

Definition of clinical gene tests

Duojiao Wu · Yunfeng Cheng  · Xiangdong Wang · CSGT group

Received: 10 January 2019 / Accepted: 28 January 2019 / Published online: 11 February 2019
© Springer Nature B.V. 2019

Abstract Clinical tests of gene sequence, structure, and function are to predict, diagnose, monitor, and prognose human disease-specific phenomes, characters severities, durations, stages, and responses to therapy. The concept and content of gene tests for clinical application mainly include chromosome/chromatins, DNA, and RNA. Structures and functions of chromosomes and chromatins vary among various durations, phases, and conditions, with the priority consideration in clinical gene tests. Sequences and functions of DNA and associated regulators are an important partial of clinical gene test. Another large group of RNA and RNA-associated factors also contribute to gene expression, regulation, and function. DNA/RNA sequencing is used to measure tumor mutation and heterogeneity, recategorize molecular phenomes and types of cancer, or guide and predict target-based therapies. The structure and function of genome dimensions and regulations as well as various factor involvement and contributions should be seriously considered in clinical gene tests, although there are a number of challenges to be overcome, e.g., method sensitivity, specificity, stability, analysis, and clinical significance. It is also critical to have the national and international standardization, guideline, and consortium of sample handling, experimental operation, quality

control, data analysis, and clinical interpretation, when clinical gene tests are developed and applied for clinical application. Thus, there is an urgent need to discover and validate those gene tests according to disease phenomes, subtypes, severity, duration, phase, progression, prognosis, and response to therapy.

Keywords Gene · Sequencing · Chromosome · DNA · RNA

There is a rapid increasing in understanding of gene structure and function and clinical application of gene tests for decades, with the development of biotechnologies. Clinical gene tests are here defined as clinical tests of gene sequence, structure, and function to predict, diagnose, monitor, and prognose human disease-specific phenomes, characters severities, durations, stages, and responses to therapy. Of a large number of gene measurements, the clinical gene test is often considered as gene sequencing to illustrate the potential and significance of particular mutations of genes in diseases for selection of therapeutic strategy. For example, alterations of lung cancer gene mutation and heterogeneity were suggested as an important evidence and criteria to screen and select characteristics of target drugs, and define molecular characterizations, severity, stage, duration, therapy, carcinogenesis, and prognosis of disease (Wang et al. 2018a). A large amount of gene mutations and heterogeneity are measured in tissue samples harvested from patients with cancer or mined and selected as cancer-specific targets in pan-cancer. Some of gene

D. Wu · Y. Cheng (✉) · X. Wang (✉)
Zhongshan Hospital Institute of Clinical Science, Zhongshan Hospital, Shanghai Institute of Clinical Bioinformatics Shanghai, Fudan University Medical School, Shanghai, China
e-mail: yfcheng@fudan.edu.cn
e-mail: Xiangdong.wang@clintransmed.org

mutations were found to clearly link to their functions and sensitivities, and clinically used as the indications of target-based therapies and responses to target-based drug efficacies. It was proposed that the precise self-validation system based on gene changes, mutations, and heterogeneity could be programmed and automated with a learning function and become an intelligent assistant for decision-making during clinical applications. Clinical exome sequencing was recently suggested as a diagnostic approach to define patients with genetic diseases, e.g., myeloproliferative neoplasms as an example of a heterogeneous disorder (Al-Dewik et al. 2018). Of tumor exomes, about 3% driver genes were related with locations and characters of cancer cells using multiple computational tools, about 40% driver gene mutations were detected using sequence- and structure-based analyses of which 60–85% might occur in driver genes, and about 60% had clinically actionable potentials (Bailey et al. 2018). However, those gene mutations and disorders presented with gene sequencing are the part of clinical gene tests, even though the gene sequencing is more and more frequently applied in clinical practice and has the impact on patient therapy (Marcell Szasz et al. 2018). Clinical gene tests are here proposed to cover three portions, i.e., chromosomes, deoxyribonucleic acid (DNA), and ribonucleic acid (RNA), during gene expression on the basis of element locations, functions, and constructions.

