Methods in Molecular Biology 1326

Springer Protocols

Anthony W. Ryan *Editor*

Celiac Disease

Methods and Protocols

Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor
John M. Walker
School of Life and Medical Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes:
http://www.springer.com/series/7651

Celiac Disease

Methods and Protocols

Edited by

Anthony W. Ryan

Department of Clinical Medicine, Trinity College Dublin; Institute of Molecular Medicine, Trinity Centre for Health Sciences, St James's Hospital, Dublin, Ireland.

Humana Press

Editor
Anthony W. Ryan
Department of Clinical Medicine
Trinity College Dublin
Dublin, Ireland

Institute of Molecular Medicine
Trinity Centre for Health Sciences
St James's Hospital
Dublin, Ireland

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-2838-5 ISBN 978-1-4939-2839-2 (eBook)
DOI 10.1007/978-1-4939-2839-2

Library of Congress Control Number: 2015944312

Springer New York Heidelberg Dordrecht London

Printed on acid-free paper

Humana Press is a brand of Springer
Springer Science+Business Media LLC New York is part of Springer Science+Business Media (www.springer.com)

Preface

Recent decades have seen considerable advances in our understanding of celiac disease. The condition, once thought to be limited to individuals of European ancestry, has been discovered at varying prevalence in North Africa, the Middle East, India, and China. The precipitating auto-antigen has been characterized. The genetic association of the HLA region was discovered early and refined in the years that followed. However, conclusive identification of non-HLA genetic risk proved elusive until the advent of genome-wide association studies, which have extended our understanding of the genetic component far beyond what could have been envisaged a short time ago. Current estimates suggest that more than 50% of the population variability associated with celiac disease risk can be explained by known genetic loci. Despite these advances, there is still a great deal to be learned, both about the nature of genetic risk and the functional genomic consequences of the established risk factors.

Building on this knowledge will require detailed molecular analysis of the associated pathways and many cell types involved in the disease, as well as embracing new technologies such as next-generation sequencing. At the same time, long established molecular and immunology methods will continue to have a place for some time to come. This book brings together novel and more traditional methods in molecular biology and immunology, in order to provide a tool-kit for all stages of celiac disease research, from the practicalities of obtaining high-quality samples, to molecular analysis and bioinformatics.

Part I of this book sets the background with a number of reviews to describe the history and nature of the disease, its diagnosis, the role of animal models, and study designs for investigating genetic susceptibility. Part II describes the molecular techniques, including tissue culture, isolation and cloning of relevant cell types, high content analysis of biopsies, and HLA genotyping. Subsequent chapters describe analyses of gene expression and functional analysis of genetic variants: detecting allelic expression imbalance, reporter assays, and siRNA knockdown. The final chapter in this part describes a number of protocols for epigenetic analysis, which has attracted little attention until recently.

A great deal of modern molecular biology relies on high-throughput automated analyses, which produce vast quantities of data. Coming to grips with this flow of information represents a new set of challenges, dealt with in Part III. The final three chapters describe analysis pipelines for bioinformatic prediction of antigenicity, quality control and analysis of GWAS data, and transcriptome analysis by next-generation sequencing. These techniques require a certain level of scripting skills, not commonly held by laboratory based researchers. For this reason, Part III begins with an outline of scripting for data management, focusing on tools which are freely available to researchers who wish to explore them.

To conclude, I would like to thank John Walker and Humana Press for their guidance and assistance throughout this project, and I extend my gratitude to all contributors to the book, who took time out from their busy research and clinical schedules to write their chapters.

Dublin, Ireland ***Anthony W. Ryan***

Contents

Contributors

Richard J.L. Anney • *Institute of Molecular Medicine, Trinity Centre for Health Sciences, St James's Hospital, Dublin, Ireland; Department of Psychiatry, Trinity College Dublin, Dublin, Ireland*

Renata Auricchio • *Department of Translational Medicine, Section of Pediatrics, University of Naples Federico II, Via S. Pansini 5, Naples, Italy; European Laboratory for the Investigation of Food Induced Diseases (ELFID), University of Naples Federico II, Via S. Pansini 5, Naples, Italy*

Anne-Marie Baird • *Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland; Thoracic Oncology Research Group, Institute of Molecular Medicine, Trinity Centre for Health Science, St. James's Hospital, Dublin, Ireland; Cancer and Aging Research Program, Queensland University of Technology, Brisbane, Australia*

