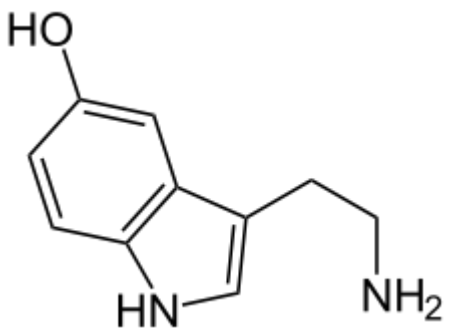
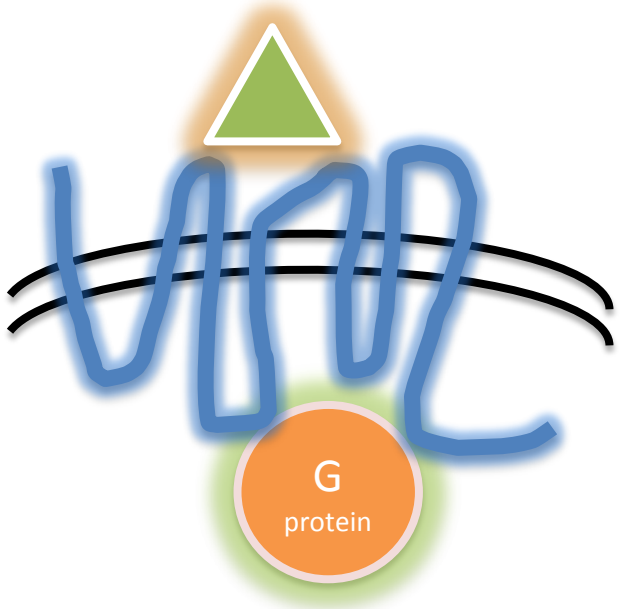
A trial of Bioinformatic analysis in Serotonin Receptor gene family

杨威
王魏然
刘绍峰

1. Background： Serotonin Receptors





neurotransmitter serotonin

GPCR structured serotonin

I. Importance:

- Serotonin receptors regulate neuronal signaling by exciting or inhibiting its neurotransmitter release.
- There are several animal neuronal activity coupled with serotonin receptors, including pain perception, angry expression etc.

1. Background： Serotonin Receptors

Family	Type	Mechanism	Potential
5-HT_1	G_i/G_o-protein coupled.	Decreasing cellular levels of cAMP.	Inhibitory
5-HT_2	G_q/G_{11}-protein coupled.	Increasing cellular levels of IP_3 and DAG.	Excitatory
5-HT_3	Ligand-gated Na^+ and K^+ cation channel.	Depolarizing plasma membrane.	Excitatory
5-HT_4	G_s-protein coupled.	Increasing cellular levels of cAMP.	Excitatory
5-HT_5	G_i/G_o-protein coupled.[6]	Decreasing cellular levels of cAMP.	Inhibitory
5-HT_6	G_s-protein coupled.	Increasing cellular levels of cAMP.	Excitatory
5-HT_7	G_s-protein coupled.	Increasing cellular levels of cAMP.	Excitatory

Eur J Pharmacol. **361** (2–3): 299–309.

II. Diversity:

- There are 15 types of serotonin receptors, 14 of which are GPCRs. It is complicated that some serotonin receptor excite neuronal activity while others inhibit, which can be reflected by their protein diversity in sequence and protein structure.
- Our interests lie in the relationship between the two.

2. Procedures

1. List all members of serotonin receptor family;
2. Evolutionary distribution of serotonin family members;
3. Analysis of PTM source of serotonin receptor functional diversity;
4. GPCR TM analysis of serotonin receptor to reveal its bifurcation on G protein subtype selection and the output distinction in nervous system

Background:

Structural Overview of GPCR-members of Serotonin Receptors

2RH1

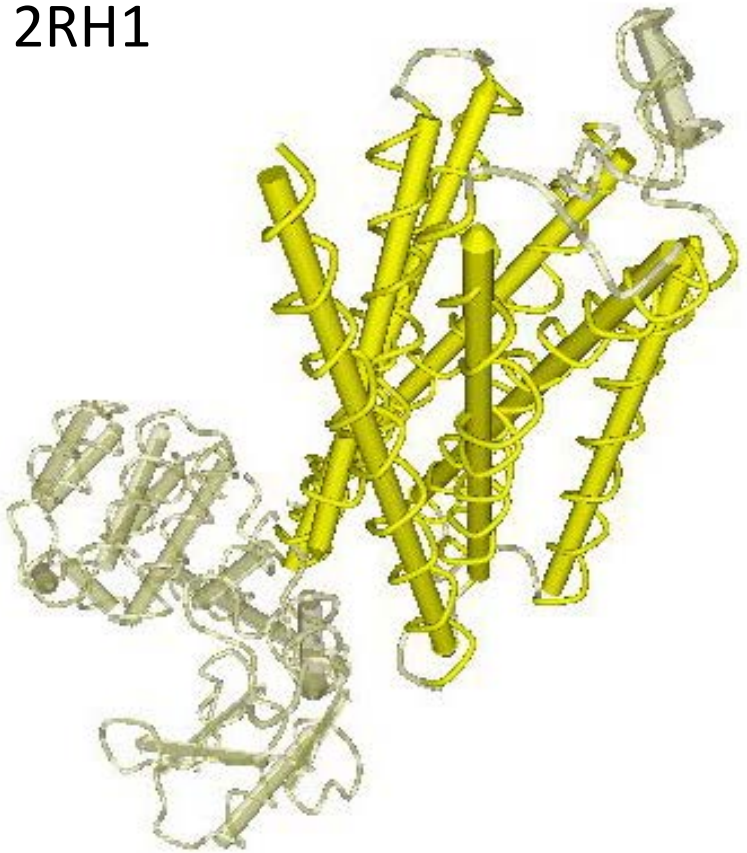


Fig. beta-2 adrenergic receptor (2^{nd} crystalized GPCR)

Most Serotonin Receptors are GPCRs, which are integral membrane protein that transduce signals inward.

1. N-terminal:

- 7 helical, hydrophobic TM;
- 3 extra- and 3 intra-cellular loops;
- barrel-like ligand binding domain;
- second extra-cellular loop serves as 'lid' for the ligand binding cavity;
- 2^{nd} and 3^{rd} intracellular loops are within the GEF domain, which are key to G-protein association
- At rest, inactive GDP-Galpha associates with GEF domain
- Upon ligand bound, a conformational change leads to GDP-GTP exchange of G protein.

Background:

Structural Overview of GPCR-members of Serotonin Receptors

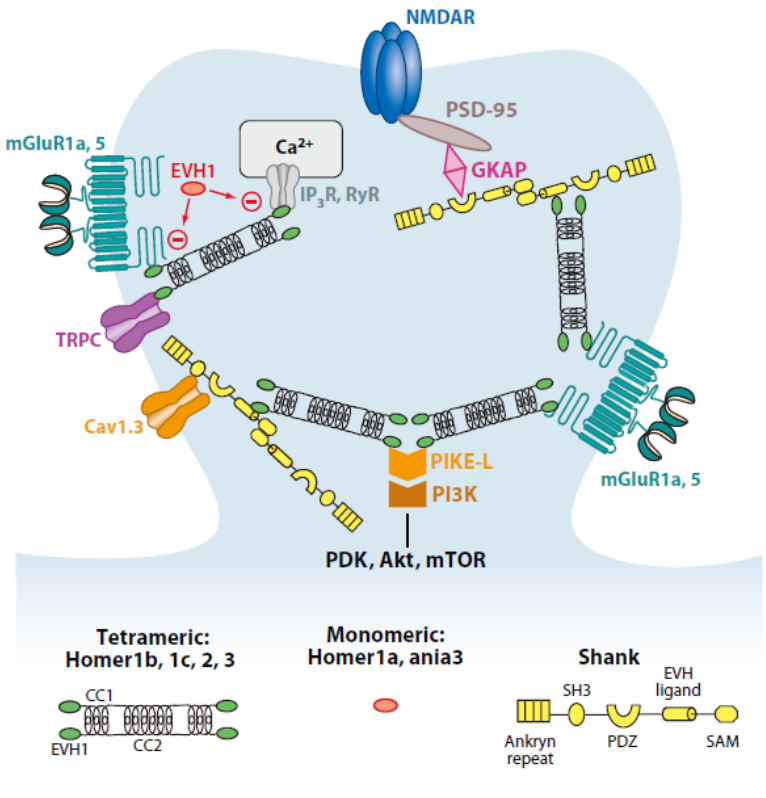


Fig. highly diverse C-termini function of GPCR(2nd crystalized GPCR)

N-termini are highly variable in seq, but diverse in its interaction with intracellular protein network, by virtue of its rich PTMs.

