Target Identification and Animal Models

According to optimistic estimations, the human genome may contain 5,000 to 10,000 new drug targets.

All applied medications that are or have been in use are aiming about 500 targets on the molecular level.

Currently all marketed drugs are aiming only 120 targets.

The top 100 of best selling medications are solely aiming 43 proteins.

Is there only a small number of so-called *valid targets*?
→ connected to a disease and producing the desired effect

Is there not enough information about so-called *drugable targets*?
→ a valid target and substances actually reach it *in vivo*

Drugs according to function

Harvoni (ledipasivir/sofosbuvir)	hepatitis C	RNA polymerase inhibitor
Revlimid (lenalidomide)	multiple myeloma	(apoptosis)
Seretide (flucticasone)	COPD	anti-inflammatory
Crestor (rosuvastatide)	enzyme	HMG-CoA inhibitor
Abilify (aripiprazole)	GPCR	(schizophrenia)
Nexium (esomeprazole)	ion channel	ATPase inhibitor
Januvia (sitagliptin)	DPP-4 inhibitor	(diabetes mellitus)
Celebrex (celecoxib)	enzyme	COX-2-inhibitor
Claritin (loratadine)	GPCR	(allergic rhinits)
Prozac (fluoxetine)	GPCR	(depression)

A selection of the best selling medications of the past years (excluding biopharmaceuticals)

Data source:
https://en.wikipedia.org/wiki/List_of_largest_selling_pharmaceutical_products

Innovation vs. „me too“

New compounds that are „frist in class drugs“ usually are aiming new targets, e.g.

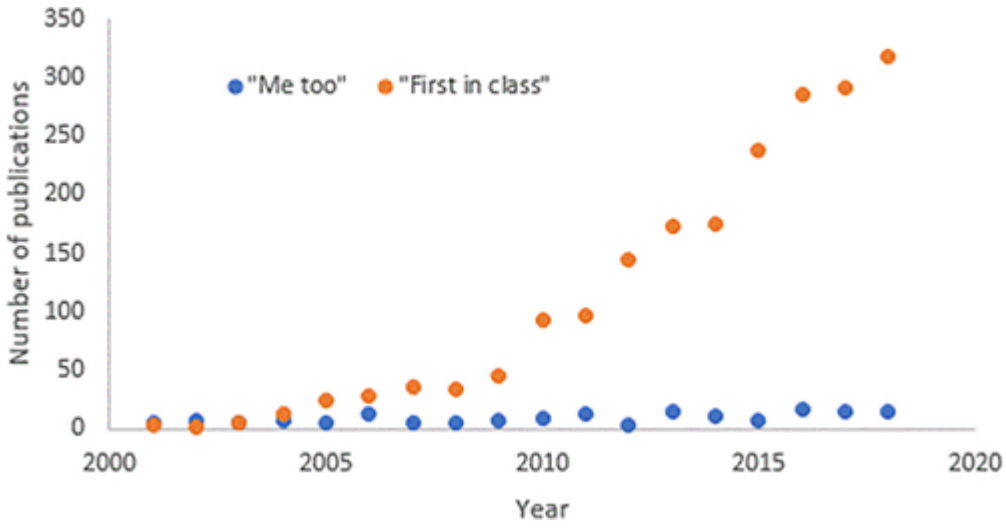
imipramine (GPCRs)
→ lofepramine

celecoxib (COX2)

sildenafil (PDE5)
→ vardenafil

imatinib (BCR-ABL)

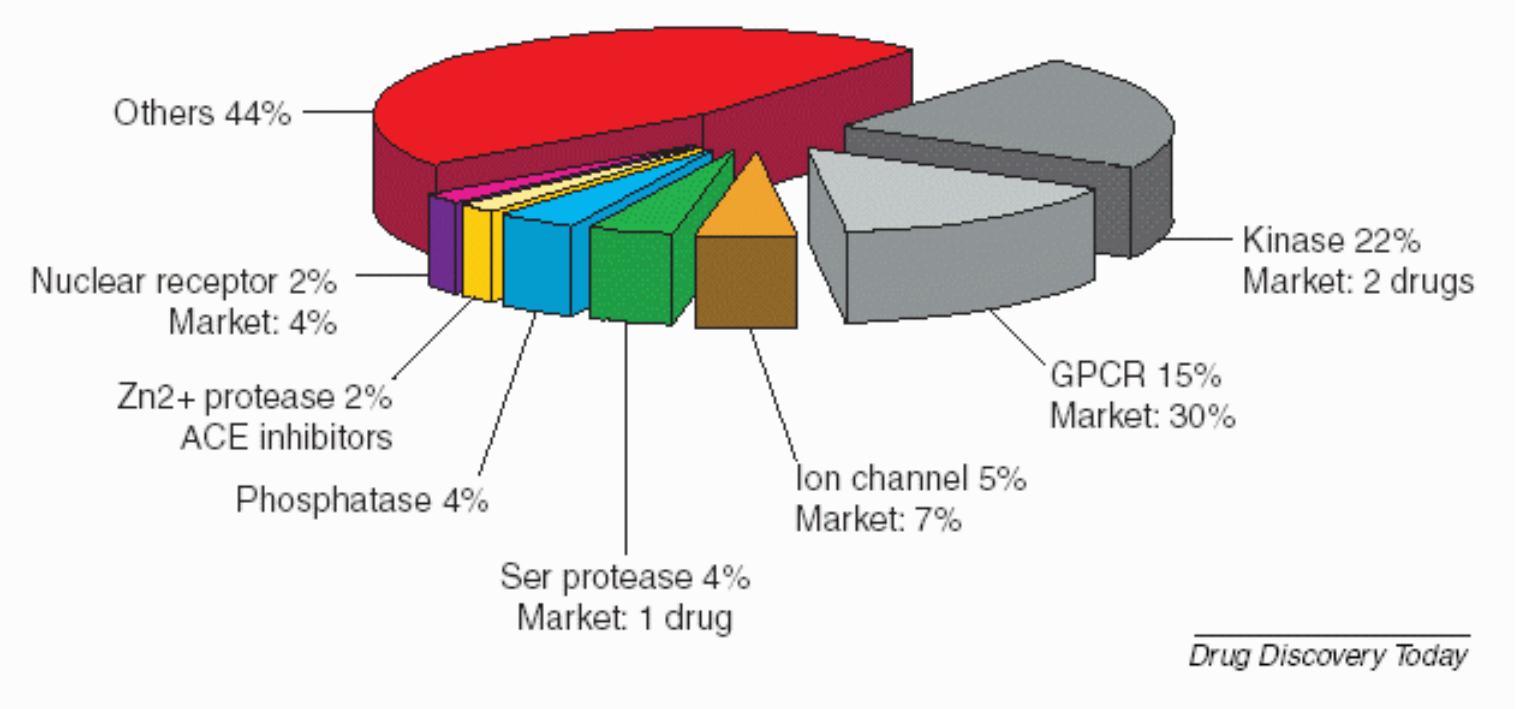
simvastatin (HMG-CoA)
→ rosuvastatin



Most new chemical entities can be regarded as „me too drugs, but they must be superior to existing medications to be approved.

Lit: J.K. Aronson & A.R. Green *Brit.J.Clin.Phramcol.* **86** (2020) 2114

Typical targets in the human genome

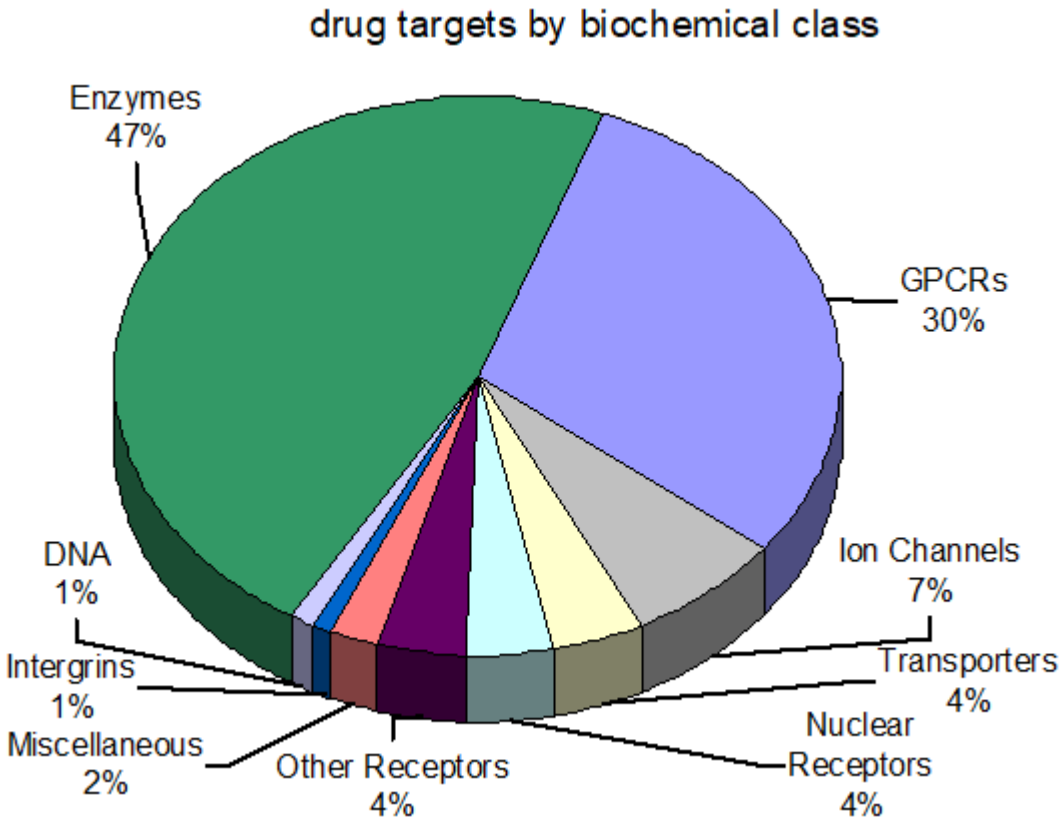


Contribution to the human genome and marketed drugs.