Greg Byrne • *School of Biological Sciences, Dublin Institute of Technology, Dublin, Ireland*

Mikaela M. Byrne • *Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland; Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland*

Ciara Coleman • *Department of Medicine, Institute of Molecular Medicine, Trinity College Dublin, College Green, Dublin, Ireland; St. James's Hospital, Dublin, Ireland*

Sarah E.J. Cooper • *Immunology Department, Trinity College Dublin, Dublin, Ireland*

Sandra D' Alfonso • *Department of Health Sciences, University of Eastern Piedmont, Novara, Italy; Interdisciplinary Research Centre of Autoimmune Disease (IRCAD), University of Eastern Piedmont, Novara, Italy*

Ennia Dametto • *Immunogenetica e Biologia dei Trapianti, AOU Città della salute e della scienza di torino, via Santena, Torino, Italy*

Anthony Mitchell Davies • *INCHSA, Institute for Molecular Medicine, Trinity College Dublin, Dublin, Ireland*

Jean Dunne • *Immunology Department, Trinity College Dublin, Dublin, Ireland; Immunology Department, St James's Hospital, Dublin, Ireland*

Margaret R. Dunne • *Department of Surgery, Trinity Centre for Health Sciences, St James's Hospital, Dublin, Ireland; National Children's Research Centre, Our Lady's Children's Hospital, Dublin, Ireland; Department of Immunology, Institute of Molecular Medicine, St. James's Hospital, Trinity College Dublin, Dublin, Ireland*

Louise Elliot • *Immunology Department, Trinity College Dublin, Dublin, Ireland*

Maria Edvige Fasano • *Immunogenetica e Biologia dei Trapianti, AOU Città della salute e della scienza di torino, via Santena, Torino, Italy*

Conleth F. Feighery • *Immunology Department, Trinity College Dublin, Dublin, Ireland; Immunology Department, St. James's Hospital, Dublin, Ireland*

Michael Freeley • *Department of Medicine, Institute of Molecular Medicine, Trinity College Dublin, St. James's Hospital, Dublin, Ireland*

Jillian M. Gahan • *Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland; Institute of Molecular Medicine, Trinity Centre for Health Sciences, St James's Hospital, Dublin, Ireland*

Martina Galatola • *Department of Translational Medicine, Section of Pediatrics, University of Naples Federico II, Via S. Pansini 5, Naples, Italy; European Laboratory for the Investigation of Food Induced Diseases (ELFID), University of Naples Frederico II, Via S. Pansini 5, Naples, Italy*

Steven G. Gray • *Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland; Thoracic Oncology Research Group, Institute of Molecular Medicine, Trinity Centre for Health Science, St. James's Hospital, Dublin, Ireland; HOPE Directorate, St James's Hospital, Dublin, Ireland*

Luigi Greco • *Department of Translational Medicine, Section of Pediatrics, University of Naples Federico II, Via S. Pansini 5, Naples, Italy; European Laboratory for the Investigation of Food Induced Diseases (ELFID), University of Naples Federico II, Naples, Italy*

Matthew Hill • *Neurosciences and Mental Health Research Institute, Cardiff University School of Medicine, Hadyn Ellis Building, Cathays, Cardiff, UK*

Karsten Hokamp • *School of Genetics and Microbiology, Smurfit Institute of Genetics, Trinity College Dublin, College Green, Dublin, Ireland*

Josephine M. Ju • *Department of Immunology, Mayo Clinic Rochester, Rochester, MN, USA*

Jacinta Kelly • *National Children's Research Centre, Our Lady's Children's Hospital, Dublin, Ireland*

Frits Koning • *Department of Immunohematology and Blood Transfusion, Leiden University Medical Centre, Leiden, The Netherlands*

Yvonne Kooy-Winkelaar • *Department of Immunohematology and Blood Transfusion, Leiden University Medical Centre, Leiden, The Netherlands*

Aideen Long • *Department of Medicine, Institute of Molecular Medicine, Trinity College Dublin, St. James's Hospital, Dublin, Ireland*

Eric V. Marietta • *Department of Immunology, Mayo Clinic Rochester, Rochester, MN, USA; Department of Gastroenterology, Mayo Clinic Rochester, Rochester, MN, USA; Department of Dermatology, Mayo Clinic Rochester, Rochester, MN, USA*

