

# Evolution Moléculaire (Masters 1 Bioinformatique / Biotechnologie)

## Evolution biologique : de la molécule à l'espèce

**Maxime Bonhomme**

Laboratoire de Recherche en Sciences Végétales

29 décembre 2020

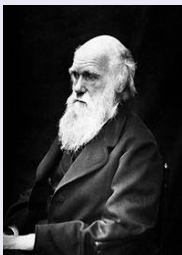
# Evolution biologique : de la molécule à l'espèce

- 1 contexte
- 2 origines du polymorphisme
  - mutation
  - recombinaison et duplication
- 3 les moteurs de l'évolution
  - Que devient une mutation ? la dérive génétique
  - Que devient une mutation ? la sélection naturelle
  - Notion d'adaptation
- 4 la spéciation
  - définition
  - la spéciation allopatrique
  - la spéciation sympatrique
- 5 références

# Evolution biologique

- la théorie de l'évolution est une explication scientifique de la diversification des formes de vie qui apparaissent dans la nature
  - l'**évolution biologique** désigne la transformation des espèces vivantes qui se manifeste par des **changements de leurs caractères génétiques au cours des générations**
  - les changements successifs peuvent aboutir à la formation de nouvelles espèces : **spéciation**
  - cette diversification depuis les premières formes de vie est à l'origine de la biodiversité actuelle
  - l'histoire des espèces peut se représenter sous la forme d'un **arbre phylogénétique**
- Charles Darwin : "On the origin of species" (1859)

Charles Darwin (1809-1882)

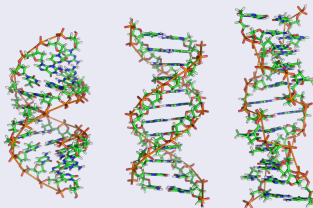


Alfred Russel Wallace (1823-1913)



# Evolution moléculaire : pierre angulaire de l'évolution biologique

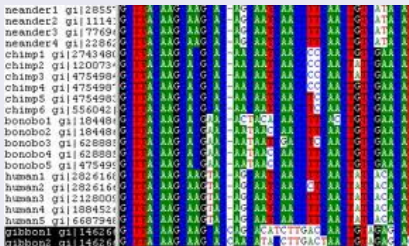
- ADN = support de l'information génétique



- hérédité = transfert de l'information à la descendance : un des prérequis pour la sélection naturelle (et l'évolution en général)
- **évolution moléculaire** = approche multidisciplinaire récente (1960...) qui vise à comprendre l'évolution de l'ADN et des génomes, de l'ARN et des protéines (structure et fonction)
- elle rassemble plusieurs disciplines : biologie moléculaire, biologie évolutive et génétique des populations
- exemples de thématiques :
  - **évolution des fonctions** enzymatiques, des gènes de l'immunité
  - **reconstruction phylogénétique**
  - concept d'"horloge moléculaire" pour étudier la divergence des espèces
  - origine et fonction de l'ADN non codant

# Evolution moléculaire : pierre angulaire de l'évolution biologique

- quelle information peut-on extraire des séquences d'ADN ?



- avancées récentes en évolution moléculaire dues aux progrès technologiques :
  - rôle des duplications de gènes dans l'émergence de nouvelles fonctions
  - **poids de l'évolution moléculaire adaptative vs neutre**
  - changements moléculaires responsables de certains phénotypes (maladies, épidémiologie, cognition)
- quelques noms :
  - Motoo Kimura : théorie neutre
  - Masatoshi Nei : évolution adaptative
  - Walter M. Fitch : reconstruction phylogénétique
  - Walter Gilbert : RNA world
  - Joe Felsenstein : méthodes phylogénétiques
  - Susumu Ohno : duplication des gènes
  - John H. Gillespie : mathématiques de l'adaptation
  - Dan Graur, Wen-Hsiung Li : modèles neutres de l'évolution moléculaire