Around 500 proteins have been used so far as *targets.*

Estimated: 10,000 potential *targets* in the genome.

typical targets (II)



Fractional content of marketed drugs according to their biochemical targets

data: Hopkins & Groom, *Nat.Rev.Drug.Disc.* **1** (2002) 727

targets according to function

enzymes: kinases, proteases, etc.

G-protein coupled receptors (GPCR)

ion channels: e.g. K-channel (hERG), Ca-channel, Na-channel

nuclear receptors

DNA itself, RNA

other receptors (e.g. hormonal, neurotransmitters)

transporters, anti-porters, proton-pumps, efflux proteins (e.g. P-gp)

targets of monoclonal antibodies (non-small molecules)

Lit: P.Imming et al. *Nature Reviews Drug Discovery* **5** (2006) 821.

Literature about GPCR and signaling networks

M.J. Marinissen & J.S. Gutkind *Trends in Pharmacological Sciences* **22** (2001) 368.

One drug, one target ?

Promiscous drugs bind to more than one target:

COX-inhibitors: COX2 vs. COX1 selectivity

Propranolol: β-adrenoceptors, phosphatidic acid phosphorylase

Omapatrilat: angiotensin converting enzyme, neutral endopetidase

Oestrogens: nuclear receptors, membrane bound receptors

Antipsychotics: multiple GPCR receptors

Kinase-inhibitors: often multiple kinases

Ibuprofen: control substance in HTS assays

„orphan“ drugs: drugs with unknown mechanism of action are frequently found in the therapeutic categories of:

Anti-bacterials, anti-malarials, inhalative anesthetics

Lit. P.Imming et al. *Nature Reviews Drug Discovery* **5** (2006) 821.

Drug Repurposing and Scaffold Hopping

Prominent examples:

sildenafil (cGMP-specific Phosphodiesterase type 5 inhibitor)

originally for erectyle disfunction, repurpose: pulmonary arterial hypertension

thalidomide (see also lenalidomide)

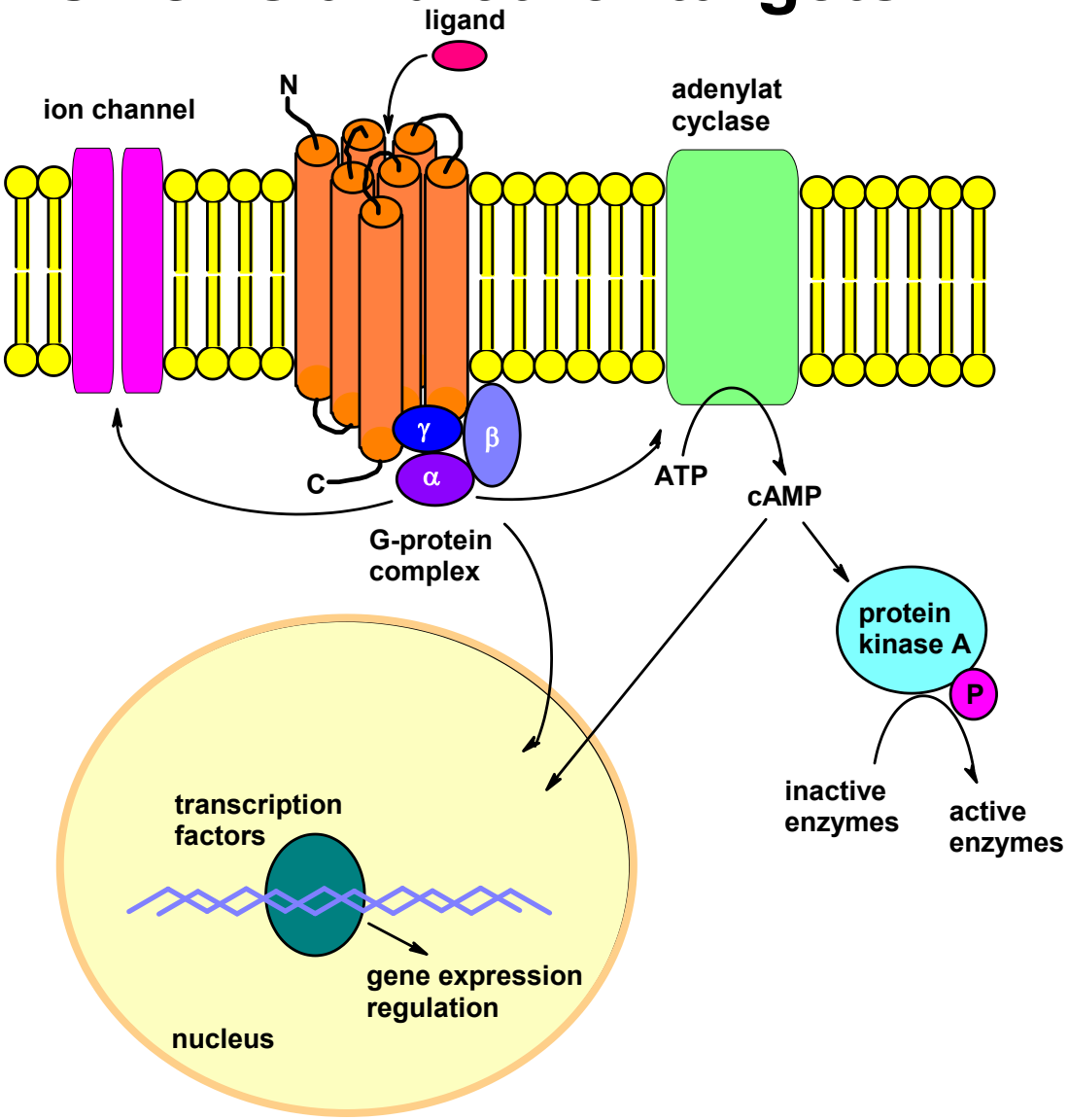
Originally as sedative, repurpose: multiple myeloma (cancer),

erythema nodosum leprosum

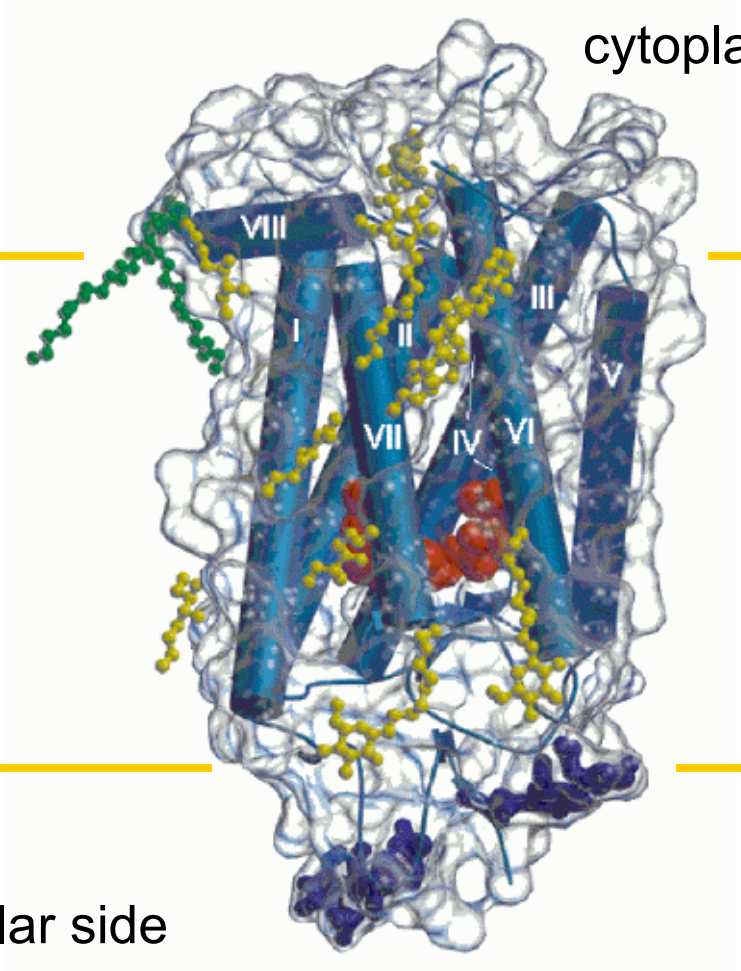
During preclinical studies, newly synthesized compounds are routinely screened in available assays → new scaffolds (common substructures) for existing targets

L.A. Tartaglia "Complementary new approaches enable repositioning of failed drug candidates„
Expert Opinion on Investigational Drugs **15** (2006) 1295.

