



Principes des techniques
de **biologie moléculaire**
et **génomique**

3^e édition revue et augmentée

Principes des techniques de biologie moléculaire et génomique

3^e édition revue et augmentée

Denis Tagu, Stéphanie Jaubert-Possamai,
Agnès Méreau, coordinateurs

Éditions Quæ

Collection *Savoir-faire*

Guide des analyses en pédologie – 3^e édition revue et augmentée
Denis Baize
2018, 328 p.

De l'analyse des réseaux expérimentaux à la méta-analyse
Méthodes et applications avec le logiciel R pour les sciences agronomiques et environnementales
D. Makowski, F. Piraux, F. Brun
2018, 162 p.

Nutrition minérale des ruminants – Nouvelle édition
F. Meschy
2017, format ebook uniquement.

Les sols
Intégrer leur multifonctionnalité pour une gestion durable
A. Bispo, C. Guellier, E. Martin, J. Sapijanskas, H. Soubelet, C. Chenu, coord.
2016, 384 p.

Protection agroécologique des cultures
J.-P. Deguine, C. Gloanec, P. Laurent, A. Ratnadass, J.-N. Aubertot, coord.
2016, 288 p.

En couverture : hélice d'ADN © Zffoto/fotolia.com ; bactéries © fotolixrender/fotolia.com ; matériel d'électrophorèse © science photo/fotolia.com.

Éditions Quæ
RD 10, 78026 Versailles Cedex, France
www.quae.com

© Éditions Quæ, 2018

ISBN (papier) : 978-2-7592-2885-0
ISBN (PDF) : 978-2-7592-2886-7
ISBN (epub) : 978-2-7592-2887-4

ISSN : 1952-1251

Le Code de la propriété intellectuelle interdit la photocopie à usage collectif sans autorisation des ayants droit. Le non-respect de cette disposition met en danger l'édition, notamment scientifique, et est sanctionné pénalement. Toute reproduction, même partielle, du présent ouvrage est interdite sans autorisation du Centre français d'exploitation du droit de copie (CFC), 20 rue des Grands-Augustins, Paris 6^e.

Sommaire

Avant-propos	7
Remerciements	9
Liste des rédacteurs	11

Partie 1 Les bases

Fiche 1. Structure et expression d'un gène eucaryote codant un ARNm et une protéine.....	18
Fiche 2. Définition des ARN non codants.....	20
Fiche 3. Paramètres de description d'un génome	22
Fiche 4. Enzymes de restriction	24
Fiche 5. Électrophorèse des acides nucléiques	26
Fiche 6. PCR (<i>Polymerase Chain Reaction</i>)	28
Fiche 7. Description d'un plasmide et d'un phagemide	30
Fiche 8. Description d'un YAC et autres grands vecteurs.....	34
Fiche 9. Clonage moléculaire	36
Fiche 10. Transformation génétique des bactéries et des levures.....	39
Fiche 11. Construction de banques d'acides nucléiques.....	41
Fiche 12. Extraction d'ADN	45
Fiche 13. Extraction d'ARN.....	48
Fiche 14. Marquage des acides nucléiques (ADN et ARN).....	51
Fiche 15. Hybridation moléculaire.....	54
Fiche 16. Hybridation <i>in situ</i> d'ARN	58
Fiche 17. La cytogénétique moléculaire : FISH, GISH.....	61
Fiche 18. Électrophorèse en champ pulsé	65

Partie 2 Caractérisation d'un gène

Fiche 19. Séquençage d'ADN par la technique Sanger	70
Fiche 20. RT-PCR (<i>Reverse Transcription - Polymerase Chain Reaction</i>)....	75
Fiche 21. PCR quantitative	78
Fiche 22. Amplification rapide d'extrémités d'ADNc : RACE.....	86
Fiche 23. Transcription <i>in vitro</i>	89
Fiche 24. Détermination du site d'initiation de la transcription	91
Fiche 25. Gel-retard	93
Fiche 26. Production de protéines recombinantes	95
Fiche 27. Système du double hybride	105