# Mutation de l'ADN

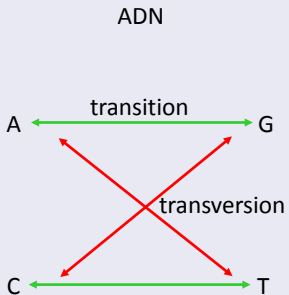
## Types de mutations

- changement héréditaire dans le matériel génétique.
- source fondamentale de variation génétique.
  - **mutations dans la séquence ADN** : changement dans la séquence nucléotidique (**substitutions nucléotidiques**, insertions / délétions, courtes répétitions en tandem - ou microsatellites -, (retro)transposons).
  - mutations chromosomiques : réarrangements chromosomiques (inversions, translocations).
  - mutations génomiques : polyploïdisation
- distinguer mutations somatiques (cancers, ...) de germinales (cellules sexuelles, donc transmissibles).
- distinguer mutations neutres (sans impact sur le phénotype par rapport aux autres allèles) de favorable/délétère (dépend des conditions du milieu)

## Fréquence des substitutions dans l'ADN

- événement rare.
  - $10^{-8}$  à  $10^{-9}$  nouvelles mutations par nucléotide par génération.
  - $10^{-4}$  à  $10^{-6}$  nouvelles mutations par copie de gène par génération.
- innovation génétique.
  - nombre de mutations par génération =  $2Nu$ , non négligeable si population grande.
  - nombre important de gènes dans les génomes (ex : 20 000 - 30 000 chez l'homme), donc plusieurs gènes mutés chez un zygote.

# Types de substitutions : ADN et acides aminés



VSCCQSTQYCQP

VSCCPTPQYCQP

- $\mathbb{P}(\text{transition}) > \mathbb{P}(\text{transversion})$

## Types de substitutions : ADN et acides aminés

- sur un cadre de lecture, les substitutions peuvent conduire ou non à un changement d'a.a (substitutions non synonymes, synonymes - troisième position du codon)

**le code génétique**

		Deuxième lettre											
		U			C			A			G		
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U	Troisième lettre (côté 3')			
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C				
	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop	A				
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp	G				
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U				
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C				
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A				
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G				
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U				
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C				
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A				
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg	G				
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U				
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C				
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A				
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G				
		codon d'initiation						codon de terminaison					

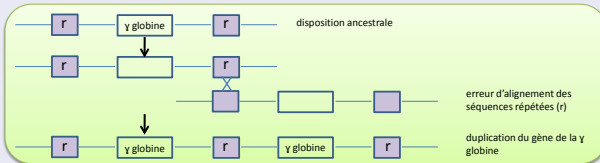
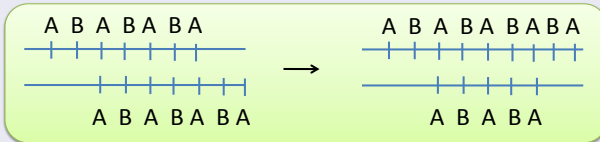
- existence de biais d'usage du codon = utilisation préférentielle d'un des triplets :
  - hypothèse neutraliste : **biais mutationnel**, lié aux probabilités de transitions et transversion
  - hypothèse sélectionniste : **sélection traductionnelle**, sélection de certains codons (dits optimaux) parce qu'ils permettraient une plus grande efficacité de réplication et de traduction. Chez la drosophile, le nématode et l'arabette, on observe que les codons optimaux sont très souvent ceux dont les ARNt complémentaires sont les plus fréquents dans la cellule





# Duplications de gènes

- recombinaison inégale : erreur d'alignement des deux chromatides lors du crossing-over

