Scandinavian Journal of Gastroenterology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/igas20

Application of next-generation sequencing in the diagnosis of gastric cancer

Narges Moradi, Solmaz Ohadian Moghadam & Siamak Heidarzadeh

To cite this article: Narges Moradi, Solmaz Ohadian Moghadam & Siamak Heidarzadeh (2022) Application of next-generation sequencing in the diagnosis of gastric cancer, Scandinavian Journal of Gastroenterology, 57:7, 842-855, DOI: 10.1080/00365521.2022.2041717

To link to this article: https://doi.org/10.1080/00365521.2022.2041717

Published online: 16 Mar 2022.

SCANDINAVIAN JOURNAL OF GASTROENTEROLOGY
2022, VOL. 57, NO. 7, 842–855
https://doi.org/10.1080/00365521.2022.2041717

Taylor & Francis
Taylor & Francis Group

REVIEW

Application of next-generation sequencing in the diagnosis of gastric cancer

Narges Moradi[a], Solmaz Ohadian Moghadam[b] and Siamak Heidarzadeh[c]

[a]Department of Life Technologies, University of Turku, Turku, Finland; [b]Uro-Oncology Research Center, Tehran University of Medical Sciences, Tehran, Iran; [c]Department of Microbiology and Virology, School of Medicine, Zanjan University of Medical Sciences, Zanjan, Iran

ABSTRACT

Objectives: Gastric cancer (GC) is a disease with high mortality, poor prognosis and numerous risk factors. GC has an asymptomatic nature in early stages of the diseases, making timely diagnosis complicated using common conventional approaches, namely pathological examinations and imaging tests. Recently, molecular profiling of GC using next generation sequencing (NGS) has opened new doors to efficient prognostic, diagnostic, and therapeutic strategies. The current review aims to thoroughly discuss and compare the current NGS techniques and commercial platforms utilized for GC diagnosis and treatment, highlighting the most recent NGS-based GC studies. Furthermore, this review addresses the challenges of clinical implementation of NGS in GC.

Materials and methods: This review was conducted according to the eligible studies identified via search of Web of Science, PubMed, Scopus, Embase and the Cochrane Library. In the present study, data on gastric cancer patients and NGS methods used to diagnose the disease were reviewed.

Conclusion: Given the ever-rising advancements in NGS technologies, bioinformatics, healthcare guidelines and refined classifications, it is hoped that these technologies can actualize their advantages and optimize GC patients' experience.

ARTICLE HISTORY
Received 22 September 2021
Revised 6 February 2022
Accepted 8 February 2022

KEYWORDS
Gastric cancer; next-generation sequencing; whole exome sequencing; whole genome sequencing

Gastric cancer

Gastric cancer (GC) is a disease with high overall mortality and poor prognosis with an annual diagnosis rate of over one million cases, mainly involving Asians and South Americans [1]. It is the fourth and seventh most commonly occurring cancer in men and women, respectively [2]. Until the mid-1990s, GC was the most common cause of cancer deaths all over the world. With 8.3% of all cancer-related mortality being associated with GC, it is currently the third leading mortal malignancy, globally [3].

Presented in Figure 1, GC has numerous risk factors including genetics, lifestyle-associated risk factors including diet, alcohol and smoking, infectious agents such as *Helicobacter pylori* and Epstein-Barr virus (EBV) infection, gastric ulcer, gastroesophageal reflux disease (GERD), pernicious anemia, gastric surgery, radiation, chemical exposure, race, sex, blood group, and socioeconomic status [1,4,5].

Helicobacter pylori infection is a major cause of gastric neoplasm with a global GC and cancer burden of 89% and 5.5%, respectively [6]. Therefore, specific analysis of its virulence factors can improve prognosis and early diagnosis of the associated disorders. Recent studies have shown that positive *vac*A *s1m*1 and *cag*A strains are strongly associated with gastric tumorigenesis making gene screening a helpful predictor of *H. pylori* induced GC [7].

EBV infection can also contribute to a particular molecular subtype of GC which includes 10% of all GC molecular subtypes. Having distinctive genetic and epigenetic characteristics, EBV^+ GC is predominant in males and younger patients [8].

According to the Lauren classification developed in the 1960s and based on structural cellular components, stomach cancer has three major subtypes including intestinal type, diffuse type, and mixed type [9]. WHO has also introduced a specific classification for the disease consisting of four subtypes including papillary, tubular, signet ring, and mucinous GC [10]. GC has been divided into four subtypes by The Cancer Genome Atlas (TCGA) project: chromosomal instability (CIN), genomically stable (GS), microsatellite instability (MSI), and EBV positive [11].

Qualitative, staging and location diagnosis of GC majorly includes pathological examinations and imaging tests. Biochemistry tests, physical examination, endoscopy, laparoscopy, and intra-peritoneal fluid examination are also among the most common diagnostic strategies in GC [12].

The gold-standard diagnosis test for GC is histopathological examination, which is an essential requirement prior to taking therapeutic initiatives. Computed tomography (CT) is primarily applied in staging GC prior to treatment. Metastatic regions such as systemic, liver, and peritoneal metastases can be diagnosed using CT alternatives namely positron emission tomography (PET) scan, laparoscopic exploration, and magnetic resonance imaging (MRI). The postoperative histopathological diagnosis (pTNM) provides the medical staff with information on histological subtype of

CONTACT Siamak Heidarzadeh heidarzadehsiamak@gmail.com Department of Microbiology and Virology, School of Medicine, Zanjan University of Medical Sciences, Zanjan, Iran

the disease and complete evaluation of the tumor including location and the status of the lymph nodes. These key features are vital in developing personalized treatment strategies. Currently, human epidermal growth factor receptor 2 (HER2) expression in the cancerous tissue is the basis of GC molecular classification. HER2 assessment should be performed for all patients who are pathologically diagnosed with gastric adenocarcinoma [12].

GC has an asymptomatic nature in early stages of the diseases, making timely diagnosis very challenging when using common conventional approaches. Next Generation Sequencing (NGS) is a novel molecular method which has expanded the existing knowledge of molecular pathogenesis of GC and is considered as a promising diagnostic tool providing the scientists and clinicians with comprehensive genomic and genetic data about gastric tumorigenesis [13].

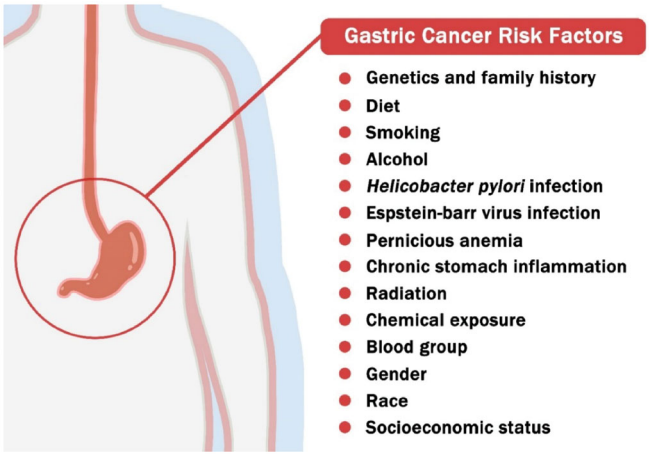

Figure 1. Gastric cancer risk factors.

Next-generation sequencing techniques

NGS is a familiar name heard from almost every department in medical sciences community. Researchers and clinicians have been using these developing comparatively new molecular technologies for numerous years resulting in obtaining brand new data about human normal genome and candidate disease-causing alterations. NGS plays an essential role in identifying health-related genomic alterations including the ones that lead to tumorigenesis and cancer including GC. Various techniques have been applied to vindicate the significant usefulness of these still-developing tools in the GC field. A brief comparison between these techniques is presented in Table 1 highlighting their main advantages and disadvantages.

Whole exome sequencing (WES)

WES or exome-sequencing is a widely applied sequencing laboratory technique used to sequence protein-coding regions in an individual's genome. Genetic variants that alter coding sequences are identified with this approach.

The clinical setting has increasingly benefited from WES. Exome sequencing reveals useful information about allelic variations which lead to anomalies and diseases [14–17]. Acquiring maximal genetic information and shifting from former expensive, time consuming and invasive strategies along with stressful medical crisis, laboratories have offered WES for diverse diagnostic purposes including cases of rare disorders and genomic variants in humans [15,17].

WES is often less costly than sequential genetic testing for genetic heterogeneity conditions since a large number of gene evaluations are necessitated [17,18]. Time is precious to concerned families considering additional babies or chances of recurrence. WES approaches have been more promising compared to traditional tests such as karyotype or microarray

Table 1. Comparison of advantages and disadvantages between NGS techniques.

Genomic technique	Advantages	Disadvantages
Whole exome sequencing (WES)	• Identifying variations across a broad diversity of applications • Covering coding regions comprehensively with a relatively greater depth • Being a more affordable substitute for whole genome sequencing • Being fast • Easier data analysis and management due to smaller amount of data	• Not covering the whole genome including gene expression regulatory regions • Higher rates of biases • Harder identification of structural variations and copy number alterations
Whole genome sequencing (WGS)	• Having a high variety of outputs • Identifying variations in both coding and non-coding areas. • Identifying variations that weren't previously detected • Having lower rates of GC bias • Having more consistent and uniform coverage	• Harder interpretation of data due to huge amount of identified genomic variants • Needing strong bioinformatics • High cost (gradually decreasing) • Higher turnaround time
Targeted sequencing	• Sequencing regions of interest with high depth at high coverage • Being cost-effective and fast • Lower turnaround time • Easier interpretation and management of data	• Identifying variations that weren't previously detected is not likely. • Not covering the whole genome
RNA sequencing (RNA-seq)	• Covering the whole transcriptome • Detecting expressions • Negligible background noise • Not needing prior knowledge of gene sequencing	• Harder interpretation of data • Time consuming • Needing strong bioinformatics • High cost (gradually decreasing)

in prenatal genetic diagnosis [17,19]. Therefore, the development of WES as a genomic test is appealing.

