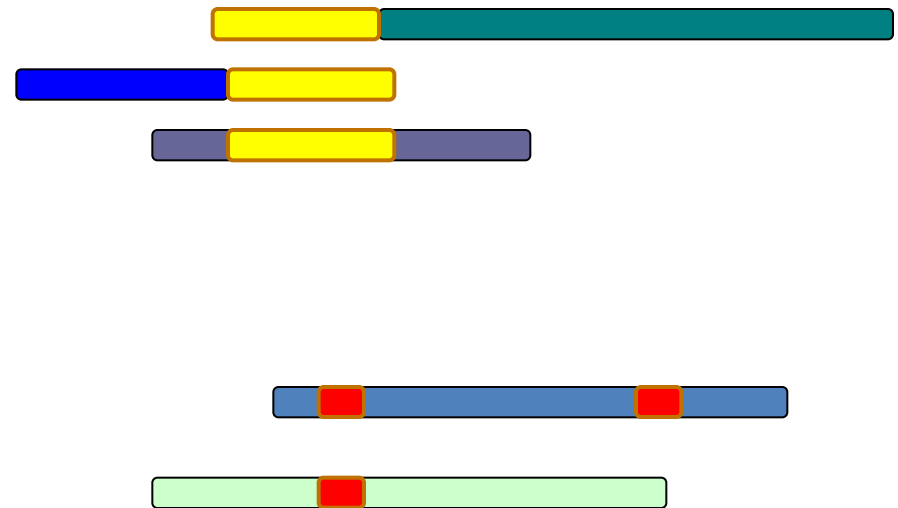


Быстрые пути эволюции белков.
Домен.
БД PFAM, InterPro.

Первые же попытки классификации коллекций аминокислотных последовательностей показали существование

семейств гомологичных белков

консервативных мотивов в негомологичных белках



- Домен – относительно консервативный в процессе эволюции непрерывный и достаточно длинный фрагмент последовательности.
- Последовательность может содержать один или несколько доменов



Домены, найденные в последовательностях, часто, но далеко не всегда совпадают со структурными доменами.



Словарик

ТИПЫ МОТИВОВ	Типы подписей
<u>Сайт</u> (site) -	Паттерн (pattern) –
<i>Мотив (motif) –</i>	Профиль–PSSM –
<u>Повтор (repeat)–</u>	Профиль–НММ -
<u>Домен (domain) –</u>
<i>Семейство –</i>	
<i>Суперсемейство -</i>	

Простой пример:

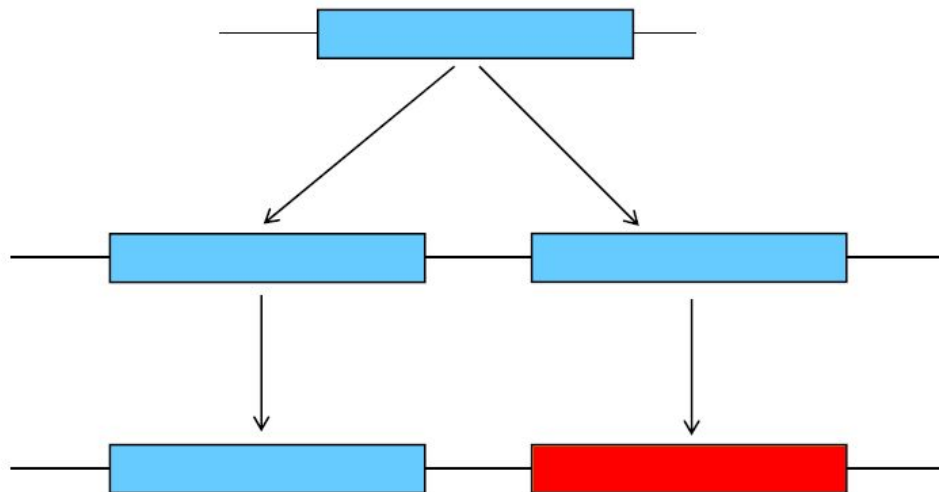
ССНС- цинксвязывающий
МОТИВ

Подпись типа паттерн –
CX2CX4HX4C

Почему это интересно?

