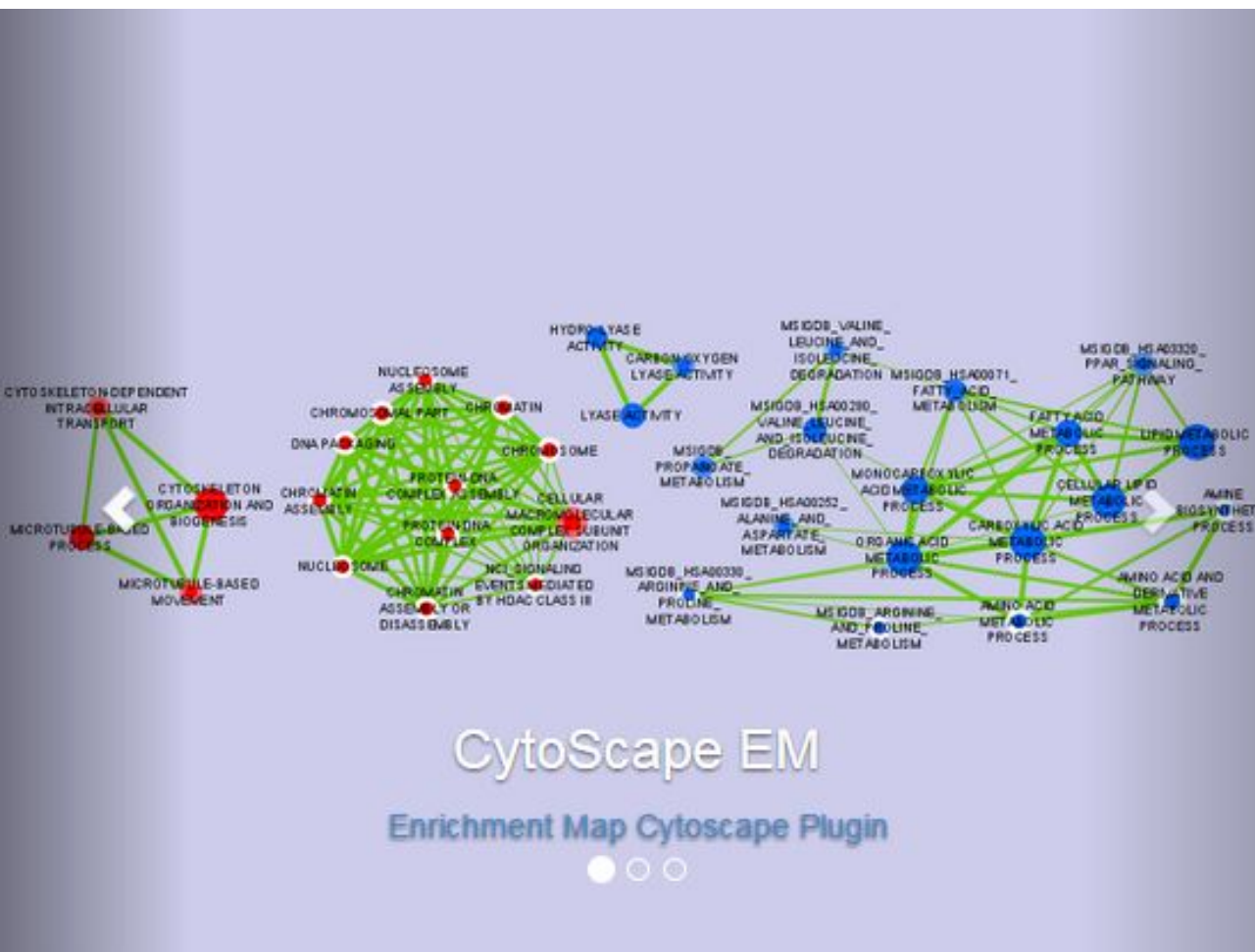


Базы данных Gene Ontology

GO - Gene Ontology consortium database



Gene Ontology Consortium



Целью создателей было установление контроля за единообразием в описаниях функций, биологических процессов и клеточных компонентов, относящихся к продуктам генов, чтобы унифицировать описания в различных базах данных и облегчить поиск в них нужного гена.

Базы данных Gene Ontology

1. Molecular Function

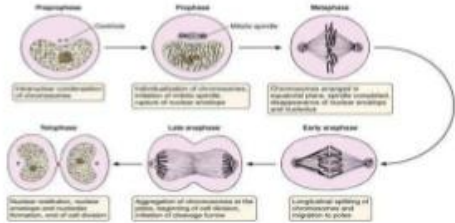
An elemental activity or task or job



- protein kinase activity
- insulin receptor activity

2. Biological Process

A commonly recognized series of events



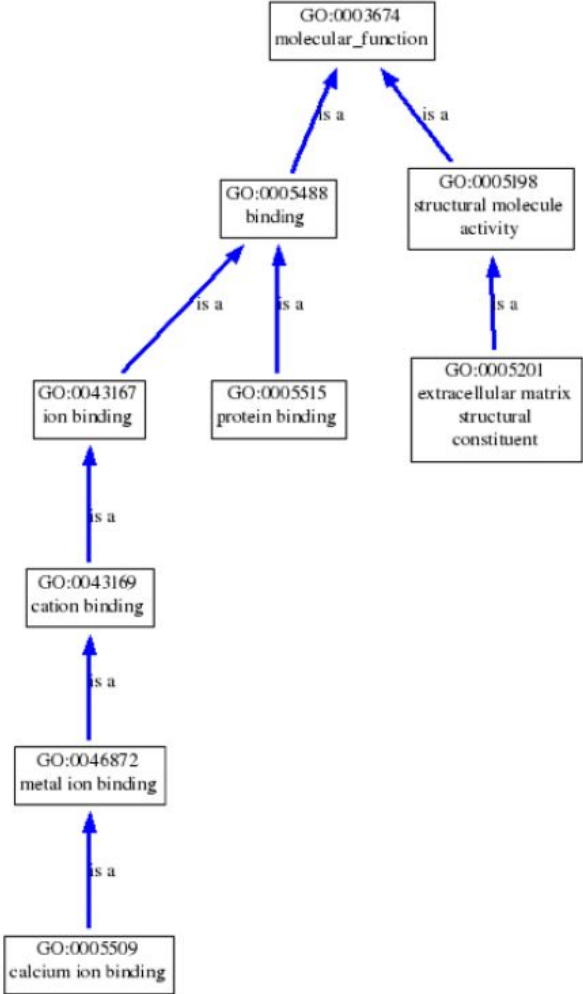
- cell division

3. Cellular Component

Where a gene product is located



- mitochondrion
- mitochondrial matrix
- mitochondrial inner membrane



Структурные базы данных

PDB (Brookhaven Protein DataBank) – коллекция экспериментально определенных 3D-структур биологических макромолекул. Раньше содержала и теоретические модели, но начиная с июля 2002 года в основном депозитории хранятся только экспериментально определенные структуры (рентгеноструктурным, ядерно-магнитнорезонансным и др. методами).

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB Login

PROTEIN DATA BANK

Advanced Search | Browse by Annotations | Search History (3) | Previous Results (2)

PDB-101 WORLDWIDE PDB PROTEIN DATA BANK EMDatabank NUCLEIC ACID DATABASE StructuralBiology Knowledgebase

f t v a i o

Summary 3D View Sequence Annotations Seq. Similarity 3D Similarity Literature Biol. & Chem. Methods Links

Crystal Structure of an Odorant-binding Protein from the Southern House Mosquito Complexed with an Oviposition Pheromone

3OGN

- Display Files
- Download Files
- Download Citation

DOI:10.2210/pdb3ogn/pdb

Primary Citation

Crystal and solution structures of an odorant-binding protein from the southern house mosquito complexed with an oviposition pheromone.

Mao, Y., Xu, X., Xu, W., Ishida, Y., Leal, W.S., Ames, J.B., Clardy, J.

