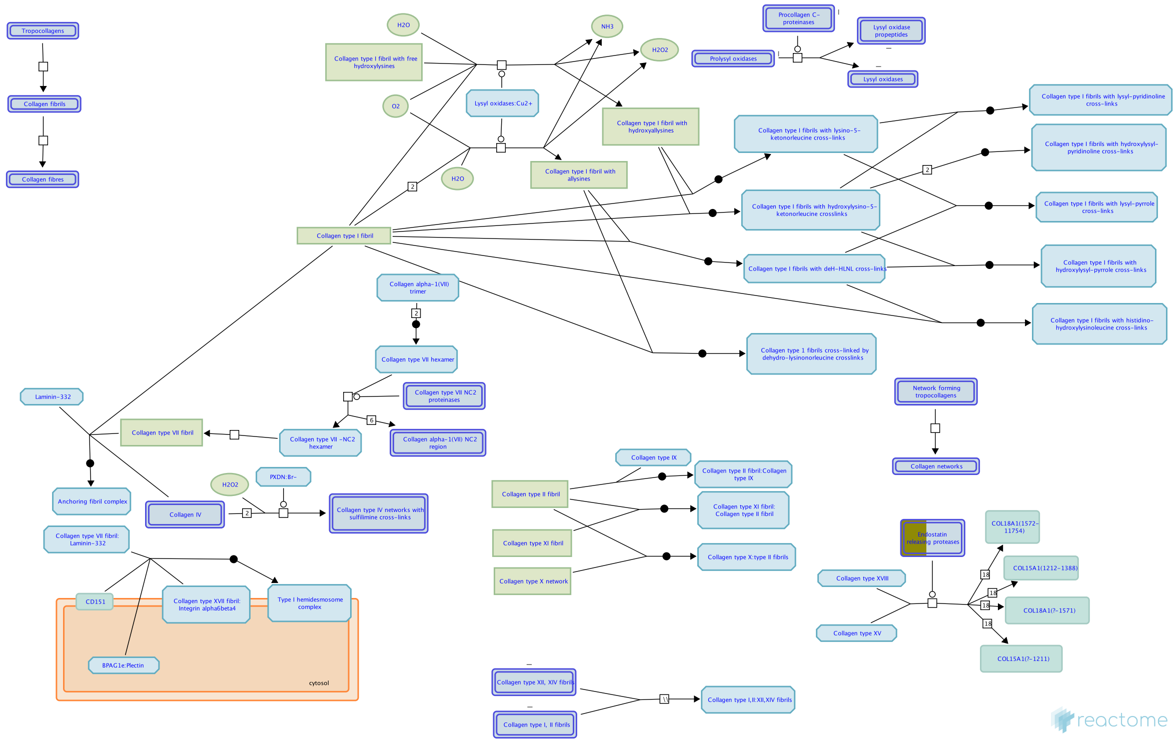
Date	Action	Author
2016-09-14	Authored	Orlic-Milacic M
2016-09-16	Created	Orlic-Milacic M
2016-12-20	Reviewed	Ito Y, Chuang LS
2017-05-09	Edited	Orlic-Milacic M

Date	Action	Author
2021-03-06	Modified	Shorser S

Entities found in this pathway (1)

Input	UniProt Id
CTSL	O60911, P07711

10. Assembly of collagen fibrils and other multimeric structures (R-HSA-2022090)



Collagen trimers in triple-helical form, referred to as procollagen or collagen molecules, are exported from the ER and trafficked through the Golgi network before secretion into the extracellular space. For fibrillar collagens namely types I, II, III, V, XI, XXIV and XXVII (Gordon & Hahn 2010, Ricard-Blum 2011) secretion is concomitant with processing of the N and C terminal collagen propeptides. These processed molecules are known as tropocollagens, considered to be the units of higher order collagen structures. They form within the extracellular space via a process that can proceed spontaneously, but in the cellular environment is regulated by many collagen binding proteins such as the FACIT (Fibril Associated Collagens with Interrupted Triple helices) family collagens and Small Leucine-Rich Proteoglycans (SLRPs). The architecture formed ultimately depends on the collagen subtype and the cellular conditions. Structures include the well-known fibrils and fibres formed by the major structural collagens type I and II plus several different types of supramolecular assembly (Bruckner 2010). The mechanical and physical properties of tissues depend on the spatial arrangement and composition of these collagen-containing structures (Kadler et al. 1996, Shoulders & Raines 2009, Birk & Bruckner 2011).

Fibrillar collagen structures are frequently heterotypic, composed of a major collagen type in association with smaller amounts of other types, e.g. type I collagen fibrils are associated with types III and V, while type II fibrils frequently contain types IX and XI (Wess 2005). Fibres composed exclusively of a single collagen type probably do not exist, as type I and II fibrils require collagens V and XI respectively as nucleators (Kadler et al. 2008, Wenstrup et al. 2011). Much of the structural understanding of collagen fibrils has been obtained with fibril-forming collagens, particularly type I, but some central features are believed to apply to at least the other fibrillar collagen subtypes (Wess 2005). Fibril diameter and length varies considerably, depending on the tissue and collagen types (Fang et al. 2012). The reasons for this are poorly understood (Wess 2005).

Some tissues such as skin have fibres that are approximately the same diameter while others such as tendon or cartilage have a bimodal distribution of thick and thin fibrils. Mature type I collagen fibrils in tendon are up to 1 cm in length, with a diameter of approx. 500 nm. An individual fibrillar collagen triple helix is less than 1.5 nm in diameter and around 300 nm long; collagen molecules must assemble to give rise to the higher-order fibril structure, a process known as fibrillogenesis, prevented by the presence of C-terminal propeptides (Kadler et al. 1987). In electron micrographs, fibrils have a banded appearance, due to regular gaps where fewer collagen molecules overlap, which occur because the fibrils are aligned in a quarter-stagger arrangement (Hodge & Petruska 1963). Collagen microfibrils are believed to have a quasi-hexagonal unit cell, with tropocollagen arranged to form supertwisted, right-handed microfibrils that interdigitate with neighbouring microfibrils, leading to a spiral-like structure for the mature collagen fibril (Orgel et al. 2006, Holmes & Kadler 2006).

Neighbouring tropocollagen monomers interact with each other and are cross-linked covalently by lysyl oxidase (Orgel et al. 2000, Maki 2006). Mature collagen fibrils are stabilized by lysyl oxidase-mediated cross-links. Hydroxylysyl pyridinoline and lysyl pyridinoline cross-links form between (hydroxy) lysine and hydroxylysine residues in bone and cartilage (Eyre et al. 1984). Arginoline cross-links can form in cartilage (Eyre et al. 2010); mature bovine articular cartilage contains roughly equimolar amounts of arginoline and hydroxylysyl pyridinoline based on peptide yields. Mature collagen fibrils in skin are stabilized by the lysyl oxidase-mediated cross-link histidino-hydroxylysinonorleucine (Yamauch et al. 1987). Due to the quarter-staggered arrangement of collagen molecules in a fibril, telopeptides most often interact with the triple helix of a neighbouring collagen molecule in the fibril, except for collagen molecules in register staggered by 4D from another collagen molecule. Fibril aggregation in vitro can be unipolar or bipolar, influenced by temperature and levels of C-proteinase, suggesting a role for the N- and C- propeptides in regulation of the aggregation process (Kadler et al. 1996). In vivo, collagen molecules at the fibril surface may retain their N-propeptides, suggesting that this may limit further accretion, or alternatively represents a transient stage in a model whereby fibrils grow in diameter through a cycle of deposition, cleavage and further deposition (Chapman 1989).

In vivo, fibrils are often composed from more than one type of collagen. Type III collagen is found associated with type I collagen in dermal fibrils, with the collagen III on the periphery, suggesting a regulatory role (Fleischmajer et al. 1990). Type V collagen associates with type I collagen fibrils, where it may limit fibril diameter (Birk et al. 1990, White et al. 1997). Type IX associates with the surface of narrow diameter collagen II fibrils in cartilage and the cornea (Wu et al. 1992, Eyre et al. 2004). Highly specific patterns of crosslinking sites suggest that collagen IX functions in interfibrillar networking (Wess 2005). Type XII and XIV collagens are localized near the surface of banded collagen I fibrils (Nishiyama et al. 1994). Certain fibril-associated collagens with interrupted triple helices (FACITs) associate with the surface of collagen fibrils, where they may serve to limit fibril fusion and thereby regulate fibril diameter (Gordon & Hahn 2010). Collagen XV, a member of the multiplexin family, is almost exclusively associated with the fibrillar collagen network, in very close proximity to the basement membrane. In human tissues collagen XV is seen linking banded collagen fibers subjacent to the basement membrane (Amenta et al. 2005). Type XIV collagen, SLRPs and discoidin domain receptors also regulate fibrillogenesis (Ansorge et al. 2009, Kalamajski et al. 2010, Flynn et al. 2010).

Collagen IX is cross-linked to the surface of collagen type II fibrils (Eyre et al. 1987). Type XII and XIV collagens are found in association with type I (Walchli et al. 1994) and type II (Watt et al. 1992, Eyre 2002) fibrils in cartilage. They are thought to associate non-covalently via their COL1/NC1 domains (Watt et al. 1992, Eyre 2002).

Some non-fibrillar collagens form supramolecular assemblies that are distinct from typical fibrils. Collagen VII forms anchoring fibrils, composed of antiparallel dimers that connect the dermis to the epidermis (Bruckner-Tuderman 2009). During fibrillogenesis, the nascent type VII procollagen molecules dimerize in an antiparallel manner. The C-propeptides are then removed by Bone morphogenetic protein 1 (Rattenholl et al. 2002) and the processed antiparallel dimers aggregate laterally. Collagens VIII and X form hexagonal networks and collagen VI forms beaded filament (Gordon & Hahn 2010, Ricard-Blum et al. 2011).

References

Kadler KE, Holmes DF, Trotter JA & Chapman JA (1996). Collagen fibril formation. Biochem J, 316, 1-11.

Orgel JP, San Antonio JD & Antipova O (2011). Molecular and structural mapping of collagen fibril interactions. Connect. Tissue Res., 52, 2-17.

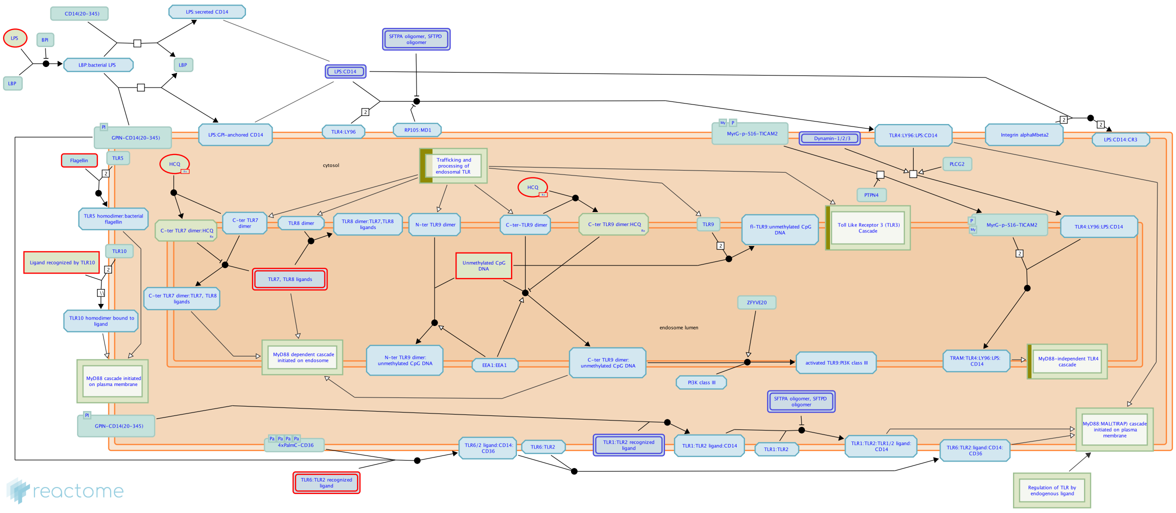
Edit history

Date	Action	Author
2011-08-05	Authored	Jupe S
2011-11-25	Created	Jupe S
2012-10-08	Reviewed	Kalamajski S, Raleigh S
2012-11-12	Edited	Jupe S
2012-11-19	Reviewed	Ricard-Blum S
2021-03-06	Modified	Shorser S

Entities found in this pathway (2)

Input	UniProt Id	Input	UniProt Id
CTSB	P07858	CTSL	O60911, P07711

11. Toll-like Receptor Cascades (R-HSA-168898)



In human, ten members of the Toll-like receptor (TLR) family (TLR1-TLR10) have been identified (TLR11 has been found in mouse, but not in human). All TLRs have a similar Toll/IL-1 receptor (TIR) domain in their cytoplasmic region and an Ig-like domain in the extracellular region, where each is enriched with a varying number of leucine-rich repeats (LRRs). Each TLR can recognize specific microbial pathogen components. The binding pathogenic component to TLR initializes signaling pathways that lead to induction of Interferon alpha/beta and inflammatory cytokines. There are two main signaling pathways. The first is a MyD88-dependent pathway that is common to all TLRs, except TLR3; the second is a TRIF(TICAM1)-dependent pathway that is peculiar to TLR3 and TLR4. TLR4-mediated signaling pathway via TRIF requires adapter molecule TRAM (TRIF-related adapter molecule or TICAM2). TRAM is thought to bridge between the activated TLR4 complex and TRIF.(Takeda & Akira 2004; Akira 2003; Takeda & Akira 2005; Kawai 2005; Heine & Ulmer 2005). This pathway is organized as trafficking and processing of TLR, various TLR cascades (TLR10,TLR3,TLR5,TLR7/8,TLR9,TLR4,TLR2) and their regulation.