WES technique suggests diagnostic, therapeutic and monitoring methods for various malignancies including GC. Genomic alteration in multiple cancers can be characterized by enrolling sequencing approaches [20]. WES is applied in intratumor genetic heterogeneity (ITGH) detection and several studies on the same person's different metastatic lesions have used WES on biopsies [21].

As an advancing technology, germ line mutations identified by multi-region sequencing have been used as a therapeutic guide for cancer patients. Schneider et al. conducted a retrospective cohort analysis on 1028 metastatic cancer patients. They analyzed pathogenic/likely pathogenic variants in the patients using WES. The study revealed that 84% of cases carried one or multiple mutations. Pathogenic variants were most frequently observed in non-cancer predisposition genes. In 12.8% of cases nevertheless, pathogenic variant were identified in a cancer predisposition gene [22].

WES is primarily advantageous in detecting copy number variations (CNVs) as a major factor in cancer-leading genomic diversities. Identifying somatic CNVs plays a significant role in cancer prognosis and therapy advance [23].

Several trials suggest that immunotherapy response prediction benefits from tumor mutation load (TML) as a worthy biomarker and tumor biopsy WES is the gold standard method in determining the mutation load [24].

WES approaches are beneficial in cancer prognosis, detection and significantly in personalized treatment and are therefore, an interesting advancement in the biology field.

Whole genome sequencing (WGS)

NCI Dictionary of Genetics Terms describes WGS as a laboratory process applied to determine almost every single nucleotide in an individual's complete DNA coding or non-coding sequence. WGS has been applied as a research and subsequently, clinic tool for its limitless comprehensive coverage of genetic variations [25–29].

Numerous studies indicate that providing better diagnosis, gapless WGS has more uniform gene coverage and is more powerful and optimistic than WES and targeted sequencing when it comes to detection rate [15,23,29–32].

The capacity of WGS in identifying non-coding variations is not its only excellency. With no capturing needed, novel PCR-free WGS is more likely to cover exomic regions especially in a GC-rich region and become more affordable as time goes by [30,33].

Over 17 million people are expected to be diagnosed with different types of cancer each year (Cancer Research UK). This major health problem occurs after a gene variation-leading mutation which can be detected by sequencing techniques. Searching the human genome to find reasons behind this chronic illness and approaching prevention, diagnosis and treatment methods are of great importance. Despite the limited amount of information on non-coding regions in cancer genomes, when we study the genome selectively, structural variants (SV) such as translocation, inversions, deletions, duplications, insertions and virus integrations remain unexplored, while WGS provides us with a gapless uniform coverage [29,30]. Wide observation of somatic SVs is commonplace in different cancers including GC.

When WGS is applied along with analyzing mathematical data, clarification of the fundamental carcinogenesis and molecular sub-classification is possible in cancer which leads to development of new biomarkers and personalized medicine approaches [29].

Copy number alterations are commonplace as cancer genome landmarks and are able to activate oncogenes. According to a recent linked read WGS study, copy number alteration of FGFR2 has valid metastatic potential in GC [34].

WGS includes the analysis of non-coding mitochondrial genome which is known for the higher chances of alteration occurrence compared to nuclear DNA [35–37]. Mitochondrial dysfunction, heteroplasmy and subsequently cellular deregulation, can originate from the mentioned cancer-related alterations. A recent NGS-based study has compared the whole mt-Genome in GC-dealing patients and healthy individuals revealing that GC tumors presented variations in MT-DLOOP2, MT-DLOOP1, MT-RNR1, MT-ND2, MT-ND4, MT-ND5 and MT-CO1 regions [35]. Using WGS, Broad analysis of *H. pylori* genome is useful in diagnosing and treating the infection [38]. As a virus that have been linked to gastric adenocarcinoma, EBV genome can be detected by NGS platforms including WGS providing a more accurate alternative for traditional detection techniques such as *in Situ* Hybridization (ISH) [39].

Targeted sequencing

Targeted gene sequencing is a time-saving economical tool which is used to analyze specific mutations in specific genes or gene regions suspected for associations with the phenotype of interest. Targeted sequencing has made assessing multiple genes in samples and managing data easier due to smaller amount of information. Application of this NGS technique is of great importance in identifying genomic variations leading to GC development.

Diffuse GC has a strong genetic association with *CDH1* gene variations. *CDH1* encodes cell adhesion protein E-cadherin which is responsible for upholding epithelial tissues. Variations in *CDH1* can be monitored by targeted sequencing [40].

A targeted sequencing based study on 121 individuals with advanced gastric cancer, revealed the following genes were the most frequent (8–36%) in mutation rate: *PIK3CA*, *ARID1A LRP1B*, *CDH1*, *SYNE1*, *CSMD3* and *TP53*. Copy number variations ranging 0.8–20% also had an important role in GC-leading variations [41].

A multigenic study by Mafficini et al. disclosed the 60% mutation rate of *TP53* and 40% mutation rate of *PIK3CA* genes using targeted sequencing. This study aimed to evaluate intra-tumor heterogeneity which plays an important role in histopathological diagnosis [42].

This inexpensive personalized medicine technique improves GC diagnosis and provides an acceptable sequencing depth.

Gene expression profiling

Gene expression profiling is a molecular technique applied to measure expression of different genes by assessing the mRNA levels. mRNA expression can be evaluated in multiple experimental conditions to find out in which, a particular gene has been expressed the most. Theses profiles provide a functional landscape of the genome and therefore, can be used to monitor cellular dysfunctions such as proliferative problems leading to different types of cancer. Gene expression profiling could also evaluate cellular reactions to a specific therapeutic strategy.

Several gene expression profiling studies have facilitated treatment and diagnosis of GC by seeking the candidate responsible genes using different types of microarrays. These studies mainly have compared gene expression in GC and normal tissues to monitor probable down-regulations and up-regulations [43–45].

Using this tool provides novel diagnosis and therapeutic approaches since new GC molecular biomarkers are discovered.

RNA sequencing

RNA sequencing or RNA-Seq technology is a molecular NGS technique that is able to disclose whether or not a specific RNA is present is a biological sample. RNA-Seq can consistently detect cellular transcriptome alterations and is applied to measure the quantity of RNA at a significant time. Clinical and research application of RNA sequencing or cDNA sequencing in recent years has made several advancements in transcription characterization [46].

Molecular classification plays an important role in prognosis, diagnosis and treatment of GC since it is a heterogeneous disease and patients could benefit from personalized medicine approaches. To make the clinical application of this tool possible, it is necessary to provide a particular screening platform. Accessing transcriptome libraries, GC and its subtypes can be detected in early stages [47]. Diagnosis and treatment of GC has a significant association with the cancer stages. RNA-Seq is also a potential tool in detecting GC stages [48].

A recent study has approved of RNA-Seq as an approach in declaring downregulation of *IRF*-8 gene in $CD8^+$ T-cells in GC. According to these findings, targeted immunotherapy methods can be designed for GC patients [49].

Fundamentals of NGS platforms

DNA sequencing has had its ups and downs since its first introduction to the biology world from the human genome project and first generation sequencers to novel developing single molecule sequencing platforms. A summary of three DNA sequencing generations is shown in Figure 2. Soon enough, commercial companies and corporations were constituted presenting several platforms with a variety of characteristics. A comprehensive comparison between some of the most important platforms is presented in Table 2 (www.Illumina.com) [50–52]. Many disease-linked variations have been identified using these platforms and GC is no exception. The percentage of application of each of these technologies in the field of GC have relatively changed during the last few years with Illumina platforms gaining more popularity due to their wide range of application and affordability. Mildly elevating application of third-generation single molecule sequencing platforms such as HeliScope (Helicos Biosciences) and PacBio (Pacific BioSciences) is still in active development but less affordable and has considerable error rate despite enormous outcome. Mediocre characteristics of 454 and SOLiD platforms has led them to have a decreased usage percent during recent years. The mentioned data are visualized in Figure 3.

Roche/454 life sciences

Among all of the commercial next generation sequencing platforms, Roche 454 sequencing system was the first. Roche acquired this 454 sequencing from a high-throughput sequencing biotechnology company called 454 Life sciences in 2007. Even though Roche decided to shut down its 454 sequencing business because of noncompetitive technology, it took until mid-2016 to completely stop the production (www.fiercebiotech.com). In this period of time, a lot of studies benefited from this technology and numerous ones were conducted on cancer and oncology.

Roche 454 sequencing system functions in exclusive steps including library preparation, DNA amplification and pyrosequencing. While constructing the library, different DNA samples are broken into 300–800bp fragments. Specific primers are used to amplify denaturized DNA and clonal amplifications take place and eventually, library of single stranded DNA is constructed. The amplification system in Roche 454 fixes DNA strands in emulsion overwhelmed beads. This emulsion PCR approach is beneficiary due to its capacity for independent reactions and various beats are separated using emulsion characteristics. The system amplifies all of the fragments about one million times. The last step of pyrosequencing is based on identifying the emitted light of a chain reaction. Molecular mechanism of pyrosequencing is depicted in Figure 4. A high average read length of 400 bp and inaccuracy in assessing homopolymer length are relatively the significant advantage and disadvantage of Roche 454 sequencing system (www.creative-biogene.com). New molecular markers can lead us to develop personal treatments and faster and more accurate diagnosis of GC.