2. C-terminal:

Highly variable in sequence and PTM;

- Glycosylation

  Important for trafficking etc.

- Phosphorylation

  alternative pathway switch from G-protein type to beta-arrestin pathways

- Palmytoylation or lipidation

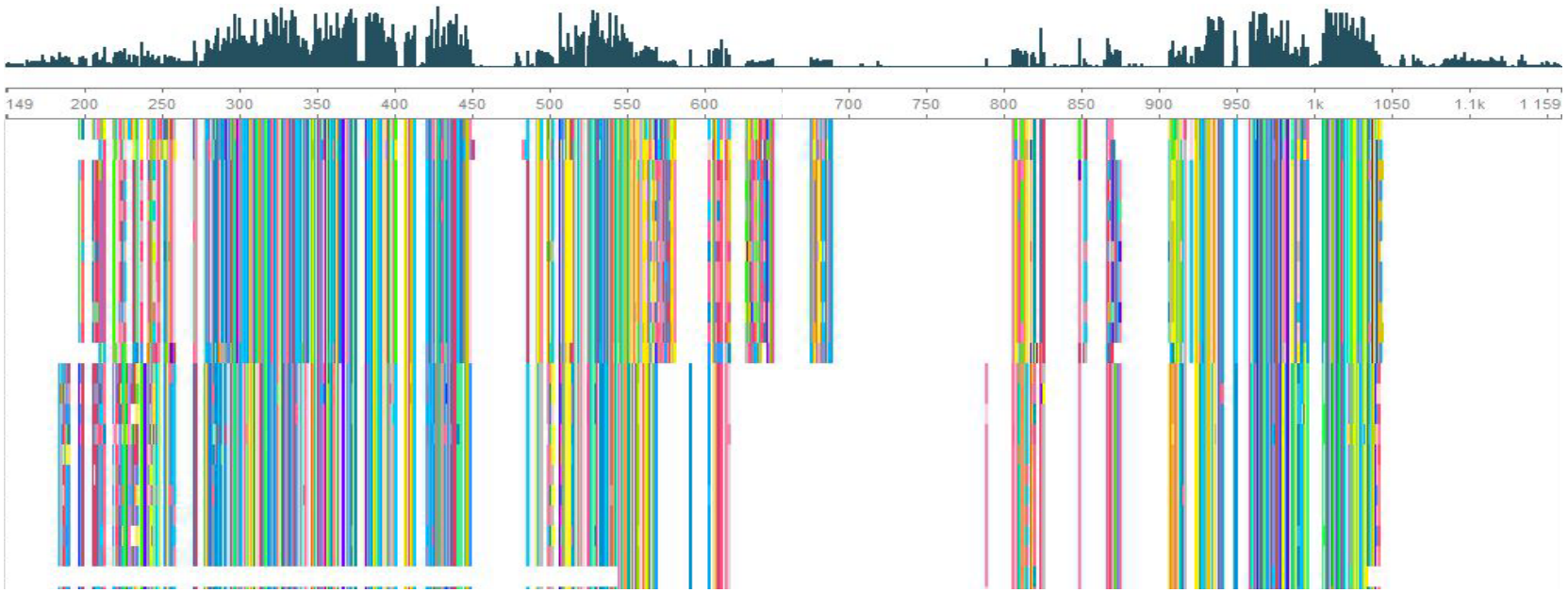
  hydrophobic acryl group addition onto cystein, critical for signal complex stabilization.

First question of interest:

The sequence source of Serotonin receptor functional diversity

Our first question of interest is:

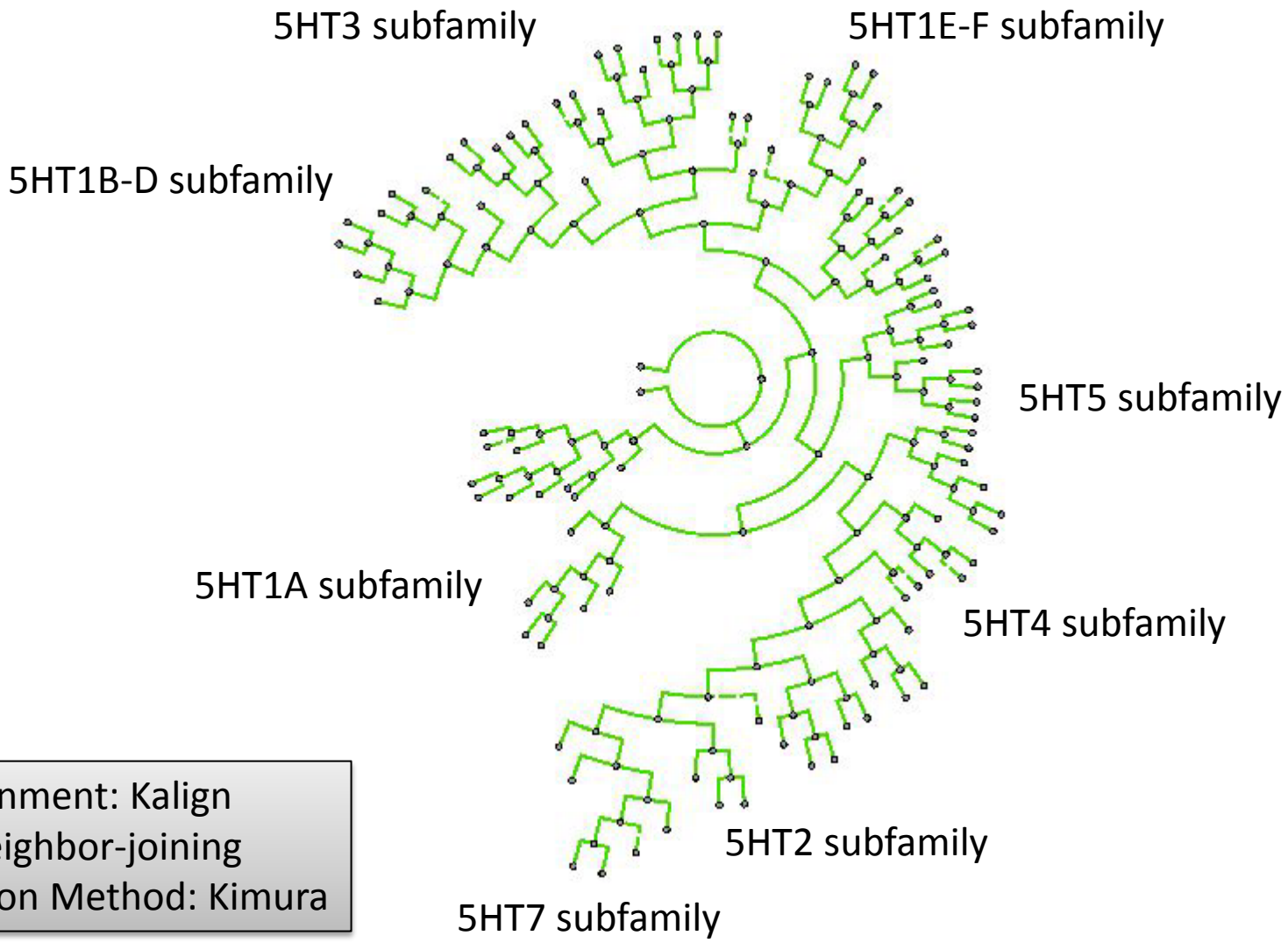
What is the source of functional diversity of serotonin receptor family?



