

TP : calcul de l'accord inter-annotateurs

Karën Fort

karen.fort@paris-sorbonne.fr

14 octobre 2016



Exercice

Utilisez les données de la campagne d'annotation de relations de renommage de noms de gènes et :

1. calculez les coefficients :

- S
- κ
- π

2. que remarquez-vous ?

Rappel des formules

		Annotateur A		
		Oui	Non	Total
Annotateur B	Oui	24	8	32
	Non	14	24	38
	Total	38	32	70

- Accord observé : $A_o = \frac{24+24}{70} = 0,68$

- Accord attendu :

1. $A_e^S = \frac{1}{2} = 0,5$

2. $A_e^\pi = \frac{\left(\frac{(38+32)}{2}\right)^2 + \left(\frac{(32+38)}{2}\right)^2}{70^2} = 0,5$

3. $A_e^\kappa = \frac{\frac{(32 \times 38)}{70} + \frac{(38 \times 32)}{70}}{70} = 0,49$

- Coef. :

1. $S = \frac{0,68 - 0,5}{1 - 0,5} = 0,36$

2. $\pi = \frac{0,68 - 0,5}{1 - 0,5} = 0,36$

3. $\kappa = \frac{0,68 - 0,49}{1 - 0,49} = 0,37$

Renommage de noms de gènes

*"The *yppB* :cat and *ypbC* :cat null alleles rendered cells sensitive to DNA-damaging agents, impaired plasmid transformation (25- and 100-fold), and moderately affected chromosomal transformation when present in an otherwise Rec+ *B. subtilis* strain. The *yppB* gene complemented the defect of the *recG40* strain. <former-name1 id="1">**yppB**</former-name1> and <former-name2 id="2">**ypbC**</former-name2> and their respective null alleles were termed " <new-name1 id="1">**recU**</new-name1>" and "recU1" (*recU* :cat) and " <new-name2 id="2">**recS**</new-name2>" and "recS1" (*recS* :cat), respectively. The *recU* and *recS* mutations were introduced into *rec*-deficient strains representative of the alpha (*recF*), beta (*addA5 addB72*), gamma (*recH342*), and epsilon (*recG40*) epistatic groups."*

Renommage de noms de gènes

		A1			
		Former	New	Nothing	Total
A2	Former	71	13	23	107
	New	8	69	15	92
	Nothing	7	8	18 840	18 855
	Total	86	90	18 878	19 054

Changer de point de vue

		A1			
		Former	New	Nothing	Gene names
A2	Former	71	13	23	107
	New	8	69	15	92
	Nothing	7	8	951	966
	Gene names	86	90	989	1 165