Genomics Proteomics Bioinformatics 13 (2015) 242–249

HOSTED BY

ELSEVIER

Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com

REVIEW

Applications of Next-generation Sequencing in Systemic Autoimmune Diseases

Yiyangzi Ma [1,a], Na Shi [1,b], Mengtao Li [2,c], Fei Chen [3,d], Haitao Niu [1,*,e]

[1] *Institute of Laboratory Animal Sciences, Chinese Academy of Medical Sciences and Comparative Medicine Center, Peking Union Medical Collage, Beijing 100021, China*
[2] *Department of Rheumatology and Clinical Immunology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical Collage, Beijing 100730, China*
[3] *CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China*

Received 19 April 2015; revised 4 September 2015; accepted 14 September 2015
Available online 15 October 2015

Handled by Quan-Zhen Li

KEYWORDS

Next-generation sequencing;
Intestinal microbiome;
Susceptibility genes;
Systemic autoimmune diseases

Abstract **Systemic autoimmune diseases** are a group of heterogeneous disorders caused by both genetic and environmental factors. Although numerous causal genes have been identified by genome-wide association studies (GWAS), these **susceptibility genes** are correlated to a relatively low disease risk, indicating that environmental factors also play an important role in the pathogenesis of disease. The **intestinal microbiome**, as the main symbiotic ecosystem between the host and host-associated microorganisms, has been demonstrated to regulate the development of the body's immune system and is likely related to genetic mutations in **systemic autoimmune diseases**. **Next-generation sequencing** (NGS) technology, with high-throughput capacity and accuracy, provides a powerful tool to discover genomic mutations, abnormal transcription and **intestinal microbiome** identification for autoimmune diseases. In this review, we briefly outlined the applications of NGS in **systemic autoimmune diseases**. This review may provide a reference for future studies in the pathogenesis of **systemic autoimmune diseases**.

Introduction

Since the inception of cyclic array-based next-generation sequencing (NGS) in 2005 [1], application of this high-throughput technology has shown exponential increase in related biomedical studies. NGS can be applied to sequence analysis on any part of the genome and the resulting transcriptome, including the whole genome, exons, and other interesting regions, and accordingly can be roughly classified as whole-genome sequencing (WGS), whole-exome sequencing

* Corresponding author.
E-mail: htniu1@yahoo.com (Niu H).
[a] ORCID: 0000-0002-9681-9162.
[b] ORCID: 0000-0003-1021-6940.
[c] ORCID: 0000-0003-4252-2889.
[d] ORCID: 0000-0003-3895-6318.
[e] ORCID: 0000-0002-8442-7444.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

http://dx.doi.org/10.1016/j.gpb.2015.09.004

(WES), RNA sequencing (RNA-seq), and DNA methylation sequencing [2–4].

Systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), multiple sclerosis (MS), ankylosing spondylitis (AS), and Sjögren's syndrome (SS) are typical systemic autoimmune diseases, which affect multiple organs and exhibit inherited susceptibility. Multiple causal genes have been identified that influence the development of autoimmune disorders by genome-wide association studies (GWAS) [5]. However, each of such genes is generally associated with only a relatively low risk of autoimmune disease occurrence, indicating that presence of the identified susceptibility genes is not a definitive pre-requisite for the disease development [6,7]. According to the "hygiene hypothesis", environmental pressure affects genetic alleles, rendering the body's immune system to adapt to the environmental impact, including the presence of microorganisms [8]. Therefore, the simultaneous presence of susceptibility genes and the gut microbiome is likely to coordinate synergistically to promote the systemic autoimmune disease progression.

The human intestinal contains a vast and diverse microbial ecosystem, consisting of 10^{14}–10^{15} microorganisms, colonizing the human intestinal tract shortly after birth, and remaining there throughout an individual's life. Both the quantities and the species composition of intestinal microbiota are closely related to human health [9]. Each person possesses millions of microbial genes, which are around 100-fold greater than the number of human genes [10]. This microbial gene pool comprises genes from hundreds of microbial species. The majority of these bacteria fall into four phyla: Actinobacteria, Firmicutes, Proteobacteria, and Bacteroidetes [10,11].