Candidtate answers:

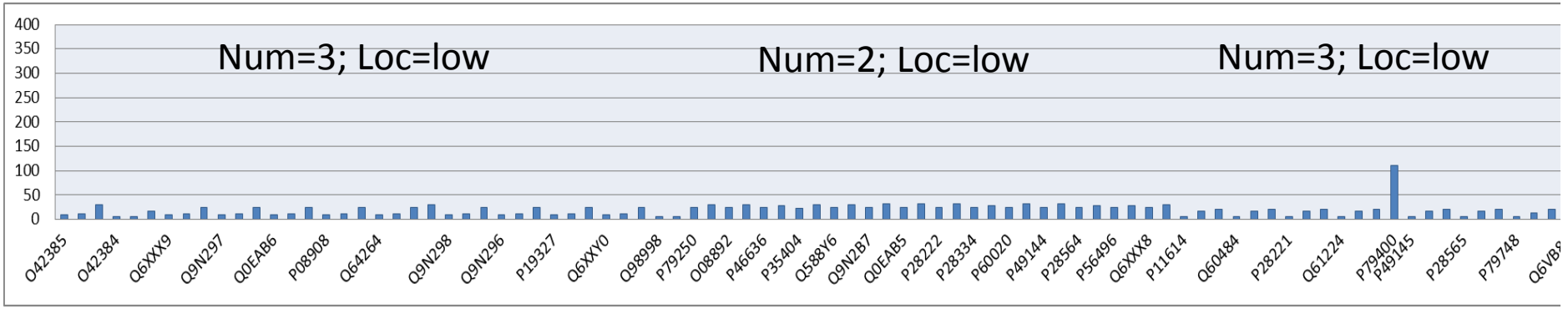
1. PTMs;
2. Conformational distinctiveness;
3. special sequence features.

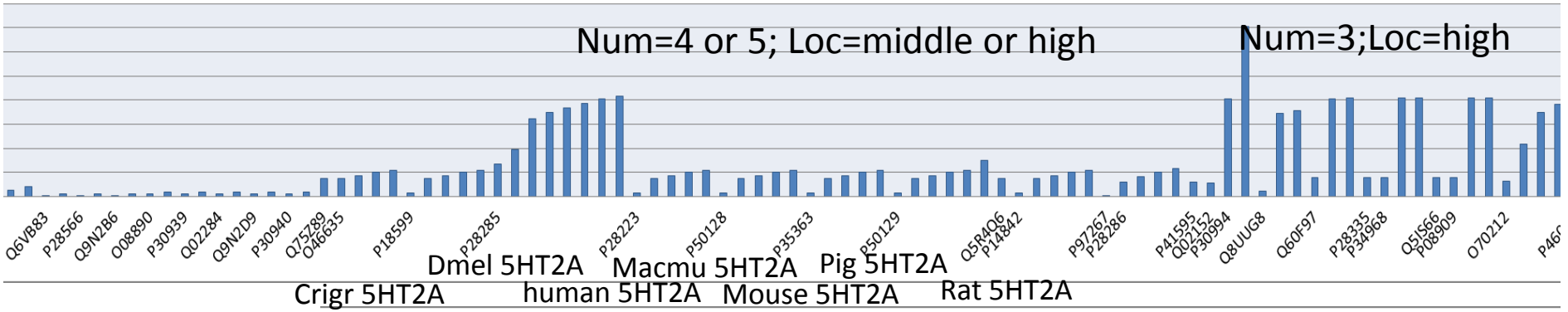
K-alian (local mode)
Gap Open: 53.90
Gap Extension: 8.52
Terminal Gap: 4.42
Bonus Score: 0.02



Multi-alignment: Kalign
PHYLIP neighbor-joining
Substitution Method: Kimura

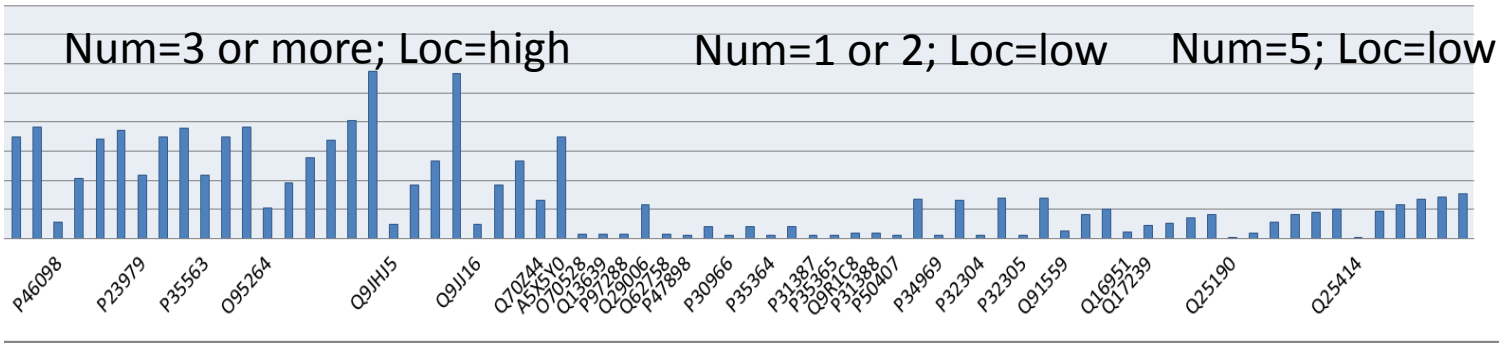
Serotonin Receptor Feature: Glycosylation diversity





Two parameters:

1. number;
2. locations



Serotonin Receptor Feature: phosphorylation

phosphoTyrosine:

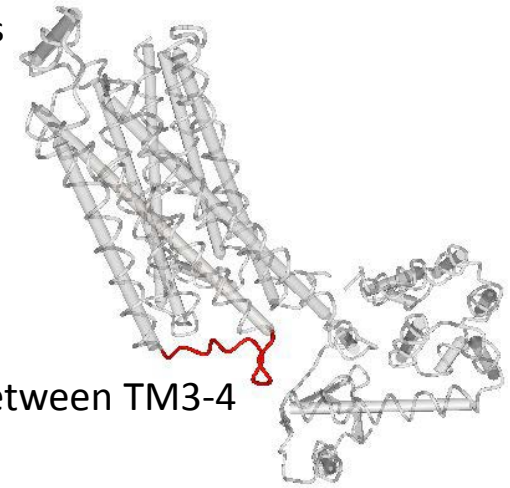
Nm/len	featr	loc
P50406/440	pY	125
Q9R1C8/440	pY	125
Q5IS65/440	pY	125
P31388/436	pY	125

MUSCLE Multi-align

Annotation

- [ ] Topological domain
- [ ] Chain
- [x] Modified residue
- [ ] Sequence conflict
- [x] Transmembrane
- [x] Glycosylation
- [ ] Disulfide bond