Numerous studies have directly or indirectly benefited from 454 sequencing technology, analyzing carcinogenesis of *H. pylori* including Cytotoxin-Associated Genes, CAG pathogenicity island genes and longitudinal studies [53–55].

There are also 454 sequencing-based studies indicative of effects of gastric microbiota other than *H. pylori* such as *pseudopneumoniae*, *parasanguinis*, and *Streptococcus oralis* on

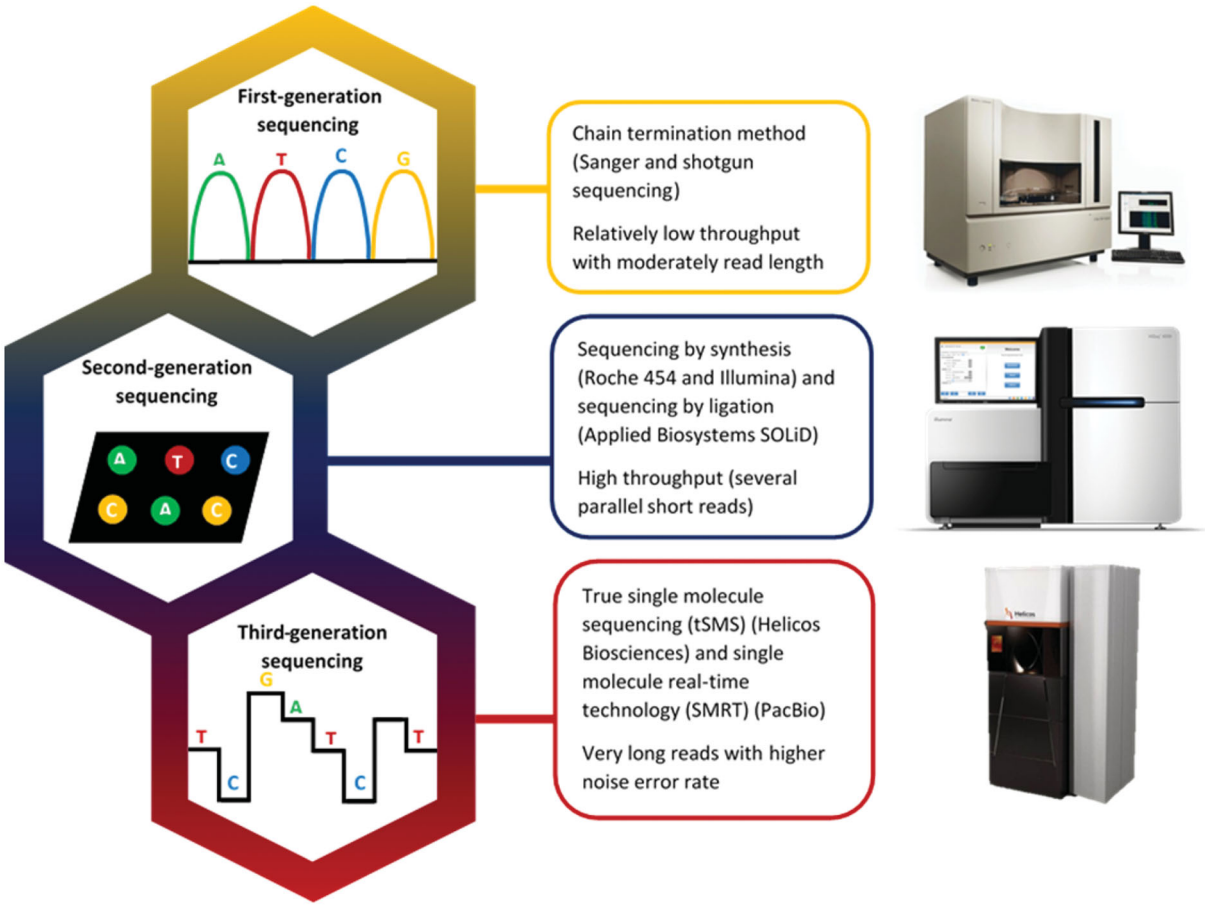

Figure 2. Sequencing generations.

Table 2. A comprehensive comparison between NGS platforms.

NGS platform	Roche/454 life sciences (GS FLX+)	Illumina/Solexa (HiSeq X ten)	Applied Biosystems/ SOLiD (v4)	Helicos Biosciences (HeliScope)
Website	www.lifescience.roche.com	www.illumina.com	www.appliedbiosystems.com	www.helicosbio.com
NGS Chemistry	Pyrosequencing	Sequencing by synthesis	Sequencing by ligation	Single molecule sequencing (SMS)
Library preparation method	Emulsion PCR	Solid phase amplification	Emulsion PCR for fragment	Not Available
Read length (bp)	700	2×150	50	32
Run time (d)	1	<3	6–7	≤ 1
Base/run (Gb)	0.7	1600–1800	120	20–30
Advantages	First commercially available NGS technology, relatively high average read length, relatively short run time	Relatively low cost, high throughput, suitable for wide range of applications and more frequently used, relatively low error rate	High accuracy, low error rate, multiple applications	Large number of reads, first commercially available SMS platform, amplification-free, relatively short run time
Disadvantages	Platform support ended in mid-2016, inaccuracy in assessing homopolymer length, expensive reagents	Relatively short reads, relatively long run time (recently improved)	Long run time (the longest of all), relatively short reads, relatively expensive	High possibility of noise errors, high cost (the most expensive)

individuals' health deficiencies including GC [56]. A study conducted in 2016 by Sung et al. compared the microbial compositions in gastric juice and mucosa with 454 pyrosequencing tool suggesting that gastric mucosa in more rich in bacteria so it should be considered for meaningful pyrosequencing [57].

In 2014, a study by Dong Soo Han et al. compared diverse microbial communities in intestinal metaplasia, chronic gastritis and GC using 454 high-throughput sequencing technology. They collected samples of gastric mucosal biopsy during endoscopy and after PCR and pyrosequencing, they came to conclusion that GC microbial communities are significantly different from microbial compositions in chronic gastritis and intestinal metaplasia [58].

Illumina/solexa

Illumina's dye sequencing was firstly developed in Cambridge University and subsequently Solexa Company which was acquired by Illumina later. Using reversible dye-terminators technology, Illumina-Solexa sequencing platforms

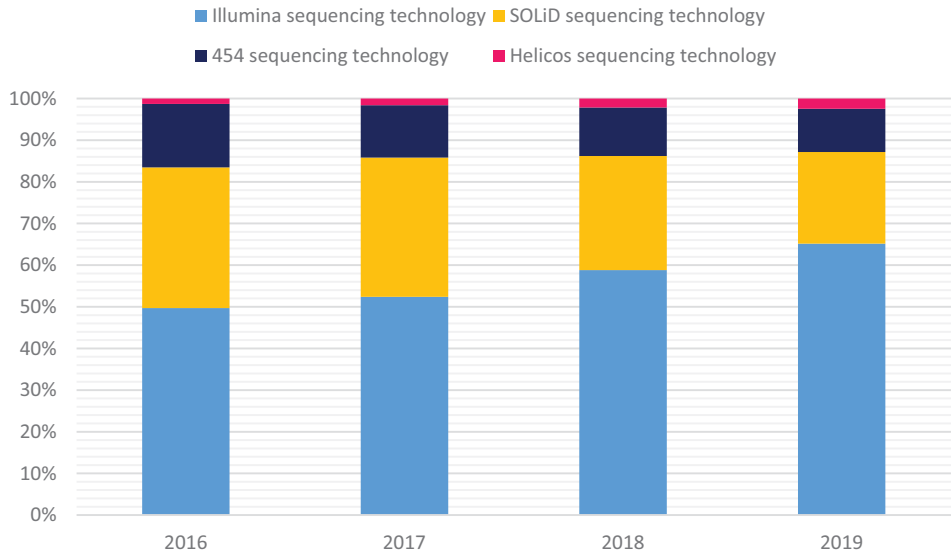

Figure 3. Application of 4 significant sequencing platforms in GC-related studies conducted from 2016 to 2019 (google scholar). Illumina has been the most frequently applied platforms gaining more popularity over time. Despite their great part, Roche and Applied Biosystems platforms have lost popularity. Helicos Biosciences platforms have an increasing application gradient however they are still developing.