Ross McManus • *Department of Medicine, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland*

Bashir M. Mohamed • *Immunology Department, Trinity College Dublin, Dublin, Ireland; Thoracic Oncology, St James's Hospital, Dublin, Ireland*

Ben Molloy • *Department of Medicine, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland*

Ross T. Murphy • *Department of Cardiology, St James's Hospital, Dublin, Ireland; Institute of Cardiovascular Science, St. James's Hospital, Dublin, Ireland*

Joseph A. Murray • *Department of Immunology, Mayo Clinic Rochester, Rochester, MN, USA; Department of Gastroenterology, Mayo Clinic Rochester, Rochester, MN, USA; Department of Gastroenterology and Hepatology, Mayo Foundation, Rochester, MN, USA*

Åsa Torinsson Naluai • *Institute of Biomedicine, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden*

Cathal P. O'Brien • *St James's Hospital, Dublin, Ireland; Trinity College Dublin, Dublin, Ireland*

Antoinette S. Perry • *Prostate Molecular Oncology, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland*

Emma M. Quinn • *Department of Clinical Medicine, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland*

Anthony W. Ryan • *Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland; Institute of Molecular Medicine, Trinity Centre for Health Sciences, St James's Hospital, Dublin, Ireland*

Graham D. Turner • *Department of Clinical Medicine, Trinity College Dublin, Dublin, Ireland; Institute of Molecular Medicine, Trinity Centre for Health Sciences, St. James's Hospital, Dublin, Ireland*

Sharon Wilson • *Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland*

Part I

Background

Chapter 1

Celiac Disease: Background and Historical Context

Graham D. Turner, Margaret R. Dunne, and Anthony W. Ryan

Abstract

Medical descriptions of celiac disease date to the first century BC, and the first modern description was published in 1888. Further insights were gained throughout the 1900s, culminating in the identification of the dietary component, the major genetic determinant, and the autoantigen by the turn of the century. Understanding of the age of onset, population prevalence, and the extent of subclinical celiac disease developed in tandem. Thanks to advances in genomics, currently established loci account for over 50 % of the genetic risk. Nonetheless, much remains to be discovered. Advances in high-throughput genomic, biochemical, and cell analyses, as well as the bioinformatics needed to process the data, promise to deepen our understanding further. Here we present a primer of celiac disease, viewing the condition in turn from the historical, epidemiological, immunological, molecular, and genetic points of view. Research into any ailment has specific requirements: study subjects must be identified and relevant tissue samples collected and stored with the appropriate timing and conditions. These requirements are summarized. To conclude, a short discussion of future prospects is presented.

Key words Celiac disease, Genetic risk, HLA, Gluten, Transglutaminase, T cell

1 A Brief Description of Celiac Disease

Celiac disease (CD) is an autoimmune inflammatory condition which primarily affects the small intestine. Common symptoms include bloating, diarrhea, and, particularly in children, "failure to thrive" due to malabsorption of dietary nutrients [1]. In a subset of cases the disease also manifests as dermatitis herpetiformis, an inflammatory, blistering condition of the skin [2]. At the molecular level, CD is triggered by an immune reaction to gluten, a protein present in wheat, barley, and rye. Virtually all sufferers of CD carry at least one copy of either HLA-DQ2 or HLA-DQ8, two genetic variants of the human leukocyte antigen (HLA, also termed major histocompatibility complex, MHC) Class II molecules [3]. However, these variants are common in many populations [4], while only about 1 % develop CD. As such, possession of one or both of the variants is necessary, but not sufficient, for onset of the disease.

Anthony W. Ryan (ed.), *Celiac Disease: Methods and Protocols*, Methods in Molecular Biology, vol. 1326,
DOI 10.1007/978-1-4939-2839-2_1,

Other factors, both genetic and environmental, come into play in its development.

The gold standard for diagnosis of CD requires endoscopic intestinal biopsy and serological detection of anti-tissue transglutaminase (anti-tTG) and/or anti-endomysial antibodies (anti-EMA). At the histological level, CD leads to flattened microvilli in the small intestine of sufferers. This villous atrophy, which may be graded according to the Marsh classification [5], is accompanied by an increase in the number of intraepithelial lymphocytes, indicative of the immune component of the disease. CD may be accompanied by additional comorbidities, including osteoporosis, other autoimmune diseases, and psychiatric disorders [6]. The characteristic villous atrophy explains the malabsorption associated with the disease, which may also lead to secondary lactose intolerance, as the cells that normally produce the lactase enzyme are damaged by the formation of the lesion. CD is treated by strict adherence to a gluten-free diet, which usually results in complete remission [6]. While barley, rye, and wheat undoubtedly contain immunogenic proteins, there is some evidence that oats, prepared in wheat-free facilities, may be safe for consumption by celiac sufferers [7].

