Review Article

The role of Next-Generation Sequencing in tumoral radiosensitivity prediction

Jean-Emmanuel Bibault a,b,⇑, Ingeborg Tinhofer c

a Radiation Oncology Department, Georges Pompidou European Hospital, Assistance Publique – Hôpitaux de Paris, Paris Descartes University, Paris Sorbonne Cité, Paris, France
b INSERM UMR 1138 Team 22: Information Sciences to Support Personalized Medicine, Paris Descartes University, Sorbonne Paris Cité, Paris, France
c Department of Radiooncology and Radiotherapy, Charité University Hospital Berlin, Berlin, Germany

ABSTRACT

Technological advances have led to more precise radiation delivery, which has resulted in significant clinical gains. A better understanding of tumoral radiosensitivity is still needed to develop strategies and further personalize radiation treatments. Next-Generation Sequencing (NGS) and system biology have significantly transformed the field of oncology in the last two decades, but have only a few clinical applications in radiation oncology. This review describes the technical aspects and evolutions of NGS and discusses the latest clinical applications of genomics to predict tumoral radiosensitivity.

© 2017 The Authors. Published by Elsevier Ireland Ltd on behalf of European Society for Radiotherapy and Oncology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

Introduction .................................................................................................................. 16
NGS: concepts and definitions .................................................................................. 17
NGS data analysis ....................................................................................................... 18
The complexity of radiation response prediction ...................................................... 18
An example of the successful use of genomics to predict treatment response ....... 18
Head and neck cancer ................................................................................................ 18
Precision oncology: pushing forward in radiobiology research ............................ 19
Conclusion ................................................................................................................ 19
Author contributions ................................................................................................. 19
Competing interest statement .................................................................................... 19
Search strategy and selection criteria ....................................................................... 19
References .................................................................................................................... 19

Introduction

Each patient has an individual set of molecular abnormalities responsible for their disease or correlated with treatment response and clinical outcome. The concept of personalized treatments relies on identifying and leveraging these aberrations for each patient. Molecular oncology has driven cancer research in the last 20 years and has seen significant progress in poor-prognosis diseases such as non-small cell lung cancer, through the use of EGFR inhibitors [1]. These personalized targeted treatments will rely heavily on Next-Generation Sequencing (NGS). While the cost associated with the first Human Genome Project was 3 billion dollars, it is now considered that a whole genome can now be sequenced for around a thousand dollars [2–4] (Fig. 1). This technique allows for the exploration of many genes and mechanisms that could be used to unravel the complexity of the molecular circuits involved in tumoral radiation response [5]. Predicting normal tissue toxicity with NGS could be used to personalize treatments...
Several studies have also established that key driver genes or genomic signatures can predict treatment response after radiation. In this review, we will give an overview of the technical aspects of NGS and its possibilities. We will also give two examples of the use of genomics to predict treatment response in head and neck [10] and prostate cancer [11].

**NGS: concepts and definitions**

In 1975, Sanger and Coulson described the first method to sequence DNA [12]. During Sanger sequencing, DNA polymerases copy single-stranded DNA templates by adding nucleotides to a growing chain, selected by base-pair matching to the template. But DNA polymerases also incorporate dideoxynucleotides, analogues of nucleotides. When dideoxynucleotides are incorporated at the 3' end of the growing chain, chain elongation is terminated selectively at A, C, G, or T. This generates different copies of the original DNA template at all possible lengths, which are separated by capillary electrophoresis. A read of the chain terminating A, C, G or T, labeled with radioactivity or fluorescence, can then be made to determine the sequence. Sanger sequencing was the most widely used sequencing method until NGS became available. It is now used for smaller-scale project or NGS results validation.

NGS has many advantages over Sanger sequencing: the automated, high-throughput allows for increased speed and resolution, with a lower cost [3]. NGS can be used to perform a comprehensive analysis of the genome at different scale: it can be focused on specific genes, or to sequence all coding genes (exome), or whole genomes. But it can also be used to explore epigenetic mechanisms (methylation, histone modifications) [13], to sequence the transcriptome (RNA-seq) [14], or to screen protein-DNA interactions (ChIP-seq) [15]. There are several NGS methods available (Table 1): single molecule real-time sequencing, ion semiconductor, pyrosequencing, sequencing by synthesis, or sequencing by ligation. In the recent years, ion semiconductor (Ion Torrent®, Thermo Fisher Scientific, Waltham, Massachusetts, USA) and sequencing by synthesis (Illumina®, San Diego, California, USA) have dominated NGS.

