



Kyoto Encyclopedia of Genes and Genome

Minoru Kanehisa

Institute for Chemical Research, Kyoto University

HFSP Workshop, Strasbourg, November 18, 2016

The KEGG Databases

Category	Database	Content
Systems information	PATHWAY	KEGG pathway maps
	BRITE	BRITE functional hierarchies
	MODULE	KEGG modules
Genomic information	KO (KEGG ORTHOLOGY)	KO groups for functional orthologs
	GENOME	KEGG organisms, viruses and addendum
	GENES / SSDB	Genes and proteins / sequence similarity
Chemical information	COMPOUND	Chemical compounds
	GLYCAN	Glycans
	REACTION / RCLASS	Reactions / reaction classes
	ENZYME	Enzyme nomenclature
Health information (KEGG MEDICUS)	DISEASE	Human diseases
	DRUG / DGROUP	Drugs / drug groups
	ENVIRON	Health-related substances
	JAPIC	Japanese drug labels
	DailyMed	FDA drug labels

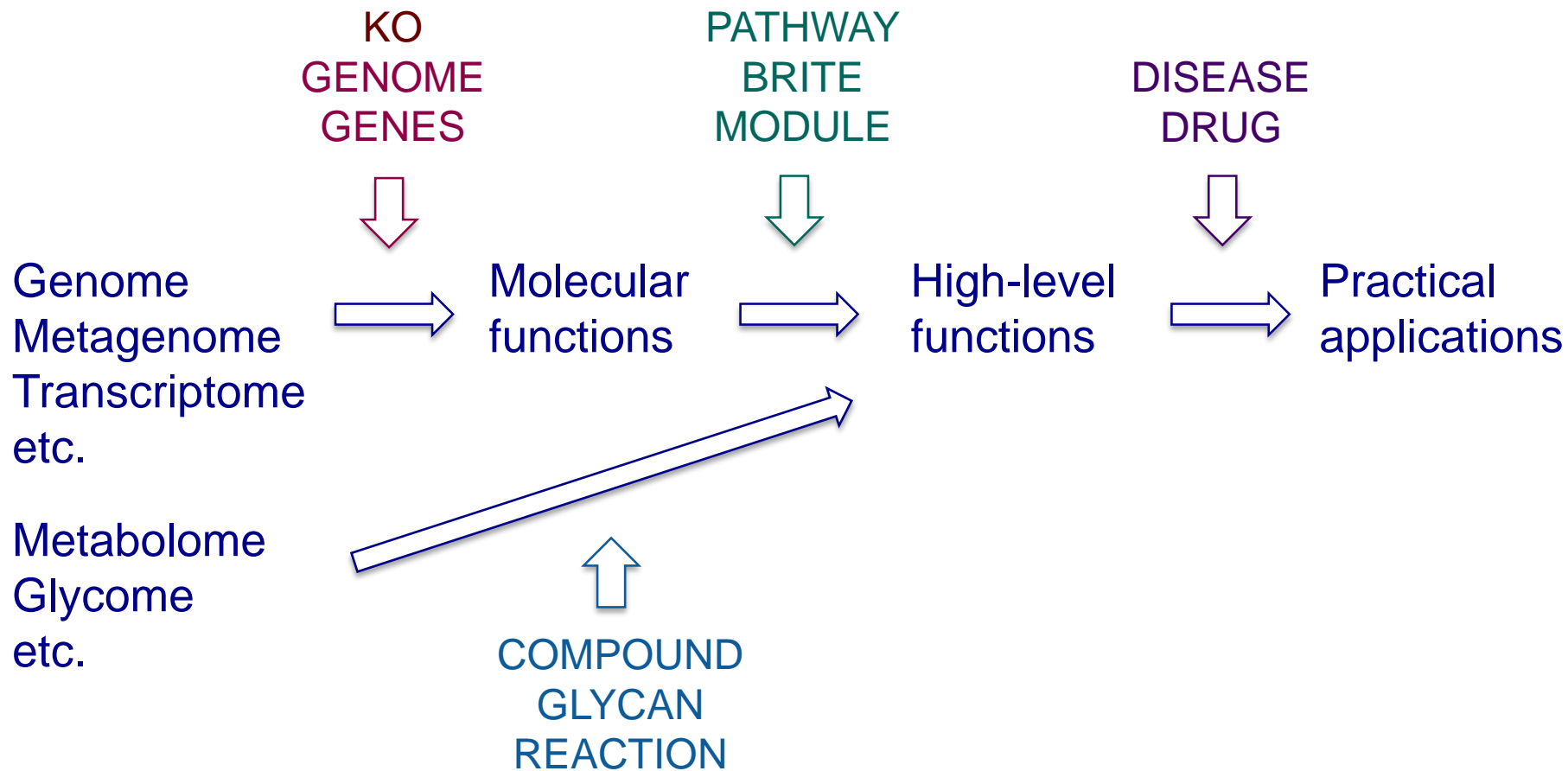
12 manually curated original DBs

3 DBs taken from outside sources and given original annotations (GENOME, GENES, ENZYME)

1 computationally generated DB (SSDB)