The value of NGS has been demonstrated in identifying susceptibility genes associated with systemic autoimmune diseases [11–13]. Its efficacy has also been demonstrated in the characterization of intestinal microbiotas [12–14]. Our goal is to provide an overview of the current applications of NGS to better understand the pathogenesis of systemic autoimmune diseases.

NGS technology is widely applied in biomedical research

With the development of NGS technology, WGS, WES, and RNA-seq are widely used to study the genetic mutations and gene expression (**Figure 1**). The most popular NGS method is WES. WES facilitates the capture of all coding exons in the genome. The term "exome" refers to all the exons in the genome, which covers approximately 1% of the genome in human. Nonetheless, approximately 85% of disease-related mutations are found in the exome [15,16]. WES is an efficient way to detect novel disease-causing genes, such as *MALT1* and

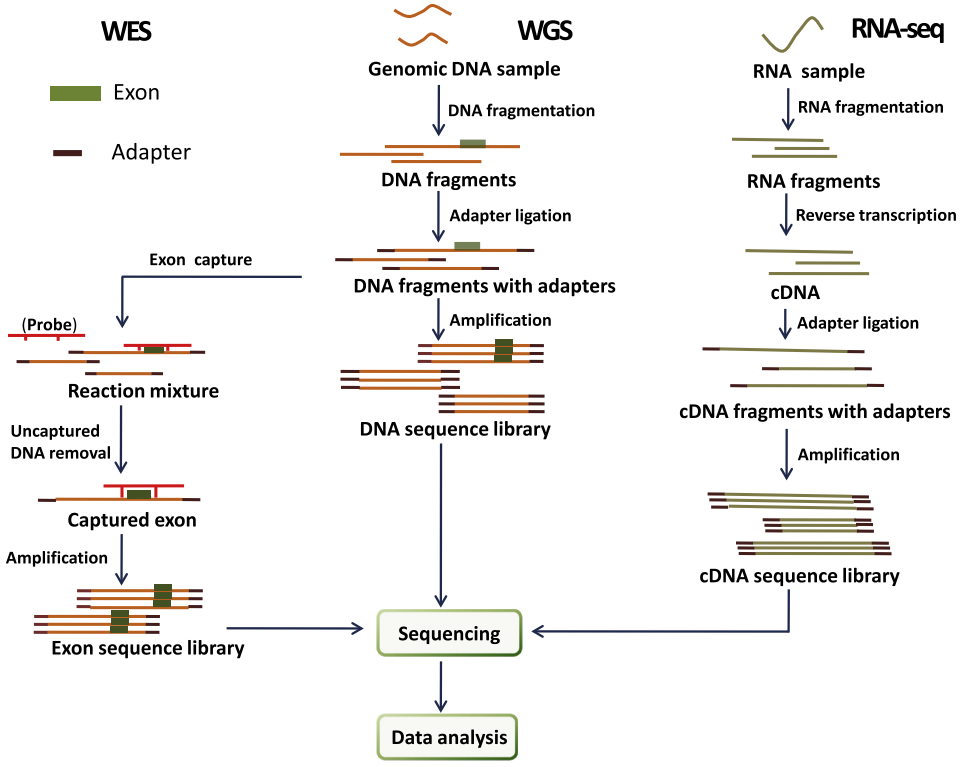

Figure 1 Flow chart of common next generation sequencing approaches

WES: Workflow for WES. Genomic DNA samples are broken up into short-length fragments. After adaptor ligation, exonic DNA fragments are captured with exon-specific probes and amplified by PCR to prepare the exon sequence library. **WGS:** Workflow for WGS. Genomic DNA samples are broken up into short-length fragments. After adaptor ligation, the DNA fragments with adaptor are amplified by PCR to prepare the DNA sequence library. **RNA-seq:** Workflow for RNA-seq. RNA samples are randomly broken up into short-length fragments. After reverse transcription, cDNA fragments are ligated to adaptors and amplified by PCR to prepare the cDNA sequence library. After the sequence library is set up for each approach, sequencing is then performed on designated sequencers followed by computational analysis. **WES**, whole-exome sequencing; **WGS**, whole-genome sequencing.