DNA can be extracted from all types of samples, including formalin-fixed, paraffin-embedded tissue [16] or even circulating tumor DNA [17], as long as the minimum quantity of DNA is present in the sample (250 ng for a whole exome for example). DNA is then tagged and fragmented into short segments to create a library. Primers of the sequence of interest are added at the end of each fragment. The library is then amplified using PCR. The last step requires a sequencer that will read each fragment of DNA in a parallel fashion. The most common sequencers will detect the fluorescent signal through digital imaging (sequencing by synthesis) or the hydrogen ions through an IFSET ion sensor (ion semiconductor) released during DNA synthesis. A run can take between 30 min and two weeks, depending on the technology used, and generate between one million and three billion reads [3]. The raw data (FASTQ files) will need bioinformatics analysis, which can be in some cases automated in custom pipelines [18]. A sequencing run quality will depend on the coverage of the region of interest (ROI), the depth (number of reads that correctly align to the ROI) and the reads’ length and accuracy.

**Table 1**

A comparison of sequencing methods.

| Method                     | Read length | Accuracy   | Reads per run | Time per run | Cost per 1 million bases | Advantages                                      | Disadvantages                                      |
|----------------------------|-------------|------------|---------------|--------------|--------------------------|-------------------------------------------------|---------------------------------------------------|
| Sanger                     | 400 to 900 bp | 99.9%      | –             | 30 min to 3 h | 2200 €                  | Long reads, gold standard for accuracy          | Expensive, time consuming                         |
| Single molecule real time  | 3 Kb        | 85%        | 75000         | 30 min to 4 h | 0.5 €                   | Longest read length, uniform coverage           | Moderate throughput, expensive equipment          |
| Ion semiconductor          | Up to 400 bp | 98%        | Up to 80 million | 2 h         | 0.9 €                   | Cheap, fast.                                    | Homopolymer errors                                |
| Pyrosequencing              | 700 bp      | 99.9%      | 1 million     | 24 h         | 9 €                      | Long read size, fast                            | Expensive, Homopolymer errors                     |
| Sequencing by synthesis     | 75–300 bp   | 99.9%      | 1 to 3 billion | 1 to 11 days | 0.04 € to 0.13 €        | Expensive equipment. Requires high concentrations of DNA |
| Sequencing by ligation      | 50 + 35 or 50 + 50 bp | 99.9% | 1.2 to 1.4 billion | 1 to 2 weeks | 0.11 €                   | Cheap                                           | Errors when sequencing palindromic sequences     |
NGS data analysis

Data analysis is a time-consuming process and is often considered as the “bottle-neck” of NGS. Next generation sequencing generates huge amounts of data that often necessitates the use of powerful computing facilities and algorithms.

The sequencing data analysis starts from files containing DNA sequences and quality values for each base. Depending on the size and the depth of the sequencing, a FASTQ file size can be several hundred gigabytes. These files need to be transferred from the sequencer into a workstation for analysis and also stored in a dedicated data warehouse. For each NGS project, a custom pipeline must be created to automate the analysis, in order to standardize results and avoid errors. A typical variant calling pipeline includes an assessment of the quality of sequencing, an assembly of the reads (alignment or de novo) and variant detection. For RNA-seq and Chip-seq, specific pipelines must be created [19].

The complexity of radiation response prediction

Tumor response to radiation depends on both treatment-related factors (total dose, dose per fraction) and tumor-intrinsic features factors, such as the tumor molecular profile. A validated predictor should guide different treatment options in order to either maximize the antitumor effect (e.g.: determine whether a patient should be treated, whether the ionizing radiation dose might be escalated or whether other antineoplastic drug should be combined), or avoid unnecessary toxicity without jeopardizing the patient’s outcome (e.g.: decide whether we should decrease the dose or the volume).

The antitumor effect of radiotherapy is not restricted to the generation of DNA damage but also involve many other pathways such as angiogenesis, hypoxia metabolism, glucose and lipid metabolism, immunological response, invasiveness and metastatic processes, and cell death (apoptosis, mitotic catastrophe, autophagy), among others [20,21].

For elucidating the role of somatic mutations in radiosensitivity, NGS was first applied in bacteria [22]. In a model of cellular adaptation to irradiation, radiosensitive E. coli strains were generated by repetitive cycles of increasing irradiation doses. Whole genome sequencing revealed a large number of genomic alterations associated with a radiosensitive phenotype. Only few were recurrent mutations were detected, suggesting that multiple mechanisms can contribute to radiation resistance. Despite this heterogeneity, clear genetic patterns also emerged. Not unexpectedly, mutations clustered more frequently in genes of DNA double strand break repair [22].