GPCRs and other targets

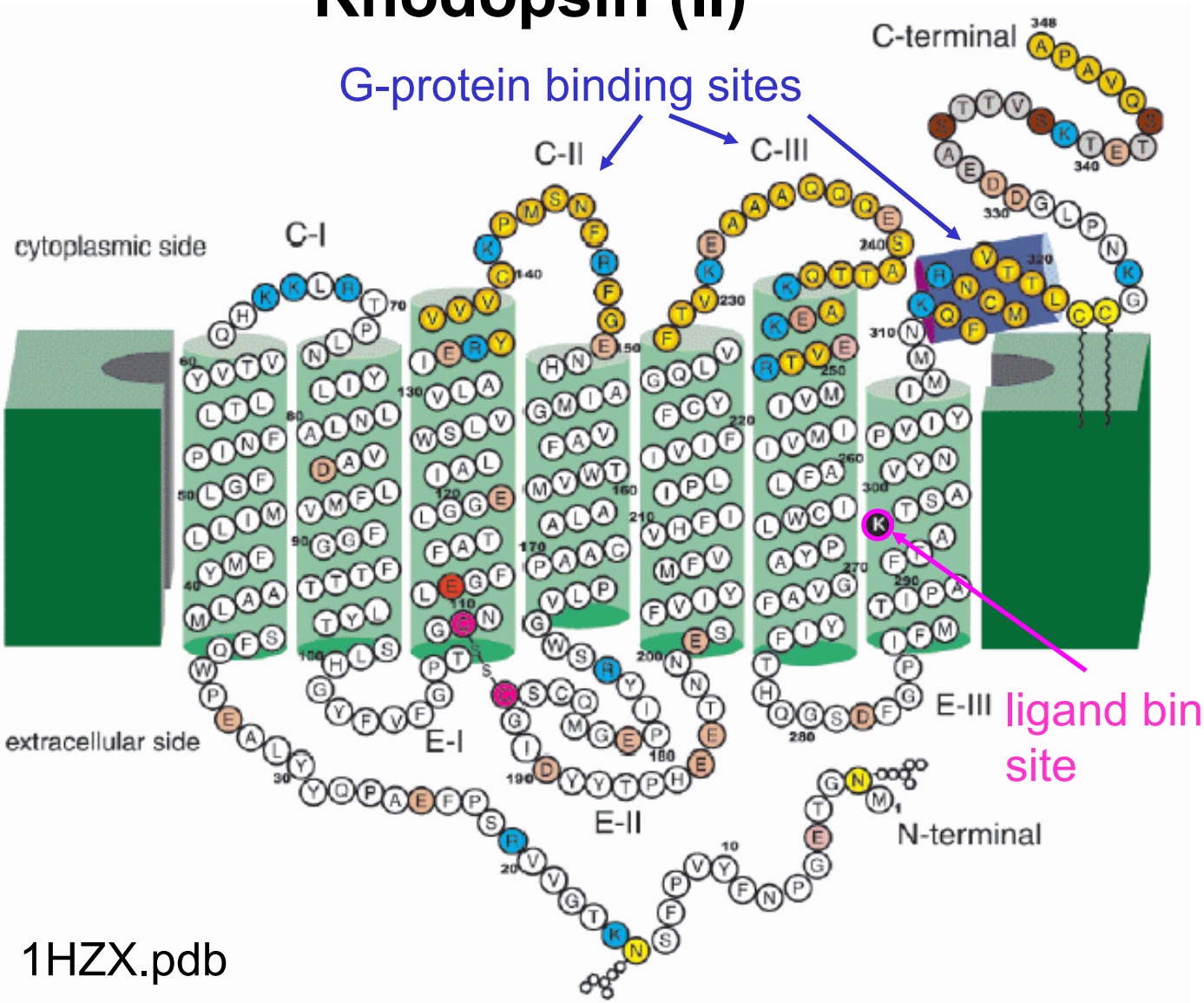


Rhodopsin (I)



Lit: D.C. Teller et al. *Biochemistry* **40** (2001) 7761
1HZX.pdb

Rhodopsin (II)



G-protein coupled receptors

G-protein coupled receptors comprise a large super-familiy of enzymes that are located at the cell surface. They transfer a number of signals forward into the cell, e.g. hormonal, visual, and neuronal. GPCRs are currently grouped into 6 large families:

class A: rhodopsin-like or adrenergic-receptor-like

class B: glucagon-receptor-like or secretin-receptor-like

class C: metabotropic-glutamate-receptor-like

class D: fungal mating pheromone receptors

class E: cyclic AMP receptors

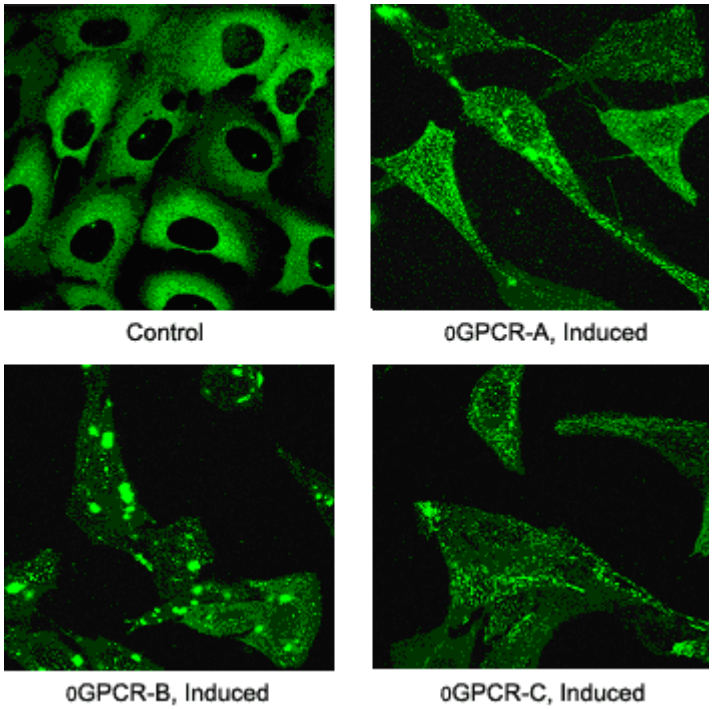
class F: frizzled / smoothened

Alternatively, the GRAFS classification system has been suggested for vertebrate GPCRs.

Orphan GPCRs

Designation for G-protein coupled receptors that have been identified in the genome, but (still) have unknown ligands (≈150).

Endeavors to find according ligands, e.g. by screening are called *deorphanizing.*



Picture source: www.moleculardevices.com

Validation of targets

When is a target that has been identified on the gene level of practical use ?