References

Takeda K & Akira S (2004). TIR domains, which are conserved among all TLRs. Recent accumulating. Semin Immunol, 16, 3-9.

Akira S (2003). Toll-like receptor signaling. J Biol Chem, 278, 38105-8.

Takeda K & Akira S (2005). TIRAP/Mal, TRIF and TRAM. Differential utilization of these TIR. Int Immunol, 17, 1-14.

Kawai T (2005). Pathogen recognition with Toll-like receptors. Curr Opin Immunol, 17, 338-44.

Heine H & Ulmer AJ (2005). Recognition of bacterial products by toll-like receptors. Chem Immunol Allergy, 86, 99-119.

Edit history

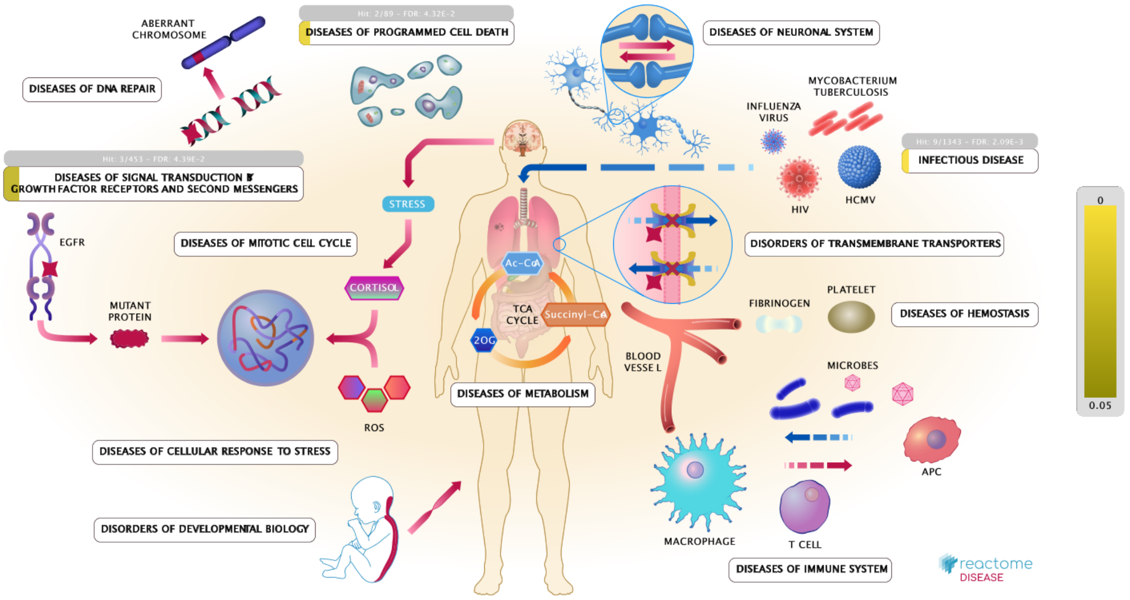
Date	Action	Author
2005-11-22	Created	de Bono B
2006-10-31	Reviewed	D'Eustachio P, Gay NJ, Gale M Jr
2021-03-02	Authored	Luo F, Gillespie ME, Gay NJ, de Bono B

Date	Action	Author
2021-03-06	Modified	Shorser S

Entities found in this pathway (3)

Input	UniProt Id	Input	UniProt Id	Input	UniProt Id
CTSB	P07858	CTSL	O60911, P07711	RIPK1	Q13546

12. Disease (R-HSA-1643685)



Biological processes are captured in Reactome by identifying the molecules (DNA, RNA, protein, small molecules) involved in them and describing the details of their interactions. From this molecular viewpoint, human disease pathways have three mechanistic causes: the inclusion of microbially-expressed proteins, altered functions of human proteins, or changed expression levels of otherwise functionally normal human proteins.

The first group encompasses the infectious diseases such as influenza, tuberculosis and HIV infection. The second group involves human proteins modified either by a mutation or by an abnormal post-translational event that produces an aberrant protein with a novel function. Examples include somatic mutations of EGFR and FGFR (epidermal and fibroblast growth factor receptor) genes, which encode constitutively active receptors that signal even in the absence of their ligands, or the somatic mutation of IDH1 (isocitrate dehydrogenase 1) that leads to an enzyme active on 2-oxoglutarate rather than isocitrate, or the abnormal protein aggregations of amyloidosis which lead to diseases such as Alzheimer's.

Infectious diseases are represented in Reactome as microbial-human protein interactions and the consequent events. The existence of variant proteins and their association with disease-specific biological processes is represented by inclusion of the modified protein in a new or variant reaction, an extension to the 'normal' pathway. Diseases which result from proteins performing their normal functions but at abnormal rates can also be captured, though less directly. Many mutant alleles encode proteins that retain their normal functions but have abnormal stabilities or catalytic efficiencies, leading to normal reactions that proceed to abnormal extents. The phenotypes of such diseases can be revealed when pathway annotations are combined with expression or rate data from other sources.

Depending on the biological pathway/process immediately affected by disease-causing gene variants, non-infectious diseases in Reactome are organized into diseases of signal transduction by growth factore receptors and second messengers, diseases of mitotic cell cycle, diseases of cellular response to stress, diseases of programmed cell death, diseases of DNA repair, disorders of transmembrane transporters, diseases of metabolism, diseases of immune system, diseases of neuronal system, disorders of developmental biology, disorders of extracellular matrix organization, and diseases of hemostatis.

References

Edit history

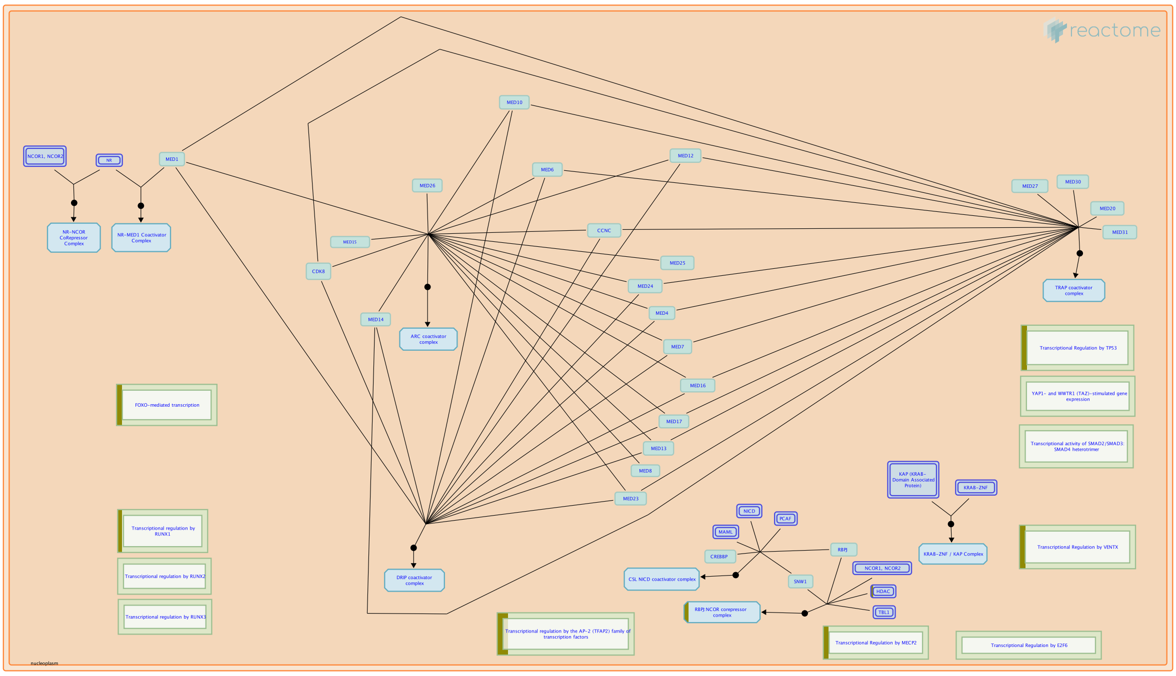
Date	Action	Author
2011-10-11	Created	Matthews L
2020-08-24	Modified	Orlic-Milacic M
2020-08-24	Edited	Orlic-Milacic M

Entities found in this pathway (9)

Input	UniProt Id	Input	UniProt Id	Input	UniProt Id
CTSL	P07711	EGFR	P00533, P04626	HDAC2	Q92769
IL6	P05231	IMPDH2	P12268	POLA1	P09884
RIPK1	Q13546	SIGMAR1	Q99720	VEGFA	P15692-4

Input	Ensembl Id
IL6	ENSG00000136244

13. Generic Transcription Pathway (R-HSA-212436)



OVERVIEW OF TRANSCRIPTION REGULATION:

Detailed studies of gene transcription regulation in a wide variety of eukaryotic systems has revealed the general principles and mechanisms by which cell- or tissue-specific regulation of differential gene transcription is mediated (reviewed in Naar, 2001. Kadonaga, 2004, Maston, 2006, Barolo, 2002; Roeder, 2005, Rosenfeld, 2006). Of the three major classes of DNA polymerase involved in eukaryotic gene transcription, Polymerase II generally regulates protein-encoding genes. Figure 1 shows a diagram of the various components involved in cell-specific regulation of Pol-II gene transcription.

Core Promoter: Pol II-regulated genes typically have a Core Promoter where Pol II and a variety of general factors bind to specific DNA motifs:

i: the TATA box (TATA DNA sequence), which is bound by the "TATA-binding protein" (TBP).

ii: the Initiator motif (INR), where Pol II and certain other core factors bind, is present in many Pol II-regulated genes.

iii: the Downstream Promoter Element (DPE), which is present in a subset of Pol II genes, and where additional core factors bind.

The core promoter binding factors are generally ubiquitously expressed, although there are exceptions to this.

Proximal Promoter: immediately upstream (5') of the core promoter, Pol II target genes often have a Proximal Promoter region that spans up to 500 base pairs (b.p.), or even to 1000 b.p.. This region contains a number of functional DNA binding sites for a specific set of transcription activator (TA) and transcription repressor (TR) proteins. These TA and TR factors are generally cell- or tissue-specific in expression, rather than ubiquitous, so that the presence of their cognate binding sites in the proximal promoter region programs cell- or tissue-specific expression of the target gene, perhaps in conjunction with TA and TR complexes bound in distal enhancer regions.