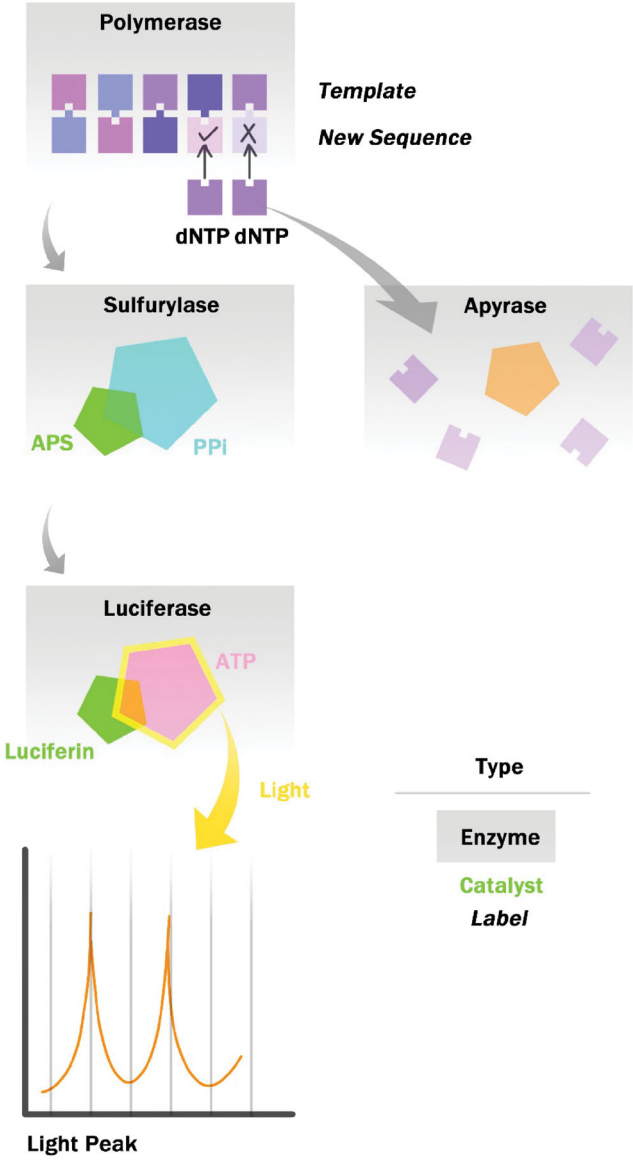

Figure 4. Molecular mechanism of pyrosequencing.

are applied to sequence many genomes rapidly in order to provide us with huge amounts of sequencing data. This sequencing-by-synthesis (SBS) data is frequently indicative of wellbeing or disease [59]. Illumina sequencing has a multi-step procedure including DNA purification, tagmentation, three steps of amplification, sequencing by synthesis and eventually data analysis. Figure 5 depicts Illumina's sequencing technique (www.Illumina.com).

This sequencing technique is advantageous over conventional sequencing tools including Sanger sequencing. Illumina dye sequencing is naturally automated and has quick outcome. Illumina sequencing only requires DNA polymerase while other sequencing techniques such as pyrosequencing require various expensive enzymes [60].

Brief and accurate prognosis, diagnosis and treatment of GC as a major health problem is profoundly associated with developing new molecular strategies including NGS tools such as Illumina platforms.

A recent study by Asakawa et al. has benefited from Illumina's sequencing tool to create a library to identify ATF3 as a candidate transcription factor (TF) that is meaningfully upregulated in gastric cancer cases with an EBV infection background [61].

In 2019, an Illumina-sequencing-based study by Ito et al. aimed to identify mutations that were causative in gastric malignancies aiming to provide more appropriate prognosis and therapeutic strategies for the disease. A total number of three mutations were observed in *KRAS*, *TP53* and *APC* genes in half of the patients after sequencing formalin-fixed paraffin-embedded (FFPE) specimens of four patients [62].

In 2017, Castaño-Rodríguez et al. evaluated microbiome activity by an Illumina MiSeq platform which leaded to observation of candidate microbial rRNA transcripts and prediction of carcinogenesis-associated changes. Their main findings included increased number of lactic acid producing bacteria, increased oral bacteria species and fatty acid and

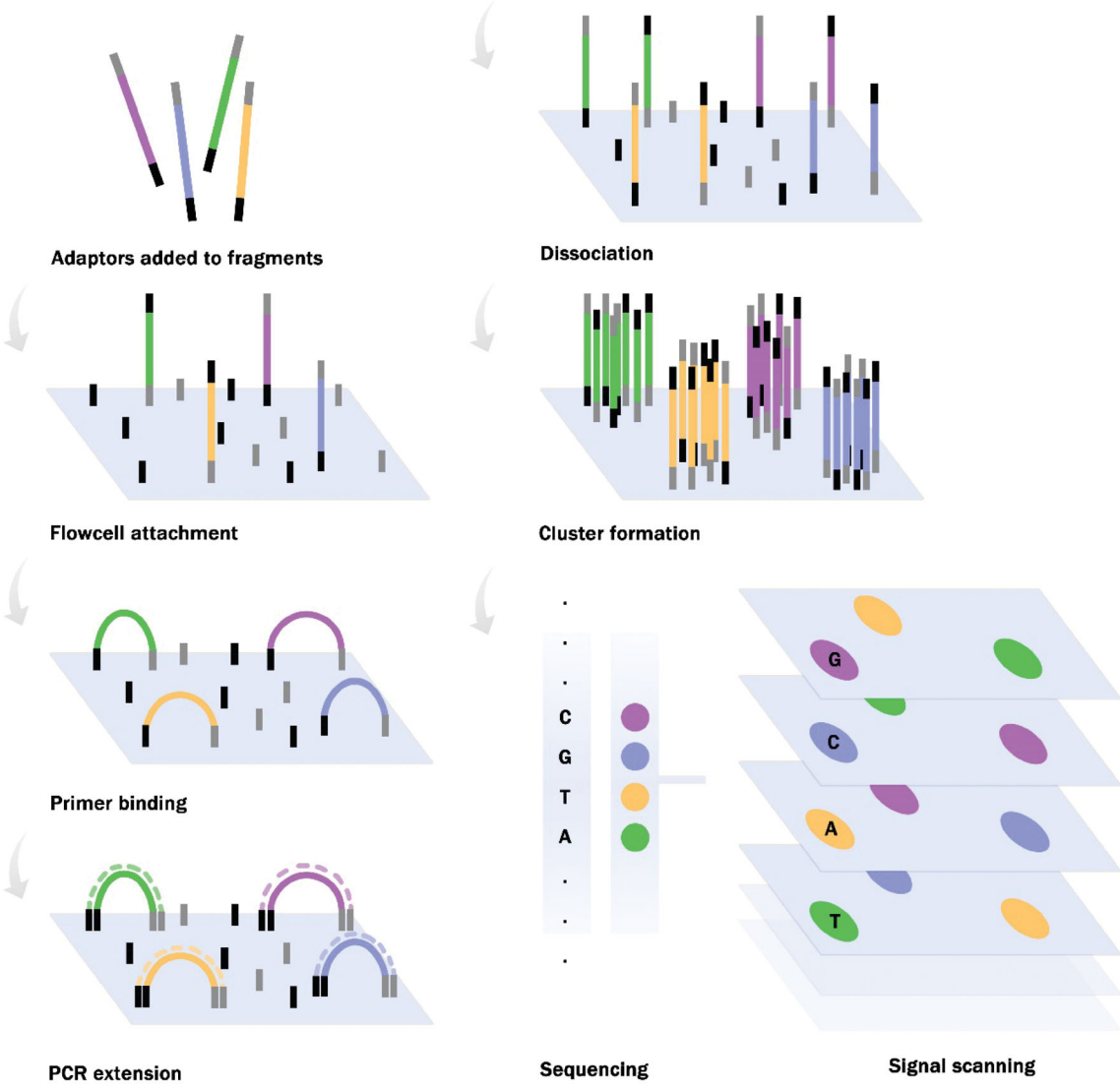

Figure 5. Molecular mechanism of Illumina sequencing.

carbohydrate pathways alterations as potential gastric carcinogenesis mechanisms [63].

An Illumina sequencing panel (RIP sequencing) was also of important use in a 2016 study by Qi et al. which aimed to identify significant RNAs that were associated with EZH2 in GC cells indicative of the role EZH2 plays in gastric tumorigenesis. They revealed that a specific RNA called MALAT1 binds EZH2, and leads to suppression of PCDH10 which is a tumor suppressor. This activity contributes to gastric cellular invasion and migration [64].

Applied biosystems/SOLiD

Acquired by Applied Biosystems, SOLiD (Sequencing by Oligonucleotide Ligation and Detection) is a 2-base-encoding NGS technology available since 2006. SOLiD is capable of reading 10^8–10^9 sequences at a time and has multiple applications including epigenetics and transcriptomics of cancerous tissues. This NGS tool is based on detecting different fluorophores each indicative of a specific set of two nucleotides.

The steps of DNA sample preparation in this technique are almost similar to those of 454 sequencing. The DNA sample is broken into smaller pieces at the first place and after the attachment of adaptors to the end of each piece, emulsion PCR takes place. In the sequencing phase, the sequencing primer binds to adaptor ends of each DNA fragment. Following that, a group of fluorescently labeled oligonucleotide octamers compete to bind to the primer and subsequently sequences of bases are detected according to fluorescent di-base codes [65].

In a 2017 study, Esser et al. applied a SOLiD ePCR Kit and SOLiD sequencing tool in order to obtain a comprehensive set of information about the genetic alteration background of GC ranging from single nucleotide variations to large structural variations [66].

A 2013 study by Nagarajan et al. suggested how WGS is of great importance in uncovering chromosomal and microsatellite instability using various NGS platforms including Applied Biosystems SOLiD platform which leaded to identification of *ACVR2A*, *RPL22*, *LMAN1* and *PAPPA* as mutated genes responsible for GC onset [67].

In 2010, Ribeiro-dos-Santos et al. conducted a study on expression of miRNAs in the normal stomach tissue due to its important role as a reference healthy profile in determining alterations that might lead to tumorigenesis using SOLiD ultra-deep sequencing [68].

Helicos biosciences and pacific biosciences single-molecule sequencing

Founded in 2003, Helicos Biosciences is a Cambridge-based biotechnology corporation that developed the first commercial single-molecule sequencing NGS platform in order to identify the DNA sequence, accurately. Helicos single molecule fluorescent sequencing contains the following steps: DNA fragmentation, tailing, blocking, loading samples and eventually sequencing. This sequencing tool is advantageous due to its parallelism, simpler DNA preparation procedure and data analysis even though noise errors are more likely [69].