expression

disease model

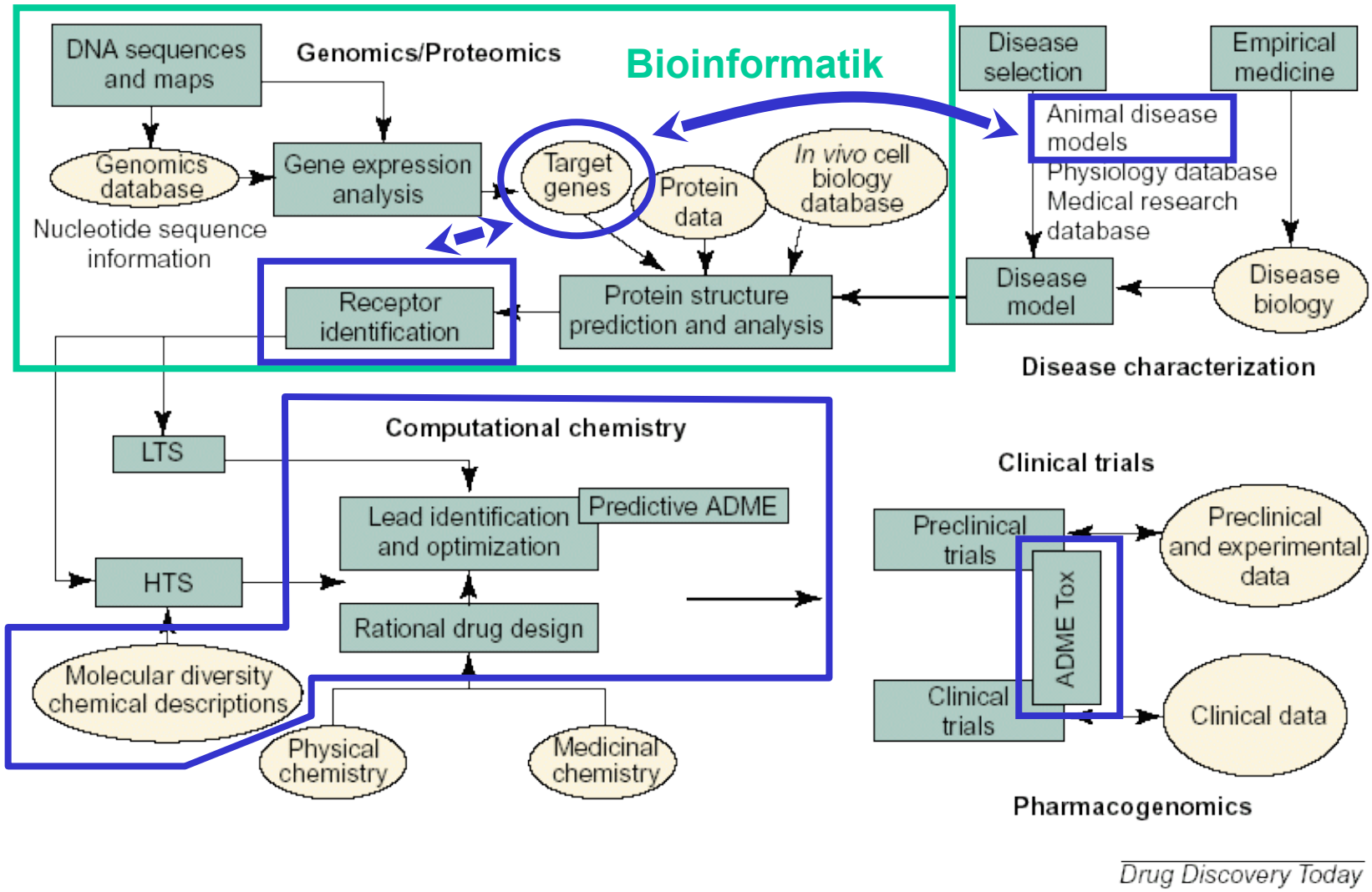
animal model

} defined physiological and clinical endpoints

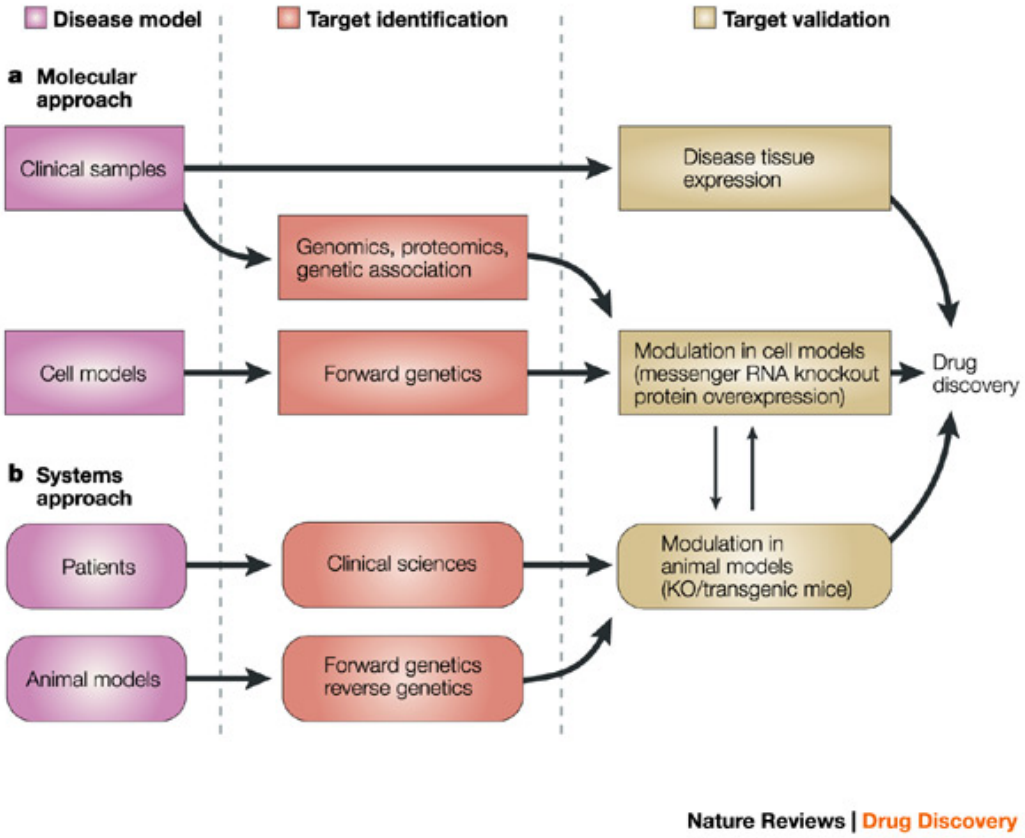
It has to be clarified if the target is suitable as a *therapeutic target* and therefore is a *valid target.*

At this stage proteomics, metabolomics, and pharmacogenetics / genomics enter.

Flow of information in a *drug discovery pipeline*



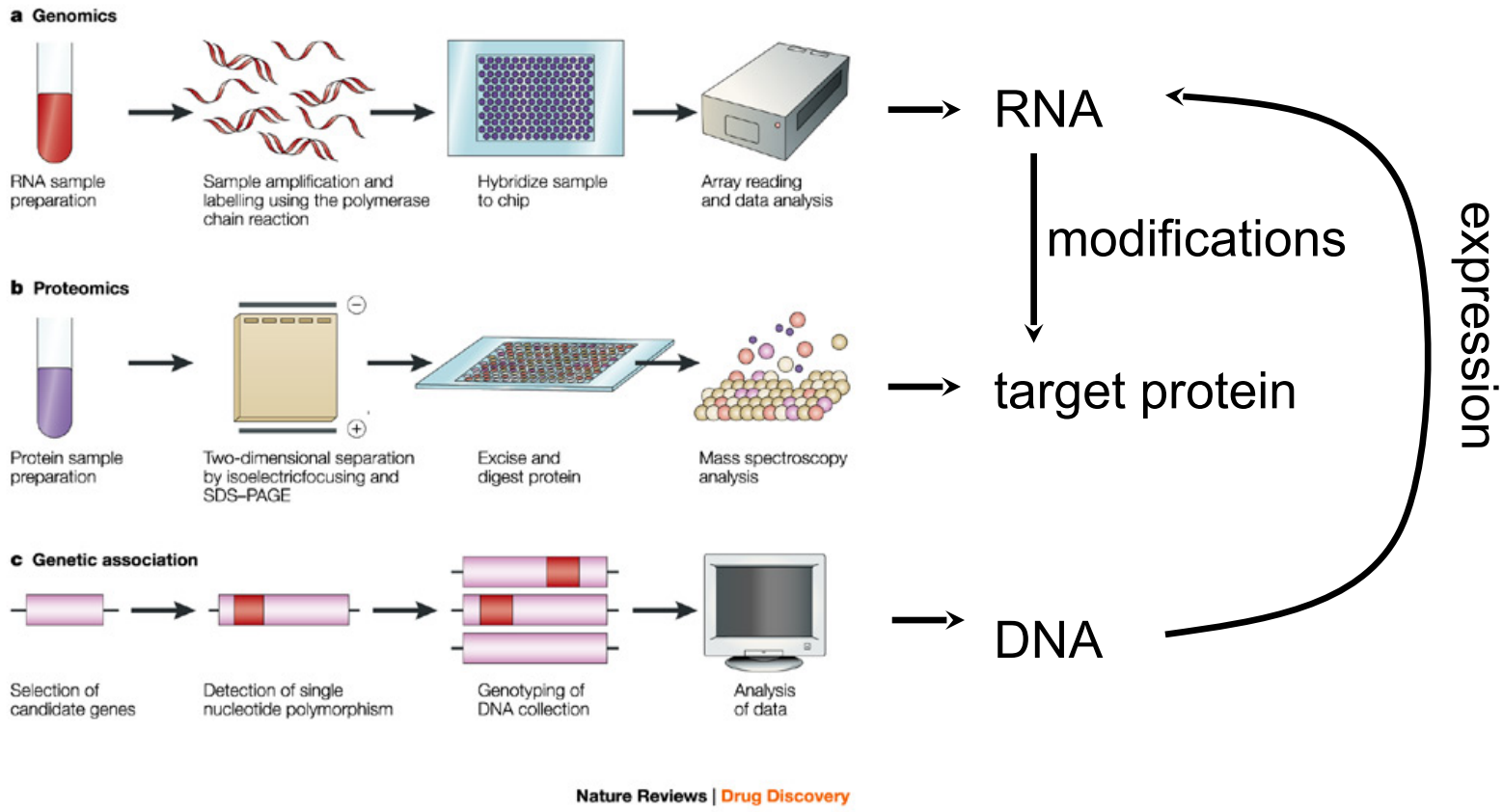
Towards the target (I)



In case of a known disease the identification of a suitable target is convergent process.

Lit: M.A.Lindsay *Nature Rev.Drug Disc.* **2** (2003) 831

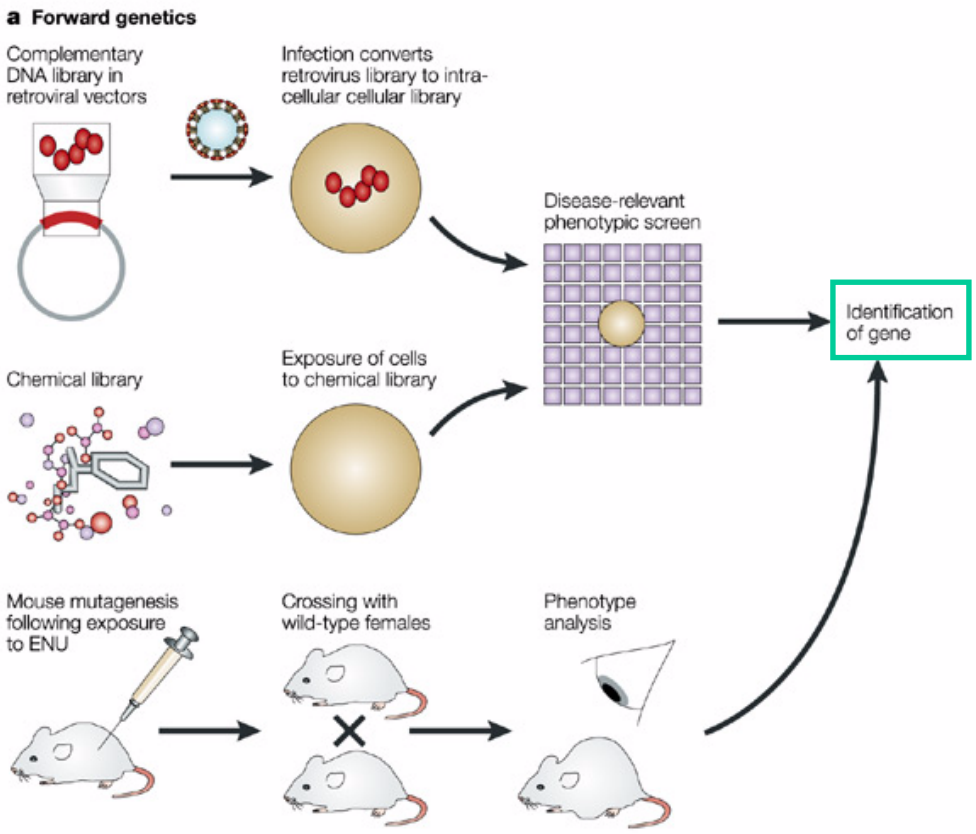
Towards the target (II)



