Developmental Biology 382 (2013) 320–329

Contents lists available at ScienceDirect

Developmental Biology

journal homepage: www.elsevier.com/locate/developmentalbiology

Contribution of rare and common variants determine complex diseases—Hirschsprung disease as a model

Maria M. Alves [a,*], Yunia Sribudiani [a], Rutger W.W. Brouwer [b,c], Jeanne Amiel [d,l], Guillermo Antiñolo [e], Salud Borrego [e], Isabella Ceccherini [f], Aravinda Chakravarti [g], Raquel M. Fernández [e], Maria-Mercè Garcia-Barcelo [h], Paola Griseri [f], Stanislas Lyonnet [d,l], Paul K. Tam [h], Wilfred F.J. van IJcken [b], Bart J.L. Eggen [i], Gerard J. te Meerman [j], Robert M.W. Hofstra [a,k]

[a] Department of Clinical Genetics, Dr. Molewaterplein, 50, Rotterdam, The Netherlands
[b] Center for Biomics, Department of Cell Biology, Erasmus University Medical Center, Dr. Molewaterplein 50, Rotterdam, The Netherlands
[c] Netherlands Bioinformatics Centre (NBIC), Nijmegen, The Netherlands
[d] INSERM U781, APHP Hôpital Necker-Enfants Malades, Paris, France
[e] Department of Genetics, Reproduction and Fetal Medicine, Institute of Biomedicine of Seville (IBIS), University Hospital Virgen del Rocío/CSIC/University of Seville, Seville, Spain
[f] Laboratorio di Genetica Molecolare, Instituto Giannina Gaslini, Genoa, Italy
[g] Department of Surgery, University of Hong Kong, Hong Kong SAR, China
[h] McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA
[i] Department of Neuroscience, Medical Physiology section, University Medical Center Groningen, University of Groningen, Antonius Deusinglaan 1, Groningen The Netherlands
[j] Department of Genetics, University Medical Centre Groningen, University of Groningen, PO BOX 30.001, Groningen, The Netherlands
[k] Neural Development Unit, UCL Institute of Child Health, 30 Guilford Street, London, UK
[l] Université Paris Descartes - Sorbonne Paris Cité, Institut Imagine, Paris, France.

ARTICLE INFO

Article history:
Received 16 November 2012
Received in revised form
13 May 2013
Accepted 15 May 2013
Available online 23 May 2013

Keywords:
Hirschsprung disease
Rare variants
GWAS
Next generation sequencing
Systems biology
Functional assays

ABSTRACT

Finding genes for complex diseases has been the goal of many genetic studies. Most of these studies have been successful by searching for genes and mutations in rare familial cases, by screening candidate genes and by performing genome wide association studies. However, only a small fraction of the total genetic risk for these complex genetic diseases can be explained by the identified mutations and associated genetic *loci*. In this review we focus on Hirschsprung disease (HSCR) as an example of a complex genetic disorder. We describe the genes identified in this congenital malformation and postulate that both common 'low penetrant' variants in combination with rare or private 'high penetrant' variants determine the risk on HSCR, and likely, on other complex diseases. We also discuss how new technological advances can be used to gain further insights in the genetic background of complex diseases. Finally, we outline a few steps to develop functional assays in order to determine the involvement of these variants in disease development.

Introduction

Complex diseases are common and a major contributor to disability and death worldwide. They are thought to arise from multiple predisposing factors, both genetic and non-genetic, and the joint effects of these factors are thought to be of key importance in disease development (reviewed by Lander and Schork, 1994; Kiberstis and Roberts, 2002).

Understanding genetically complex (polygenic) diseases has become a major topic in the field of human genetics as this may have a large impact on both the treatment and prognosis of the patients. However, finding mutations that in concert give rise to a disease is difficult (Kiberstis and Roberts, 2002). Within groups of patients with complex genetic diseases, most attention has been paid to the rare familial and syndromic cases. For these cases, linkage analysis in large multi-generational families, sib-pair analysis in smaller nuclear families, and genetically isolated populations have often been used (reviewed by Heutink and

* Correspondence to: Department of Clinical Genetics, Erasmus University Medical Center, Dr. Molewaterplein 50, 3015 GE Rotterdam, The Netherlands. Fax: +31 10 7044743.
E-mail address: m.alves@erasmusmc.nl (M.M. Alves).

0012-1606/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ydbio.2013.05.019

Oostra, 2002; Poznik et al., 2006; reviewed by Cui et al., 2010). Besides these extended family-based approaches, candidate genes were also identified based on mouse models and gene networks (McCallion et al., 2003; Ahlqvist et al., 2011; Liu et al., 2011). These lines of research resulted in a number of risk-factors and markers for various complex genetic diseases, but were unable to account for the total genetic burden.

The progress made in the Human Genome Project and the HapMap project have enabled genome-wide association studies (GWAS) (reviewed by Manolio and Collins, 2009; Witkowski, 2010). This unbiased approach has led to the identification of many novel, disease-susceptibility *loci* for a large number of multifactorial diseases (reviewed by McCarthy et al., 2008; Hindorff et al., 2009). However, recent studies make it clear that the common variants identified only explain 20–30% of the total attributable risk for these complex diseases (reviewed by Lander, 2011). Yang et al. (2010) have put forward arguments that the missing heritability can be explained as a consequence of associations that are too small to be detected with the cohort sizes in the current GWAS studies. In addition to the valid arguments raised by Yang et al., we propose that rare and (de novo) mutations also explain part of the missing heritability in complex diseases.