2 outside DBs (JAPIC, DailyMed)

KEGG is widely used for functional interpretation and practical application of genome sequences and other high-throughput data



Funding

Period	Funding source	Annual budget (USD)
1995-2010	Supported by 10+ grants from Ministry of Education, Japan Society for Promotion of Science (JSPS) and Japan Science and Technology Agency (JST)	>2 M
2011-2013	Supported by National Bioscience Database Center (NBDC) of JST	0.8 M
2014-2016	Supported by NBDC	0.5 M
2017-	?	

1995 KEGG website made freely available

1997 KEGG FTP site made freely available

2011 Plea to support KEGG

KEGG FTP academic subscription introduced

1998 First commercial licensing

1999 Pathway Solutions Inc. established

Contingency Plan

Current Model

Since 2011 Only partial funding is received from government
KEGG is mostly self-sustained by licensing fees obtained through
Pathway Solutions Inc.

Category	Funding	Website	FTP site
Systems information	None	Free	Restricted
Genomic information			
Chemical information			
Health information	NBDC		Free

KEGG website (including API)
<http://www.kegg.jp/>
<http://www.genome.jp/kegg/>

KEGG MEDICUS ftp site
<ftp://ftp.genome.jp/pub/kegg/medicus/>

Thanks to continued funding from NBDC

Computer facility Supercomputer system, Institute for Chemical Research,
Kyoto University

Lab space in Kyoto Institute for Chemical Research, Kyoto University

Thanks to Satoru Miyano

Lab space in Tokyo Human Genome Center, University of Tokyo

User community

Number of NAR papers citations (Web of Science)	16,000 total 3,000 per year
Number of NAR papers citations (Google Scholar)	24,000 total
KEGG website unique visitors per month (FY2015 average)	564,000
KEGG API unique visitors per month	7,000
KEGG API data transfer per month	200-400 GB
KEGG FTP data transfer per month (academic users only)	2,000-3,000 GB

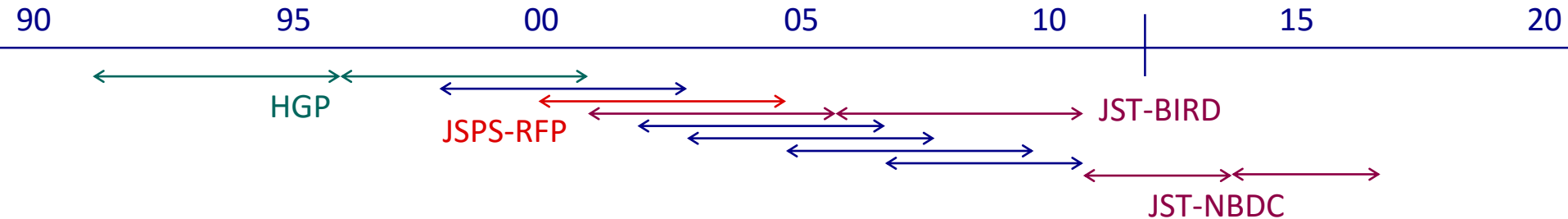
No restrictions on web access (including API)

Restricted FTP access (subscription based)