Ничто удачное не должно пропасть даром

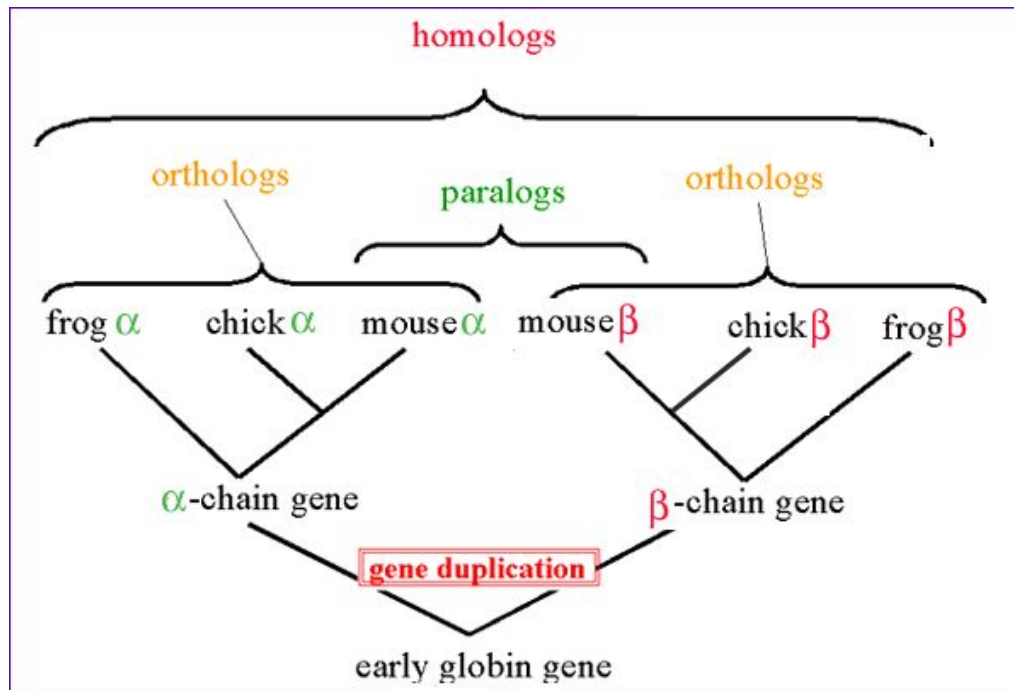
- Точечные мутации – медленный путь эволюции
- Быстрый путь – дупликация
- Дублируются **короткие фрагменты**, **домены**, гены, кластеры генов, хромосомы, геномы
- Вторая копия может приобрести новые функции



Пример 1: дупликация генов (повторяем)