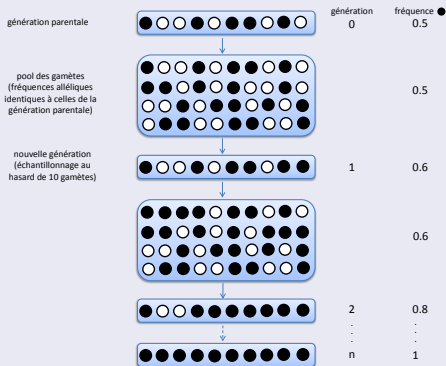


## Duplications de génomes : polypléidisation

- effets biophysiques : augmentation de la taille de la cellule, moins de cellules dans un tissu
- effets dosages : augmentation du nombre de copies d'un locus (effets additifs possibles si codominance)
- masquage des allèles délétères : expression relativement plus faible des allèles délétères récessifs chez un polypléide que chez un diploïde
- augmentation de la diversité allélique, de l'hétérozygotie
- perte ou gain d'ADN, recombinaison homéologue (génomés parentaux n'ont pas la même succession de gènes), réarrangements d'ADN

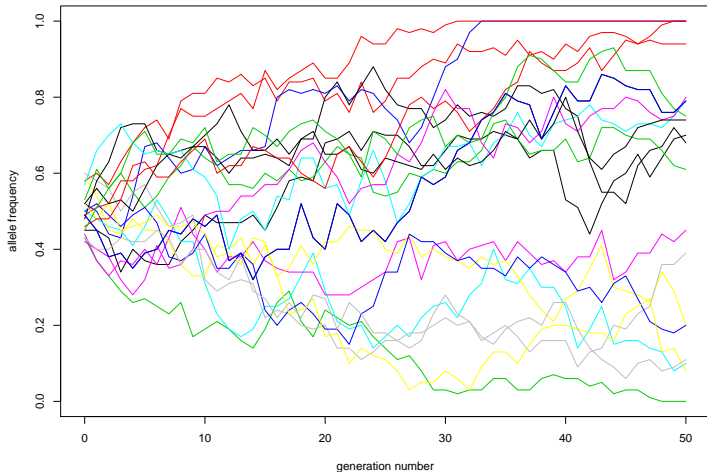
# Dérive génétique : modèle stochastique d'évolution des fréquences des mutations

Fluctuation des fréquences alléliques de générations en générations du fait d'un échantillonnage aléatoire des gamètes dans une population de taille finie (non transmission de certains allèles à la descendance, individus ne se reproduisant pas).



Echantillonnage des gamètes et changement des fréquences alléliques.

# Dérive génétique : simulations

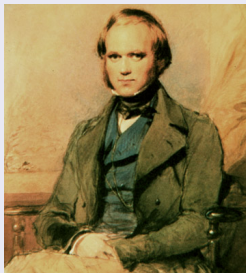


Fréquence initiale = 0.5, taille de la population = 100, 20 simulations (rouge = moyenne, bleu = variance).

# Sélection naturelle

## Principes

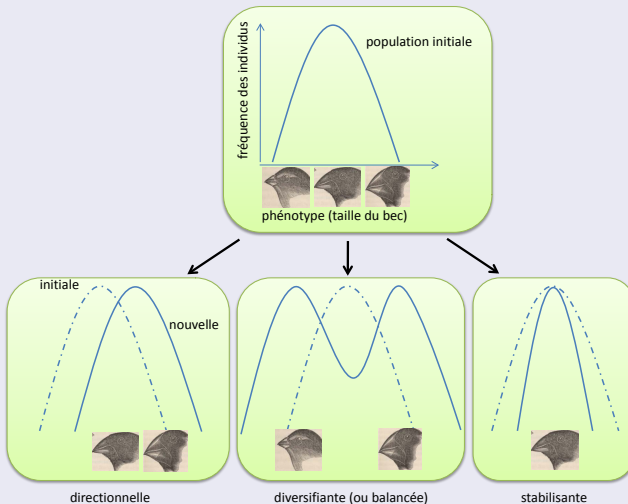
- Charles Darwin (1809 – 1882) : "l'origine des espèces" (1859)



- **Principe 1 = variabilité** : les individus d'une espèce diffèrent les uns des autres pour certains caractères (phénotypes et donc génotypes).
- **Principe 2 = adaptation / survie** : les individus les plus adaptés au milieu survivent et se reproduisent davantage.
- **Principe 3 = transmission des caractères avantageux** : les caractéristiques avantageuses doivent être héréditaires.