In this review, we focus on Hirschsprung disease (HSCR) as an example of a complex disorder and discuss expectations from technological advances in detecting rare causative mutations. Furthermore, we discuss how these technologies can be used to find new candidate genes for complex disorders.

HSCR

HSCR is a congenital disorder pathologically characterized by an absence of enteric neurons along a variable length of the gut. It is a developmental disorder that is caused by a failure in migration of enteric neural crest-derived cells (ENCDCs) into the intestinal tract, or due to a failure in survival, proliferation or differention of ENCDCs once they reach the gut (reviewed by Heanue and Pachnis, 2007). The ENCDCs form the enteric nervous system and originate mainly from vagal neural crest cells that invade the foregut and migrate in a rostral to caudal direction to colonize the entire foregut, midgut, caecum, and hindgut. These vagal neural crest cells give rise to most of the enteric nervous system. Sacral neural crest cells also make a contribution, but a much smaller one, as they migrate in opposite direction, caudal to rostral, to colonize the colon (Burns and Douarin, 1998; reviewed by Anderson et al., 2006). For more information about the enteric nervous system and the molecular mechanisms controlling the colonization of the gut by ENCDCs, please see a recent review published in a previous edition of this journal (Sasselli et al., 2012).

In HSCR, the absence of ENCDCs leads to tonic contraction of the affected segment and intestinal obstruction. This obstruction may lead to failure to pass the first stool within 48 h after birth, vomiting, and massive distension of the proximal bowel (also called megacolon) or neonatal enterocolitis. Patients are diagnosed with one of three forms: the short-segment form (S-HSCR, approximately 80% of cases) when the aganglionic segment does not extend beyond the upper sigmoid; the long-segment form (L-HSCR, approximately 15% of the cases) when aganglionosis extends proximal to the sigmoid, and total colonic aganglionosis (TCA, approximately 5% of the cases) when the absence of enteric neurons affects the whole bowel (Garver et al., 1985; Chakravarti and Lyonnet, 2001; reviewed by Amiel et al., 2008). Total intestinal aganglionosis (TIA) has also been reported in a small number of cases ($<1\%$). It is considered the most severe form of HSCR as aganglionosis extends to the small bowel, and is often lethal (Ruttenstock and Puri, 2009).

Genetics of HSCR

The genetics of HSCR resembles other complex diseases. About 20% of the cases are familial and the rest are sporadic (reviewed by Amiel et al., 2008). Furthermore, HSCR is associated with other congenital malformations in $\sim$30% of cases, but a syndrome diagnosis can only be established in a minority of these cases (reviewed by Brooks et al., 2005b; Moore, 2006; reviewed by Amiel et al., 2008). The familial and syndromic cases show a Mendelian mode of inheritance (both dominant and recessive with reduced penetrance), whereas the sporadic cases are believed to have a complex, non-Mendelian mode of inheritance (Badner et al., 1990). Since around 1990, most attention has been paid to the rare familial and syndromic cases: linkage analysis was performed in large multi-generational HSCR families, sib-pair analysis in smaller nuclear families, homozygosity mapping in consanguineous HSCR families, and genetically isolated populations were used to search for inherited ancestral alleles or shared haplotypes putatively carrying mutations (for a review see Brooks et al., 2005b).

Recently, most studies have focused on association studies using both a case-control design and a trio design, and almost all screened isolated HSCR cases (Borrego et al., 2003b; Fitze et al., 2003; Sancandi et al., 2003; Burzynski et al., 2004; Pelet et al., 2005). The association studies focused mainly on the major HSCR gene, *RET*. Haplotype sharing was found for a region of approximately 27 kb, including 4 kb of the 5′UTR, exon 1, intron 1 and exon 2 of *RET*. It is believed that one or more SNPs within intron 1 are involved in disease development (Emison et al., 2005; Garcia-Barcelo et al., 2005; Griseri et al., 2005; Burzynski et al., 2005; Emison et al., 2010; Sribudiani et al., 2011).

In an attempt to find additional *loci* that contribute to the development of HSCR, a GWAS was recently performed in Chinese cases with sporadic HSCR (Garcia-Barcelo et al., 2009). As expected, a strong association to *RET* was found. However, two additional SNPs located in intron 1 of Neuregulin-1 gene (*NRG1*) were also found to be strongly associated, pointing to *NRG1* as a plausible candidate gene. Its involvement was corroborated by the identification of coding sequence mutations in *NRG1* (Garcia-Barcelo et al., 2009, Luzon-Toro et al., 2012; Tang et al., 2012a). By performing copy number variation analysis on this data, another gene was found to be associated with HSCR, namely *NRG3*, a paralog of *NRG1*. Further validation on Chinese HSCR patients identified nine deletions and two *de novo* duplications in *NRG3*, suggesting a role of this gene in HSCR (Tang et al., 2012b).

To date, 15 genes have been implicated in HSCR development, with *RET* being the most important one (see Table 1). However, only approximately 30% of all HSCR cases have mutations in these genes, suggesting the involvement of other genes in the development of this disease.

Genetics of complex diseases: Finding common disease associated variants

To find variants in complex diseases different consortia have collected over 100,000 samples in anticipation of finding new *loci* (see Deloukas et al., 2013 as an example for coronary artery disease). In order to increase statistical power and thus identify more disease associated markers, GWAS meta-analysis studies have been performed (reviewed by Gögele et al., 2012). However, study enlargements are not without risk. To achieve a very large case cohort, consortia members need to combine their separate cohorts, which might have been diagnosed or ascertained in different ways. Moreover, small genetic variations within or between cohorts can have strong effects on the outcome and may hamper replicating the results in other cohorts (Ioannidis et al., 2006). Expanding the