Journal: (2010) Proc. Natl. Acad. Sci. USA **107**: 19102-19107

PubMed: 20956299

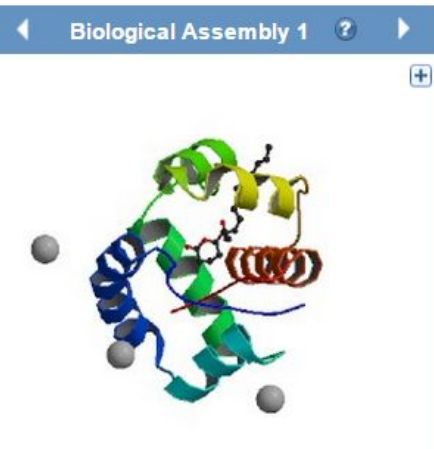
PubMedCentral: PMC2973904

DOI: 10.1073/pnas.1012274107

Search Related Articles in PubMed

PubMed Abstract:

Culex mosquitoes introduce the pathogens responsible for filariasis, West Nile virus, St. Louis encephalitis, and other diseases into



Структурные базы данных

↕ Molecular Description

Hide

Classification: Transport Protein [↗](#)

Structure Weight: 29515.32 [i](#)

Molecule: Odorant-binding protein

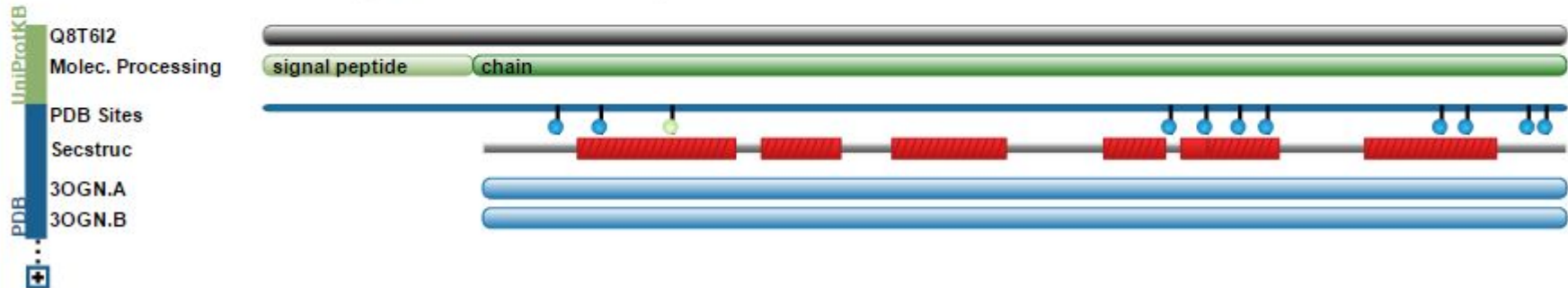
Polymer: 1 **Type:** protein **Length:** 124

Chains: A, B

Organism: *Culex quinquefasciatus* [↗](#)

Gene Name: CpipJ_CPIJ007604

UniProtKB: [Protein Feature View](#) | [Search PDB](#) [↗](#) | [Q8T6I2](#) [↗](#)



Структурные базы данных

```
HEADER      TRANSPORT PROTEIN                               17-AUG-10   30GN
TITLE       CRYSTAL STRUCTURE OF AN ODORANT-BINDING PROTEIN FROM THE SOUTHERN
TITLE       2 HOUSE MOSQUITO COMPLEXED WITH AN OVIPOSITION PHEROMONE
COMPND      MOL_ID: 1;
COMPND      2 MOLECULE: ODORANT-BINDING PROTEIN;
COMPND      3 CHAIN: A, B;
COMPND      4 SYNONYM: MOSQUITO ODORANT BINDING PROTEIN;
COMPND      5 ENGINEERED: YES
SOURCE      MOL_ID: 1;
SOURCE      2 ORGANISM_SCIENTIFIC: CULEX QUINQUEFASCIATUS;
SOURCE      3 ORGANISM_COMMON: SOUTHERN HOUSE MOSQUITO;
SOURCE      4 ORGANISM_TAXID: 7176;
SOURCE      5 GENE: CPIPJ_CPIJ007604;
SOURCE      6 EXPRESSION_SYSTEM: ESCHERICHIA COLI;
SOURCE      7 EXPRESSION_SYSTEM_TAXID: 562
KEYWDS      HELIX BUNDLE, ODORANT-BINDING PROTEIN, TRANSPORT PROTEIN
EXPDTA      X-RAY DIFFRACTION
AUTHOR      Y.MAO, J.CLARDY
REVDAT      2   15-DEC-10 30GN   1       JRNL
REVDAT      1   03-NOV-10 30GN   0
JRNL        AUTH   Y.MAO, X.XU, W.XU, Y.ISHIDA, W.S.LEAL, J.B.AMES, J.CLARDY
JRNL        TITL   CRYSTAL AND SOLUTION STRUCTURES OF AN ODORANT-BINDING
JRNL        TITL 2 PROTEIN FROM THE SOUTHERN HOUSE MOSQUITO COMPLEXED WITH AN
JRNL        TITL 3 OVIPOSITION PHEROMONE.
JRNL        REF    PROC.NATL.ACAD.SCI.USA           V. 107 19102 2010
JRNL        REFN                   ISSN 0027-8424
JRNL        PMID   20956299
JRNL        DOI    10.1073/PNAS.1012274107
REMARK      2
REMARK      2 RESOLUTION.      1.30 ANGSTROMS.
REMARK      3
REMARK      3 REFINEMENT.
REMARK      3   PROGRAM      : REFMAC 5.5.0102
REMARK      3   AUTHORS      : MURSHUDOV, VAGIN, DODSON
REMARK      3
REMARK      3   REFINEMENT TARGET : MAXIMUM LIKELIHOOD
```

Структурные базы данных

```
...
REMARK 465
REMARK 465 MISSING RESIDUES
REMARK 465 THE FOLLOWING RESIDUES WERE NOT LOCATED IN THE
REMARK 465 EXPERIMENT. (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN
REMARK 465 IDENTIFIER; SSSEQ=SEQUENCE NUMBER; I=INSERTION CODE.)
REMARK 465
REMARK 465      M RES C SSSEQI
REMARK 465      MET C      1
REMARK 465      PRO C      2
...
REMARK 470
REMARK 470 MISSING ATOM
REMARK 470 THE FOLLOWING RESIDUES HAVE MISSING ATOMS (M=MODEL NUMBER;
REMARK 470 RES=RESIDUE NAME; C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER;
REMARK 470 I=INSERTION CODE):
REMARK 470      M RES CSSEQI  ATOMS
REMARK 470      LYS A   23    CG   CD   CE   NZ
REMARK 470      LYS A   63    CE   NZ
...
SEQRES      1 A   219  GLN VAL GLN LEU GLN GLN PRO GLY ALA GLU LEU VAL LYS
SEQRES      2 A   219  PRO GLY ALA SER VAL LYS LEU SER CYS LYS ALA SER GLY
...
HET      F09   C1131      10
...
HETNAM      F09 NONAN-1-OL
...
```

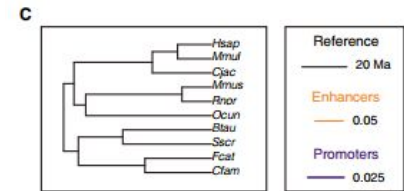
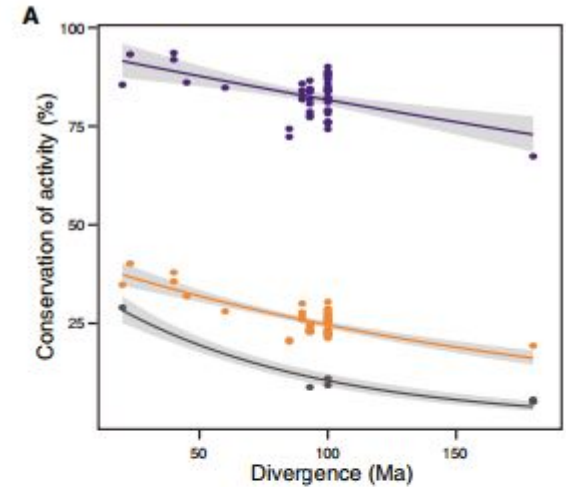
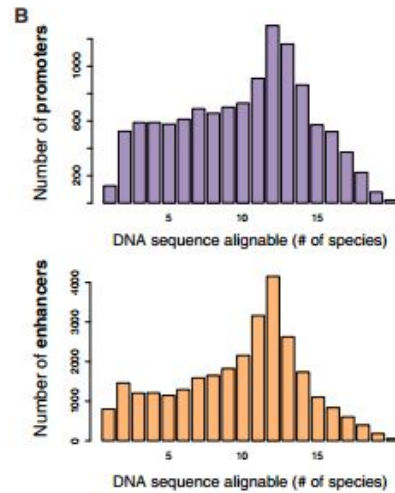
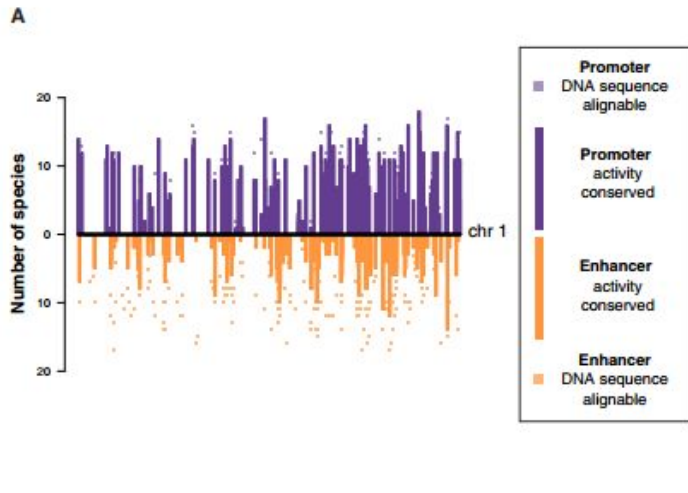
Структурные базы данных

```

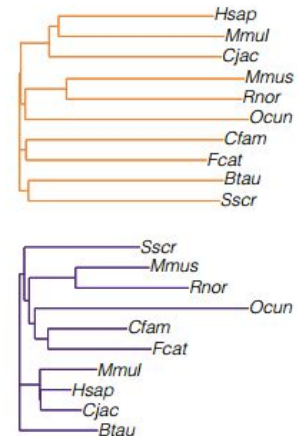
...
HELIX      1      1 THR A      87  SER A      91  5                                     5
...
HELIX     10     10 THR C      85  GLY C     123  1                                     39
SHEET      1     AA 4 LEU A      4  GLN A      5  0
...
SHEET      4     BE 4 SER B    201  ASN B    210 -1  O  SER B    201  N  HIS B    198
SSBOND     1     CYS A      22     CYS A      96                                     1555  1555
2.07
...
ATOM       22     CG  AGLN  A      3      -30.506   25.807   16.125   0.50  43.30
C
ATOM       23     CD  AGLN  A      3      -29.018   25.546   15.959   0.50  43.01
C
ATOM       24     OE1AGLN  A      3      -28.187   26.374   16.332   0.50  43.76
O
ATOM       25     NE2AGLN  A      3      -28.675   24.391   15.403   0.50  41.73
N
ATOM       26     N   BGLN  A      3      -33.005   25.819   17.852   0.50  45.16
N
ATOM       27     CA  BGLN  A      3      -32.788   24.834   16.796   0.50  44.25
C
ATOM       28     C   BGLN  A      3      -33.529   23.510   17.024   0.50  43.75
C
...
HETATM    4140     C1     F09  C1131      -21.017   -3.092   -1.563   1.00  60.56
C
HETATM    4141     C2     F09  C1131      -21.015   -1.597   -1.357   1.00  60.40