# Sélection naturelle dans une population

ce que l'on observe au niveau du phénotype (caractère)







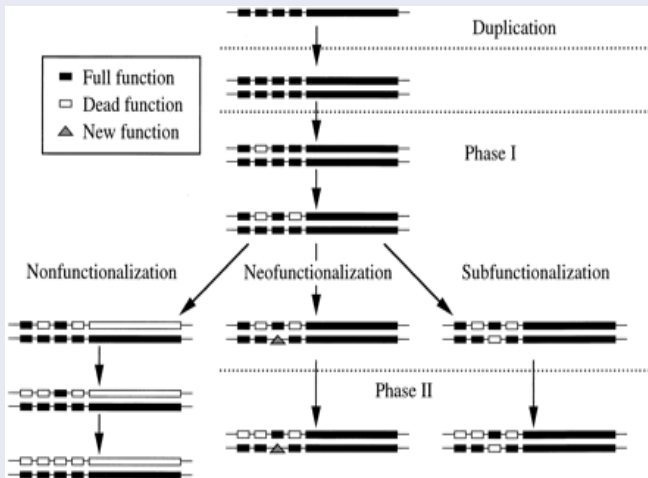
## Devenir d'une mutation EN PRATIQUE

- **mutation neutre** : le plus souvent éliminée de la population, mais peut aussi se substituer à l'allèle sauvage, à cause des effets aléatoires de la **dérive génétique** dans les petites populations.
- **mutation défavorable** : diminue en fréquence (**sélection négative**).
- **mutation favorable** : augmente en fréquence (**sélection positive**).

cf cours "évolution moléculaire : neutre ou adaptative?" de cette UE (théorie neutraliste et détection de la signature moléculaire de la sélection)

# Sélection naturelle et duplication

## rôle de la sélection dans le devenir d'une duplication



# Sélection naturelle et duplication

## rôle de la sélection dans le devenir d'une duplication

- **non fonctionalisation** : pseudogène (dérive neutre)
- **néofonctionalisation** : la duplication garde souvent la fonction ancestrale et le gène ancestral acquiert une nouvelle fonction (sélection positive, divergence fonctionnelle)
- **subfonctionalisation** : les duplications ont une activité conjointe similaire à la fonction ancestrale (sélection purifiante)

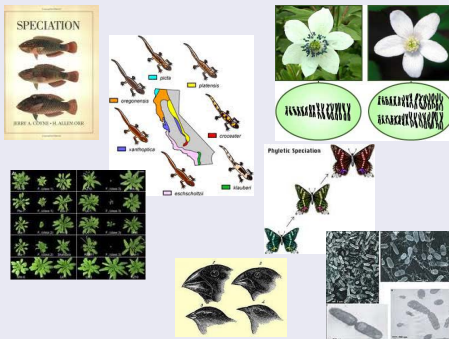
# Adaptation

## quelques points

- la sélection naturelle n'est pas le seul processus qui entraîne l'évolution, mais il est le seul qui mène à l'**adaptation**
- la sélection "bricole" sur des fonctions préexistantes (transfert de fonction, notion d'exaptation) :
  - les ailes des insectes : fonction ancestrale probable de régulation thermique, puis transfert vers le vol en milieu terrestre
- l'adaptation n'est pas une notion figée : les organismes sont perpétuellement en retard (course à l'adaptation)
  - théorie de la reine rouge (course aux armements entre les espèces)
  - Lewis Carroll (Alice au pays des merveilles) : le personnage principal et la Reine Rouge se lancent dans une course effrénée. Alice demande alors : « Mais, Reine Rouge, c'est étrange, nous courons vite et le paysage autour de nous ne change pas ? » Et la reine répondit : « Nous courons pour rester à la même place. »
- systèmes de **coévolution** : plantes-pollinisateur, symbioses plantes-microorganismes
- les imperfections des êtres vivants sont dues à des contraintes génétiques, embryologiques ou historiques, et à des compromis entre exigences adaptatives différentes (notion de "trade-off", ex : compromis reproduction/croissance)
- notion d'impasse évolutive : les sabots des chevaux = retour en arrière impossible sauf une macro-mutation