In 2017, Alarcón et al. published a book about *H. pylori* molecular pathogenesis and signal transduction. *H. pylori* is one of the most robust GC risk factors. Alarcón et al. indicated that Heliscope (a novel NGS platform of Helicos BioSciences) can be of important use in investigating the cancer-related pathogenicity of *H. pylori* [70].

Commercialized in 2011, single-molecule real-time (SMRT) sequencing is a rather novel technique for sequencing single molecule DNA. Zero-mode waveguide (ZMW) is utilized in SMRT sequencing [71]. A single DNA template joins a ZMW-affixed DNA polymerase enzyme. The platform is able to observe DNA polymerase incorporating a single nucleotide of DNA due to ZMW's small observation volume. Four different fluorescent dyes are added to each of the DNA bases. Nucleotide incorporation creates fluorescent signals, which are sensitively detected utilizing a detector leading to sequencing [72]. Figure 6 illustrates how SMRT sequencing works.

Associated with gastric neoplasm, deregulated gene expression can be identified by accurate and time-saving SMRT sequencing in terms of diagnosis and targeted therapy. In this regard, Huang et al. applied Pacific Biosciences long-read isoform sequencing (Iso-Seq) to analyze GC transcriptome across 10 associated cell lines. The study brought about a more comprehensive transcriptome database with 66% of the transcripts being novel [73]. Watanabe et al. also benefited from SMRT technology in revealing genomic factors underlying *H. pylori* induced GC. They found 702 nucleotide variants (NVs) in 275 genes which were common in three or more patients with early gastric adenocarcinoma. These variants were not seen in cancer-free patients. The study suggested that the *hopL* variant acts as an *H. pylori* virulence factor leading to GC [74].

The impact of NGS on basic research

Genomic analysis

Genomic analysis is a molecular approach used to identify biological characteristics and measure or compare features of an individual's genome including sequencing, expressions and structural variations in DNA. Bioinformatics, deep sequencing with high throughput and microarray hybridization are the most important approaches to genomic analysis. The knowledge provided from one's genomic analysis should be adjusted with a medically meaningful genomic library for further clinical intentions like prevention, diagnosis and treatment.

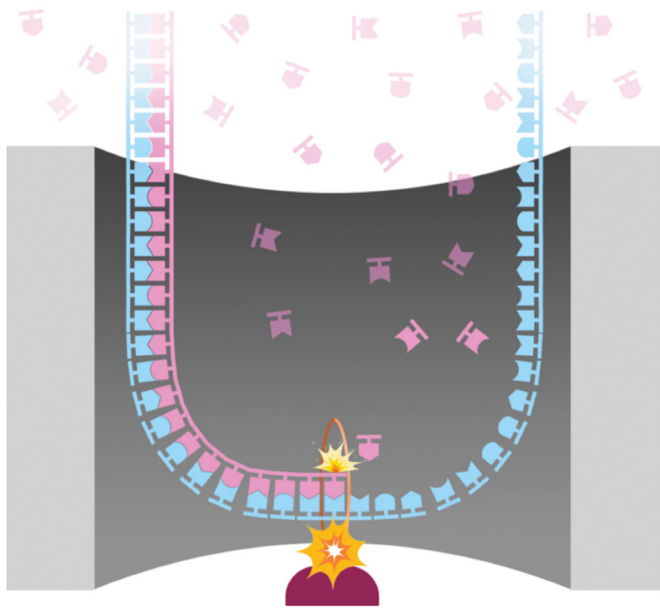

Figure 6. Molecular mechanism of Pacific Biosciences SMRT sequencing.

Genomic analysis has revolutionized cancer molecular biology by developing lots of cancer genomics programs such as The Cancer Genomic Atlas (TCGA; https://portal.gdc.cancer.gov). These programs have been helpfully applied in GC patients. Further studies and results on these subtypes are done due to the beforehand genomic analysis [75].

Many studies have revealed mutations responsible for GC using genomic analysis. Genomic analysis is a huge step forward in personalized medicine considering patients' conditions and the tumor stage [76].

Targeted genomic resequencing

Targeted resequencing is a fast and effective method that aims to analyze regions of genes of interest. Resequencing is done in order to identify genomic sequencing alterations comparing individuals' biopsies with the standard specimen. These highly sensitive and specific panels have remarkable coverage and can target from one to hundreds of genes or regions. Targeted resequencing is useful in assuring alterations identified with other NGS platforms and a potent affordable tool in screening [77].

Various NGS platforms provide researchers and clinicians with comprehensive genomic and molecular data. Targeted resequencing is especially advantageous where high-throughput sequencing is required. Targeted genomic resequencing can be of great use in many biological fields including different cancers.

CNVs have a significant association with disease-causing alterations including cancer. Targeted genomic resequencing has successfully been used in identifying cancer-related CNVs [78].

Deep targeted resequencing is able to assure us about our primary findings on oncogenes and plays an important role in validating the identification of candidate mutations in GC patients [79].

In 2011, Myllykangas et al. described oligonucleotide-selective sequencing (OS-Seq) as a potent approach to targeted genomic resequencing in which multifunctional oligonucleotide primers would play the role of sequencing substrate and a capture and they applied this approach on cancer genes [80].

In 2019, Ikari et al., used next generation resequencing to develop novel therapeutic targets and biomarkers for liver metastasis in GC. A total number of 412 cancer-associated genes in primary and metastatic tumors were sequenced using Thermo Fisher Scientific panels. TP53 mutations were significantly more frequent in patients in metastatic tumors (86.5%) compared to those without liver metastasis (40.5%). Therefore, p53 pathway can be of great importance in developing biomarkers and targeted treatments for metastases to the liver in GC [81].

Metagenomics

Metagenomics refers to the broad study of genetic material directly retrieved from a particular environment. Isolation-free metagenomics provides researchers a more accurate set of information about microbial communities. The advantage of direct examination on samples includes omitting the chances of missing any collected genomic data.

As a major health concern globally, cancer is a disease of high lethality. Combination of several effective factors such as genetics, lifestyle and environment can lead to cancer. Among all of the responsible factors, 16.1% of all cancers are evoked by micro-organisms. Metagenomics have extended the study of microbe-associated cancers by directly analyzing cancerous tissue specimens for microbial diversities using next generation sequencing platforms [82].

Metagenomics of cancerous tissue suggests new approaches in disease prognosis. It also improves cancer diagnosis field by introducing specific novel biomarkers. This method can also be useful when it comes to therapeutic strategies due to identification of potential microorganism and a better landscape of drug targets.

Although the role of microbiome is clear in occurrence of gastrointestinal diseases including gastric neoplasia, microbial dysbiosis in these conditions is not thoroughly studied. Gunathilake et al. evaluated the association between gastric microbial communities and GC. A total number of 288 controls and 268 GC patients were included in the study. After metagenomic analysis and sequencing, respective enrichments of *Streptococcus_NCVM* and *Prevotella melaninogenica* species were highly observed in cases and controls. Phylogenetic enrichment of class Bacilli was observed in cases and class Actinobacteria was associated with the controls. They came to conclusion that gastric microbial Dysbiosis leads to an increased risk of GC specifically in women [83].

In 2018, Hu et al., conducted a study on the role of gastric microbiota in superficial gastritis and GC in 11 patients using shotgun metagenomics sequencing. Their results presented novel changes in the gastric microbiome in patients with GC including depletion of 31 taxa, enrichment of 13 bacterial taxa and reduced species richness. They also reported on the depletion of *Sphingobium yanoikuyae*, a degrader of carcinogenic compounds. Studying on microbiome composition will lead to better prognosis and diagnosis of GC [84].

A 2019 study by Seol et al. assessed the differences of the gut microbiota in GC based on the existence or absence of *H. pylori* infection using 16SrRNA metagenomic sequencing of stool samples in patients with GC. The results were indicative of the dominance of four major phyla in the stool samples including, Actinobacteria, Proteobacteria, Bacteroidetes as well as Firmicutes as the most prevalent one. In general, they reported difference between *H. pylori* positive and negative patients and lower intestinal bacterial diversity in the patients infected with *H. pylori* [85].

Mapping of DNA-binding proteins and chromatin analysis

Proteins that possess DNA-binding domains and structurally and chemically interact with the DNA are known as DNA-binding proteins. These DNA-tendentious substances are a broad group of proteins including transcription factors (TFs) responsible for organizing transcription, nucleases, polymerases and chromosome-compacting histones. Particular binding of these proteins to given DNA sequences plays an important role in the function and expression of genome and human well-being so that mapping of DNA-binding proteins is necessary in modern and personalized medicine [86].

Present in eukaryotic cells nuclei, chromatin includes significant proteins with a majority of histones and nucleic acids that make up DNA. Any alterations in chromatin structure or chemistry might lead to various genomic disorders. Chromatin compaction can change due to epigenetic reasons causing gene expression swings however the primary sequence of DNA won't change [87]. Chromatin analysis is known as the study of chromatin architecture or function in order to identify probable disease-related alterations including tumorigenesis ones.

A recent study by Zhang and Yu has discussed the importance of replication protein A (RPA), a family of DNA-binding proteins, in GC prognosis concluding that the expression levels of RPA1, 2 and 3 are meaningfully elevated in GC tissues in comparison with normal tissues thus RPAs can be used as efficient prognostic factors and indicators of immunological diseases [88].

In 2019, Shi and Zhang studied the role of an important proliferation-related gene regulator DNA-binding protein, the specificity protein 1 (Sp1), on GC development using transcriptomic sequencing data. They found out about elevated levels of Sp1 expression and suggested it as a potential

Table 3. A highlight of NGS-based GC studies on GC conducted from 2016 to 2020.