Distal Enhancer(s): many or most Pol II regulated genes in higher eukaryotes have one or more distal Enhancer regions which are essential for proper regulation of the gene, often in a cell or tissue-specific pattern. Like the proximal promoter region, each of the distal enhancer regions typically contain a cluster of binding sites for specific TA and/or TR DNA-binding factors, rather than just a single site.

Enhancers generally have three defining characteristics:

i: They can be located very long distances from the promoter of the target gene they regulate, sometimes as far as 100 Kb, or more.

ii: They can be either upstream (5') or downstream (3') of the target gene, including within introns of that gene.

iii: They can function in either orientation in the DNA.

Combinatorial mechanisms of transcription regulation: The specific combination of TA and TR binding sites within the proximal promoter and/or distal enhancer(s) provides a "combinatorial transcription code" that mediates cell- or tissue-specific expression of the associated target gene. Each promoter or enhancer region mediates expression in a specific subset of the overall expression pattern. In at least some cases, each enhancer region functions completely independently of the others, so that the overall expression pattern is a linear combination of the expression patterns of each of the enhancer modules.

Co-Activator and Co-Repressor Complexes: DNA-bound TA and TR proteins typically recruit the assembly of specific Co-Activator (Co-A) and Co-Repressor (Co-R) Complexes, respectively, which are essential for regulating target gene transcription. Both Co-A's and Co-R's are multi-protein complexes that contain several specific protein components.

Co-Activator complexes generally contain at lease one component protein that has Histone Acetyl Transferase (HAT) enzymatic activity. This functions to acetylate Histones and/or other chromatin-associated factors, which typically increases that transcription activation of the target gene. By contrast, Co-Repressor complexes generally contain at lease one component protein that has Histone De-Acetylase (HDAC) enzymatic activity. This functions to de-acetylate Histones and/or other chromatin-associated factors. This typically increases the transcription repression of the target gene.

Adaptor (Mediator) complexes: In addition to the co-activator complexes that assemble on particular cell-specific TA factors, - there are at least two additional transcriptional co-activator complexes common to most cells. One of these is the Mediator complex, which functions as an "adaptor" complex that bridges between the tissue-specific co-activator complexes assembled in the proximal promoter (or distal enhancers). The human Mediator complex has been shown to contain at least 19 protein distinct components. Different combinations of these co-activator proteins are also found to be components of specific transcription Co-Activator complexes, such as the DRIP, TRAP and ARC complexes described below.

TBP/TAF complex: Another large Co-A complex is the "TBP-associated factors" (TAFs) that assemble on TBP (TATA-Binding Protein), which is bound to the TATA box present in many promoters. There are at least 23 human TAF proteins that have been identified. Many of these are ubiquitously expressed, but TAFs can also be expressed in a cell or tissue-specific pattern.

Specific Coactivator Complexes for DNA-binding Transcription Factors.

A number of specific co-activator complexes for DNA-binding transcription factors have been identified, including DRIP, TRAP, and ARC (reviewed in Bourbon, 2004, Blazek, 2005, Conaway, 2005, and Malik, 2005). The DRIP co-activator complex was originally identified and named as a specific complex associated with the Vitamin D Receptor member of the nuclear receptor family of transcription factors (Rachez, 1998). Similarly, the TRAP co-activator complex was originally identified as a complex that associates with the thyroid receptor (Yuan, 1998). It was later determined that all of the components of the DRIP complex are also present in the TRAP complex, and the ARC complex (discussed further below). For example, the DRIP205 and TRAP220 proteins were show to be identical, as were specific pairs of the other components of these complexes (Rachez, 1999).

In addition, these various transcription co-activator proteins identified in mammalian cells were found to be the orthologues or homologues of the Mediator ("adaptor") complex proteins (reviewed in Bourbon, 2004). The Mediator proteins were originally identified in yeast by Kornberg and colleagues, as complexes associated with DNA polymerase (Kelleher, 1990). In higher organisms, Adapter complexes bridge between the basal transcription factors (including Pol II) and tissue-specific transcription factors (TFs) bound to sites within upstream Proximal Promoter regions or distal Enhancer regions (Figure 1). However, many of the Mediator homologues can also be found in complexes associated with specific transcription factors in higher organisms. A unified nomenclature system for these adapter / co-activator proteins now labels them Mediator 1 through Mediator 31 (Bourbon, 2004). For example, the DRIP205 / TRAP220 proteins are now identified as Mediator 1 (Rachez, 1999), based on homology with yeast Mediator 1.

Example Pathway: Specific Regulation of Target Genes During Notch Signaling:

One well-studied example of cell-specific regulation of gene transcription is selective regulation of target genes during Notch signaling. Notch signaling was first identified in Drosophila, where it has been studied in detail at the genetic, molecular, biochemical and cellular levels (reviewed in Justice, 2002; Bray, 2006; Schweisguth, 2004; Louvri, 2006). In Drosophila, Notch signaling to the nucleus is thought always to be mediated by one specific DNA binding transcription factor, Suppressor of Hairless. In mammals, the homologous genes are called CBF1 (or RBPJkappa), while in worms they are called Lag-1, so that the acronym "CSL" has been given to this conserved transcription factor family. There are at least two human CSL homologues, which are now named RBPJ and RBPJL.

In Drosophila, Su(H) is known to be bifunctional, in that it represses target gene transcription in the absence of Notch signaling, but activates target genes during Notch signaling. At least some of the mammalian CSL homologues are believed also to be bifunctional, and to mediate target gene repression in the absence of Notch signaling, and activation in the presence of Notch signaling.

Notch Co-Activator and Co-Repressor complexes: This repression is mediated by at least one specific co-repressor complexes (Co-R) bound to CSL in the absence of Notch signaling. In Drosophila, this co-repressor complex consists of at least three distinct co-repressor proteins: Hairless, Groucho, and dCtBP (Drosophila C-terminal Binding Protein). Hairless has been show to bind directly to Su(H), and Groucho and dCtBP have been shown to bind directly to Hairless (Barolo, 2002). All three of the co-repressor proteins have been shown to be necessary for proper gene regulation during Notch signaling in vivo (Nagel, 2005).

In mammals, the same general pathway and mechanisms are observed, where CSL proteins are bifunctional DNA binding transcription factors (TFs), that bind to Co-Repressor complexes to mediate repression in the absence of Notch signaling, and bind to Co-Activator complexes to mediate activation in the presence of Notch signaling. However, in mammals, there may be multiple co-repressor complexes, rather than the single Hairless co-repressor complex that has been observed in Drosophila.

During Notch signaling in all systems, the Notch transmembrane receptor is cleaved and the Notch intracellular domain (NICD) translocates to the nucleus, where it there functions as a specific transcription co-activator for CSL proteins. In the nucleus, NICD replaces the Co-R complex bound to CSL, thus resulting in de-repression of Notch target genes in the nucleus (Figure 2). Once bound to CSL, NICD and CSL proteins recruit an additional co-activator protein, Mastermind, to form a CSL-NICD-Mam ternary co-activator (Co-A) complex. This Co-R complex was initially thought to be sufficient to mediate activation of at least some Notch target genes. However, there now is evidence that still other co-activators and additional DNA-binding transcription factors are required in at least some contexts (reviewed in Barolo, 2002).

Thus, CSL is a good example of a bifunctional DNA-binding transcription factor that mediates repression of specific targets genes in one context, but activation of the same targets in another context. This bifunctionality is mediated by the association of specific Co-Repressor complexes vs. specific Co-Activator complexes in different contexts, namely in the absence or presence of Notch signaling.

References

Näär AM, Lemon BD & Tjian R (2001). Transcriptional coactivator complexes. Annu Rev Biochem, 70, 475-501.

Kadonaga JT (2004). Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. Cell, 116, 247-57.

Maston GA, Evans SK & Green MR (2006). Transcriptional regulatory elements in the human genome. Annu Rev Genomics Hum Genet, 7, 29-59.

Barolo S & Posakony JW (2002). Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. Genes Dev, 16, 1167-81.

Roeder RG (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. FEBS Lett, 579, 909-15.

Edit history

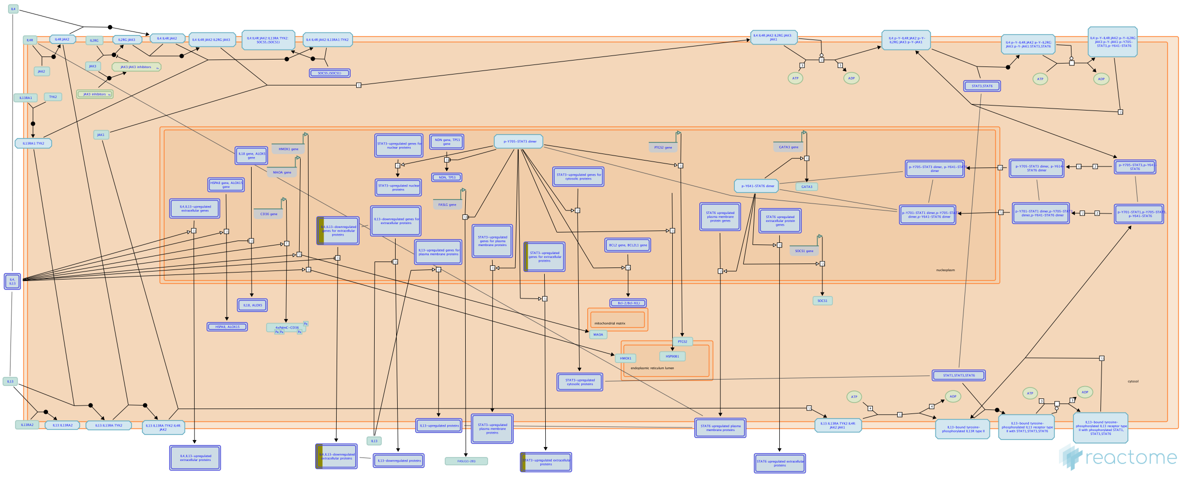
Date	Action	Author
2008-02-09	Created	Caudy M
2008-02-26	Reviewed	Freedman LP
2021-03-06	Modified	Shorser S

Entities found in this pathway (5)

Input	UniProt Id	Input	UniProt Id	Input	UniProt Id
CTSL	O60911, P07711	EGFR	P00533	HDAC2	Q92769
IL6	P05231	VEGFA	P15692		

Input	Ensembl Id	Input	Ensembl Id	Input	Ensembl Id
EGFR	ENSG00000146648	IL6	ENSG00000136244	VEGFA	ENSG00000112715

14. Interleukin-4 and Interleukin-13 signaling (R-HSA-6785807)



Interleukin-4 (IL4) is a principal regulatory cytokine during the immune response, crucially important in allergy and asthma (Nelms et al. 1999). When resting T cells are antigen-activated and expand in response to Interleukin-2 (IL2), they can differentiate as Type 1 (Th1) or Type 2 (Th2) T helper cells. The outcome is influenced by IL4. Th2 cells secrete IL4, which both stimulates Th2 in an autocrine fashion and acts as a potent B cell growth factor to promote humoral immunity (Nelms et al. 1999).

Interleukin-13 (IL13) is an immunoregulatory cytokine secreted predominantly by activated Th2 cells. It is a key mediator in the pathogenesis of allergic inflammation. IL13 shares many functional properties with IL4, stemming from the fact that they share a common receptor subunit. IL13 receptors are expressed on human B cells, basophils, eosinophils, mast cells, endothelial cells, fibroblasts, monocytes, macrophages, respiratory epithelial cells, and smooth muscle cells, but unlike IL4, not T cells. Thus IL13 does not appear to be important in the initial differentiation of CD4 T cells into Th2 cells, rather it is important in the effector phase of allergic inflammation (Hershey et al. 2003).