Partie 3 Analyse de la fonction d'un gène par mutagenèse ou transgenèse

Fiche 28. Grands principes de la transgenèse et de la mutagenèse	112
Fiche 29. Transfert de gènes	114
Fiche 30. Transgenèse chez la souris	118
Fiche 31. Analyse de la fonction d'un gène par mutagenèse ou transgenèse chez les poissons.....	123
Fiche 32. Transgenèse chez <i>Drosophila melanogaster</i>	128
Fiche 33. Transgenèse de <i>Caenorhabditis elegans</i>	135
Fiche 34. Transgenèse stable des plantes par les agrobactéries	138
Fiche 35. Techniques de ARNi (ARN interférence)	145
Fiche 36. Transformation des plantes par agro-infiltration par <i>Agrobacterium tumefaciens</i>	151
Fiche 37. Analyse fonctionnelle de promoteurs	153
Fiche 38. Mutagenèse chimique du génome de drosophile	158
Fiche 39. Mutagenèse d'insertion	161
Fiche 40. Le TILLING (<i>Targeting-Induced Local Lesions IN Genomes</i>)	166
Fiche 41. Mutagenèse dirigée par le système CRISPR-Cas9	173

Partie 4 Approches haut débit

Fiche 42. Séquençer un génome : généralités.....	184
Fiche 43. Séquençage d'ADN par la technique Illumina.....	186

Fiche 44.	Séquençage d'ADN par la technique <i>Single Molecule Real Time</i> (SMRT) : PacBio	190
Fiche 45.	Séquençage d'ADN grande longueur par la technique Nanopore	193
Fiche 46.	Séquençage d'ADN par la technique Ion Torrent™	197
Fiche 47.	Séquençage d'ADN grande longueur « synthétique » par la technique 10X Genomics®	201
Fiche 48.	Cartographie optique de génomes (<i>Optical Mapping</i>)	204
Fiche 49.	La capture d'exons	208
Fiche 50.	Métagénomique et métatranscriptomique	212
Fiche 51.	<i>Barcoding</i> et métabarcoding	215
Fiche 52.	Analyse des données RNA-seq (<i>RNA-sequencing</i>)	218
Fiche 53.	Interactions protéines-ARN à la résolution du nucléotide (iCLIP)	225
Fiche 54.	Méthodes d'analyse du traductome	229
Fiche 55.	Structure de la chromatine : introduction	232
Fiche 56.	Immunoprécipitation de la chromatine (ChIP)	234
Fiche 57.	Analyse tridimensionnelle de la chromatine par <i>Chromosome Conformation Capture</i> : les techniques 3C, 4C-seq et Hi-C	236
Fiche 58.	Techniques d'identification de régions génomiques accessibles à des régulateurs : FAIRE et ATAC	240
Fiche 59.	Méthode de détermination du positionnement des nucléosomes : MAINE	243
Fiche 60.	La méthylation de l'ADN	245

Partie 5

Étude du polymorphisme d'un génome

Fiche 61.	Les principaux types de polymorphisme	252
Fiche 62.	Microsatellites, répétitions de séquences simples : SSR	255
Fiche 63.	Génotypage SNP (<i>Single Nucleotide Polymorphism</i>)	258
Fiche 64.	Utilisation des polymorphismes : QTL, GWAS et scans génomiques	266
Fiche 65.	Détermination du polymorphisme par séquençage de nouvelle génération	271
Fiche 66.	Approches populationnelles par séquençage en pools (<i>Pool-Seq</i>)	276

Fiche 67. Détection génomique de CNV par nouvelles technologies de séquençage	280
Fiche 68. Étude du polymorphisme par séquençage d'ADN associé à des sites de restriction (RAD-seq)	286
Fiche 69. Génotypage par séquençage : GBS (<i>Genotyping By Sequencing</i>)	292

Partie 6
Bioanalyses

Fiche 70. Assemblage de génomes	298
Fiche 71. Annotation de génomes	303
Fiche 72. Galaxy	305
Fiche 73. La programmation	308

Avant-propos

Cet ouvrage présente les principales techniques utilisées dans les laboratoires pour étudier le fonctionnement du vivant *via* les acides nucléiques.