Study	Year	Sequencing technique	NGS platform	Samples	Main findings
Saranathan et al. [38]	2020	Whole genome sequencing	Illumina MiSeq	54 *H. pylori* isolates recovered from 42 gastric biopsy tissue samples	clarithromycin and levofloxacin genomic resistance in *H. pylori*
Fu et al. [49]	2020	RNA sequencing	Illumina HiSeq 4000	2 pairs of GC tissue samples, 5 peripheral blood samples from 2 healthy individuals and 3 GC patients	IRF8 downregulation in CD8+ T cells
Rizzato et al. [53]	2020	Whole genome sequencing	Illumina HiSeq 2500	*H. pylori* isolates recovered from 74 gastric cancer, chronic gastritis and intestinal metaplasia patients	Non-synonymous mutations in *cagPAI*
Asakawa et al. [61]	2020	RNA sequencing	Illumina HiSeq 1500	SNU719 EBV^+, NCC24, MKN7 EBV^- and GES1 cells obtained from cell banks	ATF3 upregulation
Cavalcante et al. [35]	2019	Whole mitochondrial genome sequencing	Illumina MiSeq	20 tumoral and 20 non-tumoral (duodenal) and 50 cancer-free gastric FFPE tissue samples	Mutations in *MT-DLOOP2, MT-DLOOP1, MT-RNR1, MT-ND2, MT-ND4, MT-ND5 and MT-CO1*
Bustos-Carpinteyro et al. [40]	2019	Targeted sequencing	Roche 454/GS Junior	20 fresh (13 diffuse and 7 mixed) gastric biopsy samples	Methylation of *CDH1* promoter and deficiency of E-cadherin protein
Heo et al. [47]	2019	RNA sequencing	Thermo Fisher Scientific Ion Torrent Proton	50 FFPE GC tissue samples provided from Asian Cancer Research Group	Highly sensitive (0.9576) and specific (0.811) identification of GC subtype with a 10-gene model
Ito et al. [62]	2019	Targeted sequencing	Illumina MiSeq	FFPE fresh GC tissues from 4 patients	*TP53* mutation
Zhou et al. [96]	2019	Whole exome sequencing	Illumina HiSeq 4000	primary FFPE tumor tissues and matched peripheral blood samples from 32 GC patients	Mutations in *PIK3CA, FAT4, BRCA2, GNAQ, LRP1B, and PREX2*
Zhang et al. [97]	2018	Whole exome sequencing	Illumina X-Ten	Gastric adenocarcina, matched normal and lymph node metastases tissue samples from 5 GC patients	Revealing 37 responsible mutations, Copy number loss in *CDKN2A* and *CDKN2B*, CNVs in *EGFR* and *ERBB2*
Pan et al. [98]	2018	Targeted sequencing	Illumina HiSeq 3000	45 FFPE gastric adenocarcinoma samples	Mutations in *P53, MLL4, ERBB3, FBXW7, MLL3, MTOR, NOTCH1, PIK3CA, KRAS, ERBB4, EGFR*
Greer et al. [34]	2017	Whole genome sequencing	Illumina HiSeq 4000 Illumina HiSeq 2500	normal stomach, gastric primary cancer and two metastases from each ovary FFPE and frozen tissue samples	*FGFR2* amplification as a metastatic factor
Sohn et al. [56]	2017	Targeted sequencing	Roche 454 sequencing	12 cancerous and non-cancerous *H. pylori* negative or positive gastric tissue samples	Frequent colonization of S. pseudopneumoniae, S. parasanguinis, and S. oralis
Castaño-Rodríguez et al. [63]	2017	Targeted sequencing	Illumina MiSeq	Tissue samples from 12 GC and 20 functional dysplasia patients	Increased number of lactic acid producing bacteria, increased oral bacteria species and fatty acid and carbohydrate pathways alterations
Esser et al. [66]	2017	Whole genome sequencing and whole exome sequencing	Illumina HiSeq 2000 Applied Biosystems SOLiD	Fresh frozen primary tumor and non-neoplastic tissue samples from 2 patients	*GNAS* mutation
Ge et al. [99]	2017	Targeted sequencing	Illumina HiSeq 2000	tumor biopsy samples from 78 advanced GC patients	*TGFBR2* mutation
Constanza Camargo et al. [39]	2016	Whole genome, whole exome, mRNA and miRNA sequencing	GATK PathSeq	295 fresh and frozen gastric adenocarcinoma tissue samples	Revealing a restricted transcription pattern of viral genes mostly encoded in *BamH1A*
Kuboki et al. [41]	2016	Targeted sequencing	Life Technologies AmpliSeq	FFPE tumor samples from 121 advanced GC patients	Mutations in *TP53* (mostly frequent), *PIK3CA, ROS1, ERBB2, EGFR, MET, FGFR2, BRAF, ALK, SYNE1, CSMD3, CDH1, ARID1A, MLH1, MSH2*
Sung et al. [57]	2016	Targeted sequencing	Roche 454 sequencing	Gastric juice and mucosa samples from 4 dyspepsia or likely malignant patients	Frequent mucosal observation of *H. pylori* and *Proteobacteria*
Qi et al. [64]	2016	RNA immunoprecipitation sequencing	Illumina HiSeq 2000	10^7 cell nuclei from MKN45 AGS GES-1 GC cell lines	Metastatic *PCDH10* repression when lncRNA MALAT1 recruits EZH2

prognostic indicator for GC which is also associated with the stage of cancer [89].

NGS application in GC precancerous lesions

GC development is a result of a series of multi-stage premalignant molecular processes, which are yet to be completely understood. Even though the knowledge on the molecular basis of GC precancerous lesions is lacking in many domains, it is of great importance to detect molecular signatures and biomarkers associated with their progression into early GC lesions. The recent emergence of novel technologies including next-generation sequencing, mass spectrometry-based proteomics, rise of novel biomarkers such as miRNAs, and high-throughput microarrays have played a significant role in bringing to light new information about progression potential of different precancerous legions according to patient follow-ups [90,91]. This has led to development of efficient risk prediction models and comprehensive molecular characterizations, followed by establishment of molecular links between malignancy risk and histopathology of the disease. Exhibiting the highest sensitivity, NGS has enabled the simultaneous exploitation of several molecular targets as complementary diagnostic biomarkers to phenotypic diagnosis of GC in precancerous lesions.

In 2014, Fassan et al. used Ion Torrent Personal Genome Machine (PGM) to differentiate between gastric high-grade intraepithelial neoplasia (HGIN), a precancerous lesion, and invasive early gastric cancer (EGC), revealing a molecular similarity and suggesting a major role for *TP53* mutation in progression toward invasive EGC [92]. In a 2021 study, Zhang et al. investigated tumorigenesis characteristics in intestinal GC and identified prognostic information by RNA expression profiling of gastric precancerous lesions (low-grade intraepithelial neoplasia [LGIN] and high-grade intraepithelial neoplasia [HGIN]) and EGC biopsies. They reported a more active immune environment in ECG compared to the precancerous lesions. Furthermore, they developed a risk scoring system based on five genes in order to analyze survival rate, which is of prognostic significance for the GC patients [93]. In general, NGS-based technologies have shown undeniable promise in demystifying the molecular basis and progression process of premalignant lesions.

Challenges of clinical implementation of NGS in GC

The genome-wide approach of NGS has undoubtedly revolutionized the paradigm of GC genomics, opening new doors to prognostic, diagnostic, and therapeutic strategies (Table 3). Nevertheless, the clinical functionality of molecular classification schemas obtained from commercialized NGS platforms in GC prognosis and therapy is yet to be determined, as only a small population of patients have benefitted from these classifications. According to a recent study on NGS-based GC classification in Japan and Korea, which have the highest global GC incidence rates, only 10–15% of the patients who underwent tumor sample NGS, ended up with receiving targeted therapies [94]. Moreover, targeted therapies have been considerably less common in GC compared to other types of cancers.

Implementation of NGS in everyday clinical practice has been facing multiple challenges. Test demand is negatively affected by lack of clinical standardization, as the vast amount of data produced through NGS techniques bring about numerous targeted options for disease prevention and treatment. The resulting complexity challenges the conventional one-biomarker-one-therapy clinical care systems. Therefore, awareness and education among both physicians and patients are key elements in making molecular profiling of GC widespread. Furthermore, limited infrastructures and reimbursements are major rationales behind limited application of NGS [95].

Even though the application of NGS in diagnosing and therapeutically targeting GC have provided a better insight on disease prevention and expanded treatment opportunities, further clinical trials should be designed considering the molecular classification of GC, so that the available molecular data and preclinical research endeavors can be fruitful. It is fundamental to address the mentioned challenges in translating the molecular profiling research into practice to facilitate precision and personalized medicine. Given the ever-rising advances in NGS technologies, bioinformatics, healthcare guidelines and refined classifications, it is hoped that these technologies can actualize their advantages and optimize GC patients' experience.

Conclusion

In conclusion, GC is one of the frequent causes of cancer-related deaths and has a poor prognosis. Therefore, application of NGS-based tools can effectively revolutionize the study of development and progress of GC and also help to further recognize the tumorigenesis mechanism and metastasis. Moreover, the developments of this technology results not only in improved classification systems for GC, but also improved diagnosis and treatment decision making as well as survival for GC patients.

Acknowledgements

The authors would like to thank Sara Yousefi for designing the scientific illustrations of the present study.

Disclosure statement

The authors have no conflicts of interest to disclose.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Siamak Heidarzadeh http://orcid.org/0000-0003-3782-2916

Data availability statement

All data used during the present study are included in this manuscript and also available from the corresponding author.