IL4 and IL13 induce “alternative activation” of macrophages, inducing an anti-inflammatory phenotype by signaling through IL4R alpha in a STAT6 dependent manner. This signaling plays an important role in the Th2 response, mediating anti-parasitic effects and aiding wound healing (Gordon & Martinez 2010, Loke et al. 2002)

There are two types of IL4 receptor complex (Andrews et al. 2006). Type I IL4R (IL4R1) is predominantly expressed on the surface of hematopoietic cells and consists of IL4R and IL2RG, the common gamma chain. Type II IL4R (IL4R2) is predominantly expressed on the surface of nonhematopoietic cells, it consists of IL4R and IL13RA1 and is also the type II receptor for IL13. (Obiri et al. 1995, Aman et al. 1996, Hilton et al. 1996, Miloux et al. 1997, Zhang et al. 1997). The second receptor for IL13 consists of IL4R and Interleukin-13 receptor alpha 2 (IL13RA2), sometimes called Interleukin-13 binding protein (IL13BP). It has a high affinity receptor for IL13 (Kd = 250 pmol/L) but is not sufficient to render cells responsive to IL13, even in the presence of IL4R (Donaldson et al. 1998). It is reported to exist in soluble form (Zhang et al. 1997) and when overexpressed reduces JAK-STAT signaling (Kawakami et al. 2001). It's function may be to prevent IL13 signalling via the functional IL4R:IL13RA1 receptor. IL13RA2 is overexpressed and enhances cell invasion in some human cancers (Joshi & Puri 2012).

The first step in the formation of IL4R1 (IL4:IL4R:IL2RB) is the binding of IL4 with IL4R (Hoffman et al. 1995, Shen et al. 1996, Hage et al. 1999). This is also the first step in formation of IL4R2 (IL4:IL4R:IL13RA1). After the initial binding of IL4 and IL4R, IL2RB binds (LaPorte et al. 2008), to form IL4R1. Alternatively, IL13RA1 binds, forming IL4R2. In contrast, the type II IL13 complex (IL13R2) forms with IL13 first binding to IL13RA1 followed by recruitment of IL4R (Wang et al. 2009).

Crystal structures of the IL4:IL4R:IL2RG, IL4:IL4R:IL13RA1 and IL13:IL4R:IL13RA1 complexes have been determined (LaPorte et al. 2008). Consistent with these structures, in monocytes IL4R is tyrosine phosphorylated in response to both IL4 and IL13 (Roy et al. 2002, Gordon & Martinez 2010) while IL13RA1 phosphorylation is induced only by IL13 (Roy et al. 2002, LaPorte et al. 2008) and IL2RG phosphorylation is induced only by IL4 (Roy et al. 2002).

Both IL4 receptor complexes signal through Jak/STAT cascades. IL4R is constitutively-associated with JAK2 (Roy et al. 2002) and associates with JAK1 following binding of IL4 (Yin et al. 1994) or IL13 (Roy et al. 2002). IL2RG constitutively associates with JAK3 (Boussiotis et al. 1994, Russell et al. 1994). IL13RA1 constitutively associates with TYK2 (Umeshita-Suyama et al. 2000, Roy et al. 2002, LaPorte et al. 2008, Bhattacharjee et al. 2013).

IL4 binding to IL4R1 leads to phosphorylation of JAK1 (but not JAK2) and STAT6 activation (Takeda et al. 1994, Ratthe et al. 2007, Bhattacharjee et al. 2013).

IL13 binding increases activating tyrosine-99 phosphorylation of IL13RA1 but not that of IL2RG. IL4 binding to IL2RG leads to its tyrosine phosphorylation (Roy et al. 2002). IL13 binding to IL4R2 leads to TYK2 and JAK2 (but not JAK1) phosphorylation (Roy & Cathcart 1998, Roy et al. 2002).

Phosphorylated TYK2 binds and phosphorylates STAT6 and possibly STAT1 (Bhattacharjee et al. 2013).

A second mechanism of signal transduction activated by IL4 and IL13 leads to the insulin receptor substrate (IRS) family (Kelly-Welch et al. 2003). IL4R1 associates with insulin receptor substrate 2 and activates the PI3K/Akt and Ras/MEK/Erk pathways involved in cell proliferation, survival and translational control. IL4R2 does not associate with insulin receptor substrate 2 and consequently the PI3K/Akt and Ras/MEK/Erk pathways are not activated (Busch-Dienstfertig & González-Rodríguez 2013).

References

Nelms K, Keegan AD, Zamorano J, Ryan JJ & Paul WE (1999). The IL-4 receptor: signaling mechanisms and biologic functions. Annu. Rev. Immunol., 17, 701-38.

Hershey GK (2003). IL-13 receptors and signaling pathways: an evolving web. J. Allergy Clin. Immunol., 111, 677-90; quiz 691.

Edit history

Date	Action	Author
2015-07-01	Authored	Jupe S
2015-07-01	Created	Jupe S
2016-09-02	Edited	Jupe S
2016-09-02	Reviewed	Leibovich SJ

Date	Action	Author
2021-03-12	Modified	Shorser S

Entities found in this pathway (2)

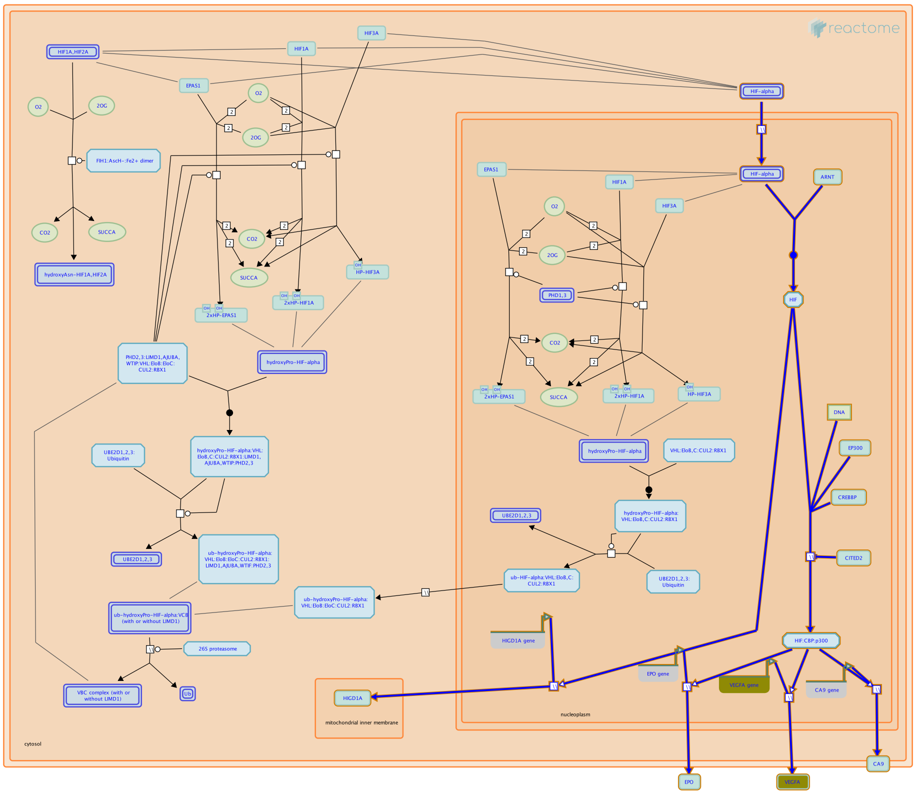
Input	UniProt Id
IL6	P05231

Input	UniProt Id
VEGFA	P15692

Input	Ensembl Id
IL6	ENSG00000136244

Input	Ensembl Id
VEGFA	ENSG00000112715

15. Regulation of gene expression by Hypoxia-inducible Factor (R-HSA-1234158)



Cellular compartments: nucleoplasm.

HIF-alpha (HIF1A, HIF2A (EPAS1), HIF3A) is translocated to the nucleus, possibly by two pathways: importin 4/7 (Chachami et al. 2009) and importin alpha/beta (Depping et al. 2008). Once in the nucleus HIF-alpha heterodimerizes with HIF-beta (ARNT) (Wang et al. 1995, Jiang et al. 1996, Tian et al. 1997, Gu et al. 1998, Erbel et al. 2003) and recruits CBP and p300 to promoters of target genes (Ebert and Bunn 1998, Kallio et al. 1998, Ema et al. 1999, Gu et al. 2001, Dames et al. 2002, Freedman et al. 2002).

References

Chachami G, Paraskeva E, Mingot JM, Braliou GG, Görlich D & Simos G (2009). Transport of hypoxia-inducible factor HIF-1alpha into the nucleus involves importins 4 and 7. Biochem Biophys Res Commun, 390, 235-40.

Kallio PJ, Okamoto K, O'Brien S, Carrero P, Makino Y, Tanaka H & Poellinger L (1998). Signal transduction in hypoxic cells: inducible nuclear translocation and recruitment of the CBP/p300 coactivator by the hypoxia-inducible factor-1alpha. EMBO J, 17, 6573-86.

Depping R, Steinhoff A, Schindler SG, Friedrich B, Fagerlund R, Metzen E, ... Köhler M (2008). Nuclear translocation of hypoxia-inducible factors (HIFs): involvement of the classical importin alpha/beta pathway. Biochim Biophys Acta, 1783, 394-404.

Tian H, McKnight SL & Russell DW (1997). Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. Genes Dev, 11, 72-82.

Jiang BH, Rue E, Wang GL, Roe R & Semenza GL (1996). Dimerization, DNA binding, and transactivation properties of hypoxia-inducible factor 1. J Biol Chem, 271, 17771-8.

Edit history

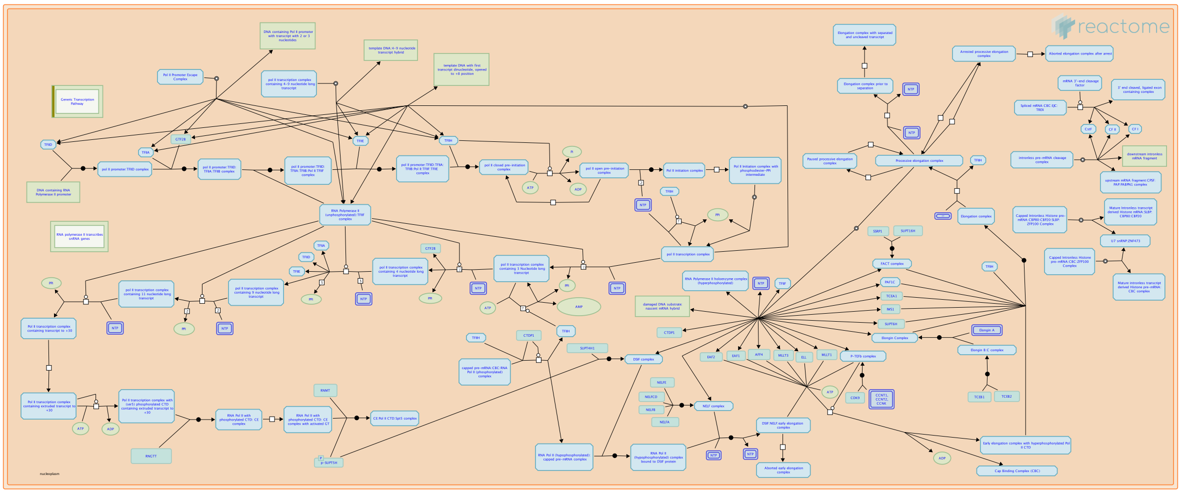
Date	Action	Author
2011-03-09	Edited	May B
2011-03-09	Authored	May B
2011-03-18	Created	May B
2012-05-19	Reviewed	Rantanen K
2021-03-12	Modified	Shorser S

Entities found in this pathway (1)

Input	UniProt Id
VEGFA	P15692

Input	Ensembl Id
VEGFA	ENSG00000112715

16. RNA Polymerase II Transcription (R-HSA-73857)



Cellular compartments: nucleoplasm.

RNA polymerase II (Pol II) is the central enzyme that catalyses DNA- directed mRNA synthesis during the transcription of protein-coding genes. Pol II consists of a 10-subunit catalytic core, which alone is capable of elongating the RNA transcript, and a complex of two subunits, Rpb4/7, that is required for transcription initiation.