N-terminus

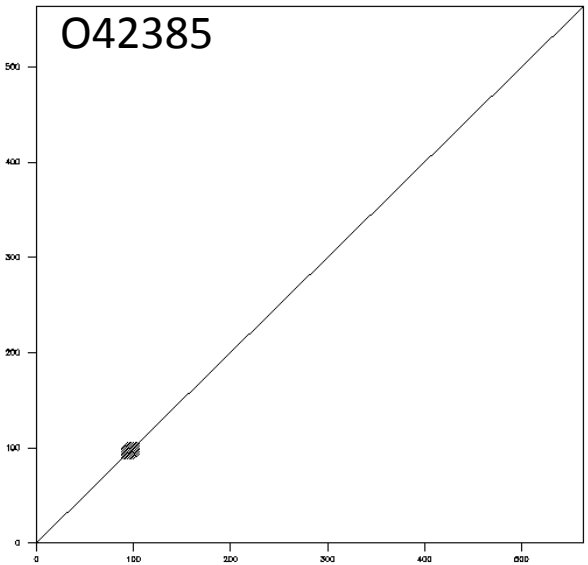


Loop between TM3-4

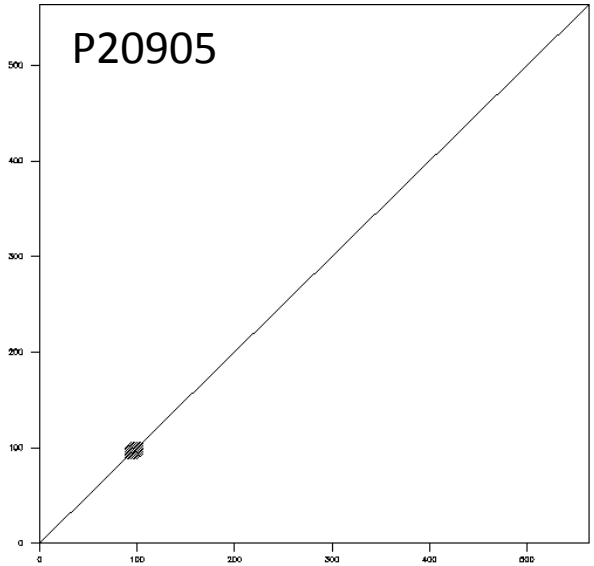
```
1    MVPEPGPTANSTPAWGAGPPSAPGGSGWVAAALCVVIALTAAANSLLIALICTQPALRNT   60  P50406   5HT6R_HUMAN
1    MVPEPGPVNSSTPAWGPGPPPAPGGSGWVAAALCVVIVLTAAANSLLIALICTQPALRNT   60  Q9R1C8   5HT6R_MOUSE
1    MVPEPGPSANSTPAWGAGPPSAPGGSGWVAAALCVVIALTAAANSLLIALICTQPALRNT   60  Q5IS65   5HT6R_PANTR
1    MVPEPGPVNSSTPAWGPGPPPAPGGSGWVAAALCVVIVLTAAANSLLIVLICTQPALRNT   60  P31388   5HT6R_RAT
     *******  .****** *** ****************.**********.***********

61   SNFFLVSLFTSDLMVGLVVMPFAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLI   120  P50406   5HT6R_HUMAN
61   SNFFLVSLFTSDLMVGLVVMPFAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLI   120  Q9R1C8   5HT6R_MOUSE
61   SNFFLVSLFTSDLMVGLVVMPFAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLI   120  Q5IS65   5HT6R_PANTR
61   SNFFLVSLFTSDLMVGLVVMPFAMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLI   120  P31388   5HT6R_RAT
     ************************************************************

121  SLDRYLLILSPLRYKLRMTPLRALALVLGAWSLAALASFLPLLLGWHELGHARPPVPGQC   180  P50406   5HT6R_HUMAN
121  SLDRYLLILSPLRYKLRMTAPRALALILGAWSLAALASFLPLLLGWHELGKARTSAPGQC   180  Q9R1C8   5HT6R_MOUSE
121  SLDRYLLILSPLRYKLRMTPPRALALVLGAWSLAALASFLPLLLGWHELGHARPPVPGQC   180  Q5IS65   5HT6R_PANTR
121  SLDRYLLILSPLRYKLRMTAPRALALILGAWSLAALASFLPLLLGWHELGKARTPAPGQC   180  P31388   5HT6R_RAT
     ******************  *****:**********************:**  .****

181  RLLASLPFVLVASGLTFFLPSGAICFTYCRILLAARKQAVQVASLTTGMA--SQASETLQ   238  P50406   5HT6R_HUMAN
181  RLLASLPYVLVASGVTFFLPSGAICFTYCRILLAARKQAVQVASLTTGTATAGQALETLQ   240  Q9R1C8   5HT6R_MOUSE
181  RLLASLPFVLVASGLTFFLPSGAICFTYCRILLAARKQAVQVASLTTGMA--SQASETLQ   238  Q5IS65   5HT6R_PANTR
181  RLLASLPFVLVASGVTFFLPSGAICFTYCRILLAARKQAVQVASLTTGTA--GQALETLQ   238  P31388   5HT6R_RAT
     *******:******:********************************* *  .** ****

239  VPRTPRPGVESADSRRLATKHSRKALKASLTLGILLGMFFVTWLPFFVANIVQAVCDCIS   298  P50406   5HT6R_HUMAN
241  VPRTPRPGMESADSRRLTTKHSRKALKASLTLGILLSMFFVTWLPFFVASIAQAVCDCIS   300  Q9R1C8   5HT6R_MOUSE
239  VPRTPRPGVESADSRRLATKHSRKALKASLTLGILLGMFFVTWLPFFVANIVQAVCDCIS   298  Q5IS65   5HT6R_PANTR
239  VPRTPRPGMESADSRRLATKHSRKALKASLTLGILLGMFFVTWLPFFVANIAQAVCDCIS   298  P31388   5HT6R_RAT
     ********:********:*******************.************.*.********

299  PGLFDVLTWLGYCNSTMNPIIYPLFMRDFKRALGRFLPCPRCPRERQASLASPSLRTSHS   358  P50406   5HT6R_HUMAN
301  PGLFDVLTWLGYCNSTMNPIIYPLFMRDFKRALGRFVPCVHCPPEHRASPASPSMWTSHS   360  Q9R1C8   5HT6R_MOUSE
299  PGLFDVLTWLGYCNSTMNPIIYPLFMRDFKRALGRFLPCPRCPRERQASLASPSLRTSHS   358  Q5IS65   5HT6R_PANTR
299  PGLFDVLTWLGYCNSTMNPIIYPLFMRDFKRALGRFLPCVHCPPEHRPALPPPPCGPLTA   358  P31388   5HT6R_RAT
     ************************************:** :** *:: :   *      :

359  GPRPGLSLQQVLPLPLPPDSDSDSDAGSGGSSGLRLTAQLLLPGEATQDPPLPTRAAAAV   418  P50406   5HT6R_HUMAN
361  GARPGLSLQQVLPLPLPPNSDSD--SASGGTSGLQLTAQLLLPGEATRDPPPPTRAPTVV   418  Q9R1C8   5HT6R_MOUSE
359  GPRPGLSLQQVLPLPLPPDSDSDSDAGSGGSSGLRLTAQLLLPGEATRDPPLPTRAAAAV   418  Q5IS65   5HT6R_PANTR
359  VP----------------DQASAC-----SRCCLCLCRQTQIQTPLQ-GAP---RACSSQ   393  P31388   5HT6R_RAT
                       :. *       . . * *  *  :        *   ** :

419  NFFNIDPAEPELRPHPLGIPTN----------------------   440  P50406   5HT6R_HUMAN
419  NFFVTDSVEPEIRQHPLGSPMN----------------------   440  Q9R1C8   5HT6R_MOUSE
419  NFFNIDPAEPELRPHPLGIPTN----------------------   440  Q5IS65   5HT6R_PANTR
394  PSFCCLERPPGTPRHPPGPPLWSTSLSQTLWSLRYGRIHSVPP   436  P31388   5HT6R_RAT
       *      *    ** * *
```

It is shown that this consensus pY is located between 3rd and 4th TM domains, the sites where they interact with Gs proteins.