References

[1] Rawla P, Barsouk A. Epidemiology of gastric cancer: global trends, risk factors and prevention. Prz Gastroenterol. 2019;14(1):26–38.

[2] Sexton RE, Al Hallak MN, Diab M, et al. Gastric cancer: a comprehensive review of current and future treatment strategies. Cancer Metastasis Rev. 2020;39(4):1179–1203.

[3] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6): 394–424.

[4] Washington MK, Goldberg RM, Chang GJ, WHO Classification of Tumours Editorial Board, et al. Diagnosis of digestive system tumours. Int J Cancer. 2021;148(5):1040–1050.

[5] Nagtegaal ID, Odze RD, Klimstra D, WHO Classification of Tumours Editorial Board, et al. The 2019 WHO classification of tumours of the digestive system. Histopathology. 2020;76(2): 182–188.

[6] de Martel C, Ferlay J, Franceschi S, et al. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. Lancet Oncol. 2012;13(6):607–615.

[7] Pormohammad A, Ghotaslou R, Leylabadlo HE, et al. Risk of gastric cancer in association with *Helicobacter pylori* different virulence factors: a systematic review and meta-analysis. Microbial Pathogenesis. 2018;118:214–219.

[8] Ignatova E, Seriak D, Fedyanin M, et al. Epstein-Barr virus-associated gastric cancer: disease that requires special approach . Gastric Cancer. 2020;23(6):951–960.

[9] Chen Y-C, Fang W-L, Wang R-F, et al. Clinicopathological variation of lauren classification in gastric cancer. Pathol Oncol Res. 2016; 22(1):197–202.

[10] Hu B, El Hajj N, Sittler S, et al. Gastric cancer: Classification, histology and application of molecular pathology. J Gastrointest Oncol. 2012;3(3):251–261.

[11] Wang Q, Liu G, Hu C. Molecular classification of gastric adenocarcinoma. Gastroenterol Res. 2019;12(6):275–282.

[12] Wang F-H, Shen L, Li J, et al. The Chinese society of clinical oncology (CSCO): clinical guidelines for the diagnosis and treatment of gastric cancer. Cancer Commun (Lond). 2019;39(1):1–31.

[13] Verma R, Sharma PC. Next generation sequencing-based emerging trends in molecular biology of gastric cancer. Am J Cancer Res. 2018;8(2):207–225.

[14] Biesecker LG, Shianna KV, Mullikin JC. Exome sequencing: the expert view. Genome Biol. 2011;12(9):128.

[15] Belkadi A, Bolze A, Itan Y, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proc Natl Acad Sci U S A. 2015;112(17): 5473–5478.

[16] Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. Genet Med. 2016; 18(7):696–704.

[17] Iglesias A, Anyane-Yeboa K, Wynn J, et al. The usefulness of whole-exome sequencing in routine clinical practice. Genet Med. 2014;16(12):922–931.

[18] Tan TY, Dillon OJ, Stark Z, et al. Diagnostic impact and cost-effectiveness of Whole-Exome sequencing for ambulant children with suspected monogenic conditions. JAMA Pediatr. 2017;171(9): 855–862.

[19] Best S, Wou K, Vora N, et al. Promises, pitfalls and practicalities of prenatal whole exome sequencing. Prenat Diagn. 2018;38(1): 10–19.

[20] Ghazani AA, Oliver NM, St. Pierre JP, et al. Assigning clinical meaning to somatic and germ-line whole-exome sequencing data in a prospective cancer precision medicine study. Genet Med. 2017;19(7):787–795.

[21] Shi W, Ng CKY, Lim RS, et al. Reliability of Whole-Exome sequencing for assessing intratumor genetic heterogeneity. Cell Rep. 2018;25(6):1446–1457.

[22] Schneider BP, Stout LA, Philips S, et al. Implications of incidental germline findings identified in the context of clinical whole exome sequencing (WES) for guiding cancer therapy. J Clin Oncol. 2019;37(15_suppl):1581–1581.

[23] Zare F, Dow M, Monteleone N, et al. An evaluation of copy number variation detection tools for cancer using whole exome sequencing data. BMC Bioinformatics. 2017;18(1):1–13.

[24] Koeppel F, Blanchard S, Jovelet C, et al. Whole exome sequencing for determination of tumor mutation load in liquid biopsy from advanced cancer patients. PLoS ONE. 2017;12(11):e0188174.

[25] Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. Nature. 2014;511(7509):344–347.

[26] Nones K, Waddell N, Wayte N, et al. Genomic catastrophes frequently arise in esophageal adenocarcinoma and drive tumorigenesis. Nat Commun. 2014;5(1):1–9.

[27] van El CG, Cornel MC, Borry P, on behalf of the ESHG Public and Professional Policy Committee, et al. Whole-genome sequencing in health care. Recommendations of the European society of human genetics. Eur J Hum Genet. 2013;21(6):580–585.

[28] Trost B, Walker S, Wang Z, et al. A comprehensive workflow for read depth-based identification of copy-number variation from whole-genome sequence data. Am J Human Genet. 2018;102(1): 142–155.

[29] Nakagawa H, Fujita M. Whole genome sequencing analysis for cancer genomics and precision medicine. Cancer Sci. 2018;109(3): 513–522.

[30] Meienberg J, Bruggmann R, Oexle K, et al. Clinical sequencing: is WGS the better WES? Hum Genet. 2016;135(3):359–362.

[31] Mattick JS, Dinger M, Schonrock N, et al. Whole genome sequencing provides better diagnostic yield and future value than whole exome sequencing. Med J Aust. 2018;209(5):197–199.

[32] Carss K, Arno G, Erwood M, et al. Comprehensive rare variant analysis via whole-genome sequencing to determine the molecular pathology of inherited retinal disease. Am J Human Genet. 2016;100(1):75–90.

[33] Davies H, Morganella S, Purdie CA, et al. Whole-Genome sequencing reveals breast cancers with mismatch repair deficiency. Cancer Res. 2017;77(18):4755–4762.

[34] Greer SU, Nadauld LD, Lau BT, et al. Linked read sequencing resolves complex genomic rearrangements in gastric cancer metastases. Genome Med. 2017;9(1):1–17.

[35] Cavalcante GC, Marinho ANR, Anaissi AK, et al. Whole mitochondrial genome sequencing highlights mitochondrial impact in gastric cancer. Sci Rep. 2019;9(1):1–13.

[36] Neiman M, Taylor DR. The causes of mutation accumulation in mitochondrial genomes. Proceedings of the Royal Society. Proc Biol Sci. 2009;276(1660):1201–1209.

[37] Stefania W, Angelo P, Stefania T. Mitochondria and familial predisposition to breast cancer. Curr Genomics. 2013;14(3):195–203.

[38] Saranathan R, Levi MH, Wattam AR, et al. *Helicobacter pylori* infections in the Bronx, New York: surveying antibiotic susceptibility and strain lineage by whole-genome sequencing. J Clin Microbiol. 2020;58(3):e01591–e01519.

[39] Camargo MC, Bowlby R, Chu A, et al. Validation and calibration of next-generation sequencing to identify Epstein-Barr virus-positive gastric cancer in the cancer genome atlas. Gastric Cancer. 2016;19(2):676–681.

[40] Bustos-Carpinteyro AR, Oliveira C, Sousa A, et al. CDH1 somatic alterations in mexican patients with diffuse and mixed sporadic gastric cancer. BMC Cancer. 2019;19(1):1–9.

[41] Kuboki Y, Yamashita S, Niwa T, et al. Comprehensive analyses using next-generation sequencing and immunohistochemistry enable precise treatment in advanced gastric cancer. Ann Oncol. 2016;27(1):127–133.