Not to police, not to enforce

History

Research grants



KEGG Project

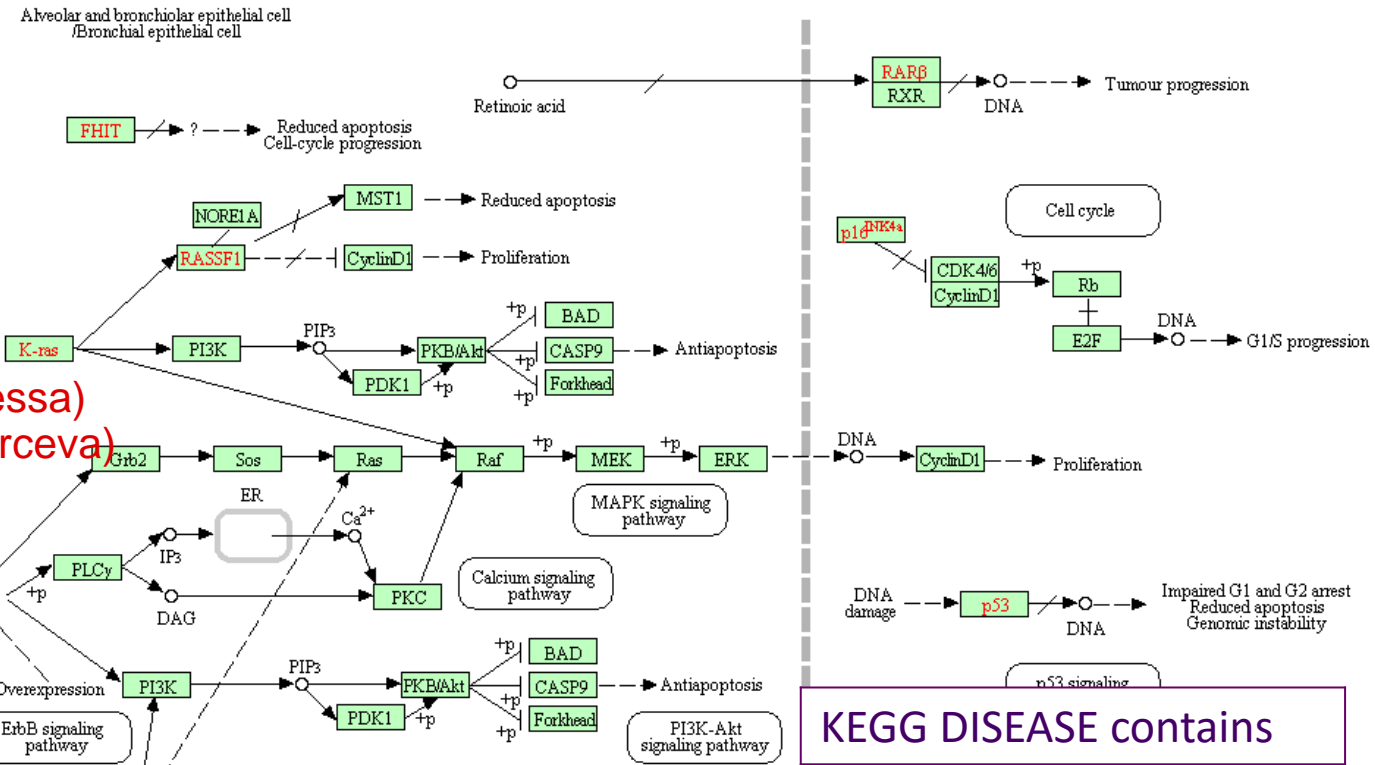
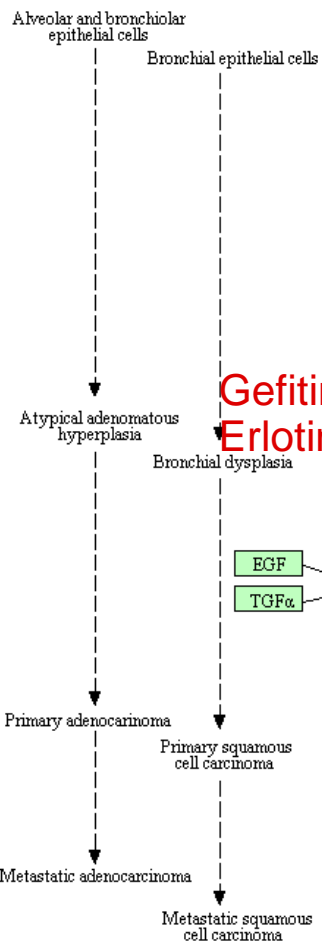
- 1991 Japanese Human Genome Project (HGP) officially started
- 1995 KEGG initiated with PATHWAY, GENES, ENZYME and COMPOUND
- 1998 REACTION
- 2000 GENOME
- 2002 KO
- 2003 GLYCAN
- 2004 RPAIR (discontinued in 2016)
- 2005 BRITE and DRUG
- 2007 MODULE
- 2008 DISEASE
- 2009 KEGG MEDICUS initiated integrating drug labels
- 2010 ENVIRON and RCLASS
- 2014 DGROUP

Perturbed molecular networks

Understanding diseases and drugs in terms of KEGG molecular networks

Non-small cell lung cancer (hsa05223)

NON-SMALL CELL LUNG CANCER



Gefitinib (Iressa)
Erlotinib (Tarceva)