Refractory CD has been postulated in a minority of cases where the gluten-free diet is unsuccessful [8]. However, it is difficult to distinguish true refractory disease from noncompliance to the rigid requirements of the diet. Recent literature has also considered the possibility of non-celiac gluten or wheat intolerance [9, 10], which is beyond the scope of this chapter.

2 Historical Aspects

There is evidence for the use of grains such as wild wheat and barley dating to 23,000 years ago in the upper Paleolithic [11, 12]. However, the use of cereals did not become widespread until after the Neolithic revolution of approximately 10,000 years ago. While natural selection may have played a role at this stage, many of the alleles associated with CD risk may have been maintained at high frequency due to their ability to confer other beneficial traits, such as resistance to bacterial infection [13]. Archaeological evidence of probable CD has been identified in 2000-year-old human remains from Italy [14]. Interestingly, there is some evidence that einkorn (*Triticum monococcum*), the earliest cultivated wheat, may be less toxic to celiac sufferers than more modern varieties [15].

CD has been recognized since ancient times. It was first described in the first century BC by the Greek physician Aretaeus of Cappadocia, whose works were translated in the 1800s [16]. Aretaeus identified CD as an affliction of later life, most commonly affecting women. The physician Samuel Gee gave the first modern description of the condition in 1888, building upon Aretaeus' observations.

However, he primarily observed the condition in infants, and considered it a disease of childhood. The "classical" picture of CD—occurring in the young, presenting with characteristic abdominal symptoms, diarrhea, and "failure to thrive"—owes itself to Gee's observations at this time [17].

A dietary, specifically carbohydrate, component to CD was long suspected. The first treatments for CD pre-date full understanding of the etiology, for example the "banana diet" [18]. However, it was not until the 1940s that the physician Wilhelm Dicke identified the ingestion of wheat as the environmental trigger, aided by the observation that reduced morbidity from CD coincided with the shortage of wheat during the Dutch Hongerwinter of 1944 [19].

3 Evolving Understanding of Onset and Prevalence

The view of CD, as described by Gee, is one of a rare illness of childhood, affecting individuals of European descent. As recently as 1985, estimates of incidence of CD placed the population frequency at 0.1 %. Even at this time however, the classical view of the disease as an affliction of childhood was beginning to be questioned, with the Celiac Society noting an increase in age of diagnosis amongst its members [20].

There is some evidence that the age of onset of adult CD may follow a bimodal distribution, with an initial peak in the fourth decade of life (mostly women) and a second, smaller peak in the sixth or seventh decade of life (predominantly men) [21]. The apparent increase in incidence amongst the older population may be due to improved screening and diagnostic techniques, leading to the recognition of CD in cases where it would previously have gone unnoticed or misdiagnosed. Additionally, individuals with "silent," or asymptomatic, disease may convert to a more aggressive phenotype with age. Increased recognition of nonclassical symptoms (active, silent, latent, and potential CD) may aid in diagnosis [22].

The prevalence of asymptomatic CD is unknown, and indeed known cases may constitute the "tip of the iceberg" [23]. Definitive diagnosis requires endoscopic biopsy, which is invasive and not conducive to population screening. Serological tests which assay for the presence of anti-tTG autoantibodies have also been developed as a diagnostic method. Many researchers have screened populations using these serological assays, which can give good estimates of the true population prevalence. However, estimates of prevalence based on serology alone, in the absence of endoscopic confirmation and HLA genotyping, will contain a proportion of false positives.

Historically, CD risk has been most comprehensively described in populations of European descent [1]. However, the DQ2 and DQ8 variants are widespread in worldwide populations, and the condition was once considered rare in some populations where it is now known to be present. Therefore, there is a precedent for underdiagnosis of CD, and future studies may reveal a more widespread distribution, particularly as the consumption of wheat increases, as observed in China [24].