Applied techniques to identify targets

Lit: M.A.Lindsay *Nature Rev.Drug Disc.* **2** (2003) 831

Towards the target (III)

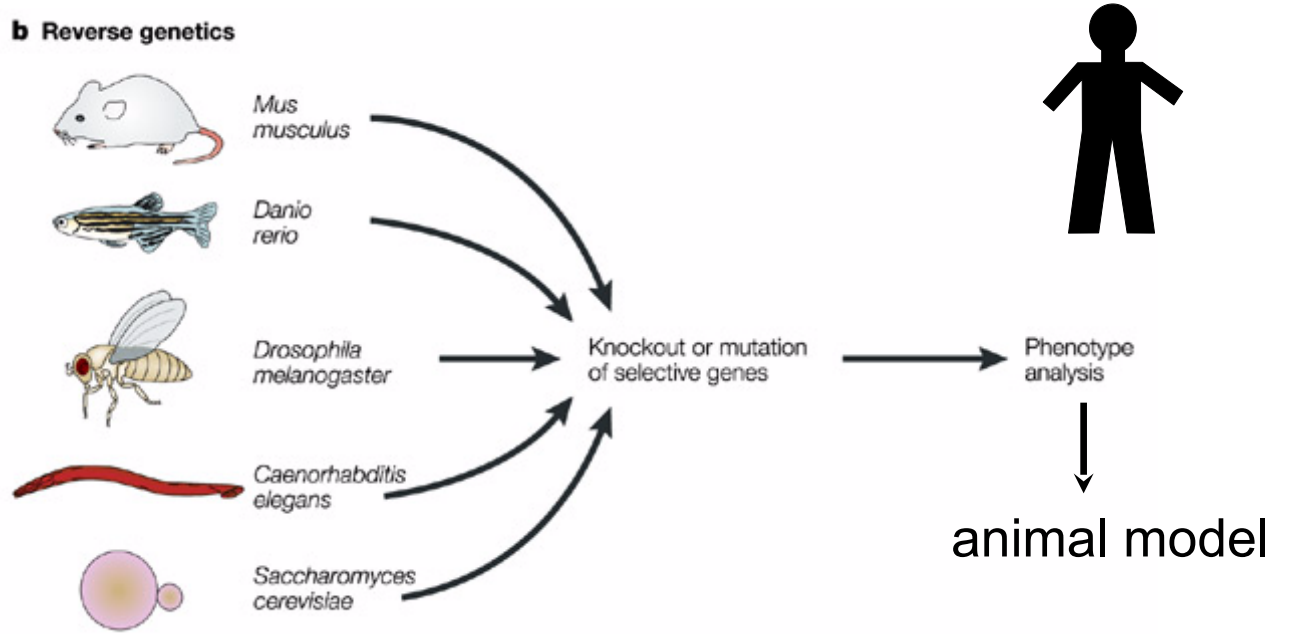


forward genetics: screening of compounds against variations of the phenotyp and mutations

Lit: M.A.Lindsay *Nature Rev.Drug Disc.* **2** (2003) 831

Towards the target (IV)

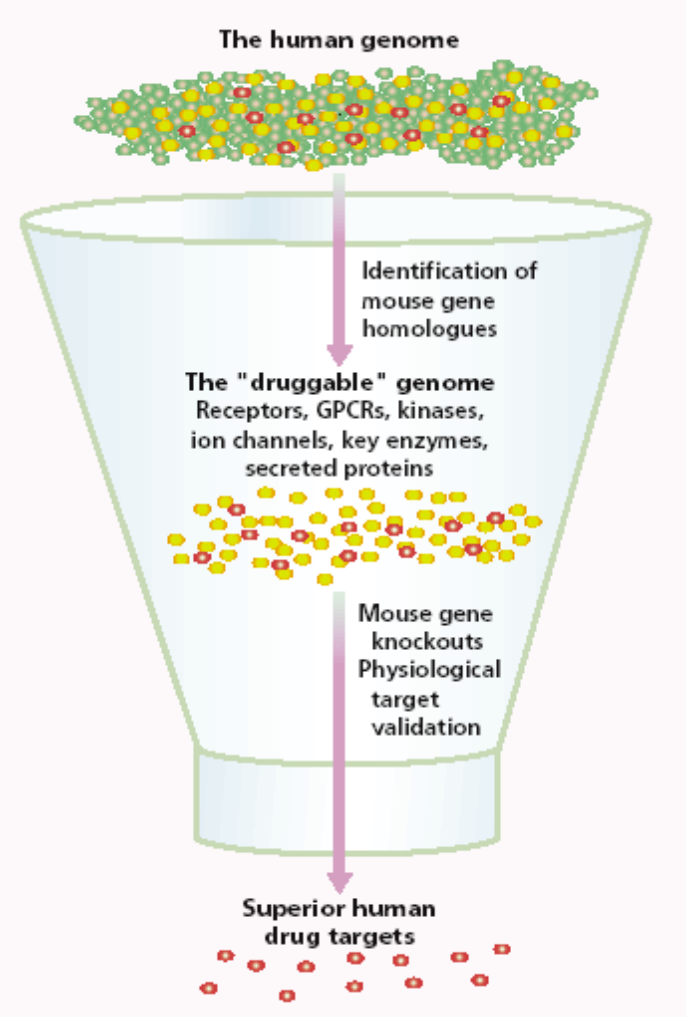
ortholog genes ← identified gene



reverse genetics: Modifications of the genotype by directed mutations

Lit: M.A.Lindsay *Nature Rev.Drug Disc.* **2** (2003) 831

Towards the target (V)



The bioinformatic approach for new targets in the ideal case (analysts scenario)

In practice there is the basic question:

„Which genes do we have to look for?“

Lit: A.T. Sands *Nature Biotech.* **21** (2003) 31

What to look for in the genome?

→ similarities to already exploited targets

Searching for targets that are so far under-represented should give the chance to find innovative targets:

kinases and proteases

Transmembrane proteins (GPCRs, ion channels, transporters)

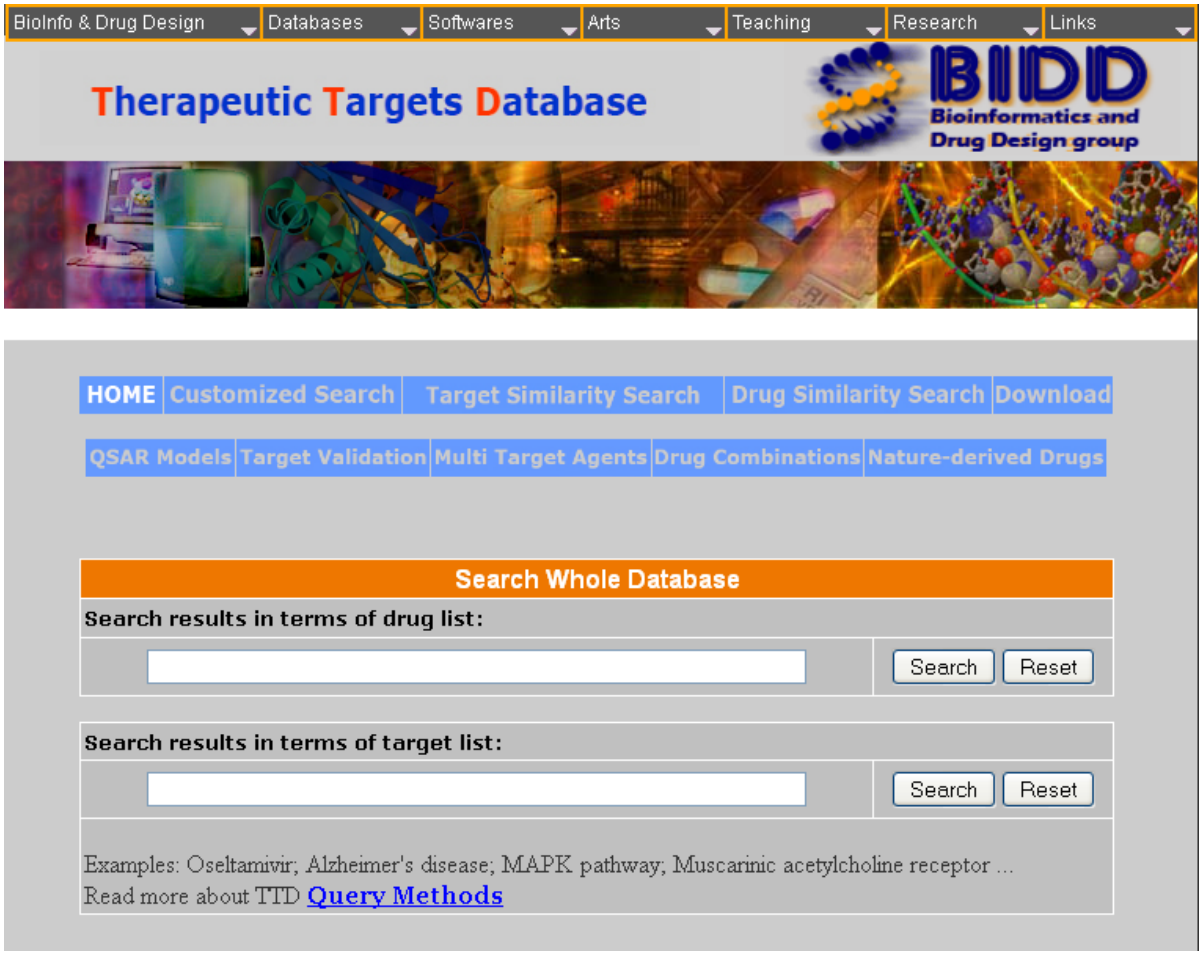
DNA and RNA binding sites, mRNA

nuclear receptors (for hormones)
(esp. *orphan nuclear receptors,*
so far only few new have been found)

According to cautious estimations there should be around 100-150 new and precious targets (valid and drugable).