The transcription cycle is divided in three major phases: initiation, elongation, and termination. Transcription initiation include promoter DNA binding, DNA melting, and initial synthesis of short RNA transcripts. The transition from initiation to elongation, is referred to as promoter escape and leads to a stable elongation complex that is characterized by an open DNA region or transcription bubble. The bubble contains the DNA-RNA hybrid, a heteroduplex of eight to nine base pairs. The growing 3-end of the RNA is engaged with the polymerase complex active site. Ultimately transcription terminates and Pol II dissociates from the template.

References

Cramer P (2004). Structure and function of RNA polymerase II. Adv. Protein Chem., 67, 1-42.

Edit history

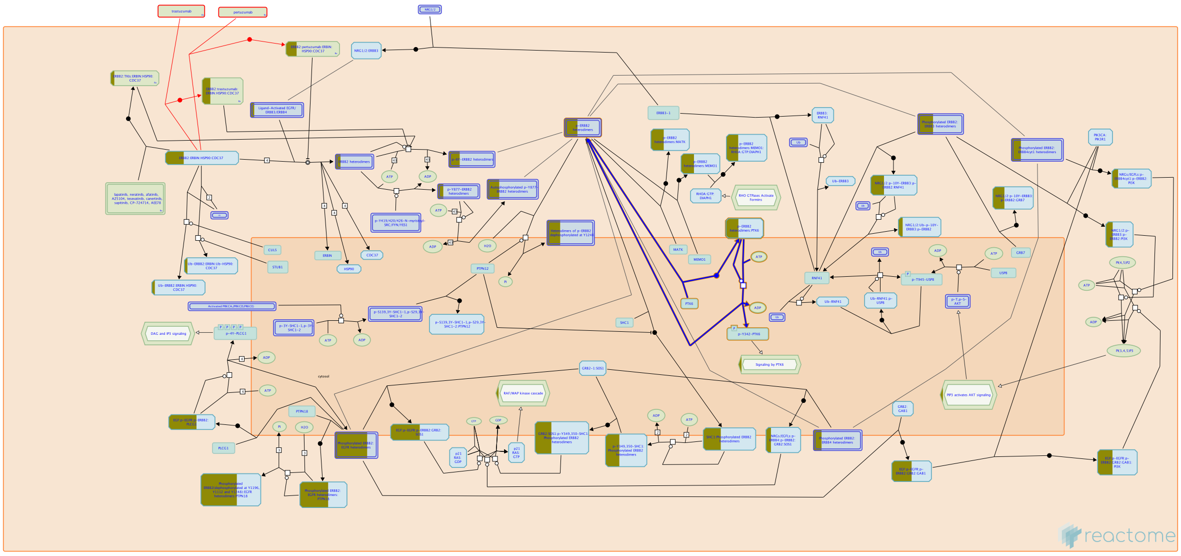
Date	Action	Author
2005-01-12	Created	Reinberg D, Timmers HTM, Conaway JW, Proudfoot NJ, Conaway RC
2021-03-06	Modified	Shorser S

Entities found in this pathway (5)

Input	UniProt Id	Input	UniProt Id	Input	UniProt Id
CTSL	O60911, P07711	EGFR	P00533	HDAC2	Q92769
IL6	P05231	VEGFA	P15692		

Input	Ensembl Id	Input	Ensembl Id	Input	Ensembl Id
EGFR	ENSG00000146648	IL6	ENSG00000136244	VEGFA	ENSG00000112715

17. ERBB2 Activates PTK6 Signaling (R-HSA-8847993)



PTK6 (BRK) is activated downstream of ERBB2 (HER) (Xiang et al. 2008, Peng et al. 2015) and other receptor tyrosine kinases, such as EGFR (Kamalati et al. 1996) and MET (Castro and Lange 2010). However, it is not clear if MET and EGFR activate PTK6 directly or act through ERBB2, since it is known that ERBB2 forms heterodimers with EGFR (Spivak-Kroizman et al. 1992), and MET can heterodimerize with both EGFR and ERBB2 (Tanizaki et al. 2011).

References

Xiang B, Chatti K, Qiu H, Lakshmi B, Krasnitz A, Hicks J, ... Muthuswamy SK (2008). Brk is coamplified with ErbB2 to promote proliferation in breast cancer. Proc. Natl. Acad. Sci. U.S.A., 105, 12463-8.

Peng M, Ball-Kell SM & Tyner AL (2015). Protein tyrosine kinase 6 promotes ERBB2-induced mammary gland tumorigenesis in the mouse. Cell Death Dis, 6, e1848.

Kamalati T, Jolin HE, Mitchell PJ, Barker KT, Jackson LE, Dean CJ, ... Crompton MR (1996). Brk, a breast tumor-derived non-receptor protein-tyrosine kinase, sensitizes mammary epithelial cells to epidermal growth factor. J. Biol. Chem., 271, 30956-63.

Castro NE & Lange CA (2010). Breast tumor kinase and extracellular signal-regulated kinase 5 mediate Met receptor signaling to cell migration in breast cancer cells. Breast Cancer Res., 12, R60.

Tanizaki J, Okamoto I, Sakai K & Nakagawa K (2011). Differential roles of trans-phosphorylated EGFR, HER2, HER3, and RET as heterodimerisation partners of MET in lung cancer with MET amplification. Br. J. Cancer, 105, 807-13.

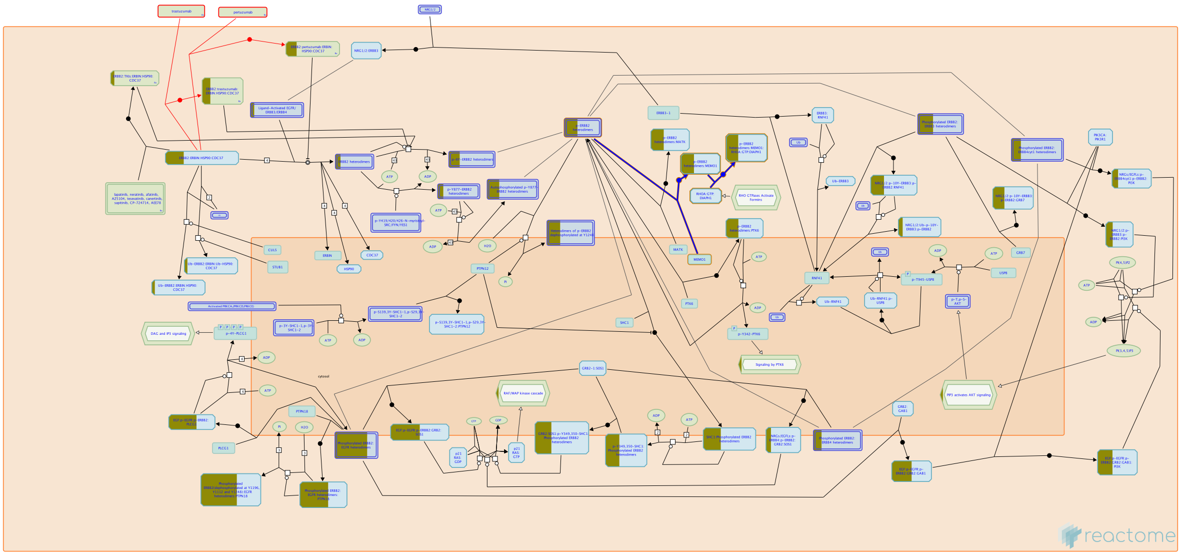
Edit history

Date	Action	Author
2015-12-04	Created	Orlic-Milacic M
2016-01-05	Edited	Orlic-Milacic M
2016-01-05	Authored	Orlic-Milacic M
2016-02-07	Reviewed	Pires IM
2021-03-12	Modified	Shorser S

Entities found in this pathway (1)

Input	UniProt Id
EGFR	P00533, P04626

18. ERBB2 Regulates Cell Motility (R-HSA-6785631)



Activated ERBB2 heterodimers regulate cell motility through association with MEMO1. MEMO1 retains activated RHOA GTPase and its associated protein DIAPH1 at the plasma membrane, thus linking ERBB2 activation with the microtubule and actin dynamics downstream of the RHOA:GTP:DIAPH1 complex (Marone et al. 2004, Qiu et al. 2008, Zaoui et al. 2008, Zaoui et al. 2010).

References

Qiu C, Lienhard S, Hynes NE, Badache A & Leahy DJ (2008). Memo is homologous to nonheme iron dioxygenases and binds an ErbB2-derived phosphopeptide in its vestigial active site. J. Biol. Chem., 283, 2734-40.

Zaoui K, Benseddik K, Daou P, Salaün D & Badache A (2010). ErbB2 receptor controls microtubule capture by recruiting ACF7 to the plasma membrane of migrating cells. Proc. Natl. Acad. Sci. U.S.A., 107, 18517-22.

Marone R, Hess D, Dankort D, Muller WJ, Hynes NE & Badache A (2004). Memo mediates ErbB2-driven cell motility. Nat. Cell Biol., 6, 515-22.

Zaoui K, Honoré S, Isnardon D, Braguer D & Badache A (2008). Memo-RhoA-mDia1 signaling controls microtubules, the actin network, and adhesion site formation in migrating cells. J. Cell Biol., 183, 401-8.

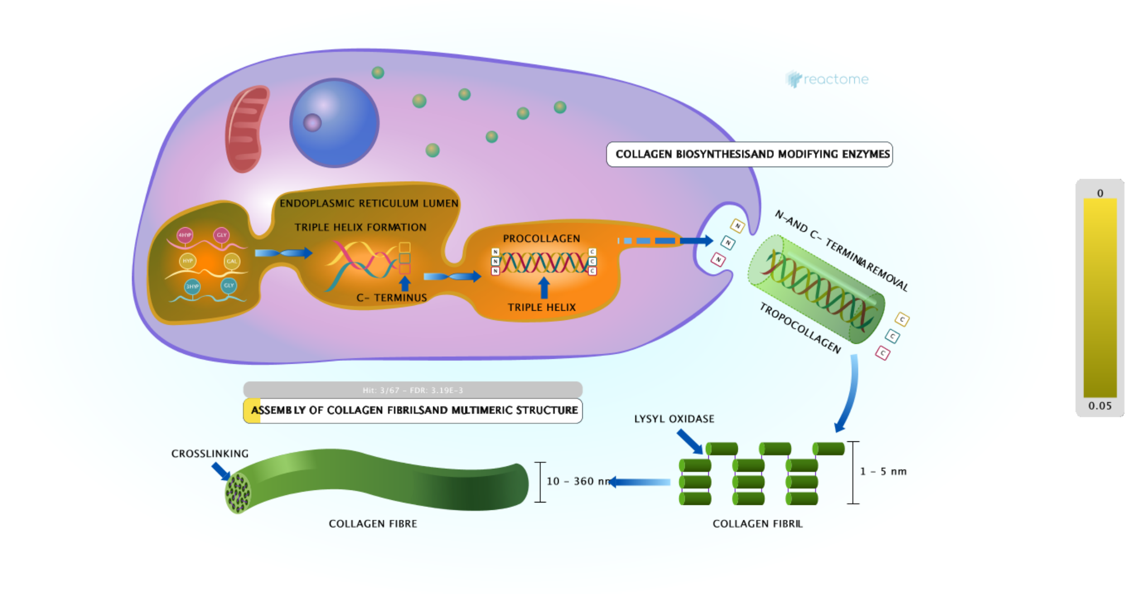
Edit history

Date	Action	Author
2015-06-29	Created	Orlic-Milacic M
2016-01-28	Edited	Orlic-Milacic M
2016-01-28	Authored	Orlic-Milacic M
2016-02-03	Reviewed	Badache A
2021-03-12	Modified	Shorser S

Entities found in this pathway (1)

Input	UniProt Id
EGFR	P00533, P04626

19. Collagen formation (R-HSA-1474290)



Collagen is a family of at least 29 structural proteins derived from over 40 human genes (Myllyharju & Kivirikko 2004). It is the main component of connective tissue, and the most abundant protein in mammals making up about 25% to 35% of whole-body protein content. A defining feature of collagens is the formation of trimeric left-handed polyproline II-type helical collagenous regions. The packing within these regions is made possible by the presence of the smallest amino acid, glycine, at every third residue, resulting in a repeating motif Gly-X-Y where X is often proline (Pro) and Y often 4-hydroxyproline (4Hyp). Gly-Pro-Hyp is the most common triplet in collagen (Ramshaw et al. 1998). Collagen peptide chains also have non-collagenous domains, with collagen subclasses having common chain structures. Collagen fibrils are mostly found in fibrous tissues such as tendon, ligament and skin. Other forms of collagen are abundant in cornea, cartilage, bone, blood vessels, the gut, and intervertebral disc. In muscle tissue, collagen is a major component of the endomysium, constituting up to 6% of muscle mass. Gelatin, used in food and industry, is collagen that has been irreversibly hydrolyzed.