# Spéciation : définition

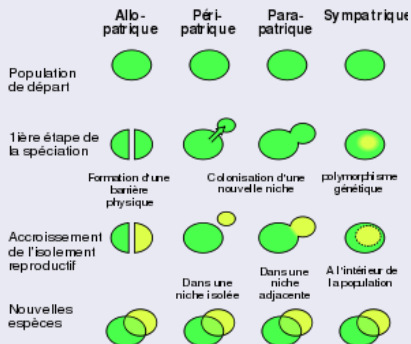
- processus évolutif par lequel de nouvelles espèces vivantes apparaissent



- la spéciation résulte de la sélection naturelle et/ou de la dérive génétique
  - les espèces s'individualisent à partir de populations appartenant à une espèce d'origine
  - évolution de populations interfécondes - composant une même espèce, par définition - en populations non-interfécondes, c'est-à-dire isolées sur le plan reproducteur
  - des barrières biologiques reproductives s'installent : **incompatibilités phénotypiques pré/post-zygotiques** dues à des modifications des génomes (mutations géniques, chromosomiques...) progressivement fixées par dérive génétique ou sélection

# Spéciation : définition

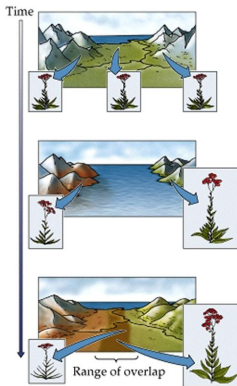
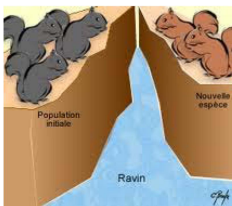
- on distingue deux grands types de spéciation, avec des intermédiaires :
  - spéciation **allopatrique** ("géographique")
  - spéciation **sympatrique** ("écologique")



- résulte de l'interaction entre les modifications au niveau moléculaire (génomés, ADN, ARN et protéines) et l'écologie des organismes (contraintes environnementales, isolement, sélection sexuelle, compétition pour les ressources)

# Spéciation allopatrique

- la **dérive génétique** joue un rôle crucial dans la divergence des populations



© 2001 Sinauer Associates, Inc.

# Spéciation sympatrique

- des populations non isolées géographiquement peuvent évoluer en espèces distinctes
- la sélection naturelle joue un rôle crucial dans la divergence des populations
- fréquent chez insectes phytophages, très fréquent chez microorganismes symbiotiques et pathogènes
  - chez pathogènes de plantes (oomycètes, champignons), la spécialisation à l'hôte (spéciation) résulte d'une perte/acquisition de gènes que la sélection naturelle favorise car permet une meilleure reproduction du microorganisme
  - la spéciation sympatrique récurrente résulte de la coévolution "hôte-pathogène"