Structures and functions of chromosomes and chromatins varied among various durations, phases, and conditions should have the priority consideration in clinical gene tests. Genome sequencing, folding, dimensioning, and spatial architecting in chromosome conformation decide the accurate of gene expression and cell fate, the development of diseases, and biological and toxicological responses of cells (Wang 2018; Wang and Wang 2018; Zhang et al. 2018). In addition to sequencing of single DNA, tandem repeats encoding rRNA, tRNA, snRNA, and histone, intervals, or satellite, small satellite, and microsatellite DNA, a large number of regulators, promoters, enhancers, suppressors, and complexes contribute to the maintenance of gene function and to the development of disease (Fig. 1A). For example, a large number of factors, e.g., cohesion, CCTC-binding factor, cohesin-loading factor, or cohesin complex component, play regulatory roles in the maintenance of three-dimensional genome organization and function by their interactions, to control chromatin structure and function, gene expression, and transcriptional

signaling (Gao et al. 2018; Zhu and Wang 2018; Wang et al. 2018b; Liu et al. 2018). It is also important to define roles of histone modification (e.g., methylation and demethylation, acetylation and deacetylation, phosphorylation and dephosphorylation, adenylation and deadenylation, ubiquitination and de-ubiquitination, or ADP ribosylation and de-ADP ribosylation) in pathogenesis of diseases and categorize them in clinical gene tests. There are great and urgent needs to discover and develop disease-specific biomarkers of clinical phenomes, progression, or responses to therapy, on the basis of chromosome/chromatin dimensions, architectures, and regulatory factor controls of gene synthesis, expression, and function. Other elements should be not ignored in key processes of genomic imprinting, X-chromosome inactivation, repression of transposable elements, aging, and carcinogenesis, including nuclear topologically associating domains, pore complex, membrane, lamina, nucleosome, chromatin loop, boundary protein, or regions of chromatin contacts.

DNA sequences and functions of DNA-associated regulators are another important component of clinical gene test, responsible for transcription of inherent messages by dynamically altering chromatin dimensions and constructions as well as function. Gene editing and modification per se can alter the dynamic three-dimensional organization and folding of genome over time during which genes are expressed and cells are functioned (Wang et al. 2018c). With rapid development of measurable and analyzable methodologies (Marti-Renom et al. 2018; Fu et al. 2018; Qian et al. 2018), the closer and more definite correlation among altered DNA sequences, multi-dimensions of genome, disease-specific phenomes, or responses to drugs are furthermore clarified and programmed in an artificial intelligence single cell as an analytic model, although there are still challenges to be overcome (Zeng et al. 2018). Marti-Renom et al. (2018) published strategic guidelines and standardizations to analyze and validate experimental 4D nucleome data sets and models. DNA sequencing has been experiencing 40-year revolutions, challenges, improvements, and progressions and is now considered as a clinical diagnostic tool to measure gene mutation and tumor heterogeneity, or to guide therapeutic strategies of target drugs (Shendure et al. 2017; Wang et al. 2018a; Hou and Wang 2018).

Clinical gene test should include measurements of structure and function of bonds, regulators, zinc fingers, variable base pairs within the core sequence, and their