```

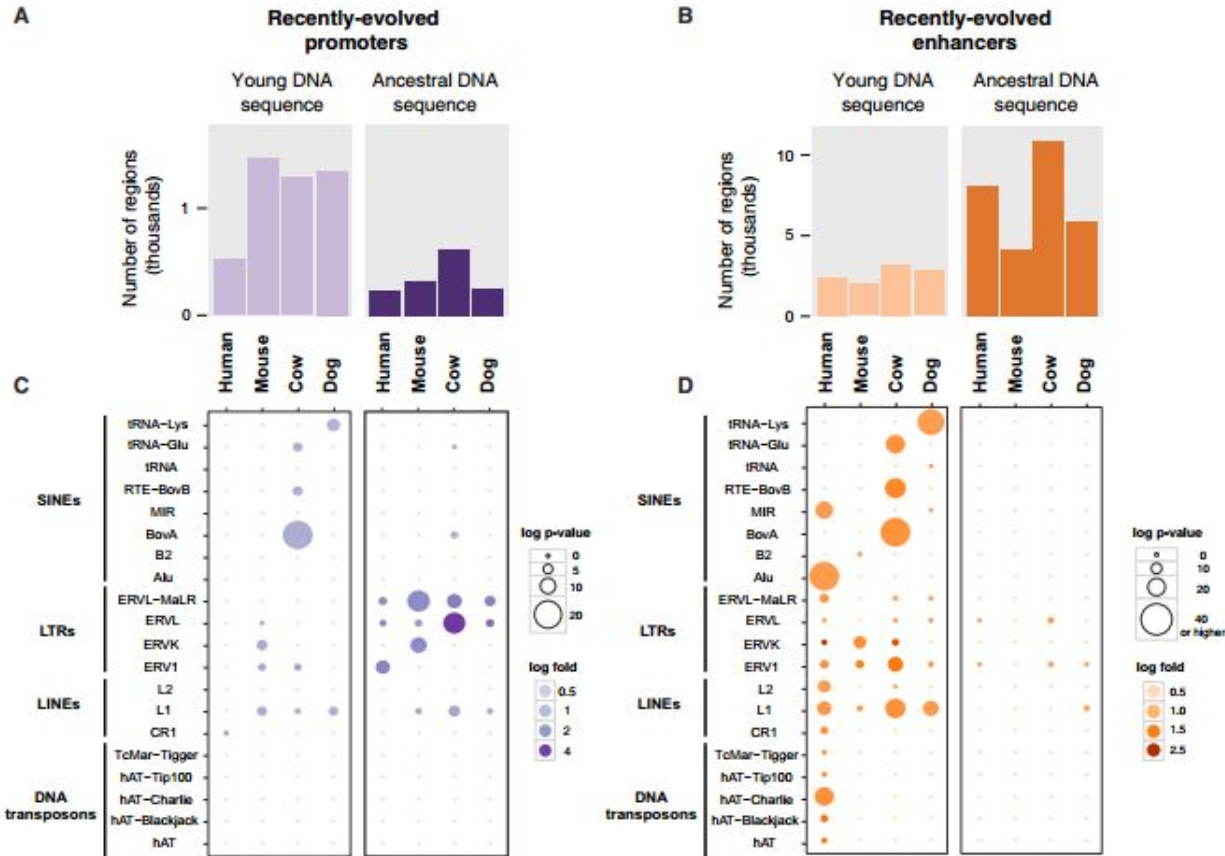
Энхансеры и промоторы



Последовательности промоторов млекопитающих более консервативны, чем энхансеры тех же самых генов