Most of the previous studies have looked at single biomarkers as predictor of radiosensitivity, and none are currently in use in the clinical daily practice [20]. However, we are still far away from that goal, and it is safe to believe that radiosensitivity cannot be predicted by a single variant, but most probably by a unique combination of variants and mechanisms. The challenge of effectively using NGS to predict tumor radiation response reveals the complexity and our limited understanding of the underlying biological mechanisms involved. A recent study published by Scott et al. used a gene-expression-based radiation-sensitivity index (GARD) for 8271 tissue samples from the TCC cohort [23]. The study was based on studies previously published by the same team that described the radiosensitivity genomic signature [24–29], an equation taking into account expression levels of ten genes (AR, cjun,STAT1, PKC, RelA, cABL, SUMO1, PAK2, HDAC1, and IRF1).

The sensitivity index was lowest for gliomas and sarcomas and highest for cervical cancer and oropharyngeal head and neck cancer. The GARD also independently predicted clinical outcome in breast cancer, lung cancer, glioblastoma, and pancreatic cancer. This index could potentially allow the personalization of radiotherapy dose to tumour radiosensitivity. These interesting results are mostly hypothesis-generating and should call for the prospective evaluation of GARD within a randomized trial if we want to achieve genome-personalized radiation dose. Finally, a recent study suggests that mutations of DNA repair genes could be key in explaining tumoral radiosensitivity: Ma et al. assessed the genomic alterations found in exceptional responders to radiotherapy and found that ATM mutations were associated with excellent radiation response [30]. This study was performed on only 9 patients, but also included an analysis of 22 DNA repair genes in The Cancer Genome Atlas (TCGA) data that revealed mutations in 16% of more than 9000 tumors across 24 cancer types, with ATM mutations being the most prevalent.

An example of the successful use of genomics to predict treatment response

Head and neck cancer

Despite significant advances in treatment by addition of concurrent chemotherapy to radiotherapy [31], diagnosis of locally advanced squamous cell carcinoma of the head and neck (HNSCC) is still associated with mediocre prognosis, with less than 50% to 60% of patients alive after 5 years. Currently, personalized treatment strategies for optimization of treatment outcome have not yet been established in routine patient care and blockade of epidermal growth factor receptor remains so far the only molecular targeting approach for radiosensitizing these tumors. Several recent genomic studies [32–35] including The Cancer Genome Atlas project [36] have profiled a large number of HNSCC to provide a comprehensive landscape of somatic genomic alterations and identify therapeutic candidate alterations. Known and novel genetic alterations were discovered and major differences in the mutational patterns of HPV – and HPV+ tumors detected. Overall, more than 15,000 genes were found to be altered but only 360 of these genes (2.5%) were affected in ≥3% of patients, underlining the complexity of the genetic profiles and the large intertumoral heterogeneity in HNSCC. These studies also revealed that the majority of genetic alterations occur in tumor suppressor genes whereas oncogenic driver gene mutations are rare, making the development of molecular targeting approaches a real challenge. Many of the affected genes have a role in cell cycle, DNA repair and cell survival under stress conditions. It seems thus very likely that the mutational profiles might influence the tumor cell response to chemoradiation and could be used for patient stratification.

Since the above-mentioned NGS studies did not include uniformly treated HNSCC patients, definite conclusions on the therapeutic relevance of individual mutations or mutational patterns are currently difficult to draw. We therefore recently initiated two projects of genetic sequencing in the setting of definitive [37] and adjuvant chemoradiation [10] where we evaluated the role of somatic mutations in well characterized and uniformly treated HNSCC patient cohorts. Our studies confirmed previous reports of poor efficacy of radiotherapy in HNSCC tumors harboring TP53 mutations [38–40].

For the first time, we could also demonstrate a difference in the role of distinct classes of TP53 mutations: In the definitive setting, nonsense/frameshift TP53 mutations [37] associated with either expression of a truncated p53 protein or complete loss of p53 expression but not missense mutations resulting in overexpression of mutated p53 were significantly associated with survival. This result was in contrast to the adjuvant setting where missense mutations associated with a gain-of-function but not
nonsense/frameshift mutations of TP53 were significantly correlated with poor outcome [10]. The reason for this discrepancy remains unclear but different biological functions under the control of (mutant) p53 might be responsible for poor outcome after definitive or adjuvant chemoradiation. While loss-of-function TP53 alterations could directly reduce intrinsic radiosensitivity of tumor cells to definitive chemoradiation, gain-of-function mutations associated with epithelial-mesenchymal transition, invasive growth and increased migratory potential [41,42] could drive tumor recurrence more indirectly in the adjuvant setting. For the first time, we also demonstrated a possible role of mutations in NOTCH1 and key driver genes (PIK3CA, KRAS, NRAS and HCRAS) [10] as well as germ-line variants of KDR [37], the gene encoding for the vascular endothelial growth factors receptor 2, as predictive biomarkers of outcome after chemoradiation.