Serotonin Receptor Feature: sequence repeats



GS repeats on N-terminus

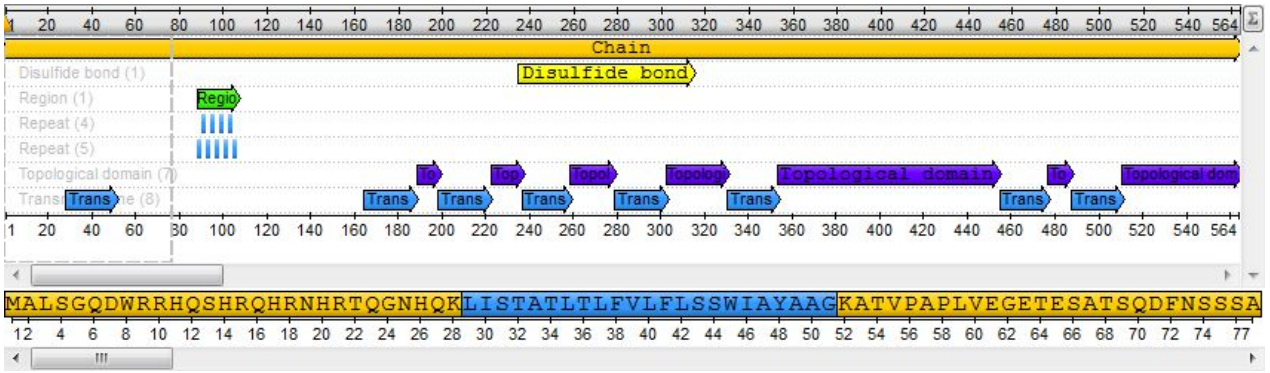


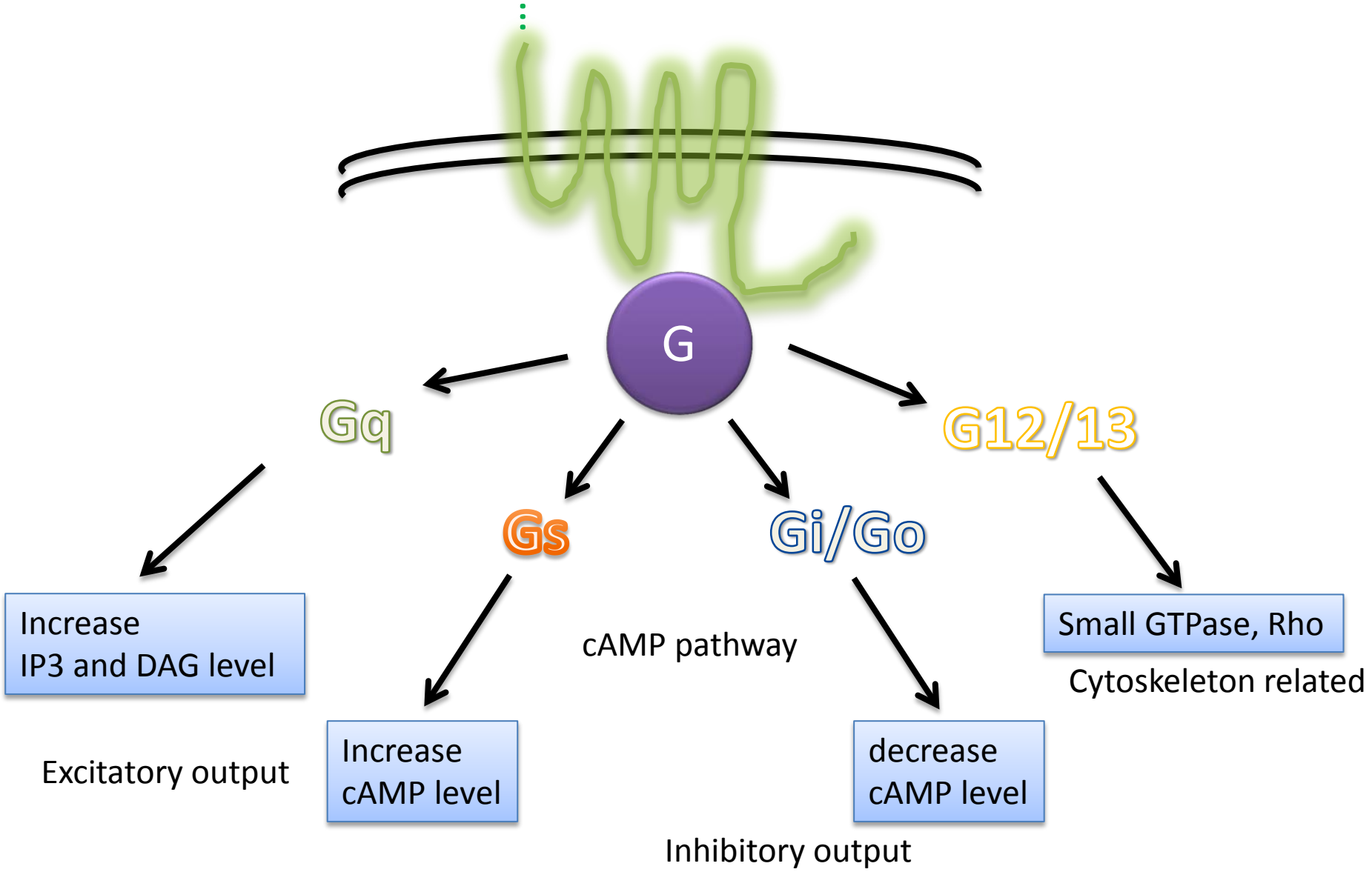
Same on the 2nd sequence of 5HT1A

...

Polydot command from Weblab of cbi: align on the whole sequence of self

GS repeats:
same loci
same numbers




G
Gq
Gs
Gi/Go
G12/13
Increase
IP3 and DAG level
Excitatory output
Increase
cAMP level
cAMP pathway
decrease
cAMP level
Inhibitory output
Small GTPase, Rho
Cytoskeleton related