On the basis of their fibre architecture in tissues, the genetically distinct collagens have been divided into subgroups. Group 1 collagens have uninterrupted triple-helical domains of about 300 nm, forming large extracellular fibrils. They are referred to as the fibril-forming collagens, consisting of collagens types I, II, III, V, XI, XXIV and XXVII. Group 2 collagens are types IV and VII, which have extended triple helices (>350 nm) with imperfections in the Gly-X-Y repeat sequences. Group 3 are the short-chain collagens. These have two subgroups. Group 3A have continuous triple-helical domains (type VI, VIII and X). Group 3B have interrupted triple-helical domains, referred to as the fibril-associated collagens with interrupted triple helices (FACIT collagens, Shaw & Olsen 1991). FACITs include collagen IX, XII, XIV, XVI, XIX, XX, XXI, XXII and XXVI plus the transmembrane collagens (XIII, XVII, XXIII and XXV) and the multiple triple helix domains and interruptions (Multiplexin) collagens XV and XVIII (Myllyharju & Kivirikko 2004). The non-collagenous domains of collagens have regulatory functions; several are biologically active when cleaved from the main peptide chain. Fibrillar collagen peptides all have a large triple helical domain (COL1) bordered by N and C terminal extensions, called the N- and C-propeptides, which are cleaved prior to formation of the collagen fibril. The intact form is referred to as a collagen propeptide, not procollagen, which is used to refer to the trimeric triple-helical precursor of collagen before the propeptides are removed. The C-propeptide, also called the NC1 domain, directs chain association during assembly of the procollagen molecule from its three constituent alpha chains (Hulmes 2002).

Fibril forming collagens are the most familiar and best studied subgroup. Collagen fibres are aggregates or bundles of collagen fibrils, which are themselves polymers of tropocollagen complexes, each consisting of three polypeptide chains known as alpha chains. Tropocollagens are considered the subunit of larger collagen structures. They are approximately 300 nm long and 1.5 nm in diameter, with a left-handed triple-helical structure, which becomes twisted into a right-handed coiled-coil 'super helix' in the collagen fibril. Tropocollagens in the extracellular space polymerize spontaneously with regularly staggered ends (Hulmes 2002). In fibrillar collagens the molecules are staggered by about 67 nm, a unit known as D that changes depending upon the hydration state. Each D-period contains slightly more than four collagen molecules so that every D-period repeat of the microfibril has a region containing five molecules in cross-section, called the 'overlap', and a region containing only four molecules, called the 'gap'. The triple-helices are arranged in a hexagonal or quasi-hexagonal array in cross-section, in both the gap and overlap regions (Orgel et al. 2006). Collagen molecules cross-link covalently to each other via lysine and hydroxylysine side chains. These cross-links are unusual, occuring only in collagen and elastin, a related protein.

The macromolecular structures of collagen are diverse. Several group 3 collagens associate with larger collagen fibers, serving as molecular bridges which stabilize the organization of the extracellular matrix. Type IV collagen is arranged in an interlacing network within the dermal-epidermal junction and vascular basement membranes. Type VI collagen forms distinct microfibrils called beaded filaments. Type VII collagen forms anchoring fibrils. Type VIII and X collagens form hexagonal networks. Type XVII collagen is a component of hemidesmosomes where it is complexed wtih alpha6Beta4 integrin, plectin, and laminin-332 (de Pereda et al. 2009). Type XXIX collagen has been recently reported to be a putative epidermal collagen with highest expression in suprabasal layers (Soderhall et al. 2007). Collagen fibrils/aggregates arranged in varying combinations and concentrations in different tissues provide specific tissue properties. In bone, collagen triple helices lie in a parallel, staggered array with 40 nm gaps between the ends of the tropocollagen subunits, which probably serve as nucleation sites for the deposition of crystals of the mineral component, hydroxyapatite (Ca10(PO4)6(OH)2) with some phosphate. Collagen structure affects cell-cell and cell-matrix communication, tissue construction in growth and repair, and is changed in development and disease (Sweeney et al. 2006, Twardowski et al. 2007). A single collagen fibril can be heterogeneous along its axis, with significantly different mechanical properties in the gap and overlap regions, correlating with the different molecular organizations in these regions (Minary-Jolandan & Yu 2009).

References

Prockop DJ & Kivirikko KI (1995). Collagens: molecular biology, diseases, and potentials for therapy . Annu Rev Biochem, 64, 403-34.

Gordon MK & Hahn RA (2010). Collagens. Cell Tissue Res, 339, 247-57.

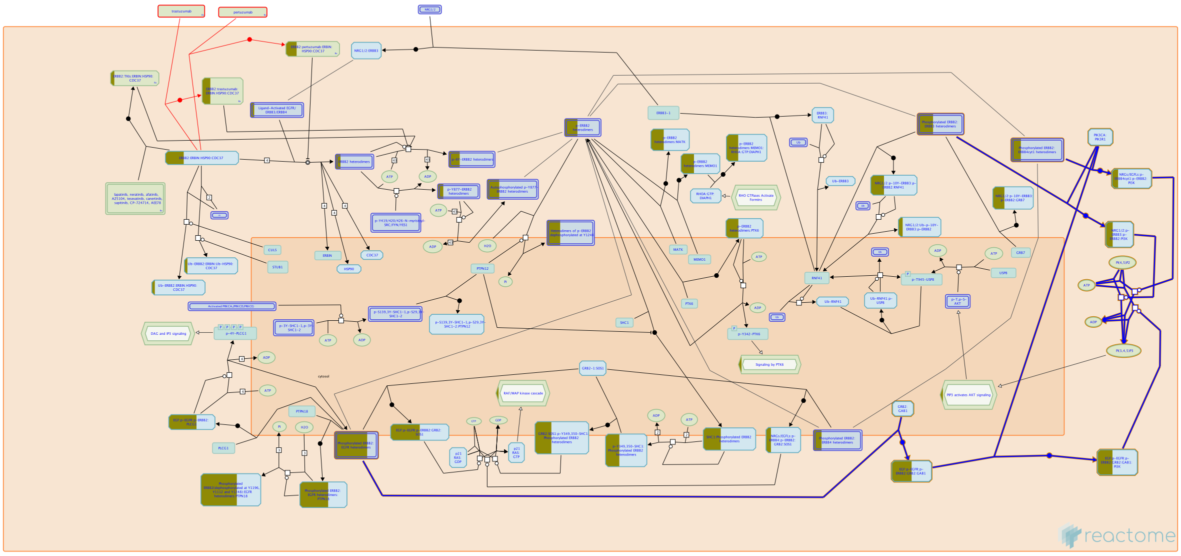
Edit history

Date	Action	Author
2011-08-05	Authored	Jupe S
2011-08-05	Created	Jupe S
2012-04-11	Edited	Jupe S
2012-05-24	Reviewed	Canty-Laird EG
2021-03-06	Modified	Shorser S

Entities found in this pathway (2)

Input	UniProt Id	Input	UniProt Id
CTSB	P07858	CTSL	O60911, P07711

20. PI3K events in ERBB2 signaling (R-HSA-1963642)



ERBB2:ERBB3 and ERBB2:ERBB4cyt1 heterodimers activate PI3K signaling by direct binding of PI3K regulatory subunit p85 (Yang et al. 2007, Cohen et al. 1996, Kaushansky et al. 2008) to phosphorylated tyrosine residues in the C-tail of ERBB3 (Y1054, Y1197, Y1222, Y1224, Y1276 and Y1289) and ERBB4 CYT1 isoforms (Y1056 in JM-A CYT1 isoform and Y1046 in JM-B CYT1 isoform). Regulatory subunit p85 subsequently recruits catalytic subunit p110 of PI3K, resulting in the formation of active PI3K, conversion of PIP2 to PIP3, and PIP3-mediated activation of AKT signaling (Junttila et al. 2009, Kainulainen et al. 2000). Heterodimers of ERBB2 and EGFR recruit PI3K indirectly, through GRB2:GAB1 complex (Jackson et al. 2004), which again leads to PIP3-mediated activation of AKT signaling.

References

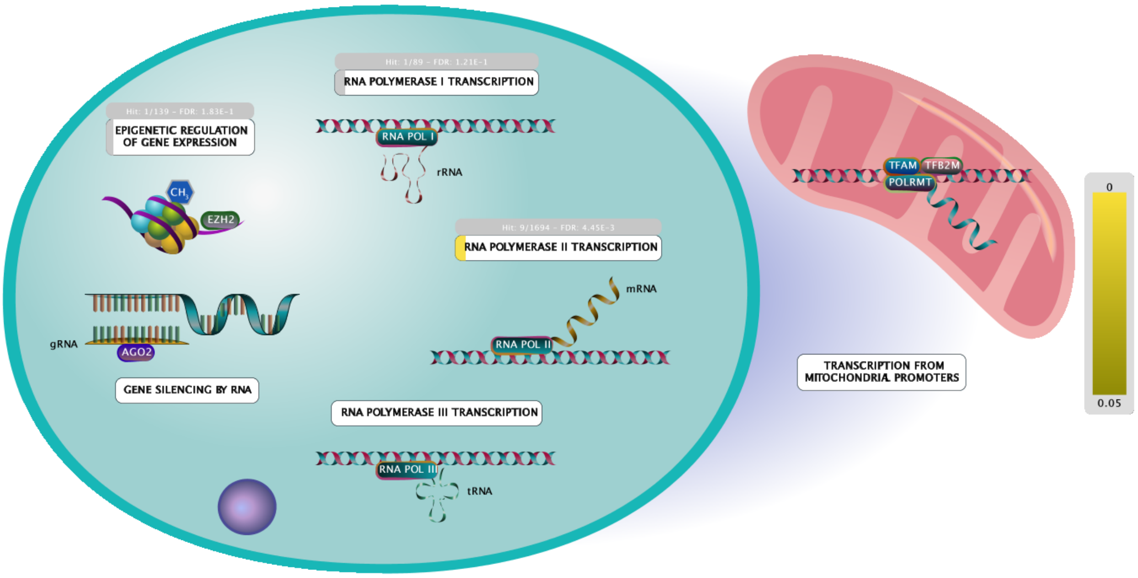
Edit history

Date	Action	Author
2011-11-04	Authored	Orlic-Milacic M
2011-11-04	Created	Orlic-Milacic M
2011-11-07	Edited	Matthews L, D'Eustachio P
2011-11-11	Reviewed	Xu W, Neckers LM
2021-03-12	Modified	Shorser S

Entities found in this pathway (1)

Input	UniProt Id
EGFR	P00533, P04626

21. Gene expression (Transcription) (R-HSA-74160)



Gene expression encompasses transcription and translation and the regulation of these processes. RNA Polymerase I Transcription produces the large preribosomal RNA transcript (45S pre-rRNA) that is processed to yield 18S rRNA, 28S rRNA, and 5.8S rRNA, accounting for about half the RNA in a cell. RNA Polymerase II transcription produces messenger RNAs (mRNA) as well as a subset of non-coding RNAs including many small nucleolar RNAs (snRNA) and microRNAs (miRNA). RNA Polymerase III Transcription produces transfer RNAs (tRNA), 5S RNA, 7SL RNA, and U6 snRNA. Transcription from mitochondrial promoters is performed by the mitochondrial RNA polymerase, POLRMT, to yield long transcripts from each DNA strand that are processed to yield 12S rRNA, 16S rRNA, tRNAs, and a few RNAs encoding components of the electron transport chain. Regulation of gene expression can be divided into epigenetic regulation, transcriptional regulation, and post-transcription regulation (comprising translational efficiency and RNA stability). Epigenetic regulation of gene expression is the result of heritable chemical modifications to DNA and DNA-binding proteins such as histones. Epigenetic changes result in altered chromatin complexes that influence transcription. Gene Silencing by RNA mostly occurs post-transcriptionally but can also affect transcription. Small RNAs originating from the genome (miRNAs) or from exogenous RNA (siRNAs) are processed and transferred to the RNA-induced silencing complex (RISC), which interacts with complementary RNA to cause cleavage, translational inhibition, or transcriptional inhibition.