Pour commencer, sont exposées quelques définitions et généralités sur ce que sont les gènes dans l'optique de cet ouvrage (il existe de nombreuses définitions, descriptions de gènes).

Les techniques de base consistent à pouvoir extraire les acides nucléiques, à les découper, en récupérer des fragments d'intérêt et les visualiser. Il s'agit également de les manipuler par les techniques de clonage et de *Polymerase Chain Reaction* (PCR). Une des applications très utilisées est le séquençage, dont les techniques ont grandement évolué en ces 15 dernières années et ne sont toujours pas stabilisées. Nous présentons dans cet ouvrage les techniques actuelles de séquençage commercialisées et accessibles *via* des plateformes publiques ou privées. Une autre grande application consiste à la manipulation des gènes par transformation génétique et mutagenèse afin d'étudier finement leur fonctionnement. Ces techniques évoluent également rapidement et nous présentons ici les techniques majeures de transformation génétique dont la technique d'édition des génomes (CRISPR-Cas9) qui est de plus en plus utilisée. Enfin, la plupart des approches font maintenant appel à des stratégies dites « haut débit » et l'analyse des données ainsi générées nécessite de la bioanalyse et de la bioinformatique : cet ouvrage ne vise pas à donner toutes les bases de ces analyses, mais quelques-unes des applications y sont décrites.

Remerciements

Cette édition est une refonte de l'ouvrage *Principes des techniques de biologie moléculaire* paru aux Éditions Inra en 2003. Nous tenons à remercier les rédacteurs de ces fiches qui ont su donner de leur temps pour que cet ouvrage existe. Les indications « d'après » se réfèrent d'ailleurs aux versions des éditions précédentes. Les nombreuses figures reprises de l'édition 2003 avaient été réalisées par Christian Moussard qui est ici chaleureusement remercié.

Liste des rédacteurs

BELCRAM Harry

Unité Génétique quantitative et Évolution
INRA, université Paris-Sud, CNRS, AgroParisTech
Ferme du Moulon, 91190 Gif-sur-Yvette

BERNARDEAU Karine

Plateforme P2R Production de protéines recombinantes, université de Nantes
CRCINA-Inserm 1232, SFR François Bonamy
8 quai Moncoussu, BP 70721, 44007 Nantes Cedex 1

BERTIN Chloé

INRA, laboratoire de Physiologie et Génomique des poissons, UR 1037
Campus de Beaulieu, Bât. 16A, 35042 Rennes Cedex

BLOYER Sébastien

Institut de biologie intégrative de la cellule (I2BC)
CNRS - université Paris-Sud
1 avenue de la Terrasse, 91198 Gif-sur-Yvette

BRETAUDEAU Anthony

INRA, UMR 1349 IGEP
BP 35327, 35653 Le Rheu Cedex
IRISA/INRIA, GenOuest Core Facility
Campus de Beaulieu, Rennes 35042

BROQUET Thomas

CNRS & Sorbonne université, UMR 7144
Station biologique de Roscoff, 29680 Roscoff

CANARD Elsa

INRA, UMR 1349 IGEP
BP 35327, 35653 Le Rheu Cedex

CORITON Olivier

Plateforme de Cytogénétique moléculaire
INRA, UMR 1349 IGEP
BP 35327, 35653 Le Rheu Cedex

DAGUIN-THIÉBAUT Claire

CNRS & Sorbonne université, UMR 7144
Station biologique de Roscoff, 29680 Roscoff

DELEURY Émeline

INRA, CNRS, université Côte d'Azur
UMR 1355 Institut Sophia-Agrobiotech
400 route des Chappes, BP 167, 06903 Sophia Antipolis

DOUCHIN Gaël

Plateforme P2R Production de protéines recombinantes, université de Nantes
CRCINA-Inserm 1232, SFR François Bonamy
8 quai Moncoussu, BP 70721, 44007 Nantes Cedex 1

