Methods xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Methods

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

Using DNase Hi-C techniques to map global and local three-dimensional genome architecture at high resolution

Wenxiu Ma [a,1], Ferhat Ay [a,2], Choli Lee [a], Gunhan Gulsoy [a], Xinxian Deng [b], Savannah Cook [c,d], Jennifer Hesson [c,d], Christopher Cavanaugh [c,d], Carol B. Ware [c,d], Anton Krumm [e], Jay Shendure [a,g], C. Anthony Blau [c,f], Christine M. Disteche [b], William S. Noble [a,*], ZhiJun Duan [c,f,*]

[a] Department of Genome Sciences, University of Washington, USA
[b] Department of Pathology, University of Washington, USA
[c] Institute for Stem Cell and Regenerative Medicine, University of Washington, USA
[d] Department of Comparative Medicine, University of Washington, USA
[e] Department of Radiation Oncology, University of Washington, USA
[f] Division of Hematology, Department of Medicine, University of Washington, USA
[g] Howard Hughes Medical Institute, Seattle, WA 98195-8056, USA

ARTICLE INFO

Article history:
Received 1 September 2017
Received in revised form 14 December 2017
Accepted 25 January 2018
Available online xxxx

Keywords:
Chromatin
Chromosome
Chromosome conformation capture (3C)
Hi-C
DNase Hi-C
Three-dimensional (3D) genome architecture

ABSTRACT

The folding and three-dimensional (3D) organization of chromatin in the nucleus critically impacts genome function. The past decade has witnessed rapid advances in genomic tools for delineating 3D genome architecture. Among them, chromosome conformation capture (3C)-based methods such as Hi-C are the most widely used techniques for mapping chromatin interactions. However, traditional Hi-C protocols rely on restriction enzymes (REs) to fragment chromatin and are therefore limited in resolution. We recently developed DNase Hi-C for mapping 3D genome organization, which uses DNase I for chromatin fragmentation. DNase Hi-C overcomes RE-related limitations associated with traditional Hi-C methods, leading to improved methodological resolution. Furthermore, combining this method with DNA capture technology provides a high-throughput approach (targeted DNase Hi-C) that allows for mapping fine-scale chromatin architecture at exceptionally high resolution. Hence, targeted DNase Hi-C will be valuable for delineating the physical landscapes of cis-regulatory networks that control gene expression and for characterizing phenotype-associated chromatin 3D signatures. Here, we provide a detailed description of method design and step-by-step working protocols for these two methods.

1. Introduction

The spatial organization of the genome has critical impacts on the various DNA-templated nuclear processes (e.g., transcription, DNA replication and repair) as well as on human development and disease [1–4]. Over the past decade, advances in technologies, especially the development of chromosome conformation capture (3C)-based high-throughput techniques (e.g., Hi-C), have yielded remarkable new insights into the principles of the three-dimensional (3D) genome organization [1,2,4–15]. It is increasingly clear that eukaryotic genomes are hierarchically organized in the nuclei [16–18]. Several conformational features corresponding to the distinct levels of genome organization have been identified, including whole chromosome territories [19], large-scale active and repressed compartments (A/B compartments) [20], specific domains, for example, topologically associated domains (TADs) [21,22], lamin associated domains (LADs) [23,24] and nucleolus associated domains (NADs) [25,26], and ultimately, individual looping interactions between specific loci [27]. The detection of these chromatin conformation signatures associated with different scales of 3D genome organization require different levels of resolution. For example, chromosome territories and A/B compartments can be identified at the megabase scale, whereas TADs, LADs and NADs are at the sub-megabase scale. However, mapping of other fine-scale chromatin signatures, such as chromatin loops between cis-regulatory elements, requires much higher resolution [27].

The regulation of eukaryotic gene expression depends on transcription factors and cofactors acting on functional DNA elements

* Corresponding authors at: Institute for Stem Cell and Regenerative Medicine, University of Washington, USA (Z. Duan).
E-mail addresses: william-noble@uw.edu (W.S. Noble), zjduan@uw.edu (Z. Duan).
[1] Present address: Department of Statistics, University of California, Riverside, USA.
[2] Present address: La Jolla Institute for Allergy and Immunology, USA.

https://doi.org/10.1016/j.ymeth.2018.01.014
1046-2023/

such as promoters, enhancers, silencers and insulators, which are situated in the 3D nuclear space [1,28]. Precisely mapping the spatial organization of cis-regulatory elements in the nucleus will lead to new insights into the principles underlying long-range gene regulation and how regulatory elements coordinate to achieve temporal and cell type-specific transcriptional regulation. Recent advances in genomic technologies have led to the systematic identification and annotation of functional DNA elements in mammalian genomes, including the human genome [29,30]. The human genome contains many thousands of cis-regulatory elements (e.g., enhancers, silencers and insulators) that, in any given human cell, are located both proximal to and far away from the genes they control [31–34]. Although the precise mechanisms for gene regulation by distant-acting cis-regulatory elements remain largely elusive, it has been shown that chromatin looping-mediated physical contacts between gene promoters and other cis-regulatory elements are essential for transcriptional activities of the human genome [31–34].

ChIA-PET [35] and 3C derivatives such as 4C [36,37], 5C [38], Hi-C [20], Capture-C [39], Capture Hi-C [40] and HiChIP/PLAC-seq [41,42] can be used to map fine-scale chromatin architecture such as the chromatin interaction landscapes of cis-regulatory elements. However, since these 3C-based methods all rely on restriction enzyme (RE)-digestion to fragment chromatin, the ultimate resolution of fine-scale chromatin architecture mapped by these 3C-based methods is limited by the local distribution of RE sites, which is uneven across the genome (Fig. 1). Ultimately, a chromatin interaction identified by RE digestion-based 3C methods can only be described as the interaction between corresponding RE fragment pairs, and thus may not necessarily refer to the precise genomic location (up to single base pair resolution) of cis-regulatory elements [43], where the physical contact occurs.