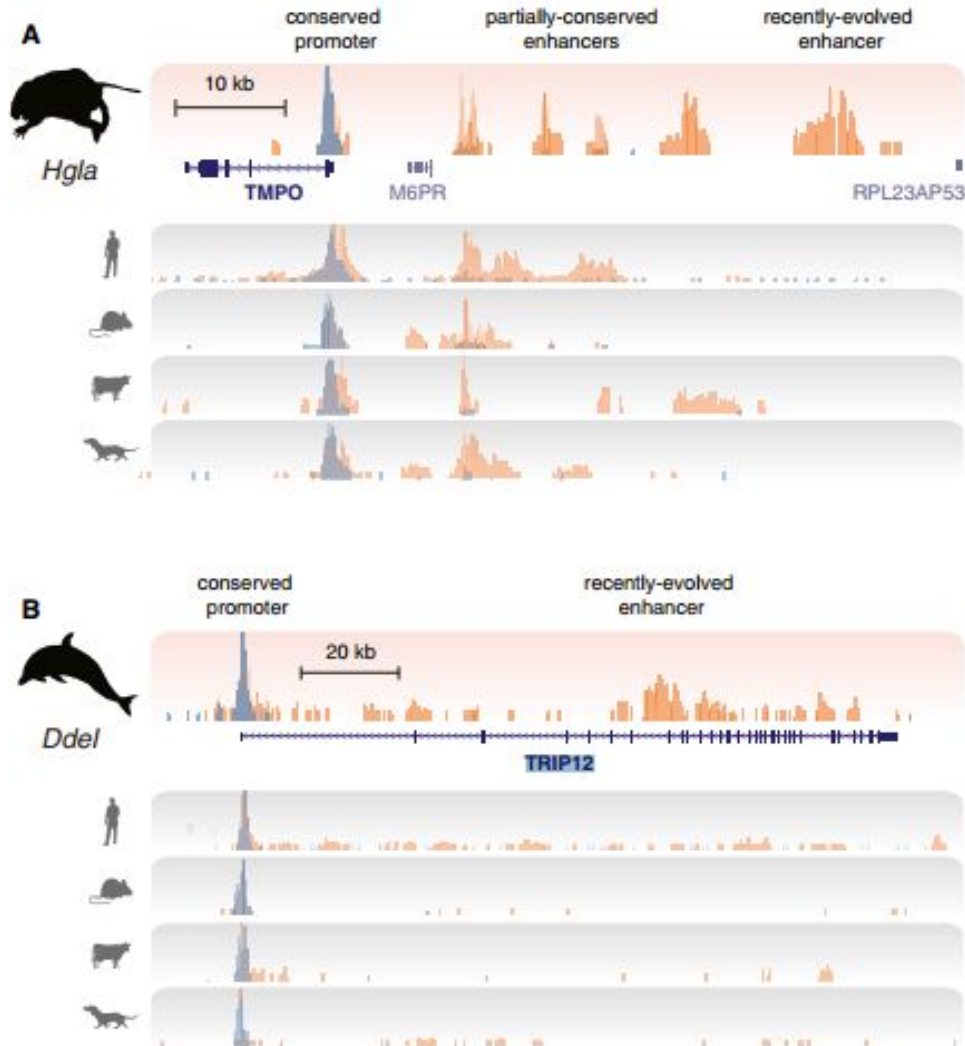


Энхансеры и промоторы



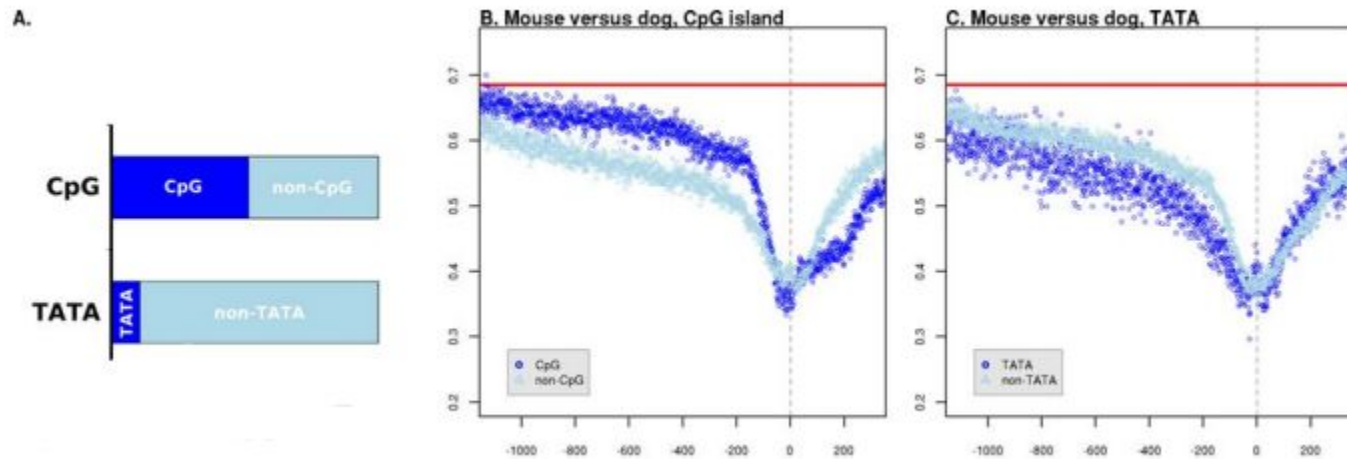
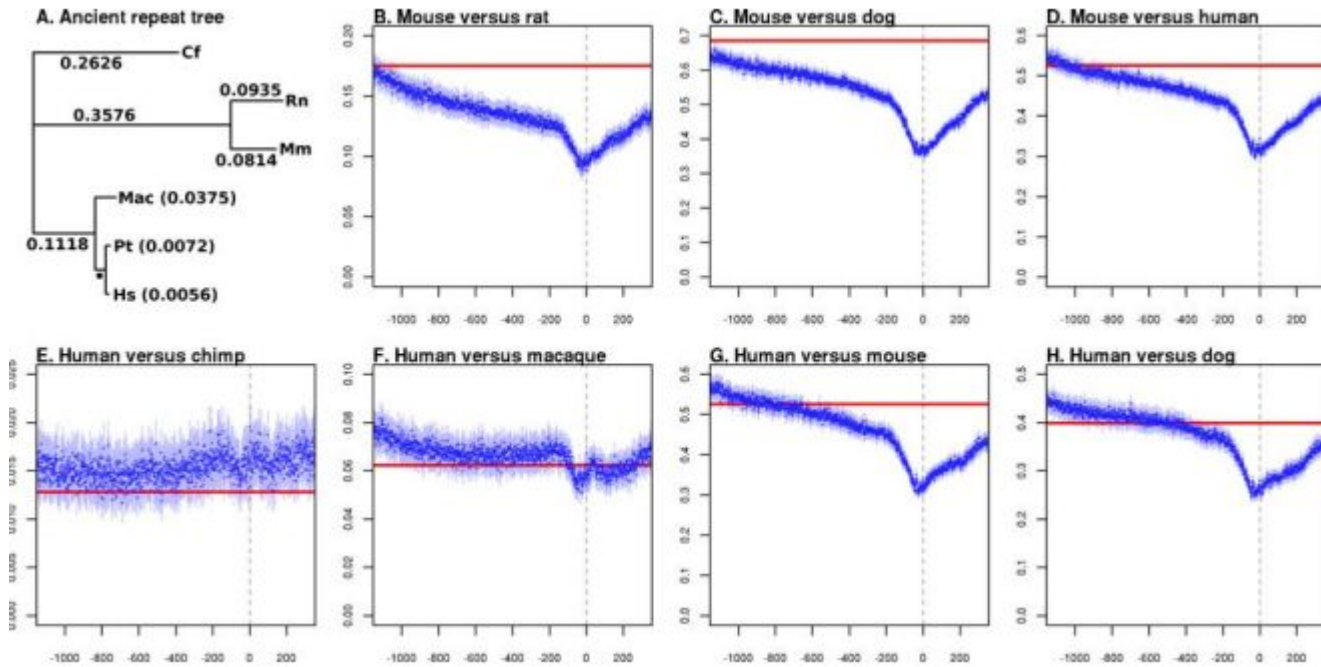
При этом, недавно изменившиеся последовательности промоторов млекопитающих большей частью являются «молодой ДНК», занесённой в проксимальную часть гена мобильными элементами, а энхансеры, наоборот, в большем числе случаев происходят в результате

Возникновение энхансеров de novo




Большинство энхансеров млекопитающих их апоморфны

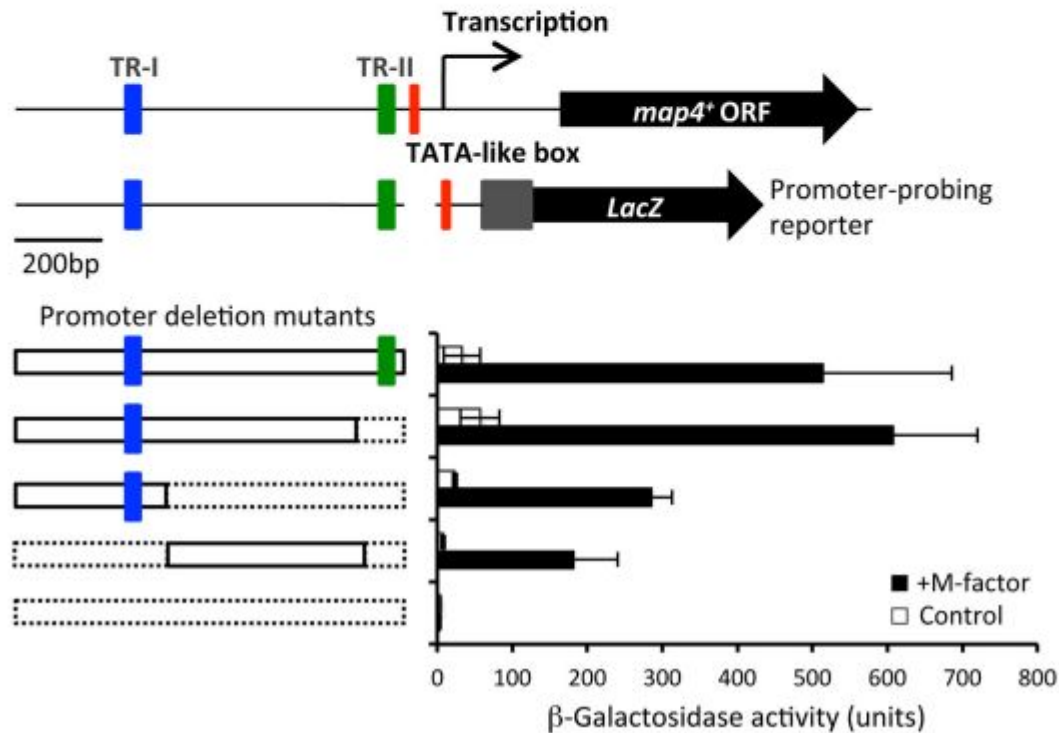
Промоторы млекопитающих

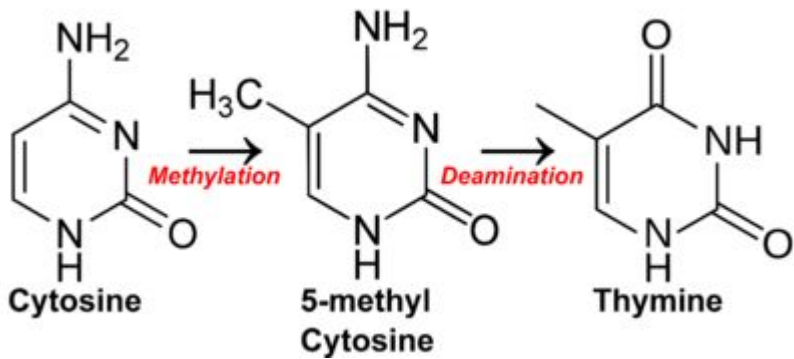


A

1 AAATGAGCTTATGTGTTTAGTTCATTCGTGTTAAAGCACAAAAGAAGTATAATCTATAAA 60
 61 AAAATATGTCGTATCATTGATCGGCTTTCGTCGCAATGATCATCATATTTGCTTGCTAAT 120
 121 GTCCATTGTTTTATATTTTTTCTACGATATACATACCTATTTTTGAATAATACGTCGATT 180
 181 TTAGGTTATGTCAGTTTGTTTTTCTTTCCTTTTTGGTATAATAAGCAGTCTTTTTTCTTT 240
 241 GTGATTGGA**AAACAAAGCA**CTTACATCTGCTACCAAAAATGGATGCTTATGAAGTTTAGA 300
 301 CGTGTGTCCACCAAAGTATTAAGGAAAAGCCTCAAAGTGCTTAATTTTCTTCATATTTAA 360
 361 AGAAAAAGCATCTTTACTAATTATATATTATTATTATTTTCTCAGCAGCTTTATTTT 420
 421 CTTTGGGCTATATGCTGTTGATATTGCCGCCTTTTCTTTATGTTTTTCATCTGAATGTTA 480
 481 CTTAGCCAACTGCTAAACGTCACCACCGTTTTTGAGCTGTCTATGCTTCTGGTATTTCAA 540
 541 AATGGATGAGTATGCATTTGTGAAATTCCTTCGACTCAATG**TTCTTTGTTG**CTGAATTTA 600
 601 **TATATAA**TGTCTAGGATTTTAGACAATAGAGC**ATTGTTCCAAACAAGGCAACTCAAGCCA** 660