References

Edit history

Date	Action	Author
2003-09-11	Authored	Larsson NG, Comai L, Gustafsson CM, Reinberg D, Timmers HTM et al.
2003-09-11	Created	Proudfoot NJ, Kornblihtt AR
2008-12-03	Authored	Caudy M, Proudfoot NJ, Kornblihtt AR, D'Eustachio P
2016-12-29	Revised	D'Eustachio P
2021-03-02	Edited	Joshi-Tope G
2021-03-02	Reviewed	Paule M, Willis I, Zhao X

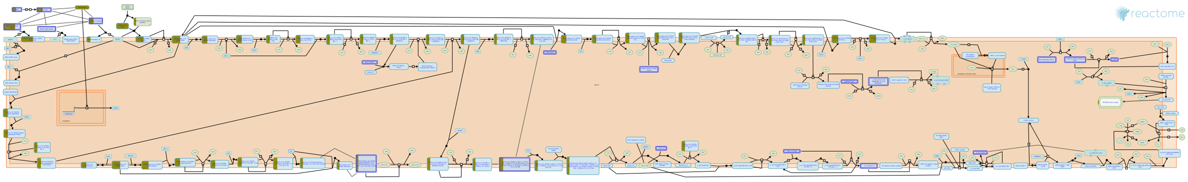
Date	Action	Author
2021-03-06	Modified	Shorser S

Entities found in this pathway (5)

Input	UniProt Id	Input	UniProt Id	Input	UniProt Id
CTSL	O60911, P07711	EGFR	P00533	HDAC2	Q92769
IL6	P05231	VEGFA	P15692		

Input	Ensembl Id	Input	Ensembl Id	Input	Ensembl Id
EGFR	ENSG00000146648	IL6	ENSG00000136244	VEGFA	ENSG00000112715

22. Signaling by VEGF (R-HSA-194138)



In normal development vascular endothelial growth factors (VEGFs) are crucial regulators of vascular development during embryogenesis (vasculogenesis) and blood-vessel formation in the adult (angiogenesis). In tumor progression, activation of VEGF pathways promotes tumor vascularization, facilitating tumor growth and metastasis. Abnormal VEGF function is also associated with inflammatory diseases including atherosclerosis, and hyperthyroidism. The members of the VEGF and VEGF-receptor protein families have distinct but overlapping ligand-receptor specificities, cell-type expression, and function. VEGF-receptor activation in turn regulates a network of signaling processes in the body that promote endothelial cell growth, migration and survival (Hicklin and Ellis, 2005; Shibuya and Claesson-Welsh, 2006).

Molecular features of the VGF signaling cascades are outlined in the figure below (from Olsson et al. 2006; Nature Publishing Group). Tyrosine residues in the intracellular domains of VEGF receptors 1, 2,and 3 are indicated by dark blue boxes; residues susceptible to phosphorylation are numbered. A circled R indicates that phosphorylation is regulated by cell state (VEGFR2), by ligand binding (VEGFR1), or by heterodimerization (VEGFR3). Specific phosphorylation sites (boxed numbers) bind signaling molecules (dark blue ovals), whose interaction with other cytosolic signaling molecules (light blue ovals) leads to specific cellular (pale blue boxes) and tissue-level (pink boxes) responses in vivo. Signaling cascades whose molecular details are unclear are indicated by dashed arrows. DAG, diacylglycerol; EC, endothelial cell; eNOS, endothelial nitric oxide synthase; FAK, focal adhesion kinase; HPC, hematopoietic progenitor cell; HSP27, heat-shock protein-27; MAPK, mitogen-activated protein kinase; MEK, MAPK and ERK kinase; PI3K, phosphatidylinositol 3' kinase; PKC, protein kinase C; PLCgamma, phospholipase C-gamma; Shb, SH2 and beta-cells; TSAd, T-cell-specific adaptor.

In the current release, the first events in these cascades - the interactions between VEGF proteins and their receptors - are annotated.

References

Shibuya M & Claesson-Welsh L (2006). Signal transduction by VEGF receptors in regulation of angiogenesis and lymphangiogenesis. Exp Cell Res, 312, 549-60.

Hicklin DJ & Ellis LM (2005). Role of the vascular endothelial growth factor pathway in tumor growth and angiogenesis. J Clin Oncol, 23, 1011-27.

Cross MJ, Dixelius J, Matsumoto T & Claesson-Welsh L (2003). VEGF-receptor signal transduction. Trends Biochem Sci, 28, 488-94.

Matsumoto T & Mugishima H (2006). Signal transduction via vascular endothelial growth factor (VEGF) receptors and their roles in atherogenesis. J Atheroscler Thromb, 13, 130-5.

Edit history

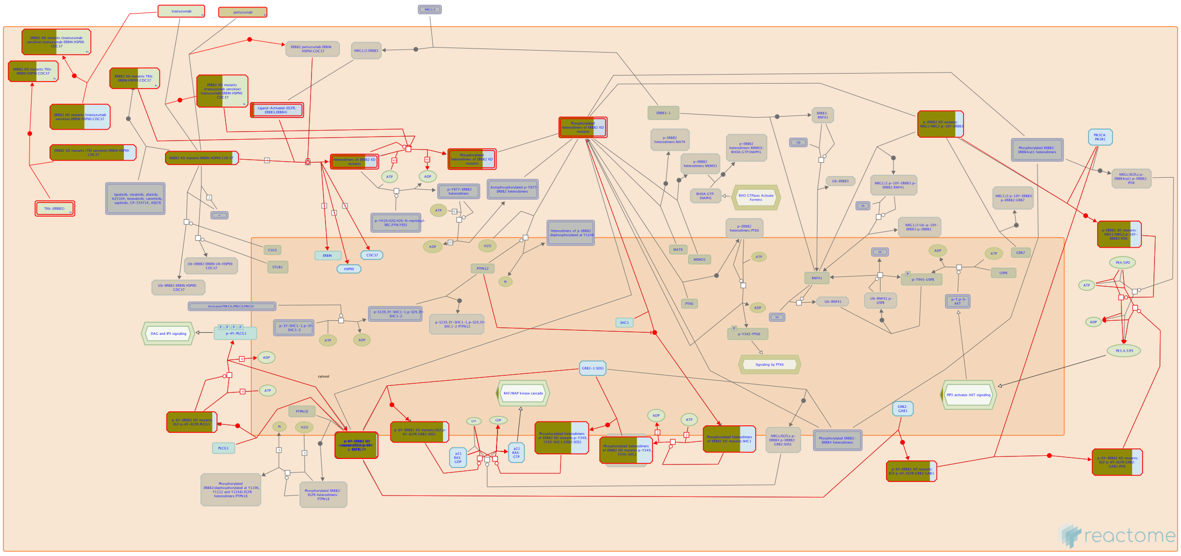
Date	Action	Author
2007-03-08	Created	Gopinathrao G

Date	Action	Author
2008-02-28	Reviewed	Claesson-Welsh L
2013-08-30	Edited	Garapati P V
2013-08-30	Authored	Garapati P V
2021-03-06	Modified	Shorser S

Entities found in this pathway (1)

Input	UniProt Id
VEGFA	P15692, P15692-4, P49765

23. Signaling by ERBB2 KD Mutants (R-HSA-9664565)



Diseases: cancer.

Mutations in the kinase domain (KD) of ERBB2 result in constitutive activation of ERBB2 signaling, facilitate heterodimerization of ERBB2 with other EGFR family members and increase the signaling intensity, leading to cellular transformation (Kancha et al. 2011).

Only a subset of potential heterodimerization partners has been tested for most ERBB2 KD mutant proteins, so our annotations here are correspondingly limited. ERBB2 L755S and ERBB2 V777L cancer variants were shown to heterodimerize with ERBB3 (HER3) at a higher rate than wild type ERBB2 (Croessmann et al. 2019). Increased activity of ERBB2 L755S, ERBB2 L755P, ERBB2 V777L, ERBB2 D769H, ERBB2 D769Y, ERBB2 V842I, ERBB2 R896C and ERBB2 P780_Y781insGSP in the presence of either EGFR (Kancha et al. 2011, Bose et al. 2013) or ERBB3 (Kancha et al. 2011, Bose et al. 2013, Collier et al. 2013) as a heterodimerization partner was also observed. The interplay of ERBB2 P780_Y781insGSP, ERBB2 I767M and ERBB2 R896C with ERBB3 has not been tested. ERBB2 L869R mutant shows increased activity in the presence of ERBB3, which is further augmented in the presence of dimerization-facilitating ERBB3 E928G mutants (Hanker et al. 2017). The interplay of ERBB2 L869R with EGFR has not been tested. Heterodimerization of ERBB2 KD mutants with ERBB4 has not been tested and ERBB4 is a candidate heterodimerization partner for these KD variants.

ERBB2 H878Y mutant has ten times higher kinase activity than the wild type ERBB2 (Hu, Wan et al. 2015; Hu, Hu et al. 2015), but its heterodimerization properties have not been studied and it is therefore annotated as a candidate.

Ligand requirements have not been studied in the context of heterodimerization of ERBB2 KD mutants, but it is assumed that ligands are required.

The signaling properties of ERBB2 L755M (Gonzalez-Alonso et al. 2015), ERBB2 L755W (COSMIC database: Forbes et al. 2017), ERBB2 V777E (Dietz et al. 2017), ERBB2 V777M (Lee et al. 2006, Ross et al. 2016, Zehir et al. 2017), ERBB2 D769N (Tschui et al. 2015), ERBB2 V842E (Siroy et al. 2015), ERBB2 R896H (Cancer Genome Atlas Research Network 2011), ERBB2 L869Q (Lee et al. 2006) and ERBB2 H878R (Trowe et al. 2008, Zehir et al. 2017) have not been experimentally tested, but they are predicted to be pathogenic (COSMIC database: Forbes et al. 2017) and they are annotated as candidates. ERBB2 T733I (Trowe et al. 2008), ERBB2 T798I (Trowe et al. 2008, Hanker et al. 2017) and ERBB2 T798M (Hanker et al. 2017) usually occur as secondary ERBB2 mutations and are responsible for treatment failure. On their own, ERBB2 T733I and ERBB2 T798I appear to be weakly transforming compared with the other ERBB2 KD mutants. As their signaling properties have been poorly studied, ERBB2 T733I, ERBB2 T798I and ERBB2 T798M are annotated as candidates.

The binding of ERBB2 KD mutants to ERBIN and the HSP90:CDC37 chaperone:co-chaperone complex has not been tested but is assumed to occur similarly to the wild type ERBB2.

Signaling by ERBB2 KD mutants has been organized into subpathways based on the current knowledge of biology of these mutants (heterodimerization, downstream signaling, drug interaction) and on the sequence similarity of their mutations.

References

Croessmann S, Formisano L, Kinch LN, Gonzalez-Ericsson PI, Sudhan DR, Nagy RJ, ... Arteaga CL (2019). Combined Blockade of Activating <i>ERBB2</i> Mutations and ER Results in Synthetic Lethality of ER+/HER2 Mutant Breast Cancer. Clin. Cancer Res., 25, 277-289.

Bose R, Kavuri SM, Searleman AC, Shen W, Shen D, Koboldt DC, ... Ellis MJ (2013). Activating HER2 mutations in HER2 gene amplification negative breast cancer. Cancer Discov, 3, 224-37.

González-Alonso P, Chamizo C, Moreno V, Madoz-Gúrpide J, Carvajal N, Daoud L, ... Rojo F (2015). Pyrosequencing-Based Assays for Rapid Detection of HER2 and HER3 Mutations in Clinical Samples Uncover an E332E Mutation Affecting HER3 in Retroperitoneal Leiomyosarcoma. Int J Mol Sci, 16, 19447-57.

Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, ... Campbell PJ (2017). COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res., 45, D777-D783.

Collier TS, Diraviyam K, Monsey J, Shen W, Sept D & Bose R (2013). Carboxyl group footprinting mass spectrometry and molecular dynamics identify key interactions in the HER2-HER3 receptor tyrosine kinase interface. J. Biol. Chem., 288, 25254-64.

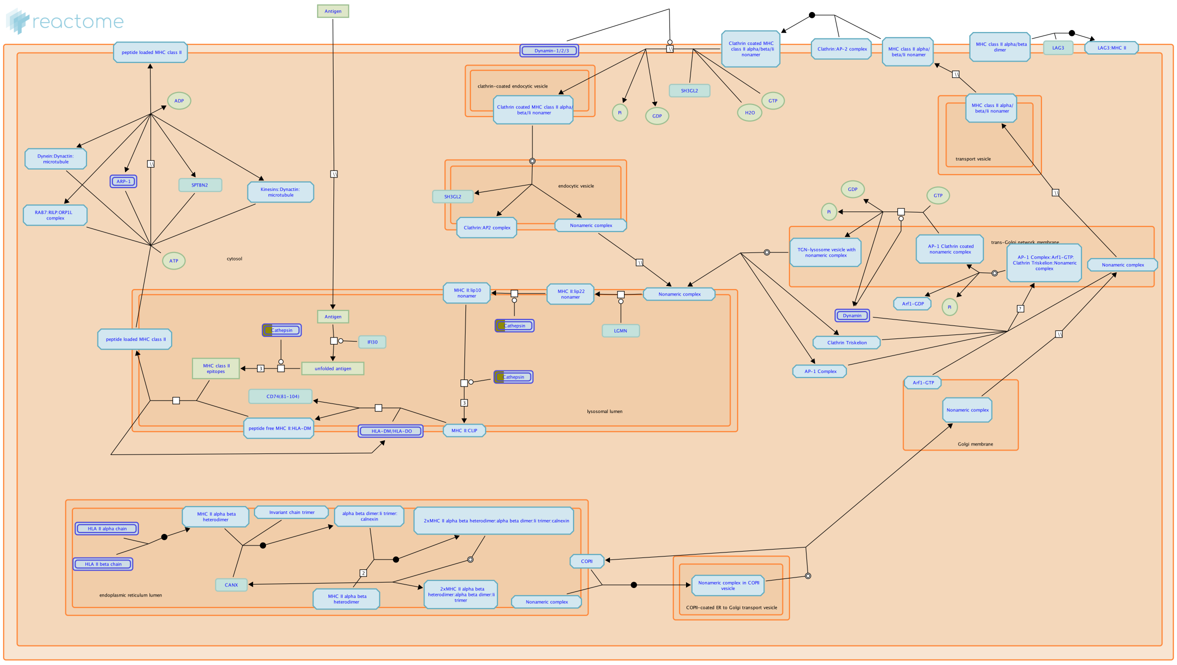
Edit history

Date	Action	Author
2019-07-31	Authored	Orlic-Milacic M
2019-09-16	Reviewed	Kancha RK
2019-10-25	Created	Orlic-Milacic M
2019-11-01	Edited	Orlic-Milacic M
2021-03-12	Modified	Shorser S

Entities found in this pathway (1)

Input	UniProt Id
EGFR	P00533, P04626

24. MHC class II antigen presentation (R-HSA-2132295)



Antigen presenting cells (APCs) such as B cells, dendritic cells (DCs) and monocytes/macrophages express major histocompatibility complex class II molecules (MHC II) at their surface and present exogenous antigenic peptides to CD4+ T helper cells. CD4+ T cells play a central role in immune protection. On their activation they stimulate differentiation of B cells into antibody-producing B-cell blasts and initiate adaptive immune responses. MHC class II molecules are transmembrane glycoprotein heterodimers of alpha and beta subunits. Newly synthesized MHC II molecules present in the endoplasmic reticulum bind to a chaperone protein called invariant (Ii) chain. The binding of Ii prevents the premature binding of self antigens to the nascent MHC molecules in the ER and also guides MHC molecules to endocytic compartments. In the acidic endosomal environment, Ii is degraded in a stepwise manner, ultimately to free the class II peptide-binding groove for loading of antigenic peptides. Exogenous antigens are internalized by the APC by receptor mediated endocytosis, phagocytosis or pinocytosis into endocytic compartments of MHC class II positive cells, where engulfed antigens are degraded in a low pH environment by multiple acidic proteases, generating MHC class II epitopes. Antigenic peptides are then loaded into the class II ligand-binding groove. The resulting class II peptide complexes then move to the cell surface, where they are scanned by CD4+ T cells for specific recognition (Berger & Roche 2009, Zhou & Blum 2004, Watts 2004, Landsverk et al. 2009).

References

Landsverk OJ, Bakke O & Gregers TF (2009). MHC II and the endocytic pathway: regulation by invariant chain. Scand J Immunol, 70, 184-93. ↗

Wolf PR & Ploegh HL (1995). How MHC class II molecules acquire peptide cargo: biosynthesis and trafficking through the endocytic pathway. Annu Rev Cell Dev Biol, 11, 267-306. ↗

Jensen PE (1997). Peptide binding and antigen presentation by class II histocompatibility glycoproteins. Biopolymers, 43, 303-22. ↗

Villadangos JA (2001). Presentation of antigens by MHC class II molecules: getting the most out of them. Mol Immunol, 38, 329-46. ↗

Ramachandra L, Simmons D & Harding CV (2009). MHC molecules and microbial antigen processing in phagosomes. Curr Opin Immunol, 21, 98-104.

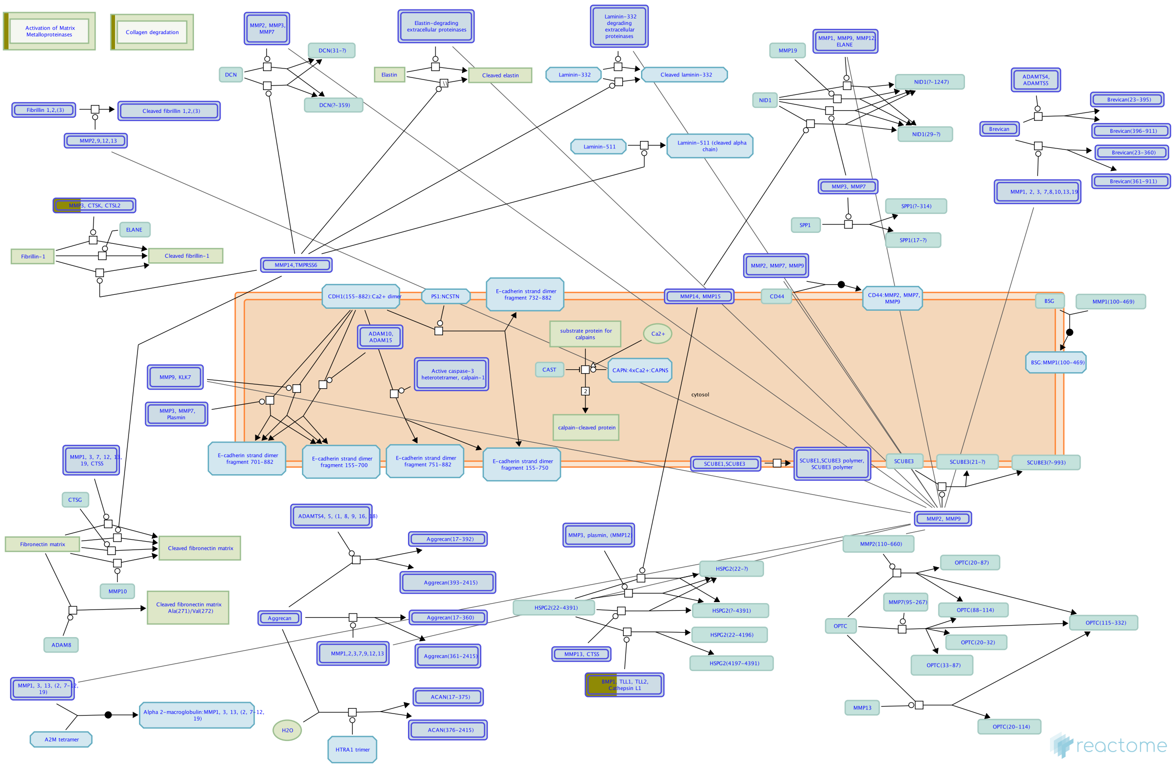
Edit history

Date	Action	Author
2012-02-21	Edited	Garapati P V
2012-02-21	Authored	Garapati P V
2012-02-21	Created	Garapati P V
2012-05-14	Reviewed	Neefjes J
2021-03-06	Modified	Shorser S

Entities found in this pathway (2)

Input	UniProt Id	Input	UniProt Id
CTSB	P07858	CTSL	O60911, P07711

25. Degradation of the extracellular matrix (R-HSA-1474228)



Matrix metalloproteinases (MMPs), previously referred to as matrixins because of their role in degradation of the extracellular matrix (ECM), are zinc and calcium dependent proteases belonging to the metzincin family. They contain a characteristic zinc-binding motif HEXXHXXGXXH (Stocker & Bode 1995) and a conserved Methionine which forms a Met-turn. Humans have 24 MMP genes giving rise to 23 MMP proteins, as MMP23 is encoded by two identical genes. All MMPs contain an N-terminal secretory signal peptide and a prodomain with a conserved PRCGXPD motif that in the inactive enzyme is localized with the catalytic site, the cysteine acting as a fourth unpaired ligand for the catalytic zinc atom. Activation involves delocalization of the domain containing this cysteine by a conformational change or proteolytic cleavage, a mechanism referred to as the cysteine-switch (Van Wart & Birkedal-Hansen 1990). Most MMPs are secreted but the membrane type MT-MMPs are membrane anchored and some MMPs may act on intracellular proteins. Various domains determine substrate specificity, cell localization and activation (Hadler-Olsen et al. 2011). MMPs are regulated by transcription, cellular location (most are not activated until secreted), activating proteinases that can be other MMPs, and by metalloproteinase inhibitors such as the tissue inhibitors of metalloproteinases (TIMPs). MMPs are best known for their role in the degradation and removal of ECM molecules. In addition, cleavage of the ECM and other cell surface molecules can release ECM-bound growth factors, and a number of non-ECM proteins are substrates of MMPs (Nagase et al. 2006). MMPs can be divided into subgroups based on domain structure and substrate specificity but it is clear that these are somewhat artificial, many MMPs belong to more than one functional group (Vise & Nagase 2003, Somerville et al. 2003).

References

Lu P, Takai K, Weaver VM & Werb Z (2011). Extracellular matrix degradation and remodeling in development and disease. Cold Spring Harb Perspect Biol, 3.

Edit history

Date	Action	Author
2011-08-05	Created	Jupe S
2011-09-09	Authored	Jupe S
2012-02-21	Edited	Jupe S
2012-02-28	Reviewed	D'Eustachio P
2021-03-06	Modified	Shorser S

Entities found in this pathway (2)

Input	UniProt Id	Input	UniProt Id
CTSB	P07858	CTSL	O60911, P07711

6. Identifiers found

Below is a list of the input identifiers that have been found or mapped to an equivalent element in Reactome, classified by resource.

Entities (13)

Input	UniProt Id
ACE	P12821
DPP4	P27487
IL6	P05231
POLA1	P09884
VEGFA	P15692

Input	UniProt Id
CTSB	P07858
EGFR	P00533
IMPDH2	P12268
RIPK1	Q13546

Input	UniProt Id
CTSL	P07711
HDAC2	Q92769
PABPC1	P11940
SIGMAR1	Q99720

Input	Ensembl Id
EGFR	ENSG00000146648
POLA1	ENSG00000101868

Input	Ensembl Id
IL6	ENSG00000136244
VEGFA	ENSG00000112715