Altogether, these data provide support for the concept of integrating targeted NGS for improved response prediction in locally advanced HNSCC. As one example, TP53 mutational analysis could identify patients with dismal outcome after chemoradiation and highest need of novel treatment concepts. Analysis of the tumor mutational load by NGS in this high-risk patient subgroup might serve as marker for potential benefit from combination of chemoradiation with immune checkpoint inhibitors – with increased activity in tumors with higher frequency of somatic mutations [43].

**Precision oncology: pushing forward in radiobiology research**

Some oncologists are concerned that precision oncology, an approach relying on the identification of a mutation to prescribe the appropriate treatment regimen, is doomed [44]. In a recent perspective published in Nature, Vinay Prasad argues that few patients actually benefitted from targeted drug on identified mutations. He notes that the sequencing programme at the MD Anderson Cancer Center included 2600 patients and showed that only 6.4% of them received a targeted drug [45]. In the same manner, only 2% of the 795 patients enrolled in the NCI-MATCH trial were paired with a targeted therapy [46]. Even among these patients, the response rate is only around 30%, with a median progression-free survival of just 5.7 months. In line with this, the only randomized phase III trial (SHIVA) performed to compare targeted therapy paired with a mutation to standard chemotherapy selected by the physician, showed no difference in the two groups, with a median progression-free survival of 2.3 months (95% CI 1.7–3.8) in the experimental group versus 2.0 months (1.8–2.1) in the control group (hazard ratio 0.88, 95% CI 0.65–1.19, p = 0.41). These results highlight the difference between biological rationale and clinical effectiveness, but they should not be a reason to reject the precision approach. Multiple factors may contribute to the limited success of the current clinical evaluation of personalized medicine [47], including limited access to targeted agents both within and outside clinical trials and only partial inhibition of signaling pathways by most molecular targeted agents. In addition, targeting a mutation or mechanism with a drug is still very different from targeting a tumor with ionizing radiation alone or combined with a drug. As radiation oncologists, we should continue to use NGS in preclinical and translational studies to better understand our treatments in order to make them more effective and less toxic.

**Conclusion**

Next-Generation Sequencing has revolutionized the way we explore the genome and the mechanisms involved in its regulation. Being both faster and cheaper than Sanger sequencing, it could unravel the cellular mechanisms involved in radiation response and help personalize our treatments. However, to this day, no genetic biomarker is used in the daily clinical routine to tailor radiation treatments. Several studies have already showed the prognostic value of NGS, but the pursuit of rigorous and thorough radiobiology studies should be encouraged. It is also our responsibility as physicians or researchers to the patients to avoid creating false hopes or make unrealistic claims.

**Author contributions**

J.E.B. and I.T. researched data and wrote the article.

**Competing interest statement**

The authors declare no competing interests.

**Search strategy and selection criteria**

Information for this Review was compiled by searching the PubMed and MEDLINE databases for articles published between January 1980 and August 2016, including early release publications. Search terms included “radiation therapy,” “genomics” and “tumoral radiosensitivity”. Only articles published in English were considered and references were chosen based on suitability for inclusion. Full articles were obtained and the references lists were checked for additional material, when appropriate.

**References**

[1] Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGF mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304:1497–500. http://dx.doi.org/10.1126/science.1095934.

[2] Mardis ER. A decade’s perspective on DNA sequencing technology. Nature 2011;470:198–203. http://dx.doi.org/10.1038/nature09796.

[3] Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet 2010;11:31–46. http://dx.doi.org/10.1038/nrg2626.

[4] DNA sequencing costs n.d. http://www.genome.gov/sequencingcosts/ (accessed March 12, 2016).

[5] Georgakilas AG, Pavlopoulos A, Lousa M, Nikitiaki Z, Vorgias CE, Bagos PG, et al. Emerging molecular networks common in ionizing radiation, immune and inflammatory responses by employing bioinformatics approaches. Cancer Lett 2015;368:164–72. http://dx.doi.org/10.1016/j.canlet.2015.01.021.

[6] West CM, Barnett GC. Genetics and genomics of radiotherapies: towards prediction. Genome Med 2011;3:52. http://dx.doi.org/10.1186/gemmed.