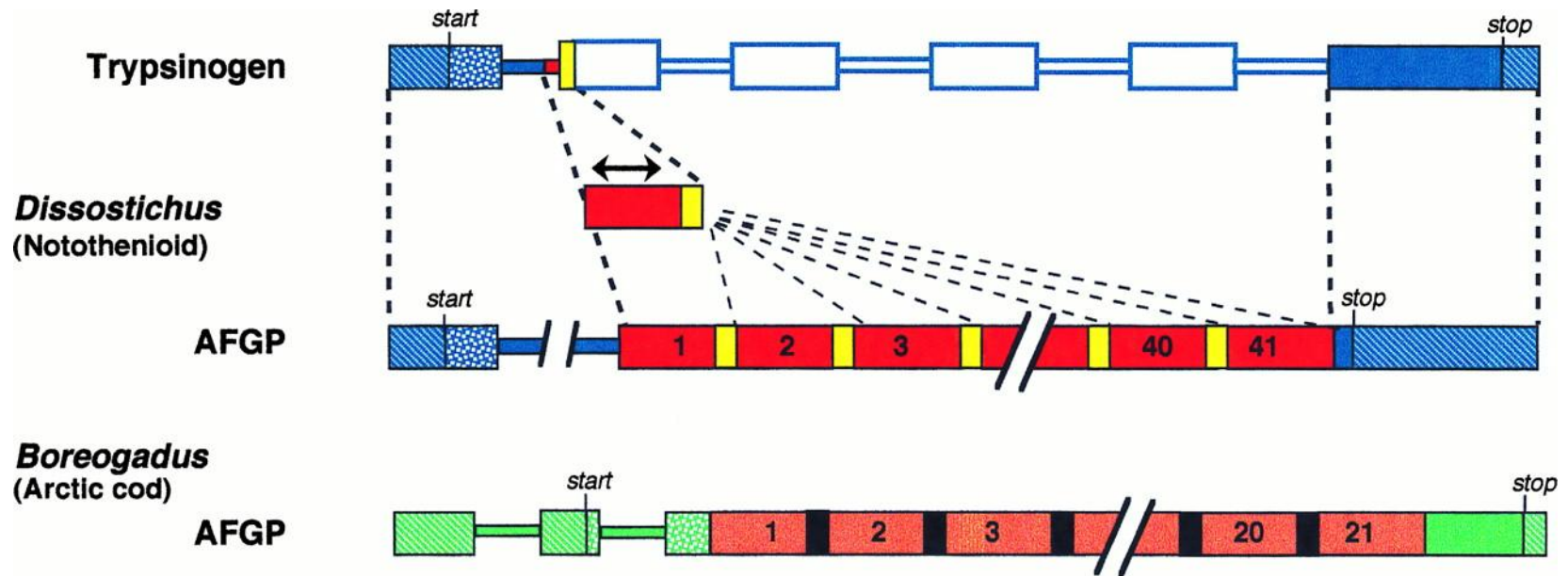
Ортологи — последовательности, возникшие из одного общего предшественника в процессе видообразования. Ортологи, как правило, имеют одну и ту же функцию

Паралоги — последовательности, возникшие из одного общего предшественника в результате дупликации одного гена в одном организме. Паралоги, как правило, имеют разные функции.



Пример 2: дупликация короткого повтора

Comparison of gene structures and their sequence similarities.



Logsdon J M , Doolittle W F PNAS 1997;94:3485-3487

Пример доменных перестроек: не только дубликации!