## Signatures of Adaptation to Obligate Biotrophy in the *Hyaloperonospora arabidopsidis* Genome

Laura Hunter,<sup>1,2</sup> Susarla Drijphty,<sup>1,2</sup> Naveed Ishaque,<sup>1,2</sup> Nico Beets,<sup>1</sup> Adriana Cubero,<sup>1</sup> Eric Kenes,<sup>1</sup> Marco Thines,<sup>1,2</sup> Audrey Ah-Fong,<sup>1</sup> Ryan Anderson,<sup>1</sup> Wale Badajoch,<sup>1</sup> Peter Blitner-Eddy,<sup>1</sup> Jeffrey L. Boore,<sup>1</sup> Marcos C. Chibazo,<sup>1</sup> Mary Coates,<sup>1</sup> Pasawee Dehal,<sup>1</sup> Kim Delehaute,<sup>1,2</sup> Suwaneng Dong,<sup>1,2</sup> Polly Downes,<sup>1</sup> Bernard Dumas,<sup>1,2,3</sup> Georgina Falco,<sup>1</sup> Cathrin Fiedler,<sup>1,1</sup> Susan I. Foweraker,<sup>1</sup> Lucinda Fulton,<sup>1</sup> Elaine Gault,<sup>1,2,3</sup> Françoise Geisen,<sup>1,1</sup> Linda Hughes,<sup>1</sup> Sean Humphrey,<sup>1,1</sup> Ravi H. V. Jang,<sup>1,2,3</sup> Howard Juddison,<sup>1</sup> Sophie Kanaan,<sup>1</sup> Kim Kyung,<sup>1,1</sup> Harald Marler,<sup>1,1</sup> Patrick Miles,<sup>1,1</sup> Paul Morris,<sup>1,1</sup> Jaesoo Nelson,<sup>1</sup> Vipa Phantanasri,<sup>1,1</sup> Shash Oshin,<sup>1,1</sup> Anne Robinson,<sup>1</sup> Benjamin Roques-Castano,<sup>1</sup> Peter Ryker,<sup>1</sup> Tedy Terto-Maldonado,<sup>1</sup> David Studholme,<sup>1</sup> Yuanhao Wang,<sup>1,1</sup> Joe Wu,<sup>1</sup> Jo Wood,<sup>1</sup> Sandro W. Gilless,<sup>1,1</sup> Jane Rogers,<sup>1</sup> Guido Van den Ackerveken,<sup>1</sup> Jonathan D. G. Jones,<sup>1,1</sup> John W. McDowell,<sup>1,1</sup> Jon Beynon,<sup>1</sup> Brett M. Tyler<sup>1,1</sup>

## Genome Expansion and Gene Loss in Powdery Mildew Fungi Reveal Tradeoffs in Extreme Parasitism

Phong D. Sparks,<sup>1</sup> James C. Abbott,<sup>1</sup> Jovita Anandaraman,<sup>1,1</sup> Jhouly A. Borge,<sup>1</sup> Darwin B. Soares,<sup>1</sup> Kurt Stöber,<sup>1</sup> Eelke Van Leren van Themaat,<sup>1</sup> James A. M. Breen,<sup>1</sup> Sarah A. Bultner,<sup>1</sup> Sarah J. Carr,<sup>1</sup> Marc-Henri Lebrun,<sup>1,1</sup> Christopher J. Ribbel,<sup>1</sup> Paul Schulze-Lenther,<sup>1</sup> Nicholas J. Talbot,<sup>1</sup> Nihal Akhmadawati,<sup>1</sup> Christian Anwar,<sup>1</sup> Gerald B. Bortz,<sup>1</sup> Huzefa Beshir,<sup>1</sup> Promyphon Haldorsen,<sup>1</sup> Laurence W. Bredschneider,<sup>1</sup> Malte Beth,<sup>1</sup> Martin T. Bremer,<sup>1</sup> Lance Cadle-Davies,<sup>1,1</sup> Blake M. Cudde-Dawson,<sup>1</sup> Jennifer Colburn,<sup>1</sup> Rainer Coaker,<sup>1</sup> Oliver Frankel,<sup>1</sup> Dale Godfrey,<sup>1</sup> James Harrison,<sup>1</sup> Claire Heide,<sup>1</sup> Brian C. King,<sup>1</sup> Sam Klages,<sup>1</sup> Jochen Klose,<sup>1</sup> Daniela Kozak,<sup>1</sup> Praxiana S. Kott,<sup>1</sup> Jonathan Kopleck,<sup>1</sup> Francisco J. Lopez-Ruiz,<sup>1</sup> Xueli Lu,<sup>1</sup> Takaki Marukawa,<sup>1</sup> Shigeharu Matsuda,<sup>1</sup> Cristina Miceli,<sup>1</sup> Michael G. Milgroom,<sup>1</sup> Giovanni Montano,<sup>1</sup> Sandra Neki,<sup>1</sup> Richard J. O'Connell,<sup>1</sup> Sitawon Oshonwaoye,<sup>1</sup> Francis Paragás,<sup>1</sup> Caetano Pereira,<sup>1</sup> Hail Quesada,<sup>1</sup> Richard Reinhardt,<sup>1</sup> Matthias Rott,<sup>1</sup> Sándor Szendrői,<sup>1</sup> Sarah M. Schindler,<sup>1</sup> Haritz Schön,<sup>1</sup> Paul Skarzynski,<sup>1</sup> Hans Sonnen,<sup>1</sup> Amber Stephens,<sup>1</sup> Hisayuki Takahara,<sup>1</sup> Hans Thordal-Christiansen,<sup>1</sup> Harshita Vignawan,<sup>1</sup> Karl Willberg,<sup>1</sup> Thomas Wilder,<sup>1</sup> Ralph Westbrook<sup>1</sup>