→ The human targetome

Therapeutic Targets Database



3730 targets (as of late 2023)

https://db.idrblab.net/ttd/

Target validation

When is a target suitable for therapeutic purposes ?

There must be sufficient and reasonable connections with the disease:

a) as enzyme, GPCR, ion channel, receptor, etc.
Verification by screening with lead compounds from *focused libraries*

b) as target on DNA, RNA, mRNA level itself
Verification by knockout mutations (see below), *single point* mutations (SNPs, see below), and *gene silencing* by RNA interference (RNAi) (see siRNA)

siRNA for target validation

Short RNA strands of 11 to 28 nucleotides length can bind to complementary mRNA and lead to degradation by RNAses.
This RNA interference (RNAi) is used in eucaryotes as protection against viral RNA.
The term small interfering RNA (siRNA) stems from this.

This effect can be exploited to shut down mRNA (*gene silencing*) and also to detect potential targets on the mRNA level.

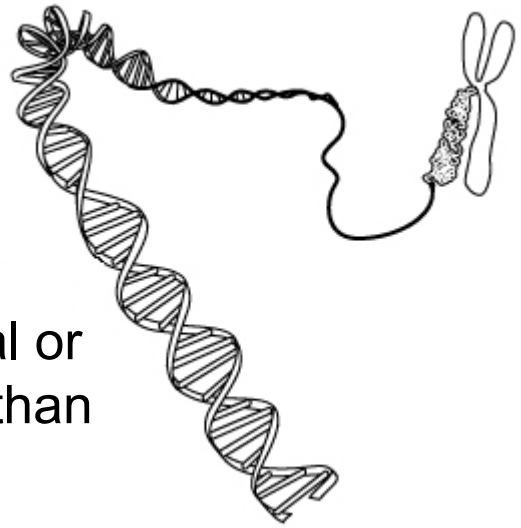
The therapeutical application of siRNAs is limited by their stability (administration) and selectivity (unspecific binding).

Lit: M.A. Lindsay *Nature Rev. Drug Disc.* **2** (2003) 831.
Y.Dorsett & T.Tuschl *ibid* **3** (2004) 318.

target characterization

There are variations in the complete (human) genome.
From the statistical point of view in

1 base pair per 1330 base pairs



which yields about $3 \cdot 10^6$ differences between two not related individual persons.

Also in the regions of genes that code potential or actual targets there are on average more than 9 exchanges of base pairs

Thus:

1. Not every variation is defective or means a predisposition (for a disease)
2. The selection of potential targets becomes even more complicated

Picture source: National Human Genome Research Institute

Pharmacogenetics & Pharmacogenomics

The causal assignment of a clinical phenotype (allel or symptom) to a genetic cause is hampered by the vast number of possible or existing variations of the genotype.

Alleles that are found in 1% or more of the population are referred to as polymorph (polymorphism). This means that these genotyps are found regularly.

In contrast, modifications of the genome that are found in less than 1% are refered to as mutations.

→ sequencing of the (eligible) genomic regions from as many individues as possible.

Lit: D.B. Goldstein et al. *Nature Rev. Genetics* **4** (2003) 937.

Single Nucleotide Polymorphism

SNPs are differences of a single DNA base that appear within a population.

The probability to find SNPs of a certain frequency can be estimated from the following table:

Number of individuals	SNP frequency >1%	>2%	>5%	>10%	>20%
2	4%	8%	19%	34%	59%
5	10%	18%	40%	65%	89%
10	18%	33%	64%	88%	99%
20	33%	55%	87%	99%	>99%
40	55%	80%	98%	>99%	>99%

source: J.J. McCarthy „Turning SNPs into Useful Markers of Drug Response“ in *Pharmacogenomics*, J.Licinio & M.-L.Wong (Eds.), Wiley-VCH (2002) pp.35-55.

Multiple SNPs

Even more complicated is the causal assignment of a reaction caused by a medication, if there are different SNPs that are independent from each other. In other words, if there is no conclusive hypothesis.

This can make the size of genetic regions that have to be sequenced becoming too large to be doable.

As examples of so-called *valid biomarkers,* the FDA has so far only precised the polymorphism of CYP2D6 (cytochrome P450) and of TPMT (thiopurine S-methyl-transferase).
Both enzymes contribute decisively to the metabolic conversion of many drugs.

More about the polymorphisms of CYP2D6 in lecture10

Lit. P.C.Sham et al. *Am.J.Hum.Genet.* **66** (2000) 1616.

R.Weinshilboum & L.Wang *Nature Rev.Drug Discov.* **3** (2004) 739.

Qualified Biomarkers

Besides *valid biomarkers*, the FDA defines *qualified biomarkers* as endpoints or valid targets.

According to their BEST glossary there are seven categories of biomarkers:

- susceptibility/risk
- diagnositic
- monitoring
- prognostic
- predictive
- pharmacodynamic/response
- safety

Susceptible genes

So far, susceptible genes have been identified in connection with the following symptoms:

sudden cardiac death

neurodegenerative diseases (dementia, Alzheimer,...)

epilepsy

schizophrenia

diabetes

arthritis

diseases of the lung (cystic fibrosis)

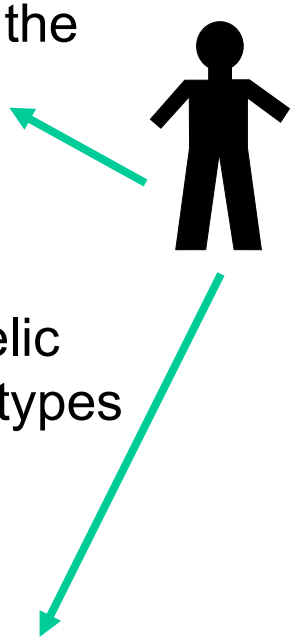
excess weight

Lit. V.D.Schmith et al. *Cell.Mol.Life Sci.* **60** (2003) 1636.

Gene Candidate Studies

Principal procedure for potential gene candidates

Selection of the pharmaceutical target gene either known target (enzyme, transporter, pathogenic gene,...) or newly identified gene from DNA-microarrays (on mRNA level), proteomics (on the protein level), bioinformatics



⇩

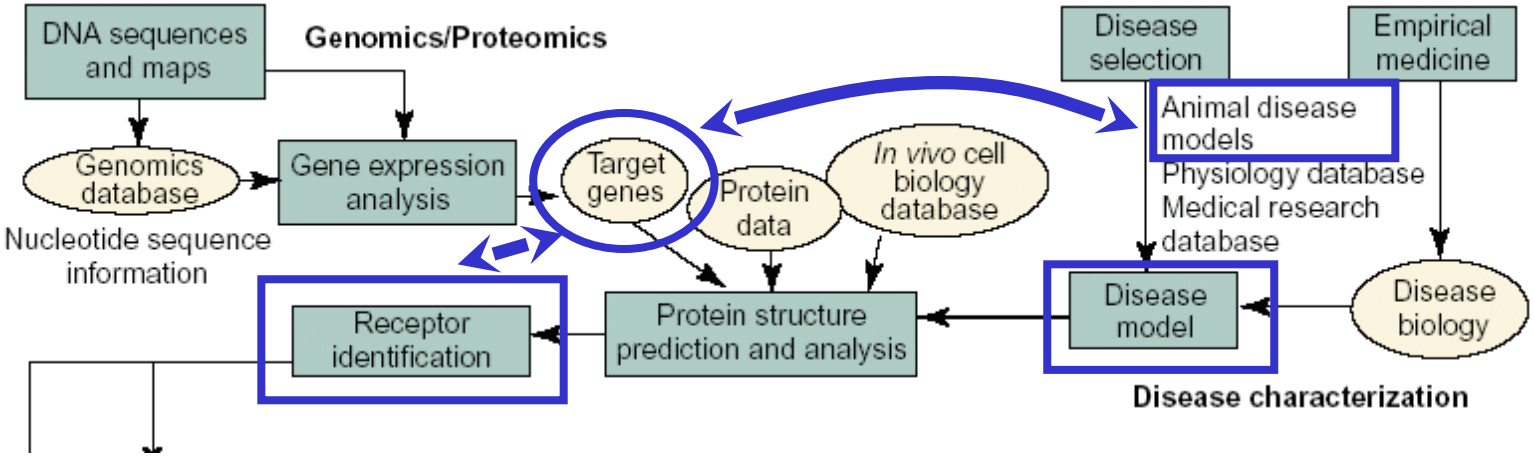
Identification of SNPs in the selected gene by SNP-mapping on a larger scale, determination of the allelic frequencies and ethnic distribution, analysis of the haplotypes

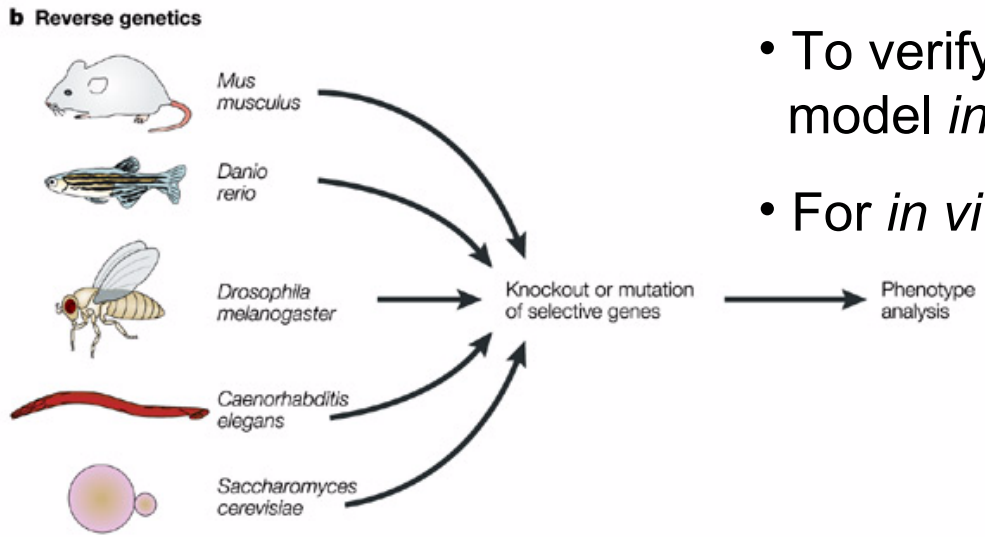
⇩

Genotyping of SNPs in clinical studies
Identification of the patient population, statistical analysis

Lit. H.Z.Ring & D.L.Kroetz *Pharmacogenomics* **3** (2002) 47-56.
highly recommended review

Why animal models ?