DUVAUX Ludovic

UMR 1202 BIOGECO
INRA Pierroton
69 route d'Arcachon, 33612 Cestas

ECHASSERIEAU Klara

Plateforme P2R Production de protéines recombinantes, université de Nantes
CRCINA-Inserm 1232, SFR François Bonamy
8 quai Moncoussu, BP 70721, 44007 Nantes Cedex 1

FALENTIN Cyril

INRA, UMR 1349 IGEPP
BP 35327, 35653 Le Rheu Cedex

FAVERY Bruno

INRA-ISA
400 route des Chappes, BP 167, 06903 Sophia Antipolis

FUDAL Isabelle

UMR Bioger, campus AgroParisTech
Avenue Lucien Brétignières, 78850 Thiverval-Grignon

GRENIER Éric

INRA, UMR 1349 IGEPP
BP 35327, 35653 Le Rheu Cedex

HERPIN Amaury

INRA, laboratoire de Physiologie et Génomique des poissons, UR 1037
Campus de Beaulieu, bât. 16A, 35042 Rennes Cedex

HUTEAU Virginie

Plateforme de Cytogénétique moléculaire
INRA, UMR 1349 IGEPP
BP 35327, 35653 Le Rheu Cedex

HUYNH Jean-René

Collège de France, CIRB, UMR7241
11 place Marcelin-Berthelot, 75005 Paris

JAOUANNET Maëlle

INRA-ISA
400 route des Chappes, BP 167, 06903 Sophia Antipolis

JAQUIÉRY Julie

INRA, UMR 1349 IGEPP
BP 35327, 35653 Le Rheu Cedex

JAUBERT-POSSAMAI Stéphanie

INRA-ISA
400 route des Chappes, BP 167, 06903 Sophia Antipolis

JUBAULT Mélanie

INRA, UMR 1349 IGEPP
BP 35327, 35653 Le Rheu Cedex

LEGEAI Fabrice

INRA, UMR 1349 IGEPP
BP 35327, 35653 Le Rheu Cedex
INRIA/IRISA, GenScale
Campus de Beaulieu, 35042 Rennes

LE PASLIER Marie-Christine

INRA, EPGV US1279
CEA/IbFJ bât. G1, 2 rue Gaston-Crémieux, CP 5708, 91057 Évry Cedex

LE TRIONNAIRE Gaël

INRA, UMR 1349 IGEPP
BP 35327, 35653 Le Rheu Cedex

LOMBAERT Éric

INRA, CNRS, université Côte d'Azur
UMR 1355 Institut Sophia-Agrobiotech
400 route des Chappes, BP 167, 06903 Sophia Antipolis

MARCEL Fabien

INRA, Institut des sciences des plantes de Paris-Saclay, UMR1403
Plateforme de biologie translationnelle
IPS2 bâtiment 630, rue de Noetzelin, plateau du Moulon, 91190 Gif-sur-Yvette

MEJIAS Joffrey

Université Nice Sophia Antipolis - ISA
400 route des Chappes, BP 167, 06903 Sophia Antipolis

MÉREAU Agnès

Institut de génétique et développement de Rennes, UMR 6290 CNRS
Université de Rennes 1, faculté de médecine
2 avenue du Professeur-Léon-Bernard, CS34317, 35043 Rennes Cedex

MONÉGER Françoise

Laboratoire de Reproduction et Développement des plantes, université de Lyon
ENS de Lyon, UCBL, INRA, CNRS
46 allée d'Italie, 69364 Lyon Cedex 07

MONTARRY Josselin

INRA, UMR 1349 IGEP
BP 35327, 35653 Le Rheu Cedex

MONTIER Tristan

UMR INSERM 1078 Génétique, Génomique fonctionnelle et Biotechnologies,
équipe Transfert de gènes et Thérapie génique
Faculté de médecine et des sciences de la santé, université de Brest, CHRU de Brest
22 rue Camille-Desmoulins, CS 93837, 29238 Brest Cedex 3

MORALES Julia

CNRS, Sorbonne université
Biologie intégrative des modèles marins, UMR 8227
Station biologique de Roscoff, CS 90074, 29688 Roscoff