## Genome Evolution Following Host Jumps in the Irish Potato Famine Pathogen Lineage

Sylvia Kulkuleli,<sup>1</sup> Riya A. Farris,<sup>1</sup> Liliana M. Cano,<sup>1</sup> David J. Studholme,<sup>1</sup> Derek Madden,<sup>1</sup> Marco Thines,<sup>1,2</sup> Ravi H. V. Jang,<sup>1</sup> Michael C. Zody,<sup>1</sup> Sridhar G. Karjane,<sup>1</sup> Nicole M. Dowrick,<sup>1</sup> Blake C. Heyen,<sup>1</sup> Chad Neidham,<sup>1</sup> Sophie Kanaan<sup>1</sup>

Linsear et al. Genome Biology 2016, 17:R212  
<http://dx.doi.org/10.1186/s13052-016-0717-0>



### RESEARCH

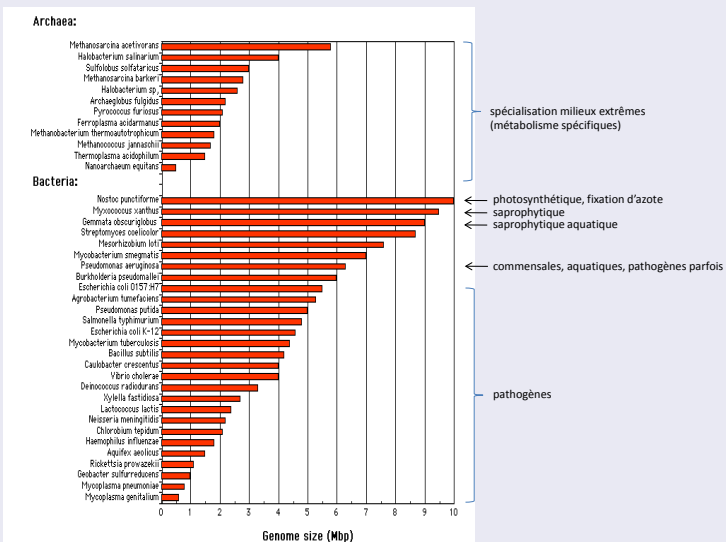
Open Access

Genome sequence of the necrotrophic plant pathogen *Pythium ultimum* reveals original pathogenicity mechanisms and effector repertoire