 Transcriptional start site

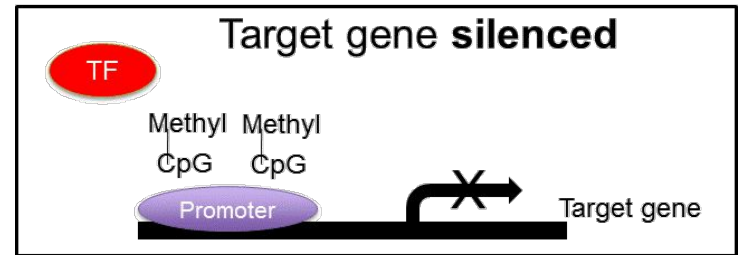
B



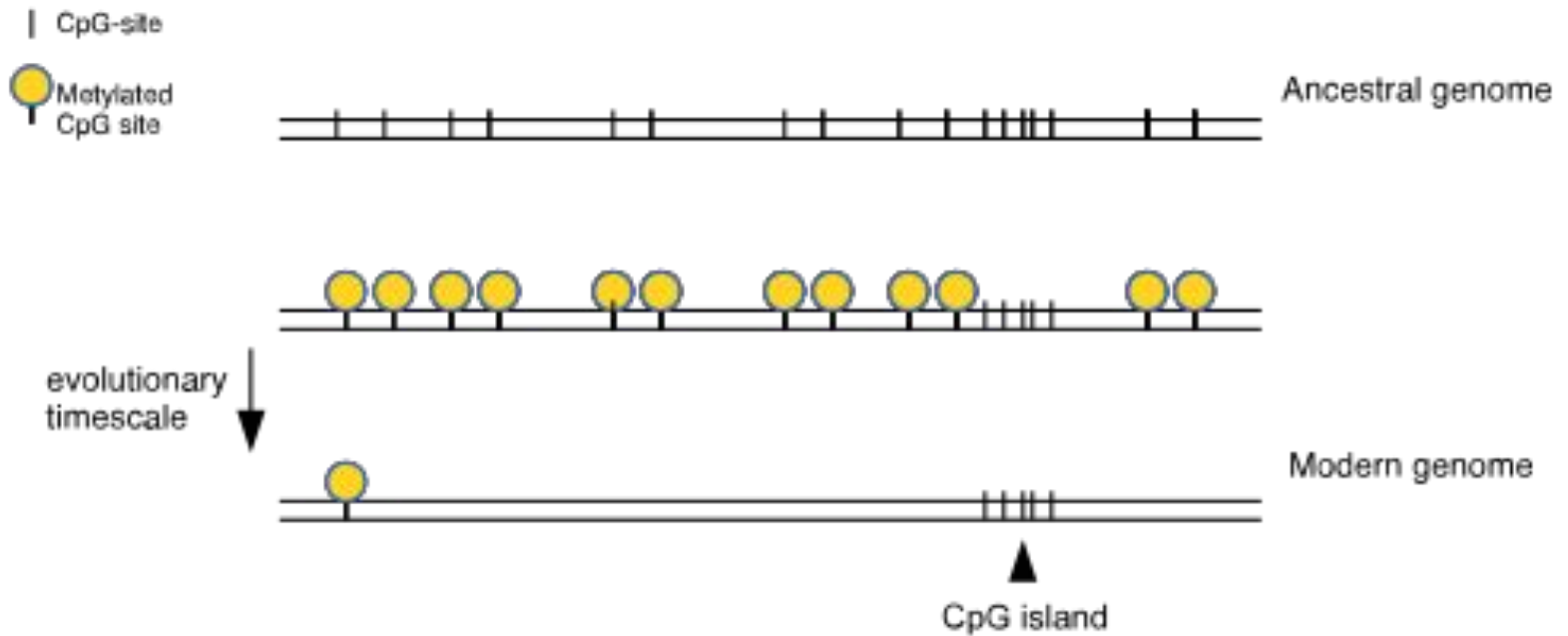
Target gene expressed



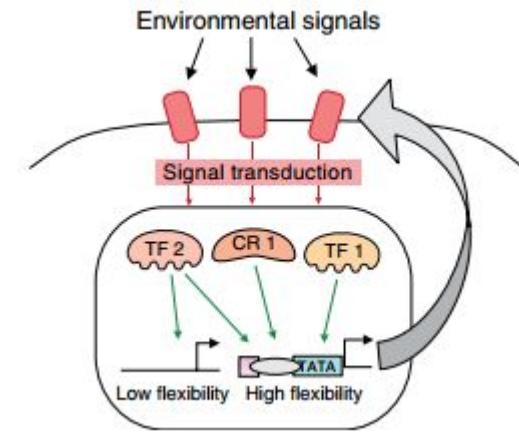
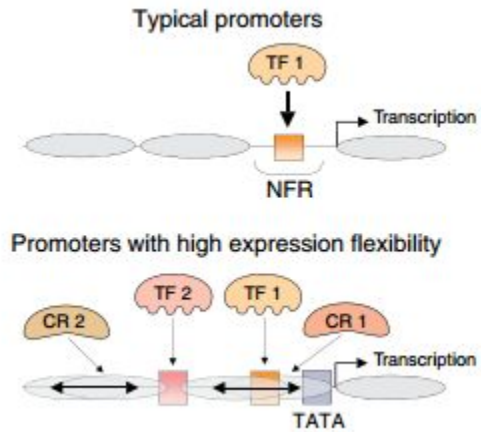
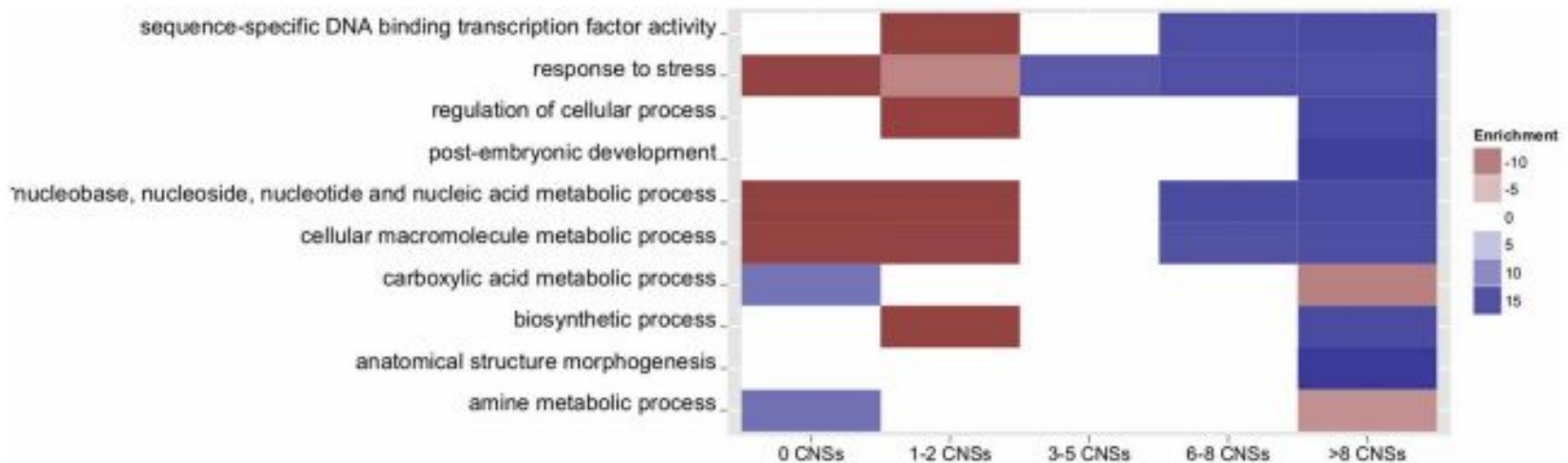
Target gene silenced