MOUGEL Christophe

INRA, UMR 1349 IGEP
BP 35327, 35653 Le Rheu Cedex

NOORDERMEER Daan

Institut de biologie intégrative de la cellule (I2BC), CNRS, université Paris-Sud
1 avenue de la terrasse, 91190, Gif-sur-Yvette

NOUHAUD Pierre

Institut für Populationsgenetik, Vetmeduni Vienna
Veterinärplatz 1, 1210 Vienne, Autriche

PECCOUD Jean

Université de Poitiers, CNRS 7267, Écologie & Biologie des interactions
Bâtiment B8-B35, 5 rue Albert-Turpain, TSA 51106, 86073 Poitiers Cedex 9

PECORARI Frédéric

Plateforme P2R Production de protéines recombinantes, université de Nantes
CRCINA-Inserm 1232, SFR François Bonamy
8 quai Moncoussu, BP 70721, 44007 Nantes Cedex 1

PONTS Nadia

INRA, UR1264-MycSA,
71 avenue Édouard-Bourlaux, CS 20032, 33882 Villenave-d'Ornon Cedex

QUENTIN Michaël

Université Nice Sophia Antipolis - ISA
400 route des Chappes, BP 167, 06903 Sophia Antipolis

REMY Séverine

Platform Transgenic Rats and ImmunoPhenomics, INSERM UMR 1064-CRTI
Université de Nantes
CHU Nantes, Hôtel Dieu, 30 bd Jean-Monnet, 44093 Nantes Cedex 01

RISPE Claude

UMR BIOEPAR, INRA
Oniris, La Chantrerie BP40706, 44307 Nantes

ROBERT Valérie J.

ENS de Lyon, CNRS UMR 5239, INSERM U1210
Laboratory of Biology and Modelling of the Cell
46 allée d'Italie, site Jacques-Monod, Lyon

ROBIN Stéphanie

INRA, UMR 1349 IGEP
BP 35327, 35653 Le Rheu Cedex
IRISA/INRIA, GenOuest Core Facility
Campus de Beaulieu, Rennes 35042

SOLON Jérôme

The Barcelona Institute of Science and Technology
Cell and Developmental Biology Programme, Centre for Genomic Regulation (CRG)
Dr. Aiguader 88, Barcelone 08003, Espagne
Universitat Pompeu Fabra (UPF)
Barcelone 08003, Espagne

STOCKINGER Petra

The Barcelona Institute of Science and Technology
Cell and Developmental Biology Programme, Centre for Genomic Regulation (CRG)
Dr. Aiguader 88, Barcelone 08003, Espagne
Universitat Pompeu Fabra (UPF)
Barcelone 08003, Espagne

TAGU Denis

INRA, UMR 1349 IGEP
BP 35327, 35653 Le Rheu Cedex

Partie 1

Les bases

Fiche 1

Structure et expression d'un gène eucaryote codant un ARNm et une protéine

Denis Tagu & Stéphanie Jaubert-Possamai

Un gène eucaryote à ARNm (ARN messenger) est codé par une séquence d'ADN double brin dont chaque brin est orienté ($5' \rightarrow 3'$). Un des deux brins de cette séquence d'ADN servira de matrice afin d'être transcrit en ARNm simple brin par l'ARN polymérase II. Ces gènes comportent une séquence codante bordée de séquences de régulation. La séquence codante est constituée d'exons et d'introns. Ces deux types de séquences sont transcrites (transcrit primaire), mais les introns sont éliminés lors de l'épissage. L'ARN est d'abord modifié en $5'$ -P (addition d'une coiffe) et en $3'$ -OH (addition de nucléotides à adénine), puis épissé (élimination des introns) avant d'être transporté dans le cytoplasme. Là, les ribosomes se fixent sur l'ARNm et, par l'intermédiaire des ARNt (ARN de transfert), l'ARNm est traduit en le polypeptide correspondant.

Les séquences de régulation servent de signaux de début ou de fin de transcription du gène par l'ARN polymérase II. Certaines de ces séquences (par exemple la TATA box du promoteur) sont reconnues par des protéines appelées « facteurs de transcription généraux » qui guident l'ARN polymérase II dans les étapes d'initiation et d'élongation de la transcription.