4 The Immunology of Celiac Disease

CD is driven by aberrant immune responses to dietary gluten in genetically predisposed individuals. Therefore, in addition to genetic and environmental components, the immune system plays an important role in CD pathogenesis.

The immune system is divided into innate and adaptive arms. The innate immune system is comprised of first-line defence barriers and rapidly responding immune cells activated by conserved molecular patterns. The adaptive immune system, comprised of T cells and B cells, can take days to mount a response but specifically recognizes peptide antigens and can develop immunological memory. Antigen-presenting cells bridge the innate and adaptive immune systems by processing and presenting peptide antigens via MHC molecules to stimulate T cell responses. In the context of CD, antigen-presenting cells present deamidated gliadin peptides bound to MHC molecules HLA-DQ2 or HLA-DQ8 to activate gluten-specific T cells.

One of the early hallmarks of CD is an influx of activated lymphocytes into the small intestine, concomitant with villous atrophy [25, 26]. Infiltrating T cells show memory and cytotoxic phenotypes and are thought to drive mucosal damage in response to gluten peptides [27, 28]. It has been postulated that gluten-specific CD4+ T cells found in the lamina propria activate intraepithelial CD8+ T cells and B cells via cross-priming and cytokine production, which further drives inflammation and tissue damage [29, 30]. Cytokines, such as interferon-γ (IFN-γ), interleukin-15 (IL-15), and IL-21 in particular, have been implicated in CD pathogenesis [31, 32]. IFN-γ, a pro-inflammatory T helper 1 (T_H1) type cytokine, increases intestinal barrier permeability and drives cytotoxic CD8+ T cell activation, whereas IL-15 further drives T_H1 functions and cytotoxic T cell survival.

Activated B cells produce antibodies specific for gluten peptides and tissue transglutaminase, the latter providing a useful diagnostic test for CD [33]. Other types of infiltrating immune cells have also been described in the celiac small intestine. Gamma delta (γδ) T cells, particularly the Vδ1 subtype, are significantly enriched in the human small intestinal epithelium, persisting even after

elimination of gluten from the diet [34, 35]. The precise role of these cells in CD remains unclear, but immunoregulatory and tissue repair functions have been described, suggesting that Vδ1 cells may aid the restoration and maintenance of the small intestinal epithelium [36, 37].

Chronic activation of T cells can result in lymphomagenesis in a small number of CD patients, leading to the development of aggressive enteropathy-associated T cell lymphomas (EATLs) with poor prognosis [38, 39].

5 Molecular Basis

Sollid et al. [40] identified the HLA-DQ variants as the main disease susceptibility factor. The contribution of these variants to celiac disease risk is considerably greater than any of the other known genetic risk factors.

Gluten, found in the endosperm of wheat, barley, and rye, is composed of various subunits. One class of these subunits, the gliadins, is capable of triggering an immunogenic response in CD [6, 41]. Gliadin is composed of more than a hundred components, which can be classified into four main types, termed ω5-, ω1,2-, α/β-, and γ-gliadins. Multiple gliadin-derived epitopes are immunogenic and toxic to CD sufferers [42]. Storage proteins with similar amino acid composition and toxicity to celiac sufferers have been identified in barley and rye (hordeins and secalines, respectively) [43].

Although many gluten-derived epitopes are immunogenic, they display a hierarchy of immunogenicity. A peptide of 33 amino acids (residues 57–89) derived from an α-gliadin fraction appears to be immune-dominant, properties attributable to its proline and glutamine-rich peptide structure. The density of proline residues increases the resistance of the peptide to gastrointestinal proteolysis, both in CD and in unaffected individuals. In addition, the resulting left-handed helical conformation strengthens binding to HLA-DQ2 and HLA-DQ8 molecules. tTG-mediated deamidation of gliadin peptides results in further enhanced immunogenicity [44]. The criteria for epitope binding to DQ2 or DQ8 are well recognized, with both DQ2 and DQ8 dimers exhibiting preferences for negatively charged residues at positions within the core peptide-binding groove [45, 46]. These negatively charged residues are the result of the deamidation of glutamine residues by tTG [47]. Therefore, the endogenous enzyme tTG, implicated in CD due to the presence of autoantibodies, is the catalyst for the deamidation and subsequent increase in epitope affinity for HLA-DQ2 and HLA-DQ8 binding sites, enhancing gliadin-specific T cell activity [48].