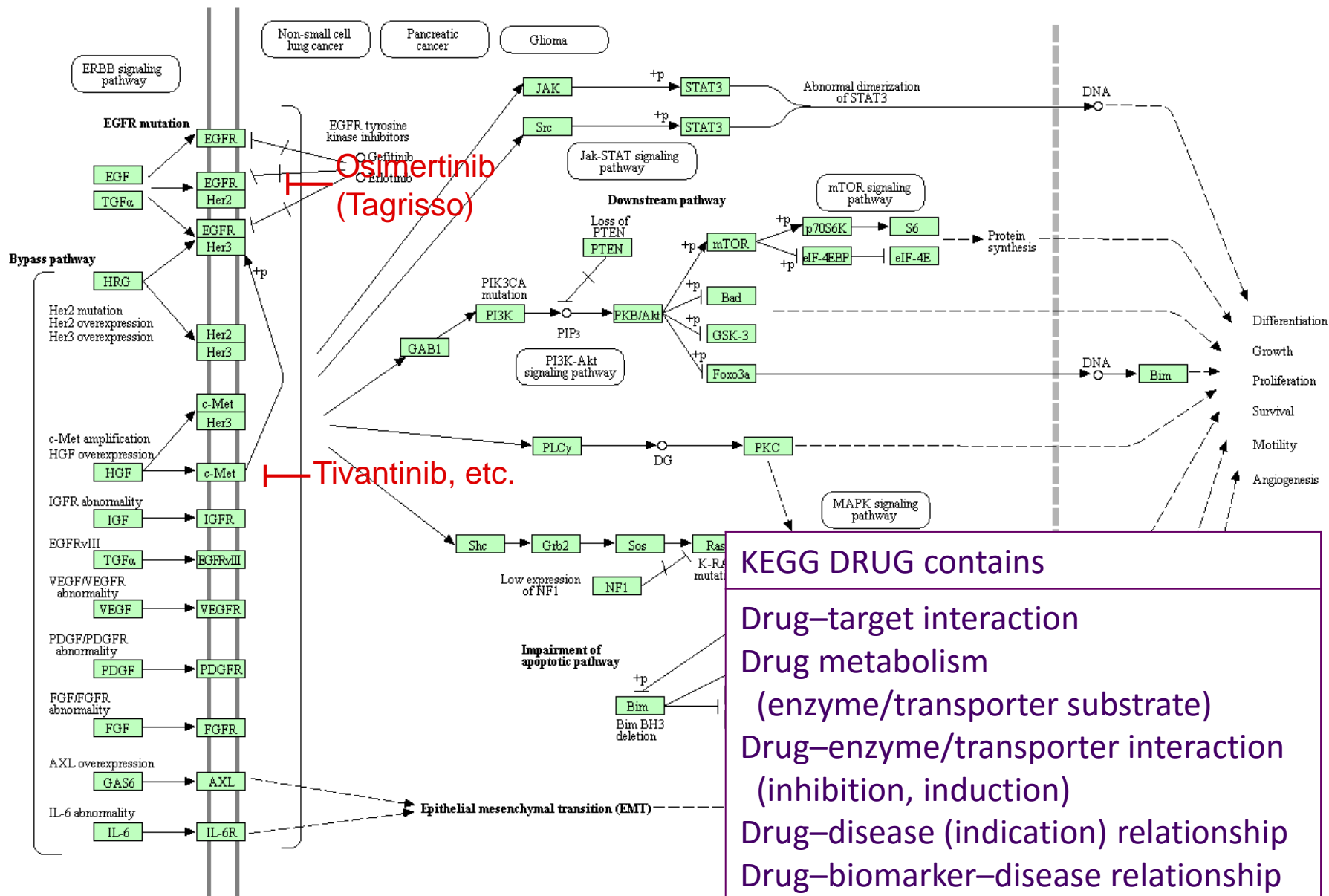
Crizotinib (Xalkori)

Genetic alterations
 Oncogene : K-Ras, EGFR, EML4-ALK
 Tumor suppressors : RARβ, FHIT, RASSF1, INK4a/ARF, p53

KEGG DISEASE contains
 Disease genes
 Environmental factors
 Carcinogens
 Pathogens
 Linked to KEGG pathways

EGFR tyrosine kinase inhibitor resistance (hsa01521)

EGFR TYROSINE KINASE INHIBITOR RESISTANCE



KEGG DRUG contains

- Drug-target interaction
- Drug metabolism (enzyme/transporter substrate)
- Drug-enzyme/transporter interaction (inhibition, induction)
- Drug-disease (indication) relationship
- Drug-biomarker-disease relationship

Before and After

Before 2011	After 2011
The entire KEGG was funded	Only the health information category (KEGG MEDICUS) is funded The other categories are self-sustained
KEGG was expanded by adding new databases	KEGG is expanded by incorporating society needs
Contents are derived mostly from published research articles	Contents are derived from both research articles and regulatory documents

Regulatory information incorporated into KEGG

- Japanese drug labels and processed data, including indications (drug-disease relationships) and drug-drug interactions associated with contraindications and precautions
- FDA drug labels
- New drug approvals (FDA, EMA, PMDA)
- Orphan drugs (FDA, PMDA)
- Excipients (Japan)