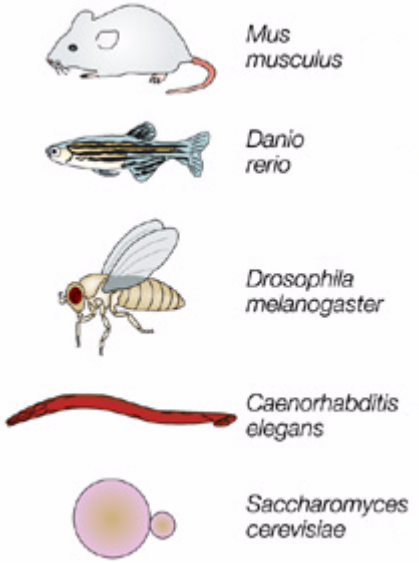




- To verify the disease model *in vivo*
- For *in vivo* screening

Model organisms

Before mice and other mammals are used for *in vivo* screening, other model organisms are used that carry according ortholog genes.



Larger number of genes being ortholog to human

Increasingly complex organisms

Increasing expense for experimental setup

literature :

R. Knippers *Molekulare Genetik* 8. Auflage
S. 498-503 Modellorganismen, Knockout Technologie

Performance of animal models

Animal models are helpful to verify a disease model *in vivo.*

1. Comparison of the target in the animal and the human genome.

2. Generation of *knockout* mutants / transgenic animals

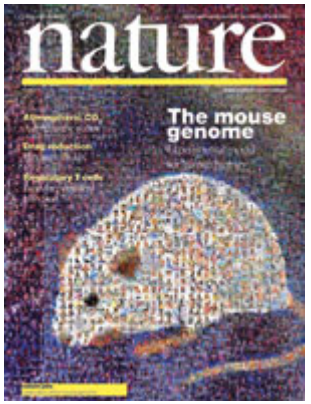
The existance of an adequate animal model is practically always the prerequisite for further development toward the clinical drug.

Literature about transgenic mice:

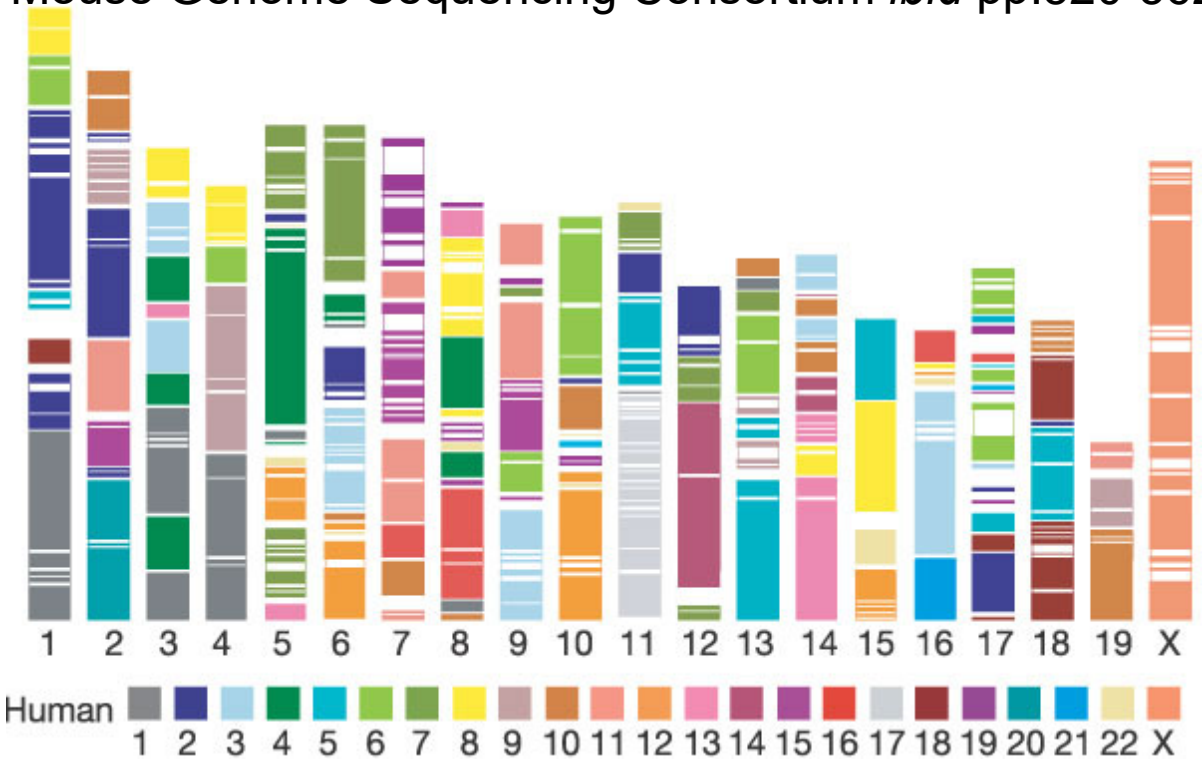
R. Knippers *Molekulare Genetik* 8. Auflage
S. 522 Textbox Plus 18.2

Why *mus musculus* as animal model? (I)

- For 99% of all mouse genes homologe or ortholog genes in human have been identified.



Lit: *Nature* 420 (2002) number 6915 of 5.12.2002
Mouse Genome Sequencing Consortium *ibid* pp.520-562.



Comparison of common elements in human and mouse chromosomes

Why *mus musculus* as animal model? (II)

• From all eligible model organisms mice are thus closest related to human among the group of mammals (rabbit, monkey, pig)

• mice propagate rapidly:

Mice become sexually mature at 10 or 12 weeks of age. 22 to 24 days after mating 4 to 8 cubs are born, upto 5 to 6 times per year. Thus a single mouse can have roughly 40 descendents within one year.

• The used breeds are rather homogenous regarding genetic aspects (high degree of inbreeding)

• The production of homozygote transgenic mice is easier than those for rats (Rattus norvegicus)

See also http://en.wikipedia.org/wiki/Mus_musculus

KO-mouse models (I)

Importance of *knockout* mouse models in the pharmaceutical area:

medical category	turnover (2001 in Mio.US$)	number of targets	number of drugs
immunology	20 000	8	15
neurology/psychatry	19 000	6	13
cardiology	13 000	6	13
gastroenterology	12 000	2	6
metabolisms	11 000	6	10
onkology	7 000	4	8
hematology	7 000	2	3

source: A.T.Sands *Nature Biotech.* **21** (2003) 31

Nobel Prize (2007) to M.R.Capecchi, M.J.Evans, O.Smithies

KO-mouse models (II)

Examples for the application of *knockout* mouse models in successful drugs:

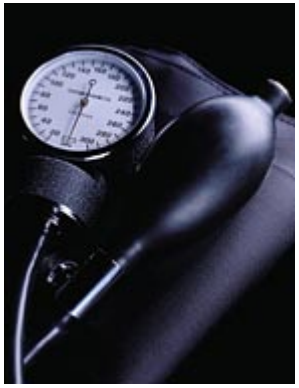
targets	drug	mouse phenotyp shows:
Proton pump	lansoprazol	neutral stomach pH
histamine H1-receptor	famotidine	repressed secretion of gastric acid
ACE	enalapril	lower blood pressure
AT_1-receptor	losartan	lower blood pressure
COX2	celecoxib	less inflammation
COX1 and COX2	diclofenac	less pain

Lit: B.P.Zambrowicz & A.T.Sands *Nature Rev.Drug Disc.* **2** (2003) 38

Model organisms for hypertension

Hypertension has not been observed in mice. The genes for the renin and angiotensin system were transfered from rat to mouse by *knock-in* mutations (cf. lecture 2)

Lit: H.Ohkubo et al. *Proc.Natl.Acad.Sci.USA* **87** (1990) 5153.



Conversely, *knockout* mice missing the ACE gene show lower blood pressure.

Lit: J.H.Krege et al. *Nature* **375** (1995) 146.

Since rats are better suited for functional studies, also transgenic rats containing the Ren-2 gene have been made. These showed strong symptoms of hypertension that could be treated with ACE-inhibitors and Angiotensin-II antagonists.

Lit: J.J.Mullins et al. *Nature* **344** (1990) 541.

Lit: Li-Na Wei *Annu.Rev.Pharmacol.Toxicol.* **37** (1997) 119.