(Fry, 2011)

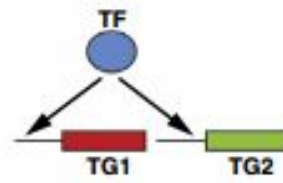


Гены, участвующих в развитии, ответе на стресс, регуляции клеточных процессов, обладают большим числом регуляторных элементов, чем гены домашнего хозяйства и некоторых метаболических процессов

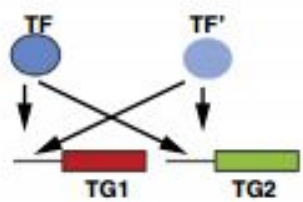


Эволюция системы транскрипционного фактора и гена-мишени

Pre-duplication

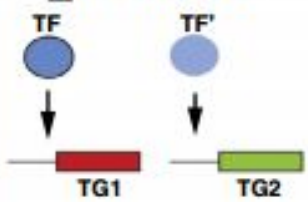


Duplication of the TF: initial redundancy



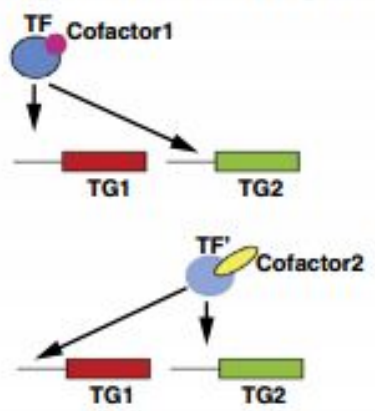
SUBFUNCTIONALIZATION

1 Division of labour

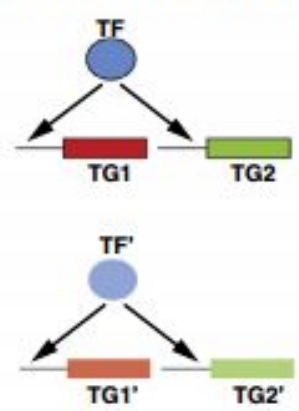


NEOFUNCTIONALIZATION

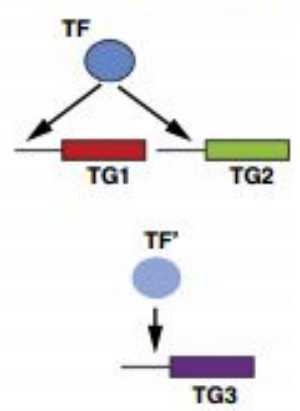
2 Divergence of a cofactor or a signal



3 Emergence of a novel circuit



4 Acquisition of a novel target



Duplication of the TF: diversification

Baker et al, 2013

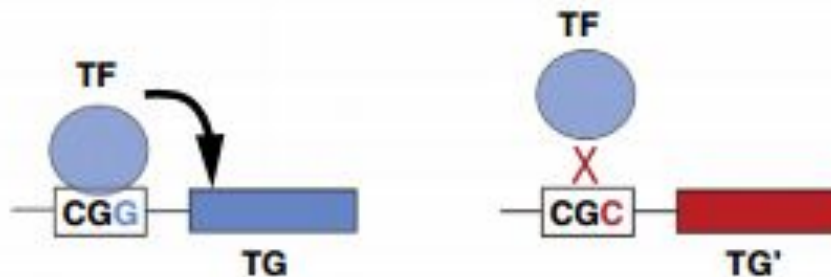
Bridgham et al, 2006

Pougach et al, 2014

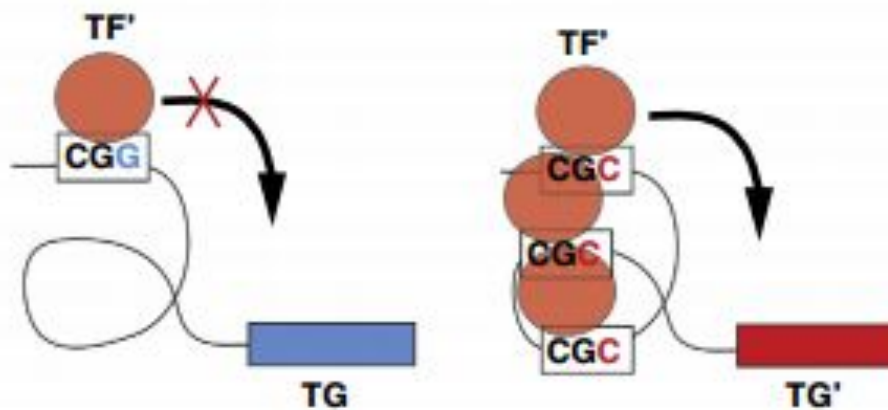
Perez et al, 2014

Эволюция системы транскрипционного фактора и гена-мишени

(b)

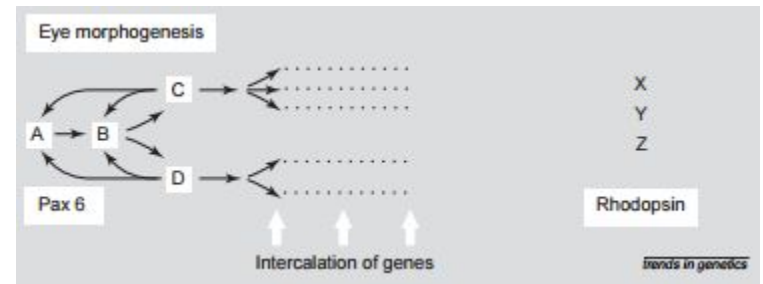
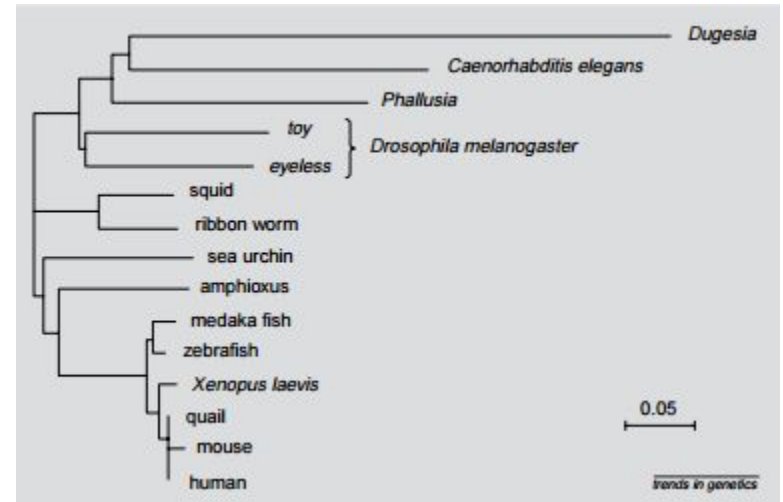
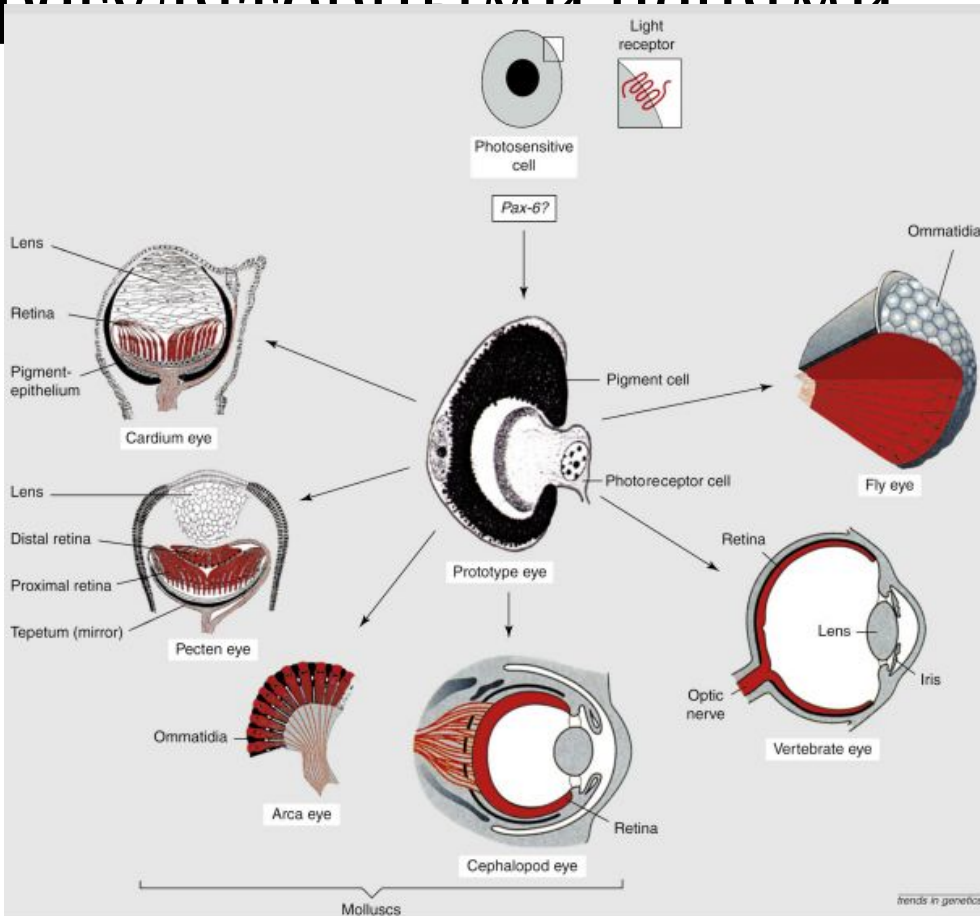