[42] Mafficini A, Amato E, Fassan M, et al. Reporting tumor molecular heterogeneity in histopathological diagnosis. PLoS ONE. 2014; 9(8):e104979.
[43] Vecchi M, Nuciforo P, Romagnoli S, et al. Gene expression analysis of early and advanced gastric cancers. Oncogene. 2007; 26(29):4284–4294.
[44] Marimuthu A, Jacob HKC, Jakharia A, et al. Gene expression profiling of gastric cancer. J Proteomics Bioinform. 2011;4(4): 74–82.
[45] Sun H. Identification of key genes associated with gastric cancer based on DNA microarray data. Oncol Lett. 2016;11(1):525–530.
[46] Chu Y, Corey DR. RNA sequencing: platform selection, experimental design, and data interpretation. Nucleic Acid Ther. 2012;22(4): 271–274.
[47] Heo YJ, Park C, Yu D, et al. Reproduction of molecular subtypes of gastric adenocarcinoma by transcriptome sequencing of archival tissue. Sci Rep. 2019;9(1):9675.
[48] Wu J, Zhao X, Shao Z. A system level analysis of gastric cancer across tumor stages with RNA-seq data. Mol BioSyst. 2015;11(7): 1925–1932.
[49] Fu K, Hui B, Wang Q, et al. Single-cell RNA sequencing of immune cells in gastric cancer patients. Aging (Albany, NY). 2020; 12(3):2747–2763.
[50] Singh BP, Gupta VK. Molecular markers in mycology. Switzerland: Springer International Publishing; 2017.
[51] Xiao W, Wu L, Yavas G, et al. Challenges, solutions, and quality metrics of personal genome assembly in advancing precision medicine. Pharmaceutics. 2016;8(2):15.
[52] Kumar S, Krishnani KK, Bhushan B, et al. Metagenomics: retrospect and prospects in high throughput age. Biotechnol Res Int. 2015;2015:1–13.
[53] Rizzato C, Torres J, Obazee O, et al. Variations in cag pathogenicity island genes of *Helicobacter pylori* from Latin American groups may influence neoplastic progression to gastric cancer. Sci Rep. 2020;10(1):6570.
[54] Rizzato C, Torres J, Plummer M, et al. Variations in Helicobacter pylori cytotoxin-associated genes and their influence in progression to gastric cancer: implications for prevention. PLoS ONE. 2012;7(1):e29605.
[55] Zheng Z, Andersson AF, Ye W, et al. A method for metagenomics of Helicobacter pylori from archived formalin-fixed gastric biopsies permitting longitudinal studies of carcinogenic risk. PloS One. 2011;6(10):e26442.
[56] Sohn S-H, Kim N, Jo HJ, et al. Analysis of gastric body microbiota by pyrosequencing: possible role of bacteria other than *Helicobacter pylori* in the gastric carcinogenesis. J Cancer Prev. 2017;22(2):115–125.
[57] Sung J, Kim N, Kim J, et al. Comparison of gastric microbiota between gastric juice and mucosa by next generation sequencing method. J Cancer Prev. 2016;21(1):60–65.
[58] Eun CS, Kim BK, Han DS, et al. Differences in gastric mucosal microbiota profiling in patients with chronic gastritis, intestinal metaplasia, and gastric cancer using pyrosequencing methods. Helicobacter. 2014;19(6):407–416.
[59] Hu Z, Cheng L, Wang H. The Illumina-Solexa sequencing protocol for bacterial genomes. In: Mengoni A, Galardini M, Fondi M, editors. Bacterial pangenomics: methods and protocols. New York (NY): Springer New York; 2015. p. 91–97.
[60] Pettersson E, Lundeberg J, Ahmadian A. Generations of sequencing technologies. Genomics. 2009;93(2):105–111.
[61] Asakawa Y, Okabe A, Fukuyo M, et al. Epstein-Barr virus-positive gastric cancer involves enhancer activation through activating transcription factor 3. Cancer Sci. 2020;111(5):1818–1828.
[62] Ito T, Matoba R, Maekawa H, et al. Detection of gene mutations in gastric cancer tissues using a commercial sequencing panel. Mol Clin Oncol. 2019;11(5):455–460.
[63] Castaño-Rodríguez N, Goh K-L, Fock KM, et al. Dysbiosis of the microbiome in gastric carcinogenesis. Sci Rep. 2017;7(1):15957.
[64] Qi Y, Ooi HS, Wu J, et al. MALAT1 long ncRNA promotes gastric cancer metastasis by suppressing PCDH10. Oncotarget. 2016; 7(11):12693–12703.
[65] Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. Clin Chem. 2009; 55(4):641–658.
[66] Esser D, Holze N, Haag J, et al. Interpreting whole genome and exome sequencing data of individual gastric cancer samples. BMC Genomics. 2017;18(1):517.
[67] Nagarajan N, Bertrand D, Hillmer AM, et al. Whole-genome reconstruction and mutational signatures in gastric cancer. Genome Biol. 2012;13(12):R115.
[68] Ribeiro-dos-Santos Â, Khayat AS, Silva A, et al. Ultra-deep sequencing reveals the microRNA expression pattern of the human stomach. PLoS ONE. 2010;5(10):e13205.
[69] Crosetto N, Mitra A, Silva MJ, et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. Nat Methods. 2013;10(4):361–365.
[70] Alarcón T, Llorca L, Perez-Perez G. Impact of the microbiota and gastric disease development by *Helicobacter pylori*. In: Tegtmeyer N, Backert S, editors. Molecular pathogenesis and signal transduction by *Helicobacter pylori*. Cham: Springer International Publishing; 2017. p. 253–275.
[71] Levene MJ, Korlach J, Turner SW, et al. Zero-mode waveguides for single-molecule analysis at high concentrations. Science. 2003; 299(5607):682–686.
[72] Eid J, Fehr A, Gray J, et al. Real-time DNA sequencing from single polymerase molecules. Science. 2009;323(5910):133–138.
[73] Huang KK, Huang J, Wu JKL, et al. Long-read transcriptome sequencing reveals abundant promoter diversity in distinct molecular subtypes of gastric cancer. Genome Biol. 2021;22(1):44.
[74] Watanabe Y, Oikawa R, Kodaka Y, et al. Cancer-related genetic variants of *Helicobacter pylori* strains determined using gastric wash-based whole-genome analysis with single-molecule real-time technology. Int J Cancer. 2021;148(1):178–192.
[75] Sohn BH, Hwang J-E, Jang H-J, et al. Clinical significance of four molecular subtypes of gastric cancer identified by the cancer genome atlas project. Clin Cancer Res. 2017;23(15):4441–4449.
[76] Yang Q, Zhu C, Zhang Y, et al. Molecular analysis of gastric cancer identifies genomic markers of drug sensitivity in Asian gastric cancer. J Cancer. 2018;9(16):2973–2980.
[77] Johansson H, Isaksson M, Sörqvist EF, et al. Targeted resequencing of candidate genes using selector probes. Nucleic Acids Res. 2011;39(2):e8.
[78] Li J, Lupat R, Amarasinghe KC, et al. CONTRA: copy number analysis for targeted resequencing. Bioinformatics. 2012;28(10): 1307–1313.
[79] Nadauld LD, Garcia S, Natsoulis G, et al. Metastatic tumor evolution and organoid modeling implicate TGFBR2 as a cancer driver in diffuse gastric cancer. Genome Biol. 2014;15(8):428.
[80] Myllykangas S, Buenrostro JD, Natsoulis G, et al. Efficient targeted resequencing of human germline and cancer genomes by oligonucleotide-selective sequencing. Nat Biotechnol. 2011;29(11): 1024–1027.
[81] Ikari N, Serizawa A, Mitani S, et al. Near-comprehensive resequencing of cancer-associated genes in surgically resected metastatic liver tumors of gastric cancer. Am J Pathol. 2019;189(4): 784–796.
[82] Banerjee J, Mishra N, Dhas Y. Metagenomics: a new horizon in cancer research. Meta Gene. 2015;5:84–89.
[83] Gunathilake M, Lee J, Choi IJ, et al. Alterations in gastric microbial communities are associated with risk of gastric cancer in a korean population: a case-control study. Cancers. 2020;12(9):2619.
[84] Hu Y-L, Pang W, Huang Y, et al. The gastric microbiome is perturbed in advanced gastric adenocarcinoma identified through shotgun metagenomics. Front Cell Infect Microbiol. 2018;8(433): 1–11.
[85] Seol M, Lee YR, Kim KM, et al. The difference of the gut microbiota of gastric cancer in relation to Helicobacter pylori negativity and positivity. J Clin Oncol. 2019;37(4_suppl):10–10.

[86] Shiroma Y, Takahashi R-U, Yamamoto Y, et al. Targeting DNA binding proteins for cancer therapy. Cancer Sci. 2020;111(4): 1058–1064.
[87] Leili Sadeghi A, Ali B, Alireza R, et al. An overview of the epigenetic modifications of gene expression in tumorigenesis. Res Mol Med (RMM). 2019;6(3):1–9.
[88] Zhang Y, Yu C. Distinct expression and prognostic values of the replication protein a family in gastric cancer. Oncol Lett. 2020; 19(3):1831–1841.
[89] Shi S, Zhang Z-G. Role of Sp1 expression in gastric cancer: a meta-analysis and bioinformatics analysis. Oncol Lett. 2019;18(4): 4126–4135.
[90] Fassan M, Baffa R, Kiss A. Advanced precancerous lesions within the GI tract: the molecular background. Best Pract Res Clin Gastroenterol. 2013;27(2):159–169.
[91] Li X, Zheng N-R, Wang L-H, et al. Proteomic profiling identifies signatures associated with progression of precancerous gastric lesions and risk of early gastric cancer. eBioMedicine. 2021;74: 103714.
[92] Fassan M, Simbolo M, Bria E, et al. High-throughput mutation profiling identifies novel molecular dysregulation in high-grade intraepithelial neoplasia and early gastric cancers. Gastric Cancer. 2014;17(3):442–449.
[93] Zhang Y, Wu X, Zhang C, et al. Dissecting expression profiles of gastric precancerous lesions and early gastric cancer to explore crucial molecules in intestinal-type gastric cancer tumorigenesis. J Pathol. 2020;251(2):135–146.
[94] Nakamura Y, Shitara K, Lee J. The right treatment of the right patient: Integrating genetic profiling into clinical decision making in advanced gastric cancer in asia. Am Soc Clin Oncol Educ Book. 2021;(41):1–e73.
[95] Horgan D, Curigliano G, Rieß O, et al. Identifying the steps required to effectively implement next-generation sequencing in oncology at a national level in Europe. JPM. 2022;12(1):72.
[96] Zhou J, Zhao W, Wu J, et al. Neoantigens derived from recurrently mutated genes as potential immunotherapy targets for gastric Cancer. Biomed Res Int. 2019;2019:8103142.
[97] Zhang J, Qiu W, Liu H, et al. Genomic alterations in gastric cancers discovered via whole-exome sequencing. BMC Cancer. 2018; 18(1):1270.
[98] Pan X, Ji X, Zhang R, et al. Landscape of somatic mutations in gastric cancer assessed using next-generation sequencing analysis . Oncol Lett. 2018;16(4):4863–4870.
[99] Ge S, Li B, Li Y, et al. Genomic alterations in advanced gastric cancer endoscopic biopsy samples using targeted next-generation sequencing. Am J Cancer Res. 2017;7(7):1540–1553.