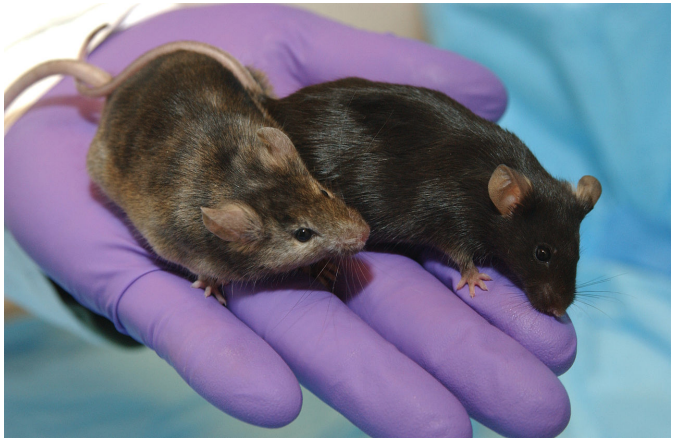
Model organisms for cancer

In cancer research two areas play a major role: The molecular mechanism of cancer origin and the therapeutic effiacy of the various medications.

Therefore a series of transgenic mouse models have been developed that show increased susceptibility for certain cancers.

In general, however, tumors seem to be the most frequent cause of death in mice if other factors during their lifespan are excluded.

The (ethnical) problematic nature of patents for transgenic animals on their own (without linking a technical use) should be mentioned for completeness.

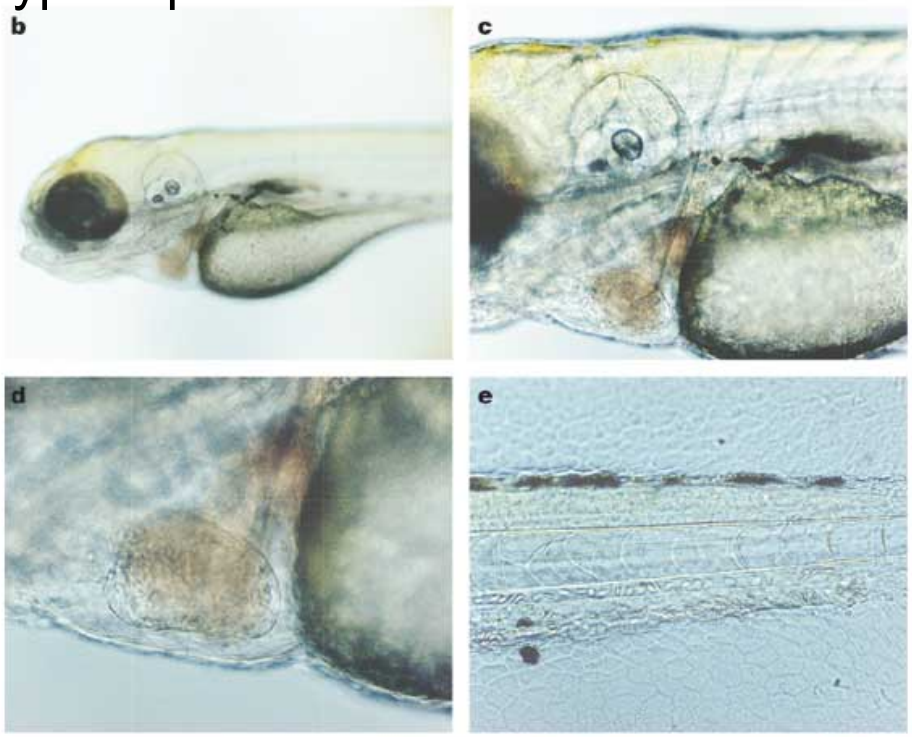


picture source: Wikipedia

Zebra fish as animal model (I)

Due to their size, zebra fish (*Danio rerio*) are easy to handle. Moreover, during their embryonal and larva stadium they are translucent, which facilitates the analysis of *in vivo* studies.

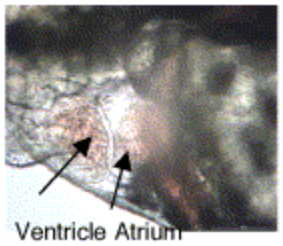
Thus High Throuput Screening regarding the consequences on the phenotype is possible.

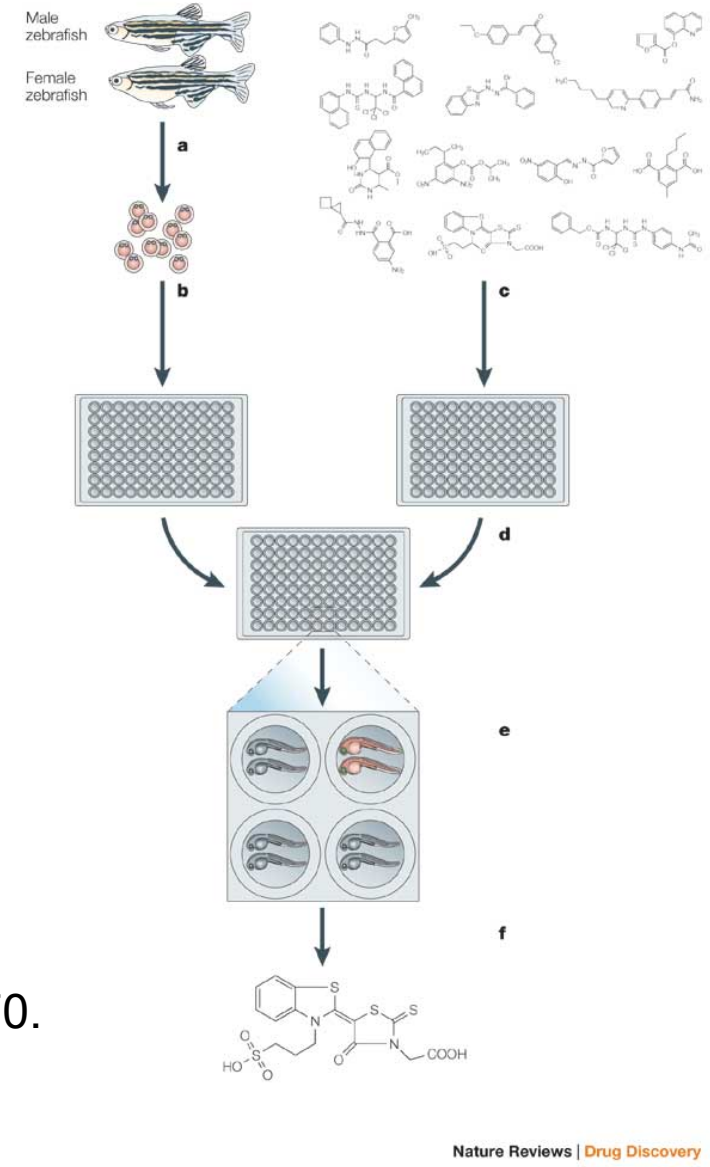


Lit: L.I.Zon & R.T.Peterson *Nat. Rev. Drug Disc.* **4** (2005) 35.

Zebra fish as animal model (II)

HTS *in vivo* screening
e.g. on QT-prolonging drugs
Zerg is the ortholog gene to hERG





Lit: L.I.Zon & R.T.Peterson
Nat. Rev. Drug Disc. **4** (2005) 35.

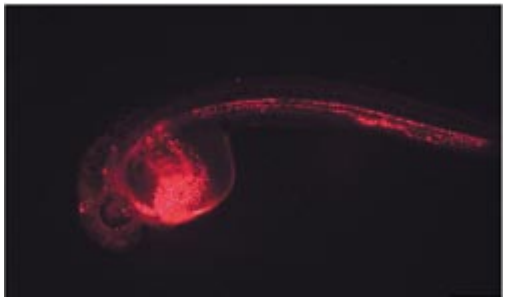
U.Langheinrich et al.
Toxicol. Appl. Pharm. **193** (2003) 370.

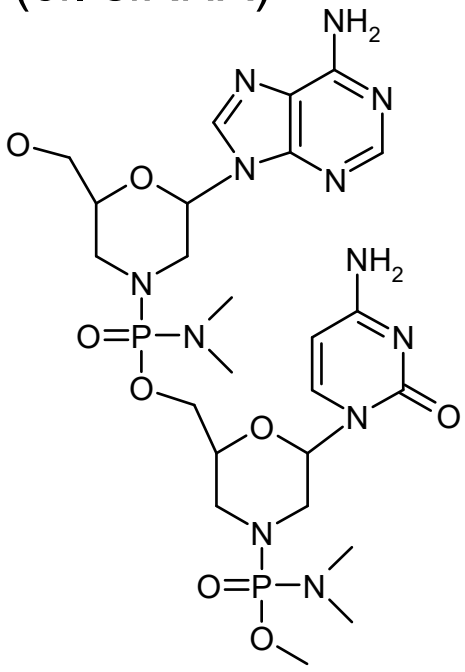
Zebra fish as animal model (III)

Furthermore there are a number of standard tools for genetic manipulations, e.g.

Knock down using morpholino oligonucleotides (cf. siRNA)

As well as the usual transgenic methods





Lit: A.Nasevicus & S.C.Ekker *Nature Genetics* **26** (2000) 216.

http://www.sanger.ac.uk/Projects/D_rerio/

Further animal models

Higher mammals such as mouse, rat, rabbit, dog, and pig are frequently being used to test metabolic and toxic properties of chemical substances.

Particularly the comparison of screening results of the metabolic conversions of drugs with those obtained from CYP P450 enzymes expressed in *E. coli* is of interest, in order to chose the most „suitable“ animal model.

Transgenic mice will be the prefered animal model in the future, not only due to financial considerations.

See also http://en.wikipedia.org/wiki/Model_organism