(c)



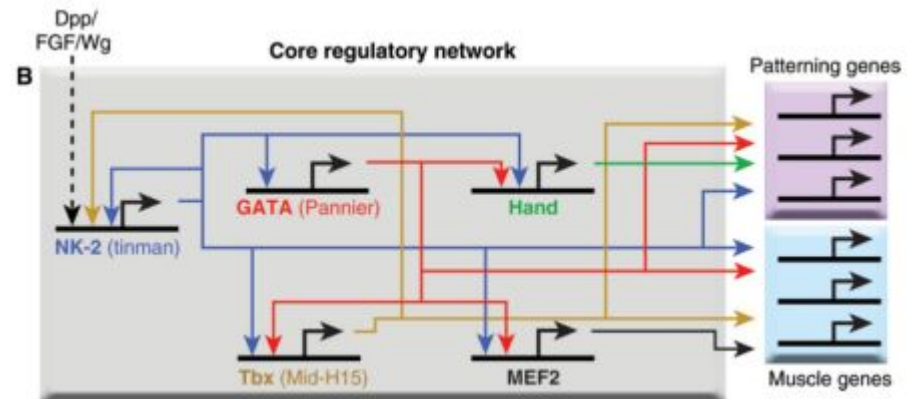
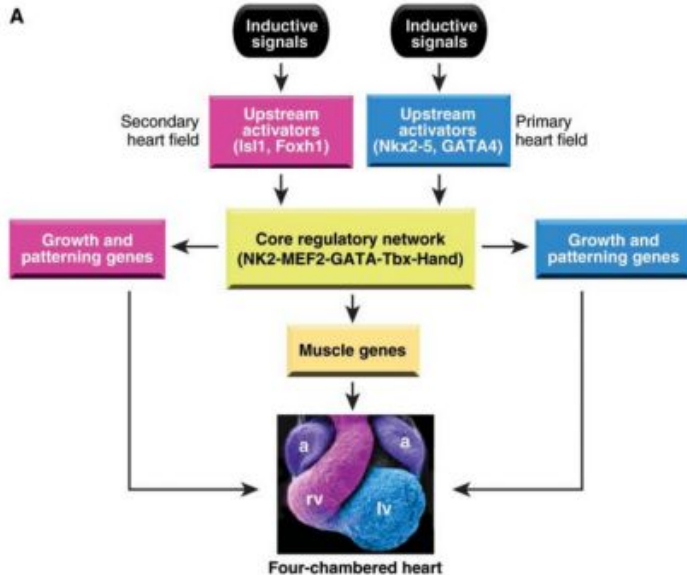
Глубокая гомология

Формирование и дифференциация многих структур, таких как глаза, конечности и сердце, управляется одним и тем же набором генов и сильно консервативными регуляторными цепями



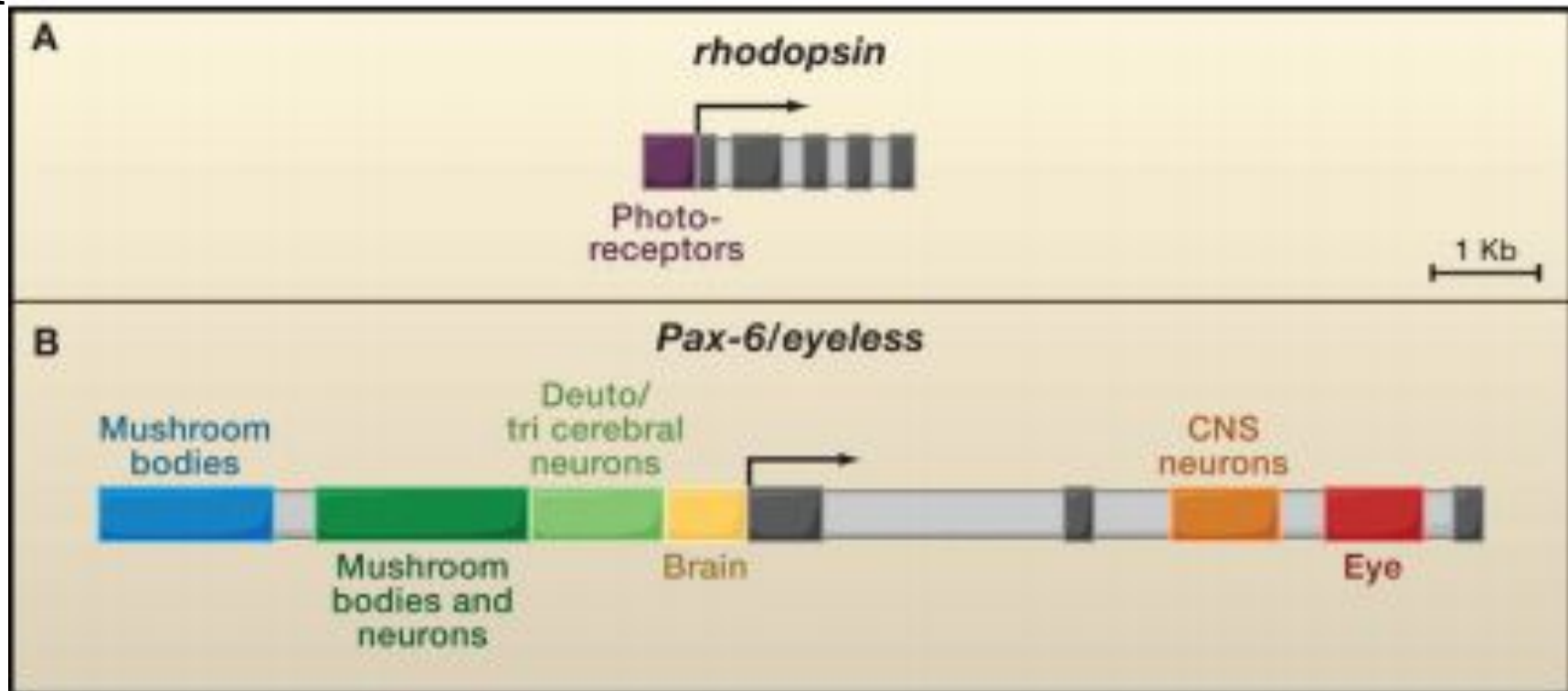
Глубокая гомология

| | Phylogenetic Group | Representative Organisms | Cardiac Structures | Cardiac Expression of Network Genes | | | | |
|---------------|--------------------|--|--|-------------------------------------|------|------|-----|------|
| | | | | NK2 | MEF2 | GATA | Tbx | Hand |
| Pre-bilateria | diploblastic | Cnidaria Hydra, jellyfish | Striated muscle cells associated with gastrodermis | 1 | 1 | ? | 1 | 0 |
| Bilateria | triploblastic | Nematodes <i>C. elegans</i> | Contractile pharynx | 1 | 1 | 2 | 1 | 0 |
| | | Anthropods <i>Drosophila</i> | Tubular heart Open circulatory system | 1 | 1 | 2 | 3 | 1 |
| | | Urochordates or tunicates Ascidians | Tubular heart, striated Open circulatory system Bidirectional pumping | 1 | 1 | 2 | ≥3 | 1 |
| | | Cephalochordates Amphioxus | Tubular heart, non-striated Closed circulatory system Unidirectional pumping | 1 | 1 | 2 | ? | 1 |
| | | Fish | Single atrium and ventricle | ≥3 | 4 | 3 | ≥4 | 1 |
| | | Amphibians | Two atria and one ventricle Pulmonary circulation | ≥3 | 4 | 3 | ≥5 | 1 |
| | | Amniotes Reptiles, birds, mammals | Four-chambered heart | ≥2 | 4 | 3 | ≥7 | 2 |



Модульность цис-регуляторных элементов

Отдельная черта toolkit-генов — большие и сложные модульные цис-регуляторные регионы.