3

бе



ЕС 4.1.2.25

~~243~~
~~50~~
белка
~~7~~



ЕС 2.7.6.3

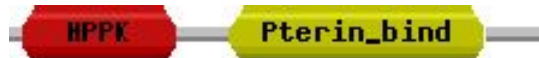
бе



ЕС 2.5.1.15

ЛК

• 25 белков
ОВ



• 9 белков



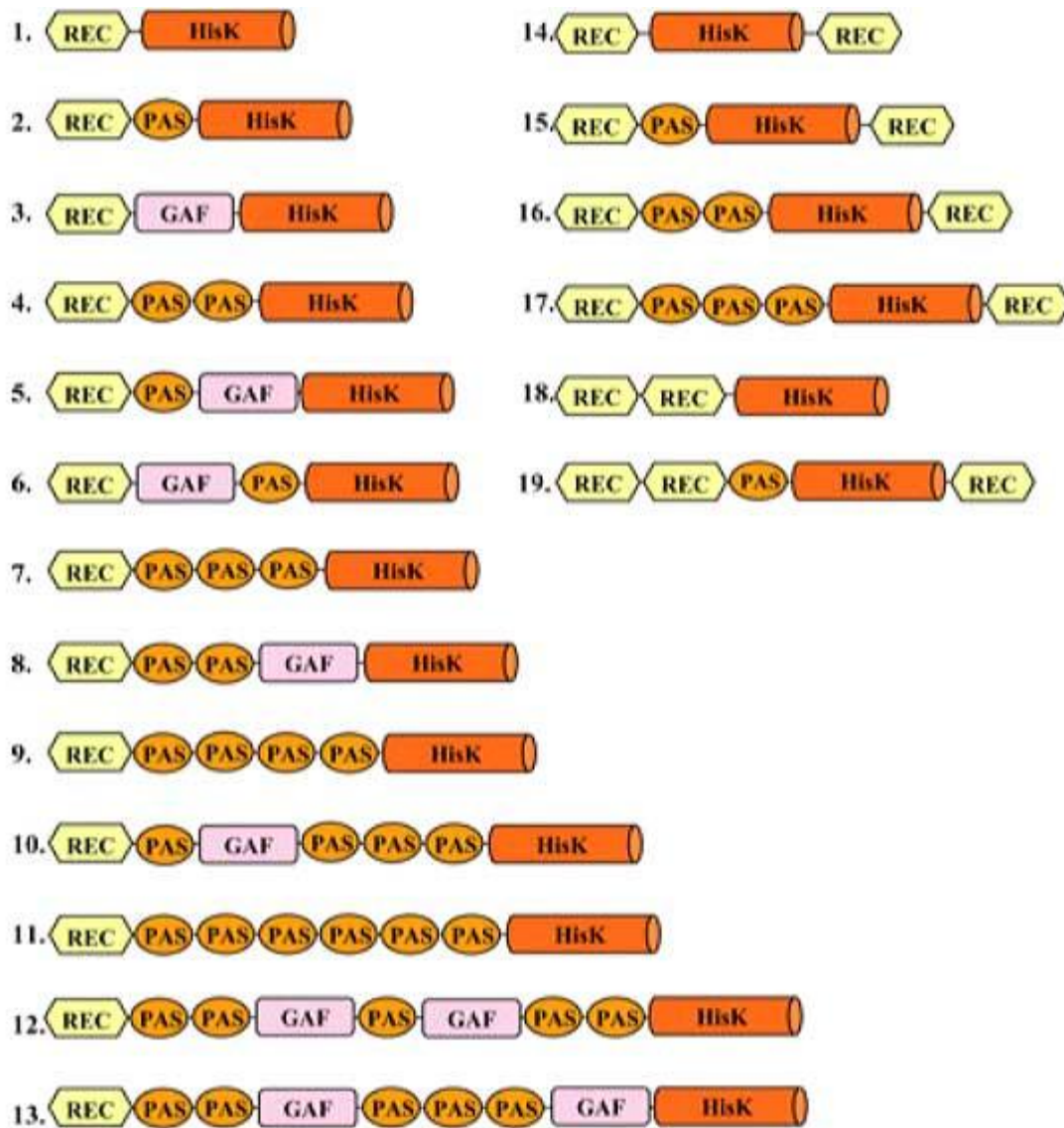
• 2 белка



• 12 белков



Пример 4



Structural Classification of Bacterial Response Regulators: Diversity of Output Domains and Domain Combinations

Домен – единица эволюции, структуры и функции белков.

Домен – компактная, относительно независимо сворачивающаяся структура, относительно консервативная в процессе эволюции.

Белок может состоять из одного или многих доменов.

Pfam



- <http://pfam.sanger.ac.uk>
- Большая коллекция семейств доменов
Для каждого семейства есть множественное выравнивание и профиль-НММ .
- Состоит из 2-х частей:
 - PfamA – курируемая часть, покрывает 76% UniProt
 - PfamB – большое число маленьких семейств из автоматически сгенерированной базы доменов, не вошедших в PfamA (раньше – ProDom, теперь – ADDA) .
- Удобна для анализа доменной структуры белков.



Язык Pfam :

