Chemogenomics and Chemical Genetics

Grenoble Sciences

The aims of Grenoble Sciences are double:

- to produce works corresponding to a clearly defined project, without the constraints of trends or programme,
- to ensure the utmost scientific and pedagogic quality of the selected works: each project is selected by Grenoble Sciences with the help of anonymous referees. Next, the authors work for a year (on average) with the members of an interactive reading committee, whose names figure in the front pages of the work, which is then co-published with the most suitable publishing partner.

Contact: Tel.: (33) 4 76 51 46 95 - E-mail: Grenoble.Sciences@ujf-grenoble.fr
website: *http://grenoble-sciences.ujf-grenoble.fr*

Scientific Director of Grenoble Sciences: **Jean BORNAREL**,
Emeritus Professor at Joseph Fourier University, Grenoble, France

Grenoble Sciences is a department of Joseph Fourier University, supported by the **French National Ministry for Higher Education and Research** and the **Rhône-Alpes Region**.

Chemogenomics and Chemical Genetics is an improved version of the original book ***Chemogénomique - Des petites molécules pour explorer le vivant*** sous la direction de Eric MARÉCHAL, Sylvaine ROY et Laurence LAFANECHÈRE, EDP Sciences - Collection Grenoble Sciences, 2007, ISBN 978 2 7598 0005 6.

The Reading Committee of the French version included the following members:

- **Jean DUCHAINE**, Principal Advisor of the Screening Platform, Institute for Research in Immunology and Cancer, University of Montreal, Canada
- **Yann GAUDUEL**, Director of Research at INSERM, Laboratory of Applied Optics (CNRS), *Ecole Polytechnique*, Palaiseau, France
- **Nicole MOREAU**, Professor at the *Ecole Nationale Supérieure de Chimie*, Pierre and Marie Curie University, Paris, France
- **Christophe RIBUOT**, Professor of Pharmacology at the Faculty of Pharmacy, Joseph Fourier University, Grenoble, France

Translation performed by ***Philip SIMISTER***

Typesetted by ***Centre technique Grenoble Sciences***

Cover illustration: ***Alice GIRAUD***
(with extracts from a DNA microarray image - *Biochip Laboratory/Life Sciences Division/CEA* - and a photograph of actin filaments array and adhesion plates in a mouse embryonic cell - *Yasmina SAOUDI, INSERM U836 Grenoble, France*)

Eric Maréchal • Sylvaine Roy • Laurence Lafanechère
Editors

Chemogenomics and Chemical Genetics

A User's Introduction for Biologists, Chemists and Informaticians

Springer

Editors
Dr. Eric Maréchal
Laboratory of Plant Cell Physiology
UMR 5168, CNRS-CEA-INRA-
Joseph Fourier University
Rue des Martyrs 17
38054 Grenoble Cedex 9
France
eric.marechal@cea.fr

Sylvaine Roy
Laboratory of Plant Cell Physiology
UMR 5168, CNRS-CEA-INRA-
Joseph Fourier University
Rue des Martyrs 17
38054 Grenoble Cedex 9
France
sylvaine.roy@cea.fr

Laurence Lafanechère
Albert Bonniot Institute
Department of Cellular Differentiation and Transformation
Rond-point de la Chantourne
38706 La Tronche Cedex
France
laurence.lafanechere@ujf-grenoble.fr

Translator:
Philip Simister
Weatherall Institute of Molecular Medicine
University of Oxford
Oxford OX3 9DS, UK

Originally published in French: Chemogénomique - Des petites molécules pour explorer le vivant sous la direction de Eric MARÉCHAL, Sylvaine ROY et Laurence LAFANECHÈRE, EDP Sciences - Collection Grenoble Sciences, 2007, ISBN 978 27598 0005 6.

ISBN 978-3-642-19614-0 e-ISBN 978-3-642-19615-7
DOI 10.1007/978-3-642-19615-7
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011930786

Cover illustration: Alice Giraud

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

CONTENTS

THIRD PART
TOWARDS AN IN SILICO EXPLORATION OF CHEMICAL AND BIOLOGICAL SPACES

PREFACE

Jean CROS

Having completed the reading of this work, one can only feel satisfied for having encouraged Laurence LAFANECHÈRE, Sylvaine ROY and Eric MARÉCHAL, who attempted and succeeded in achieving the impossible: writing, along with colleagues from the public sector, a book that will endure, concerning a technology the mastery of which had remained until this point the domain of the pharmaceutical industry. Indeed, this work has arisen from the competence and practical knowledge of fifteen or so academic scientists who, often against the tide of strategies defined by their host organisations, have established automated pharmacological screening for fundamental research ends. It is important to recall that the first book *High throughput screening*, edited in 1997 by John P. DEVLIN, which enabled all scientists to discover the importance of robotics in the discovery of new medicines, was written by about a hundred contributors, all of whom were industrial scientists involved in drug discovery.