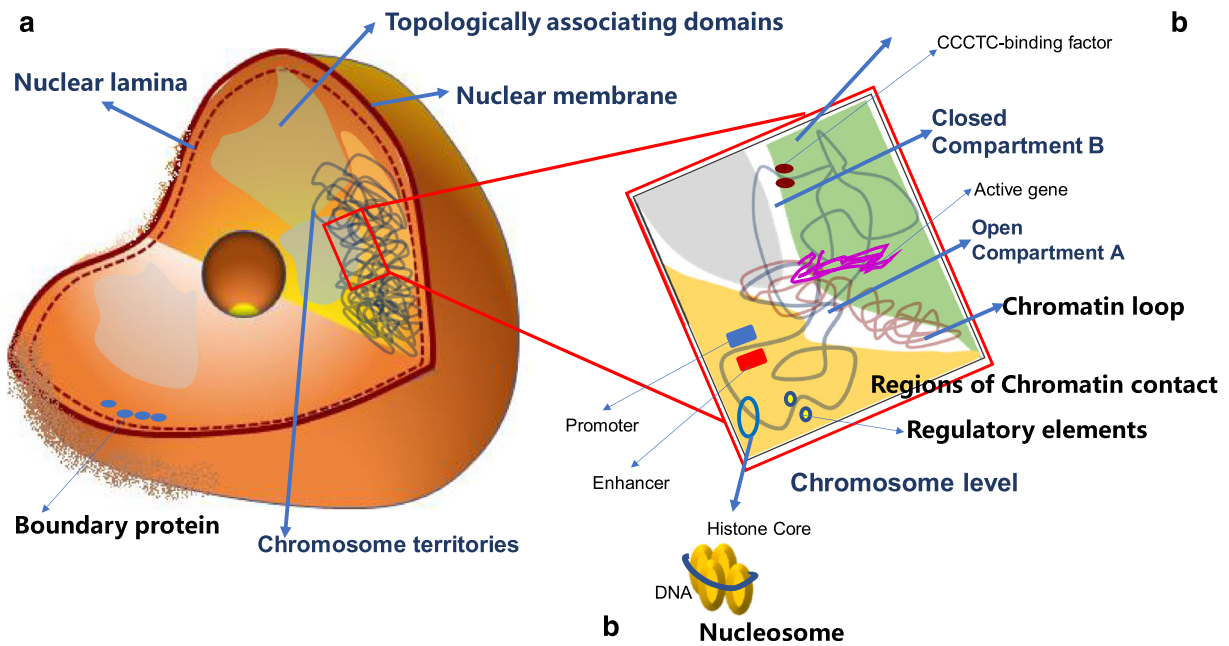


Fig. 1 There are a large number of regulators, promoters, enhancers, suppressors, and complexes, as well as factors, e.g., cohesion, CCCTC-binding factor, cohesin-loading factor, or cohesin complex component, to maintain dimension, organization and function of chromosome, chromatin, genome, and gene. A: In the nucleus, there are nuclear lamina, topologically associating

domains, nuclear membrane, boundary protein, and chromosome territories; B: In the chromosome level, there are open and closed compartments A and B, active gene, CCCTC-binding factor, chromatin loop, regions of chromatin contact, promoters, and enhancers; B: The nucleosome contains histone core and DNA

interactions, which maintain gene function and stability, although the sensitivity and specificity of those factors in diseases need to be furthermore defined. For example, DNA sequence variations, methylation patterns, and DNA-binding protein architectures and interactions, e.g., C2H2 ZF proteins, basic leucine-zipper, and basic-helix-loop-helix transcription factors together control gene expression (Hashimoto et al. 2017). CCCTC-binding factor interacts with the sequence variants consistently to compensate for deviations from the consensus sequence during epigenetic regulatory processes. Interactions among the large number of those factors, as listed in Table 1, are dependent upon the degree of sequence variations, methylations, locations, and specificity, leading to altered genome architecture from chromatin loops to the whole genome and development of diseases from mono-gene disorder to multi-gene-associated cancer (Szalaj and Plewczynski 2018; Li et al. 2018). Another example is that the function of telomeres and telomerase, e.g., dynamic properties, lengths, activities, protein, DNA, and RNA molecules, can be defined by single-molecule measurements (Parks and Stone 2017). Changes of telomeres and telomerases

can be disease-specific or associated biomarkers to predict, diagnose, and monitor a large number of diseases, including age-related diseases, Down syndrome birth, varicocele, subclinical atherosclerosis, ischemia, cancer, melanoma, and others.

The third category of clinical gene tests should be another large group of RNA and RNA-associated factors which contribute to and be critically responsible for gene expression, regulation, and function. Recently, single-cell RNA sequencing (scRNAseq) is an example to demonstrate cell heterogeneity, phenomes, transcriptomic profile and variation, cell-cell communication, or cell response to drug, as a powerful hammer to break the mystery of the cell and as a system to establish the intelligent model of single cell (Wang et al. 2018d; Wang and Wang 2017; Wang et al. 2017; Zeng et al. 2018). With rapid development and optimization of analytic protocols and methodologies with more sensitivity and accuracy (Svensson et al. 2017), new cell types and transcriptional stochasticity of cell development can be clarified and associated mechanisms by which gene expression and function are regulated by the interactions among intranuclear factors, e.g., RNA-

Table 1 Potential measures of DNA, RNA, and their related factors in clinical gene tests

DNA and related factors	RNA and related factors
Nucleotide sequences	Type
Hydrogen bonds	Ribosomal RNA, rRNA
Deoxyribose-phosphodiester bonds	Heterogeneous nuclear RNA, hnRNA
Base-stacking interactions	Small nucleolar RNA, snoRNAs
Conjugated π bonds	Small cytoplasmic RNA, scRNA
Non-canonical bases	
Enzymes	Enzymes
Restriction enzymes	RNA polymerases
Topoisomerases	Telomerase RNA component
Telomerase	Telomerases
Modifications	Ribozyme
Methyl (CH ₃)	RNA ligase
Hydroxymethyl (CH ₂ OH)	RNA phosphatase
Formyl (CHO)	
Carboxyl (COOH)	
Grooves	Function
DNA major and minor grooves	Transfer RNA, tRNA
Binding site	MicroRNAs, miRNA
Regulators	Small RNA
Transcription factors	Signal recognition particle
Promoters	Non-coding RNA, ncRNA
Enhancers	Antisense RNA, asRNA
Super-enhancers	RNA interference, RNAi
DNA temple	Small interfering RNA, siRNA
Complementary DNA	piwi-interacting RNA, piRNA
TATAAT Pribnow box	tRNA-derived small RNA, tsRNA
Base pairing	Small rDNA-derived RNA, srRNA
	Enhancer RNAs
	Regulatory RNAs
GC-content	Enzyme-dsRNA complex
DNA double-stranded (dsDNA)	RNA-induced silencing complex, RISC
DNA single-stranded DNA (ssDNA)	
Closed circular DNA	
DNA supercoiling	
Linker DNA	
Branched DNA	
Conformations	Shapes
A-DNA	Long non-coding RNA, lncRNA
B-DNA	Circular RNA, cirRNA
Z-DNA	Linear RNA
Telomeres	Double-strand RNA, dsRNA
G-quadruplex structure	
Chelation	
D-loop	