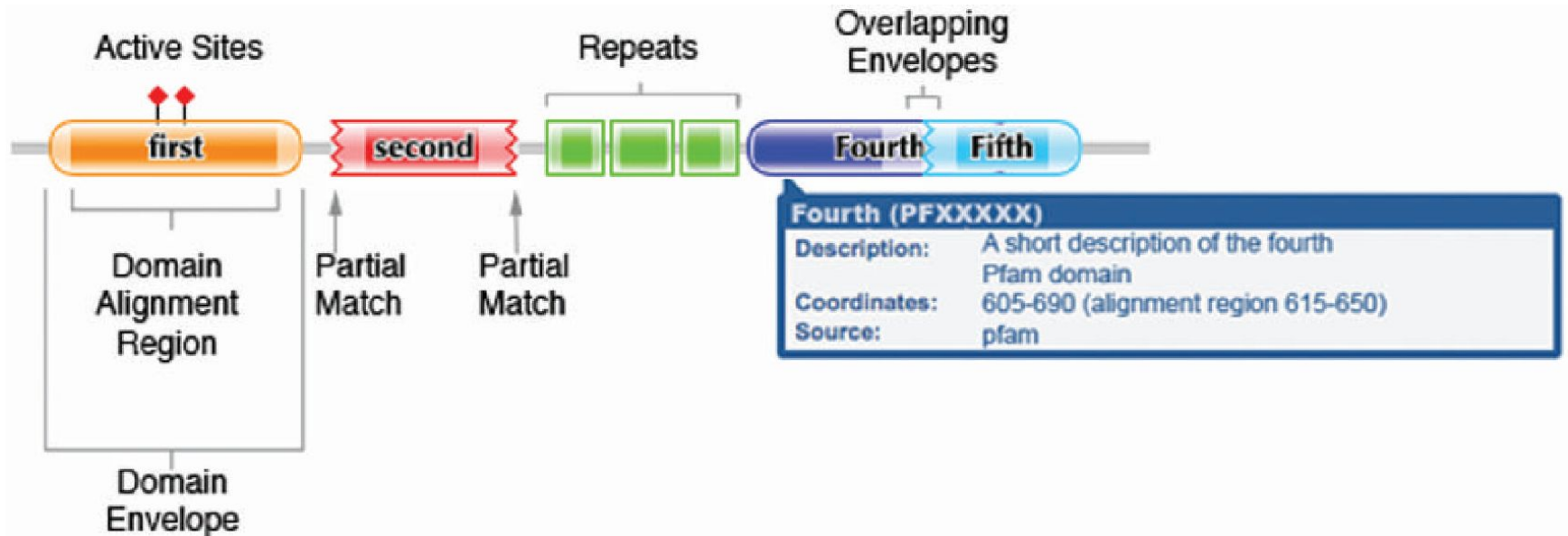
Семейство – коллекция гомологичных белков.

Домен – структурная единица, которую можно найти во множественном выравнивании.

Повтор – короткая единица, нестабильная сама по себе, но образует стабильные структуры, если есть много копий.

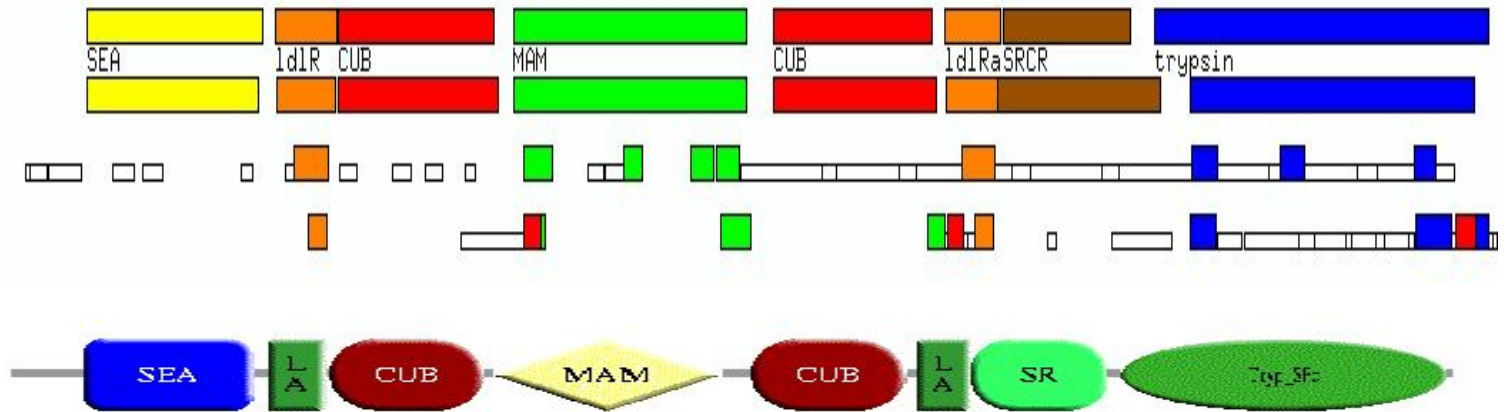
Мотив – короткая единица структуры вне глобулярных доменов.

Клан – группа родственных записей.