C. André Leung<sup>1,2</sup>, Heidi Brueker<sup>1</sup>, Liliana Cano<sup>1</sup>, John P. Hamblin<sup>3</sup>, Carole Hill<sup>4</sup>, Edgar Hübner<sup>5</sup>, Sylvain Kulkuleli<sup>6</sup>, Greg P. Robinson<sup>7</sup>, Marco Thines<sup>1,2</sup>, Joe Wu<sup>1</sup>, Marco W. Zeppe<sup>8</sup>, Gordon W. Slesar<sup>9</sup>, Jeffrey Boren<sup>10</sup>, Dani Boonij<sup>11</sup>, Jomari Dumala<sup>12</sup>, Sime Hensel<sup>13</sup>, Susan Houtenberg<sup>14</sup>, Clara Mill-Guckler<sup>15</sup>, Blake Gault<sup>16</sup>, Françoise Geisen<sup>17</sup>, Luca Garavalle-Boggi<sup>18</sup>, Neil Horne<sup>19</sup>, Jessica Hübner<sup>20</sup>, Ravi H. V. Jang<sup>1</sup>, Justin Johnson<sup>21</sup>, Theerapong Kijjanap<sup>22</sup>, Haining Lin<sup>1</sup>, Harist G. Mago<sup>23</sup>, Barry Moore<sup>24</sup>, Paul Morris<sup>25</sup>, Vignaporn Phrasman<sup>26</sup>, Daniela Posa<sup>27</sup>, Sarah Shery<sup>28</sup>, Isaac F. Staley<sup>29</sup>, Sachiro Taniyoshi<sup>30</sup>, Stephen Abbott<sup>31</sup>, Peter van Weel<sup>32</sup>, Joon H. White<sup>33</sup>, Pedro M. Guedes<sup>34</sup>, Bernard Henkel<sup>35</sup>, Frank Isenhardt<sup>36</sup>, Paul D. Thomas<sup>37</sup>, Brent W. Hale<sup>38</sup>, Ronald F. De Vries<sup>39</sup>, Saptan Kumar<sup>40</sup>, Mark Vandell<sup>41</sup>, Neil Young<sup>42</sup>, C. Robin Buell<sup>43</sup>

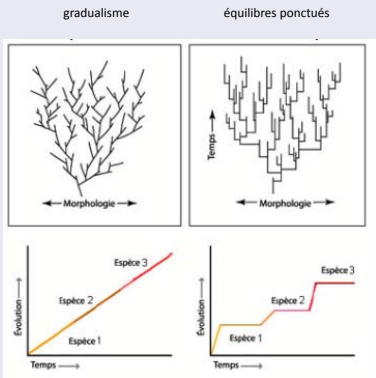


# Relation entre taille du génome et mode de vie chez les bactéries



## Vitesse de spéciation

- l'histoire de l'évolution montre que les spéciations nécessitent généralement plusieurs (centaines de) milliers d'années.
- MAIS il y a toujours beaucoup d'exceptions liées :
  - aux contingences historiques, environnementales (ex : 500 ans pour les souris introduites à Madère)
  - aux caractéristique biologiques des espèces (ex : les bactéries évoluent plus vite car elles ont un temps de génération court)
- controverse "gradualisme phylétique" vs "théorie des équilibres ponctués" (S.J Gould, N Eldredge : longues périodes d'équilibre, ponctuées de brèves périodes de spéciations / extinctions)



## Questions abordées dans l'UE Evolution Moléculaire

- quelles sont les étapes et les outils pour construire un arbre phylogénétique à partir de séquences d'ADN ? de protéines ?
- comment retracer l'histoire phylogénétique des espèces (ou plus modestement d'un groupe d'espèces) au travers d'un arbre phylogénétique ?
- comment retracer l'évolution des gènes au cours du temps ?
- comment retracer l'évolution des familles de gènes au cours du temps ?
- peut-on distinguer des gènes adaptés par rapport à des gènes non adaptés ou non fonctionnels, en analysant les séquences ?
- quel type de sélection agit sur quel type de gène ? sur quelle(s) partie(s) d'un gène ?
- .... autant de questions pour lesquelles nous tenterons de donner quelques réponses !

## References



*Evolution Biologique. Ridley, De Boeck Universié*



*Principles of Population Genetics, 4th Edition. Hartl DL, Clark AG*



*Computational Molecular Evolution, Ziheng Yang, Oxford Series in Ecology and Evolution*