RNA, RNA-DNA, RNA-factors, and factors-factors. Gene expression and function are dependent upon the structure and functional interactions of these RNA gene sequences, RNAs and RNA-protein complexes, nucleotide-nucleotide interactions within RNA molecules, and nucleotide-amino acid interactions in RNA-protein complexes.

The RNAseq has been considered as a critical tool to define variations of RNA molecules and transcriptomic profiles of cells or tissues for decades, during which the methodologies of measurements and analyses as well as data deep mining are continuously improved. The direct amplification-free sequencing of full-length RNA molecules can be more practical and applicable in clinic. In addition to the understanding of transcriptomic differentiations between the health and disease, RNA diverse functions reflected by intermolecular RNA-RNA interactions should gain more attentions as the part of clinical gene tests. Engreitz et al. (2014) found that U1 small nuclear RNA as a component of the spliceosome interacts with Malat1 as a large ncRNA in nuclear speckles through distinct targeting mechanisms. During RNA-RNA interactions, a number of factors listed in Table 1 play the important role in hybridization, intermediation, motif activation, and relocation, to activate gene expression and function and target specific pre-mRNAs and genomic sites. RNAseq can demonstrate the severity of ribosomal DNA stability and genome adaptability, RNA methylation and post-transcriptional regulation, or cell-cell communication.

In conclusion, it is important and critical to define the concept and content of gene tests for clinical application, mainly including chromosome/chromatins, DNA, and RNA. DNA/RNA sequencing becomes a more applicable tool used to measure tumor mutation and heterogeneity, recategorize molecular phenomes and types of cancer, or guide and predict target-based therapies. The structure and function of genome dimensions and regulations as well as various factor involvement and contributions should be seriously considered in clinical gene tests, although there are a number of challenges to be overcome, e.g., method sensitivity, specificity, stability, analysis, and clinical significance. It is also critical to have the national and international standardization, guideline, and consortium of sample handling, experimental operation, quality control, data analysis, and clinical interpretation, when clinical gene tests are developed and applied for clinical application. Thus, there is an urgent need to discover and validate those