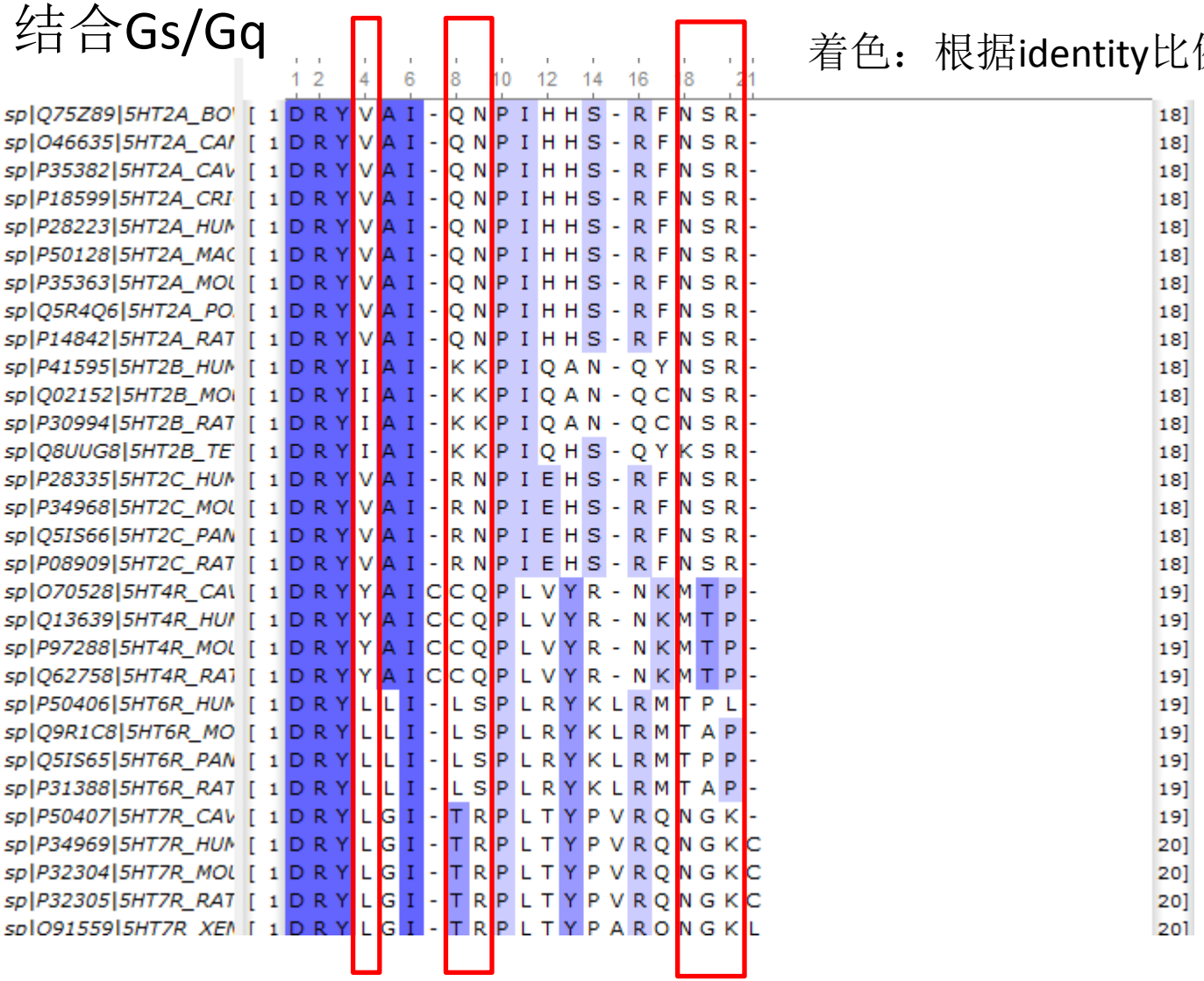
The second question of interest:
Why Galpha are different among Serotonin receptors

- Functional Selectivity theory: It is the ligand diversity that determines the distinctiveness in output types.
- However, it is difficult to apply this on case of serotonin receptors, since their ligands are the same while their outputs are not.
- In this case, some SRs are excitatory in neuronal signaling regulation while others are inhibitory.
- Therefore, our second question of interest is whether this difference could be related to the sequence bifurcation in the interface that interacts with the G proteins.
- Notice: it is only in theory prediction, and further evidence from experiments are required to demonstrate them.