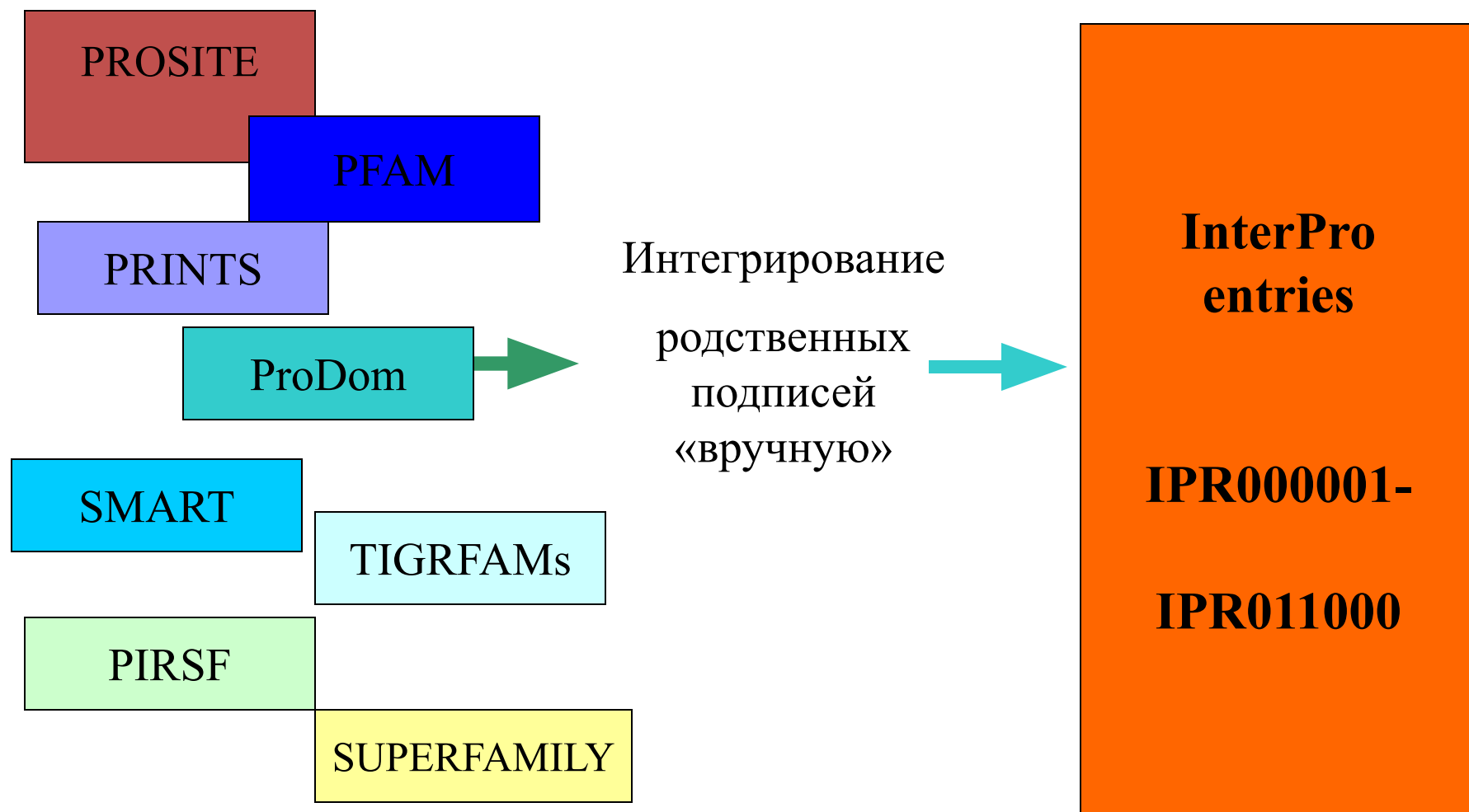
СРАВНИТЕ

Pfam
Prosite
Prints
Blocks
Smart



(ProDom, PIRaln, ProClass, Systers, Picasso etc. not shown)

Создание интегрированной базы данных InterPro



InterPro- an integrated resource of protein families, domains and functional sites.