To overcome the resolution limitation of RE-based 3C methods, and toward the development of a single-base-resolution method for mapping chromatin interactions in the future, we recently developed DNase Hi-C, a method for comprehensively mapping chromatin interactions on a whole-genome scale in living cells [44] (Fig. 2). Although, like all the traditional 3C-based methods, the core concept of DNase Hi-C is also based on proximity ligation, DNase Hi-C employs DNase I instead of an RE to fragment the crosslinked chromatin (Fig. 2). Importantly, the sequence-independent digestion of chromatin by DNase I enables DNase Hi-C to overcome the resolution limitation associated with RE-based 3C methods. Subsequently, we have combined DNase Hi-C with DNA capture technologies to develop a method for fine-mapping chromatin architecture of specific genomic regions of interest at unprecedented resolution in a massively parallel fashion, termed "targeted" DNase Hi-C [44] (Fig. 2). We have applied DNase Hi-C to a diversity of biological systems, including yeast cells [45], mouse cell lines and tissues [46], and human cell lines

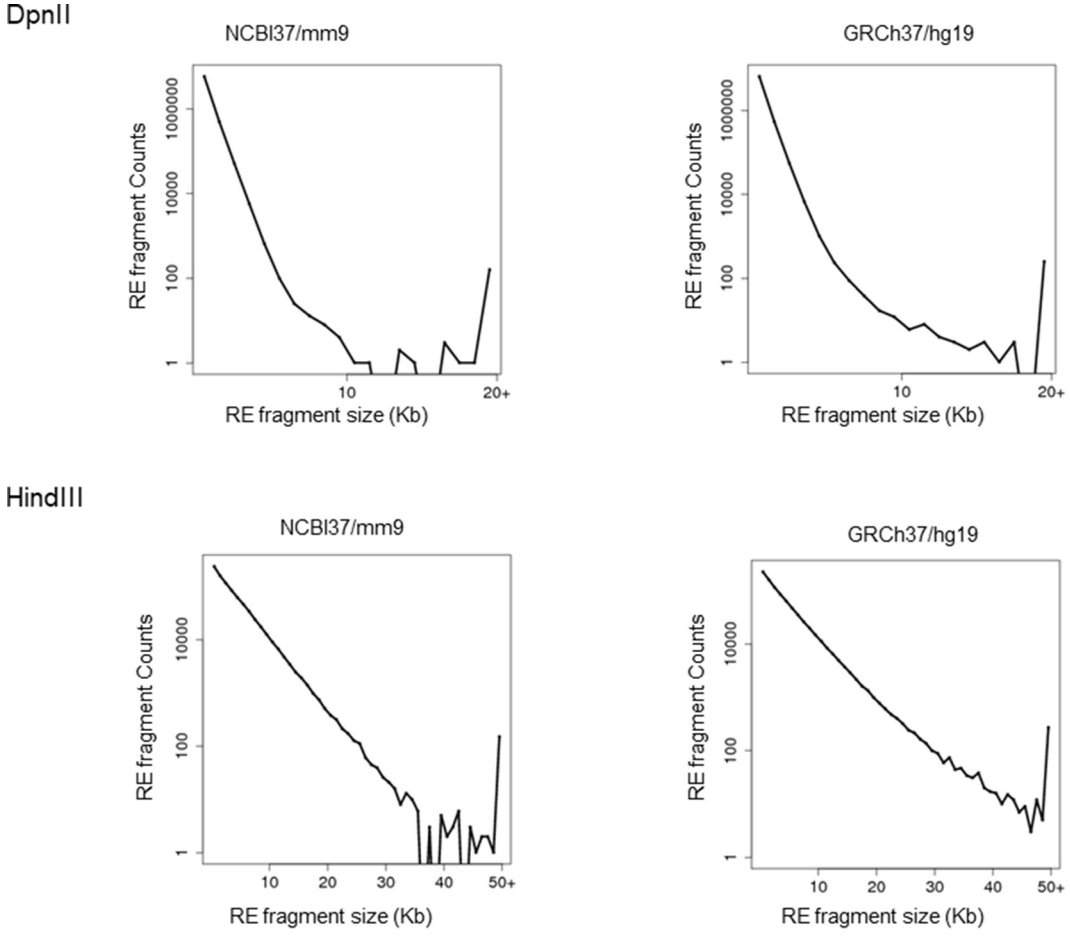

Fig. 1. Size distribution of the projected restriction fragments in mouse and human genomes after the *in silico* digestion with DpnII and HindIII, two REs frequently used in Hi-C assays. Note that for the 4 bp-cutter DpnII, about 8.73% (559,888 of 6,415,145) of fragments (RE-fragment counts) generated in the mouse genome (NCBI/mm9) and 8.58% (611,434 of 7,127,561) in the human genome (GRCH37/hg19) are larger than 1 kb. Importantly, the DpnII fragments with a size greater than 1 kb cover about 31.25% of the mouse genome and 33.28% of the human genome, respectively. For the 6 bp-cutter HindIII, about 20.38% (167,820 of 823,331) of fragments generated in the mouse genome and 22.74% (190,456 of 837,599) in the human genome are larger than 5 kb. Notably, the HindIII fragments with a size greater than 5 kb cover about 52.22% of the mouse genome and 58.33% of the human genome, respectively.

Please cite this article in press as: W. Ma et al., Methods (2018), https://doi.org/10.1016/j.ymeth.2018.01.014