Serotonin Receptor Feature: G protein interacting sequence and G subtype

结合Gs/Gq

着色：根据identity比例



IL-3

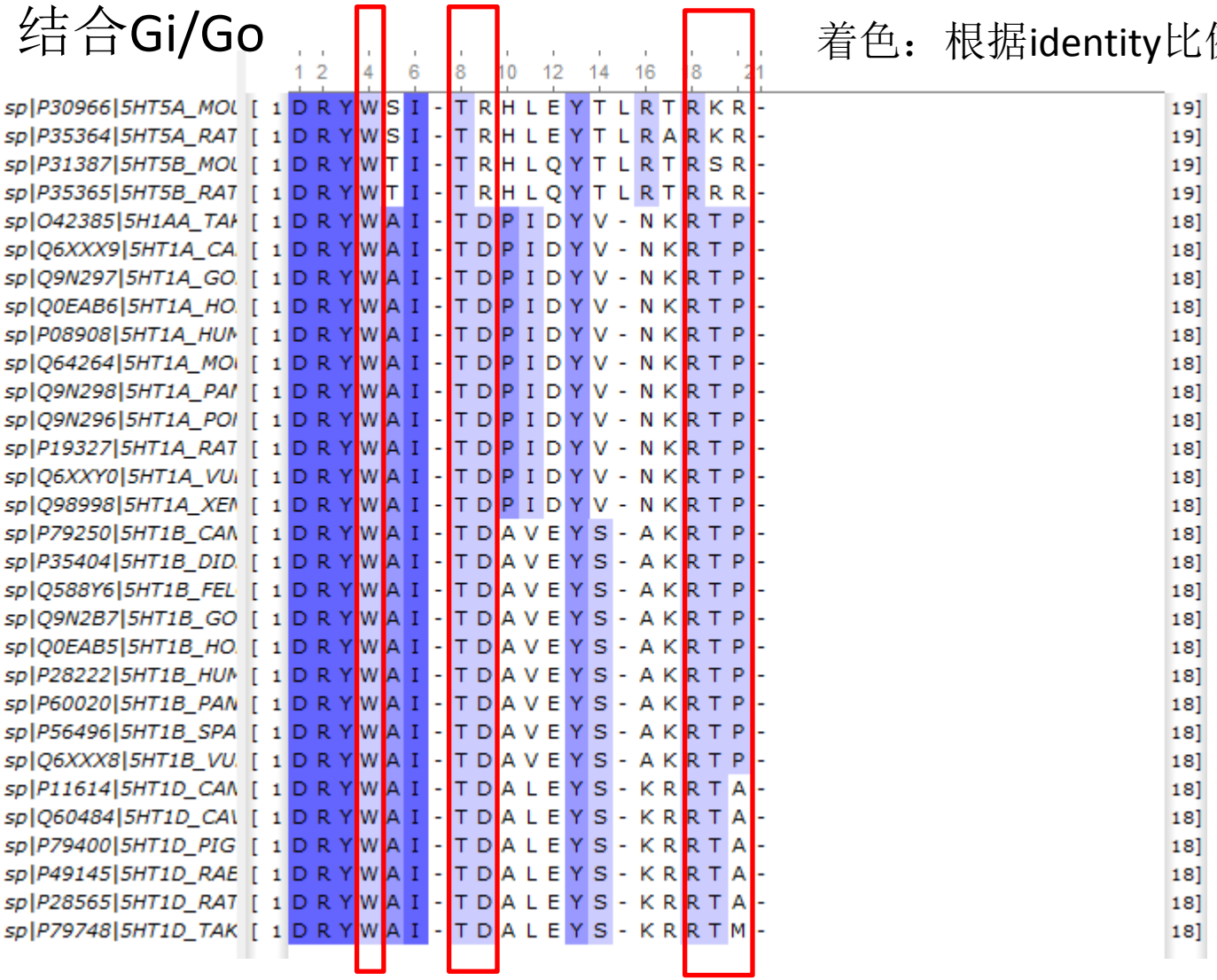
CHARGE 0 + or 0 0 0 +

POLAR N NPN

Serotonin Receptor Feature: G protein interacting sequence and G subtype

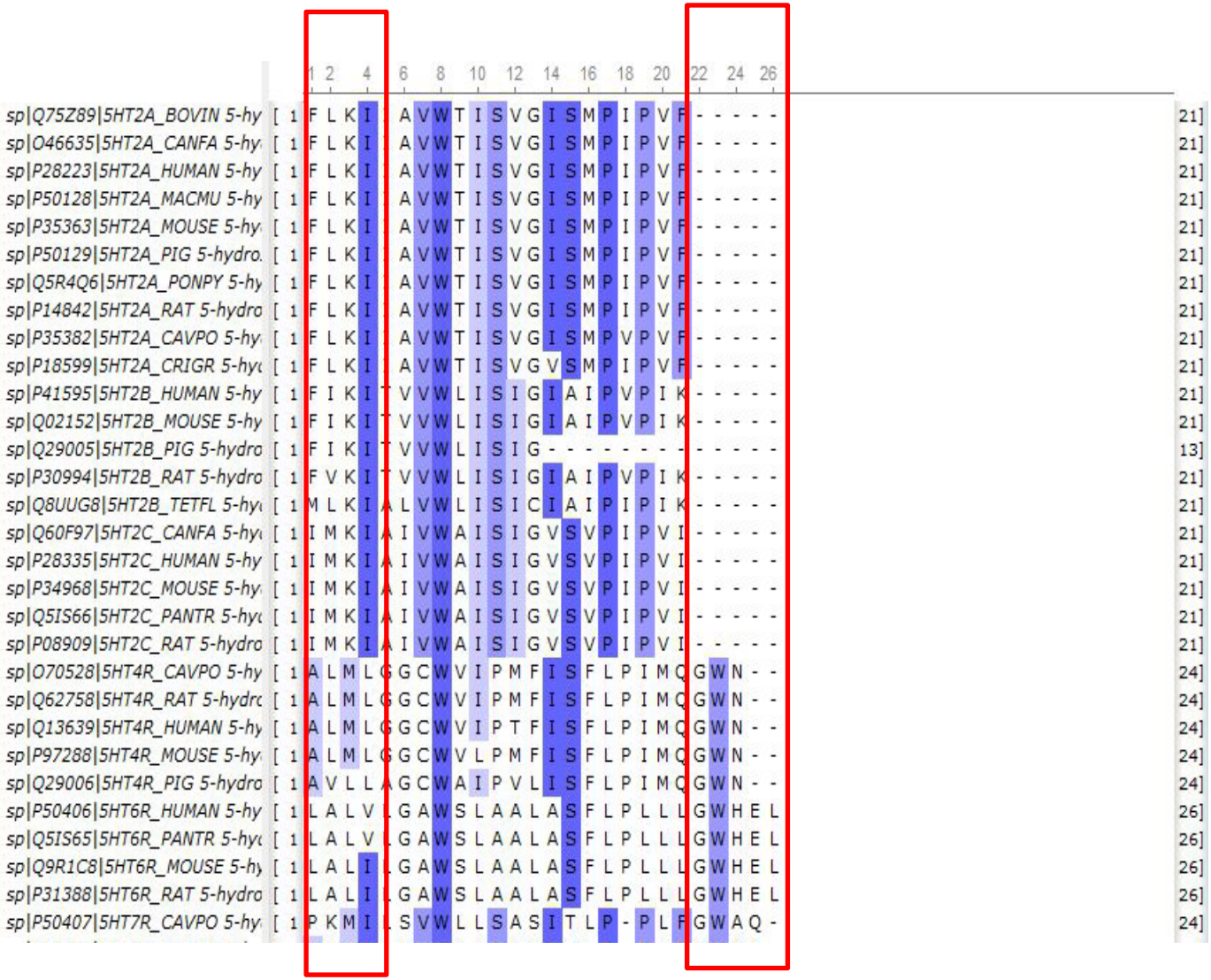
结合Gi/Go

着色：根据identity比例

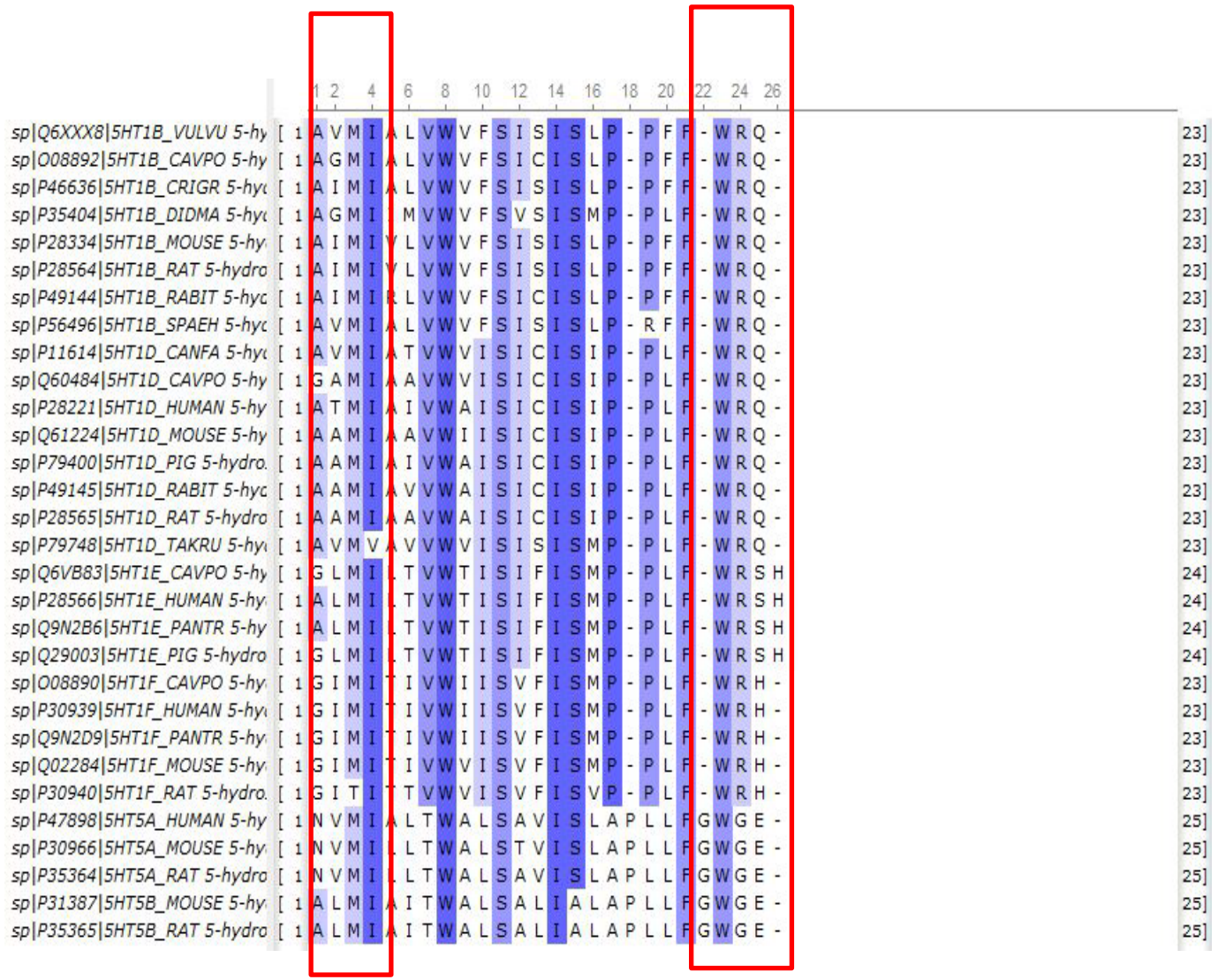


IL-3 CHARGE - 0- + 0 0

POLAR N PN NPN


1 2 4 6 8 10 12 14 16 18 20 22 24 26
sp|Q75Z89|5HT2A_BOVIN 5-hy [1 F L K I A V W T I S V G I S M P I P V F - - - - - 21]
sp|O46635|5HT2A_CANFA 5-hy [1 F L K I A V W T I S V G I S M P I P V F - - - - - 21]
sp|P28223|5HT2A_HUMAN 5-hy [1 F L K I A V W T I S V G I S M P I P V F - - - - - 21]
sp|P50128|5HT2A_MACMU 5-hy [1 F L K I A V W T I S V G I S M P I P V F - - - - - 21]
sp|P35363|5HT2A_MOUSE 5-hy [1 F L K I A V W T I S V G I S M P I P V F - - - - - 21]
sp|P50129|5HT2A_PIG 5-hydro [1 F L K I A V W T I S V G I S M P I P V F - - - - - 21]
sp|Q5R4Q6|5HT2A_PONPY 5-hy [1 F L K I A V W T I S V G I S M P I P V F - - - - - 21]
sp|P14842|5HT2A_RAT 5-hydro [1 F L K I A V W T I S V G I S M P I P V F - - - - - 21]
sp|P35382|5HT2A_CAVPO 5-hy [1 F L K I A V W T I S V G I S M P V P V F - - - - - 21]
sp|P18599|5HT2A_CRIGR 5-hy [1 F L K I A V W T I S V G V S M P I P V F - - - - - 21]
sp|P41595|5HT2B_HUMAN 5-hy [1 F I K I T V V W L I S I G I A I P V P I K - - - - - 21]
sp|Q02152|5HT2B_MOUSE 5-hy [1 F I K I T V V W L I S I G I A I P V P I K - - - - - 21]
sp|Q29005|5HT2B_PIG 5-hydro [1 F I K I T V V W L I S I G - - - - - - - - - - - 13]
sp|P30994|5HT2B_RAT 5-hydro [1 F V K I T V V W L I S I G I A I P V P I K - - - - - 21]
sp|Q8UUG8|5HT2B_TETFL 5-hy [1 M L K I A L V W L I S I C I A I P I P I K - - - - - 21]
sp|Q60F97|5HT2C_CANFA 5-hy [1 I M K I A I V W A I S I G V S V P I P V I - - - - - 21]
sp|P28335|5HT2C_HUMAN 5-hy [1 I M K I A I V W A I S I G V S V P I P V I - - - - - 21]
sp|P34968|5HT2C_MOUSE 5-hy [1 I M K I A I V W A I S I G V S V P I P V I - - - - - 21]
sp|Q5IS66|5HT2C_PANTR 5-hy [1 I M K I A I V W A I S I G V S V P I P V I - - - - - 21]
sp|P08909|5HT2C_RAT 5-hydro [1 I M K I A I V W A I S I G V S V P I P V I - - - - - 21]
sp|O70528|5HT4R_CAVPO 5-hy [1 A L M L G G C W V I P M F I S F L P I M Q G W N - - 24]
sp|Q62758|5HT4R_RAT 5-hydro [1 A L M L G G C W V I P M F I S F L P I M Q G W N - - 24]
sp|Q13639|5HT4R_HUMAN 5-hy [1 A L M L G G C W V I P T F I S F L P I M Q G W N - - 24]
sp|P97288|5HT4R_MOUSE 5-hy [1 A L M L G G C W V L P M F I S F L P I M Q G W N - - 24]
sp|Q29006|5HT4R_PIG 5-hydro [1 A V L L A G C W A I P V L I S F L P I M Q G W N - - 24]
sp|P50406|5HT6R_HUMAN 5-hy [1 L A L V L G A W S L A A L A S F L P L L L G W H E L 26]
sp|Q5IS65|5HT6R_PANTR 5-hy [1 L A L V L G A W S L A A L A S F L P L L L G W H E L 26]
sp|Q9R1C8|5HT6R_MOUSE 5-hy [1 L A L I L G A W S L A A L A S F L P L L L G W H E L 26]
sp|P31388|5HT6R_RAT 5-hydro [1 L A L I L G A W S L A A L A S F L P L L L G W H E L 26]