Over the last ten years, we have seen appear in the scientific literature much more about 'small' molecules coming from robotic screens that have been used with success in revealing new biological mechanisms. From drug candidate, the molecule has thus become a research tool.

The successful experience at Harvard is a fertile example which should serve as a model for some of our research centres: basic research in chemical genetics, discovery of new drug candidates and training of young researchers. May this book, which has developed out of training workshops organised by the CNRS, CEA and INSERM, be the stimulus for future careers in a field which is eminently multidisciplinary and which brings together biologists, chemists, informaticians and robotics specialists. The great merit of this book is to have simply, from everyday experiences, united researchers and competencies that until now had not associated with one another.

Beyond the new terms that we discover or rediscover throughout the chapters: chemical genetics, cheminformatics, chemogenomics etc., there are the techniques, certainly, but also and above all there are the scientific questions to which these technologies will henceforth help to find answers. In addition, there are the economic issues that from now on become the duty of every researcher to take into account.

Congratulations to all of the authors and editors.

INTRODUCTION

André TARTAR

Over the last two decades, biological research has experienced an unprecedented transformation, which often resulted in the adoption of highly parallel techniques, be it the sequencing of whole genomes, the use of DNA chips or combinatorial chemistry. These approaches, which have in common the repeated use of trial and error in order to extract a few significant events, have only been made possible thanks to the progress in miniaturisation and robotics informatics.

One of the first sectors to put into practice this approach was within pharmaceutical research with the systematic usage of high-throughput screening for the discovery of new therapeutic targets and new drug candidates. Academic research has for a long time remained distanced from this process, as much for financial as for cultural reasons. For several years, however, the trivialisation of these techniques has led to a considerable reduction in the cost of accessing them and has thus permitted academic groups to employ such methods in projects having generally more cognitive objectives.

Nevertheless, it is no less vital, as with all involved methods, to take into account the cost factor as a fundamental parameter in the development of an experimental protocol relative to the expected benefit. The value of a chemical library is in effect an evolving notion resulting from the sum of two values that evolve in opposite directions:

- On the one hand, the set of physical samples whose value will fatally decrease due both to its consumption in tests, but above all to the degradation of the components. The experience of the last few years also shows that it will be subjected to the effects of fashion, which will contribute rapidly to its obsolescence: no-one today would assemble a chemical library as would have been done only five years ago. Since the great numbers that dominated the first combinatorial chemical libraries, a more realistic series of criteria has progressively been introduced, bearing witness to the difficulties encountered. 'Drugability' has thus become a keyword, with LIPINSKI's rule of 5 and the 'frequent hitters' becoming the *bête noire* of screeners having given them too often cause for hope, albeit unfounded.
- On the other hand, the mass of information accumulated over the different screening tests is ever increasing and will progressively replace the physical chemical library. With a more or less distant expiry date, the physical chemical

library will have disappeared and the information that it has allowed to accumulate will be all that remains. This information can then be used either directly, constituting the 'specification sheet' of a given compound, or as a reference source in virtual screening exercises or *in silico* prediction of the properties of new compounds.

A very simple strategic analysis shows that with the limited means available to academic teams, it is easier to be competitive with respect to the second point (quantity and quality of information) than to the first (number of compounds and high thoughput). This also shows that the value of an isolated body of information is much less than that of an array organised in a logical manner based on two main dimensions: the diversity of compounds and the consistency of the biological tests.

It is in this vein that high-content screening should become established, permitting the collection and storage of the maximum amount of data for each experiment. This high-content screening will be the guarantee for the optimal evaluation of physical collections. It is interesting to note that the problem of information loss during a measurement was at the centre of spectroscopists' preoccupations a few decades ago. In the place of dispersive systems (e.g. prisms, networks) that sequentially selected each observation wavelength but let all others escape, they have substituted non-dispersive analysis techniques entrusting deconvolution algorithms and multi-channel analysers with the task of processing the global information. Biology is undergoing a complete transformation in this respect. Whereas about a decade ago one was satisfied by following the expression of a gene under the effect of a particular stimulus, today, thanks to pan-genomic chips, the expression profile of the whole genome has become accessible. It is imperative that screening follows the same path of evolution: no longer losing any information will become the rule. In the longer term, it will be necessary for this information to be formatted and stored in a lasting and reusable manner.

With this perspective, this book appears at just the right moment since it constitutes a reference tool enabling different specialists to speak the same language, which is essential to ensure the durability of the information accrued.