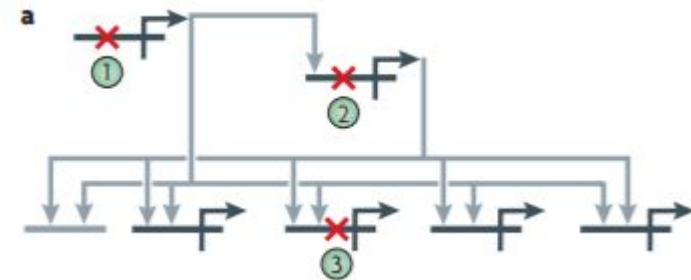
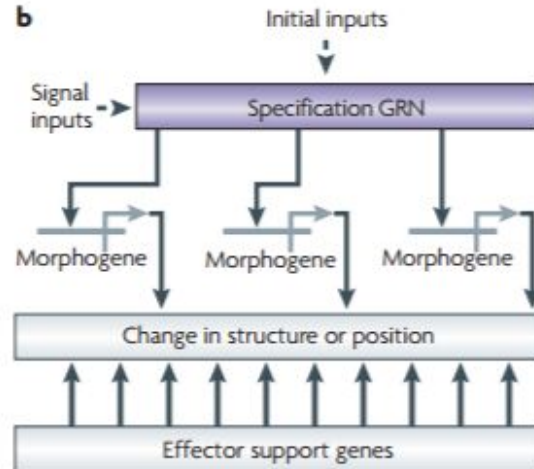
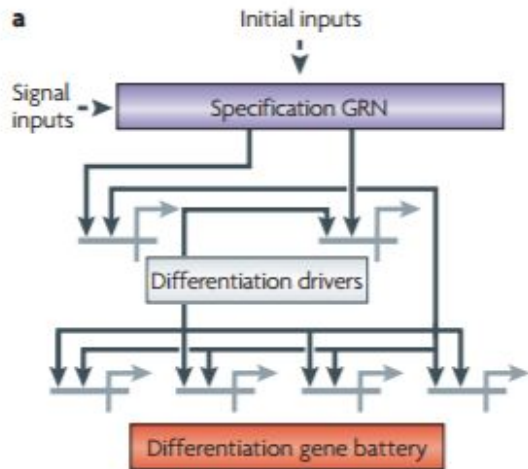


Модульность цис-регуляторных элементов

Наличие большого числа цис-регуляторных элементов у toolkit-генов открывает несколько важных для эволюции формы аспектов:

1. Множественные цис-регуляторные элементы — путь к увеличению числа функций без увеличения числа генов
2. Мутации в одном из ЦРЭ не затронут функции самого белка и функции других ЦРЭ, т. е. не будут иметь плейотропного эффекта
3. Мутации в гене с множеством ЦРЭ будут всегда иметь плейотропный эффект.

Большие генные сети

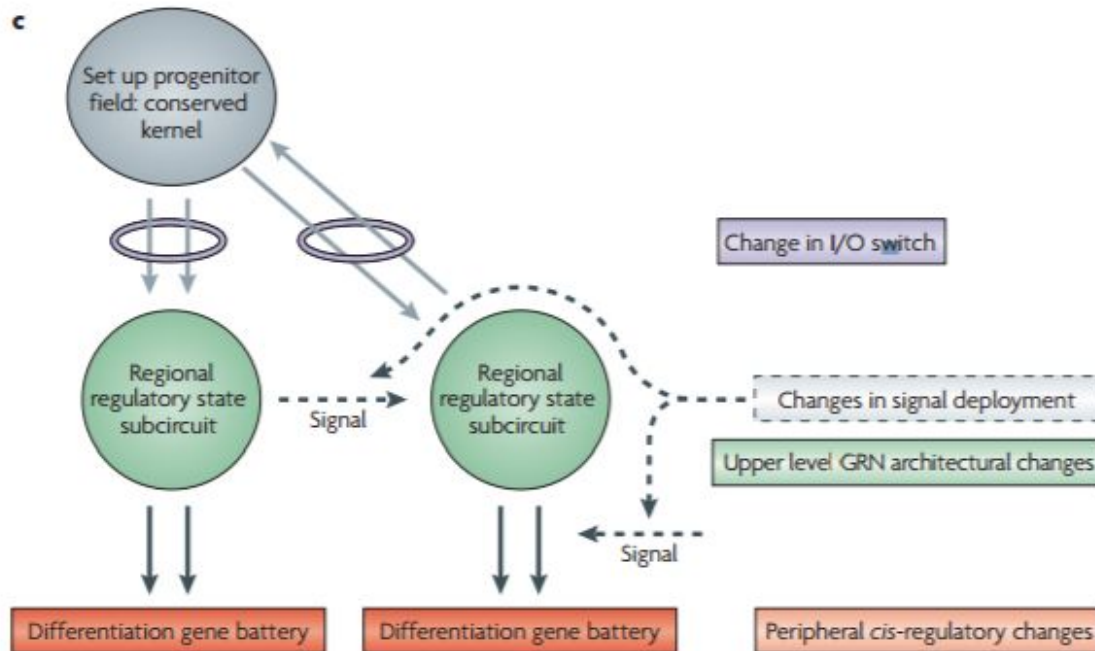


Cis-regulatory module loss

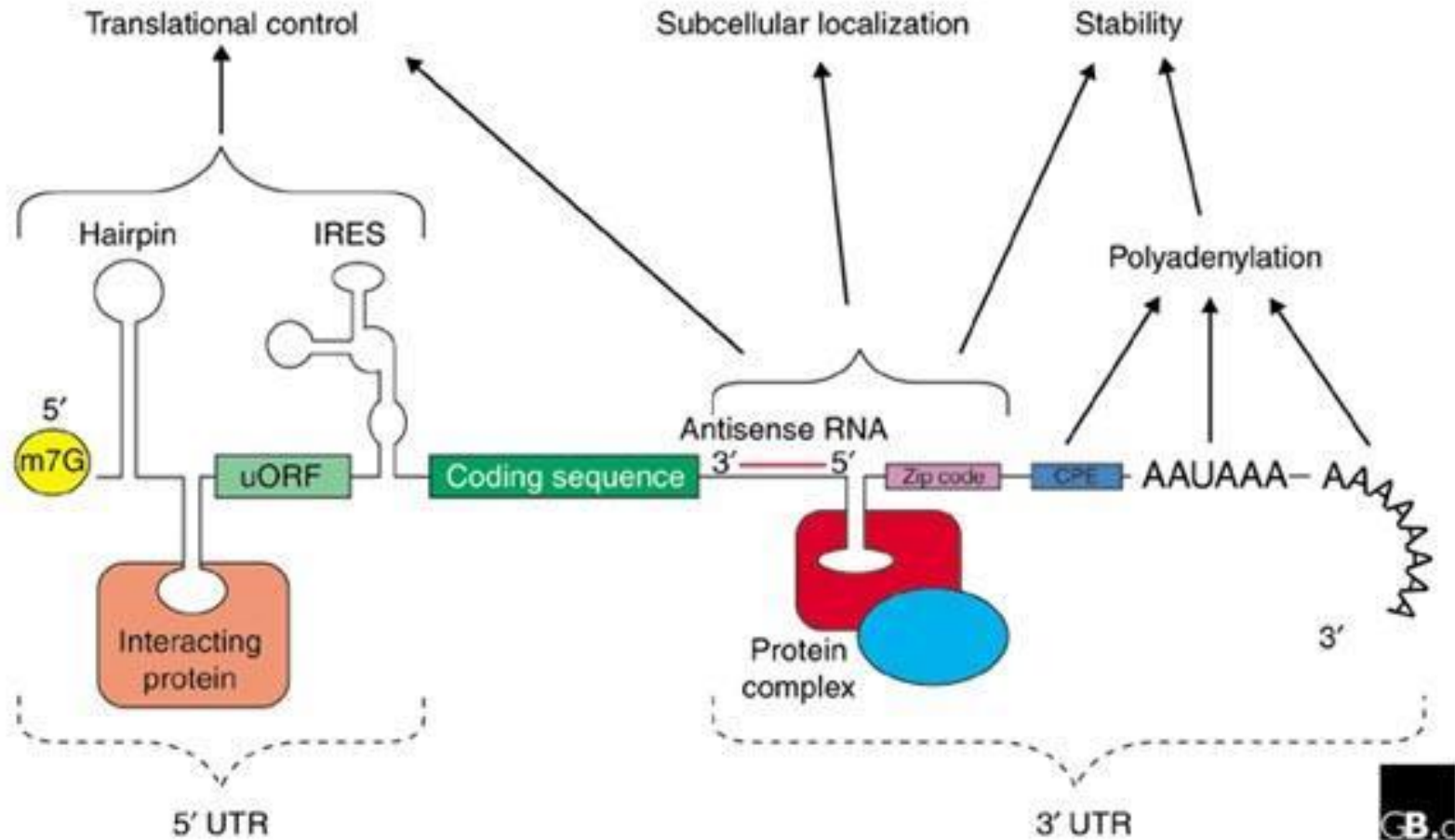
- ① Lose expression of whole battery
- ② Lose expression of whole battery
- ③ Lose expression of differentiation gene

Cis-regulatory module redeployment

- ① Battery expressed in a new developmental domain
- ② Lose expression of whole battery
- ③ Redeploy gene



Регуляция 3' и 5' UTR



Отрицательный отбор