gene tests according to disease phenomes, subtypes, severity, duration, phase, progression, prognosis, and response to therapy.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Al-Dewik N, Ben-Omran T, Zayed H, Trujillano D, Kishore S, Rolfs A, et al. Clinical exome sequencing unravels new disease-causing mutations in the myeloproliferative neoplasms: a pilot study in patients from the state of Qatar. *Gene*. 2018;689:34–42. <https://doi.org/10.1016/j.gene.2018.12.009>.
- Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive characterization of cancer driver genes and mutations. *Cell*. 2018 Aug 9;174(4):1034–5. <https://doi.org/10.1016/j.cell.2018.07.034>.
- Engreitz JM, Sirokman K, McDonel P, Shishkin AA, Surka C, Russell P, et al. RNA-RNA interactions enable specific targeting of noncoding RNAs to nascent pre-mRNAs and chromatin sites. *Cell*. 2014;159(1):188–99. <https://doi.org/10.1016/j.cell.2014.08.018>.
- Fu S, Zhang L, Lv J, Zhu B, Wang W, Wang X (2018) Two main stream methods analysis and visual 3D genome architecture. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.07.011>.
- Gao D, Zhu B, Cao X, Zhang M, Wang X. (2018) Roles of NIPBL in maintenance of genome stability. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.08.005>.
- Hashimoto H, Wang D, Horton JR, Zhang X, Corces VG, Cheng X. Structural basis for the versatile and methylation-dependent binding of CTCF to DNA. *Mol Cell*. 2017;66(5):711–720.e3. <https://doi.org/10.1016/j.molcel.2017.05.004>.
- Hou J, Wang X. The polycomb group proteins functions in epithelial to mesenchymal transition in lung cancer. *Semin Cell Dev Biol* 2018. <https://doi.org/10.1016/j.semcdb.2018.07.010>.
- Li R, Liu Y, Hou Y, Gan J, Wu P, Li C. 3D genome and its disorganization in diseases. *Cell Biol Toxicol*. 2018;34(5):351–65. <https://doi.org/10.1007/s10565-018-9430-4>.
- Liu F, Wu D, Wang X. (2018) Roles of CTCF in conformation and functions of chromosome. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.07.021>.
- Marcell Szasz A, Malm J, Rezel M, Sugihara Y, Betancourt LH, Rivas D, Gyorffy B, Marko-Varga G. (2018) Challenging the heterogeneity of disease presentation in malignant melanoma-impact on patient treatment. *Cell Biol Toxicol*. <https://doi.org/10.1007/s10565-018-9446-9>.
- Marti-Renom MA, Almouzni G, Bickmore WA, Bystricky K, Cavalli G, Fraser P, et al. Challenges and guidelines toward 4D nucleome data and model standards. *Nat Genet*. 2018;50(10):1352–8. <https://doi.org/10.1038/s41588-018-0236-3>.
- Parks JW, Stone MD. Single-molecule studies of telomeres and telomerase. *Annu Rev Biophys*. 2017;46:357–77. <https://doi.org/10.1146/annurev-biophys-062215-011256>.
- Qian M, Cheng Y, Wang X. (2018) The methodology study of three-dimensional (3D) genome research. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.07.003>.
- Shendure J, Balasubramanian S, Church GM, Gilbert W, Rogers J, Schloss JA, et al. DNA sequencing at 40: past, present and future. *Nature*. 2017;550(7676):345–53. <https://doi.org/10.1038/nature24286>.
- Svensson V, Natarajan KN, Ly LH, Miragaia RJ, Labalette C, Macaulay IC, et al. Power analysis of single-cell RNA-sequencing experiments. *Nat Methods*. 2017;14(4):381–7. <https://doi.org/10.1038/nmeth.4220>.
- Szalaj P, Plewczynski D. Three-dimensional organization and dynamics of the genome. *Cell Biol Toxicol*. 2018;34(5):381–404. <https://doi.org/10.1007/s10565-018-9428-y>.
- Wang X. (2018) Why 3D genome alternations play important roles in diseases. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.07.024>.
- Wang W, Wang X. Single-cell CRISPR screening in drug resistance. *Cell Biol Toxicol*. 2017;33(3):207–10. <https://doi.org/10.1007/s10565-017-9396-7>.
- Wang DC, Wang X. Genome dimensions control biological and toxicological functions; myth or reality? *Cell Biol Toxicol*. 2018;34:333–6. <https://doi.org/10.1007/s10565-018-9440-2>.
- Wang W, Zhu B, Wang X. Dynamic phenotypes: illustrating a single-cell odyssey. *Cell Biol Toxicol*. 2017;33(5):423–7. <https://doi.org/10.1007/s10565-017-9400-2>.
- Wang DC, Wang W, Zhang L, Wang X. (2018a) A tour of 3D genome with a focus on CTCF. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.07.020>.
- Wang W, Zhang L, Wang X, Zeng Y (2018b) The advances in CRISPR technology and 3D genome. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.07.009>.
- Wang DC, Wang W, Zhu B, Wang X. Lung cancer heterogeneity and new strategies for drug therapy. *Annu Rev Pharmacol Toxicol*. 2018c;58:531–46. <https://doi.org/10.1146/annurev-pharmtox-010716-104523>.
- Wang W, Gao D, Wang X. Can single-cell RNA sequencing crack the mystery of cells? *Cell Biol Toxicol*. 2018d;34(1):1–6. <https://doi.org/10.1007/s10565-017-9404-y>.
- Zeng Y, Chen X, Gao H, Wang X. An artificial intelligent single cell is part of the cell dream world. *Cell Biol Toxicol*. 2018;34(4):247–9. <https://doi.org/10.1007/s10565-018-9433-1>.
- Zhang L, Song D, Zhu B, Wang X. (2018) The role of nuclear matrix protein HNRNPU in maintaining the architecture of 3D genome. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.07.006>.
- Zhu Z, Wang X. (2018) Roles of cohesin in chromosome architecture and gene expression. *Semin Cell Dev Biol*. <https://doi.org/10.1016/j.semcdb.2018.08.004>.