sp|P50407|5HT7R_CAVPO 5-hy [1 P K M I L S V W L L S A S I T L P - P L F G W A Q - 24]

IL-4


1 2 4 6 8 10 12 14 16 18 20 22 24 26
sp|Q6XXX8|5HT1B_VULVU 5-hy [1 A V M I A L V W V F S I S I S L P - P F F - W R Q - 23]
sp|O08892|5HT1B_CAVPO 5-hy [1 A G M I A L V W V F S I C I S L P - P F F - W R Q - 23]
sp|P46636|5HT1B_CRIGR 5-hyd [1 A I M I A L V W V F S I S I S L P - P F F - W R Q - 23]
sp|P35404|5HT1B_DIDMA 5-hyd [1 A G M I M V W V F S V S I S M P - P L F - W R Q - 23]
sp|P28334|5HT1B_MOUSE 5-hy [1 A I M I Y L V W V F S I S I S L P - P F F - W R Q - 23]
sp|P28564|5HT1B_RAT 5-hydro [1 A I M I Y L V W V F S I S I S L P - P F F - W R Q - 23]
sp|P49144|5HT1B_RABIT 5-hyd [1 A I M I R L V W V F S I C I S L P - P F F - W R Q - 23]
sp|P56496|5HT1B_SPAEH 5-hyd [1 A V M I A L V W V F S I S I S L P - R F F - W R Q - 23]
sp|P11614|5HT1D_CANFA 5-hyd [1 A V M I A T V W V I S I C I S I P - P L F - W R Q - 23]
sp|Q60484|5HT1D_CAVPO 5-hy [1 G A M I A A V W V I S I C I S I P - P L F - W R Q - 23]
sp|P28221|5HT1D_HUMAN 5-hy [1 A T M I A I V W A I S I C I S I P - P L F - W R Q - 23]
sp|Q61224|5HT1D_MOUSE 5-hy [1 A A M I A A V W I I S I C I S I P - P L F - W R Q - 23]
sp|P79400|5HT1D_PIG 5-hydro [1 A A M I A I V W A I S I C I S I P - P L F - W R Q - 23]
sp|P49145|5HT1D_RABIT 5-hyd [1 A A M I A V V W A I S I C I S I P - P L F - W R Q - 23]
sp|P28565|5HT1D_RAT 5-hydro [1 A A M I A A V W A I S I C I S I P - P L F - W R Q - 23]
sp|P79748|5HT1D_TAKRU 5-hyd [1 A V M V A V V W V I S I S I S M P - P L F - W R Q - 23]
sp|Q6VB83|5HT1E_CAVPO 5-hy [1 G L M I L T V W T I S I F I S M P - P L F - W R S H 24]
sp|P28566|5HT1E_HUMAN 5-hy [1 A L M I L T V W T I S I F I S M P - P L F - W R S H 24]
sp|Q9N2B6|5HT1E_PANTR 5-hy [1 A L M I L T V W T I S I F I S M P - P L F - W R S H 24]
sp|Q29003|5HT1E_PIG 5-hydro [1 G L M I L T V W T I S I F I S M P - P L F - W R S H 24]
sp|O08890|5HT1F_CAVPO 5-hy [1 G I M I T I V W I I S V F I S M P - P L F - W R H - 23]
sp|P30939|5HT1F_HUMAN 5-hyd [1 G I M I T I V W I I S V F I S M P - P L F - W R H - 23]
sp|Q9N2D9|5HT1F_PANTR 5-hy [1 G I M I T I V W I I S V F I S M P - P L F - W R H - 23]
sp|Q02284|5HT1F_MOUSE 5-hy [1 G I M I T I V W V I S V F I S M P - P L F - W R H - 23]
sp|P30940|5HT1F_RAT 5-hydro [1 G I T I T T V W V I S V F I S V P - P L F - W R H - 23]
sp|P47898|5HT5A_HUMAN 5-hy [1 N V M I A L T W A L S A V I S L A P L L F G W G E - 25]
sp|P30966|5HT5A_MOUSE 5-hy [1 N V M I L L T W A L S T V I S L A P L L F G W G E - 25]
sp|P35364|5HT5A_RAT 5-hydro [1 N V M I L L T W A L S A V I S L A P L L F G W G E - 25]
sp|P31387|5HT5B_MOUSE 5-hy [1 A L M I A I T W A L S A L I A L A P L L F G W G E - 25]
sp|P35365|5HT5B_RAT 5-hydro [1 A L M I A I T W A L S A L I A L A P L L F G W G E - 25]

IL-4

Thank You for your attention !

工作分配：
杨威，五羟色胺受体和GPCR背景知识，五羟色胺受体家族同源序列的进化树分析，所有家族成员的局部比对；
王魏然，五羟色胺受体同源序列翻译后修饰（磷酸化，糖基化，GS重复序列）位点的多态性分析；
刘少锋，五羟色胺受体桐原序列的跨膜区重要序列多态性分析，G蛋白识别差异分析；

在这里再次感谢以上组员的合作与支持，预祝本次ppt展示圆满成功！