[7] Barnett GC, Coles CE, Elliott RM, Baynes C, Lucarrini C, Conroy D, et al. Independent validation of genes and polymorphisms reported to be associated with radiation toxicity: a prospective analysis study. Lancet Oncol 2012;13:65–77. http://dx.doi.org/10.1016/S1470-2045(11)70393-3.

[8] Barnett GC, Thompson D, Fachal I, Kerrins S, Talbot C, Elliott RM, et al. A genome wide association study (GWAS) providing evidence of an association between common genetic variants and late radiotherapy toxicity. Radiother Oncol J Eur Soc Thor Radiol Oncol 2014;111:178–85. http://dx.doi.org/10.1016/j.radonc.2014.02.012.

[9] Fachal L, Gómez-Gañacho A, Barnett GC, Peleteiro P, Carballo AM, Calvo-Crespo F, et al. A three-stage genome-wide association study identifies a susceptibility locus for late radiotherapy toxicity at 2q24.1. Nat Genet 2014;46:891–4. http://dx.doi.org/10.1038/ng.3020.

[10] Tinhofer I, Budach V, Saki M, Konschak R, Neirh F, Jöhrns K, et al. Targeted next-generation sequencing of locally advanced squamous cell carcinomas of the head and neck reveals druggable targets for improving adjuvant chemoradiation. Eur J Cancer Oxf Engl 1990;2016(57):78–86. http://dx.doi.org/10.1016/j.ejca.2016.01.014.

[11] Lalonde E, Ishkanian AS, Sykes J, Fraser M, Ross-Adams H, Erho N, et al. Tumour genetic and microenvironmental heterogeneity for integrated prediction of 5-year biochemical recurrence of prostate cancer: a retrospective cohort study. Lancet Oncol 2014;15:1521–32. http://dx.doi.org/10.1016/s1470-2045(14)72021-6.

[12] Sanger F, Coulson AR. A rapid method for determining sequences in DNA primed synthesis with DNA polymerase. J Mol Biol 1975;94:441–8.

[13] Sarda S, Hannenhalli S. Next-generation sequencing and epigenomics research: a hammer in search of nails. Genomics Inform 2014;12:2–11. http://dx.doi.org/10.5808/GI.2014.12.1.2.

[14] Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 2008;5:621–8. http://dx.doi.org/10.1038/nmeth.1226.

[15] Johnson DS, Mortazavi A, Myers RM, Wold B. Genome-wide mapping of in vivo protein-DNA interactions. Science 2007;316:1497–502. http://dx.doi.org/10.1126/science.1141312.
DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491–8.

Gogol-Döring A, Chen W. An overview of the analysis of next generation sequencing data. Methods Mol Biol Clifton NJ 2012;802:249–57. http://dx.doi.org/10.1007/978-1-61779-400-1_16.

Bibault J.-E, Tinhofer I. Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. Mol Oncol 2015;9:783–90. http://dx.doi.org/10.1016/j.molonc.2014.12.004.

DeProst MA, Bankes E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491–8.

Hedegaard J, Thorsen K, Lund MK, Blin A, Leroy Q, et al. Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. Mol Oncol 2015;9:783–90. http://dx.doi.org/10.1016/j.molonc.2014.12.004.

Lebofsky R, Decraene C, Bernard V, Kamal M, Blin A, Leroy Q, et al. Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. Mol Oncol 2015;9:783–90. http://dx.doi.org/10.1016/j.molonc.2014.12.004.

Jia J, Ling J, Liu Y, Guo R, Zhang B, et al. Genomic analysis of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science 2011;333:1154–7. http://dx.doi.org/10.1126/science.1206923.

Lebofsky R, Decraene C, Bernard V, Kamal M, Blin A, Leroy Q, et al. Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. Mol Oncol 2015;9:783–90. http://dx.doi.org/10.1016/j.molonc.2014.12.004.

Ma J, Setton J, Morris L, Albornoz PBC, Barker C, Lok BH, et al. Genomic analysis of radiation therapy and biomarker-driven treatment strategies: a systematic review. Cancer Metastasis Rev 2013;32:479–92. http://dx.doi.org/10.1007/s10555-013-9419-7.

Liauw SL, Connell PP, Weichselbaum RR. New paradigms and future challenges in radiation oncology: an update of biological targets and technology. Sci Transl Med 2013;5:173rv2. http://dx.doi.org/10.1126/scitranslmed.3005146.

Hedgegaard J, Thorsen K, Lund MK, Blin A, Leroy Q, et al. Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. Mol Oncol 2015;9:783–90. http://dx.doi.org/10.1016/j.molonc.2014.12.004.

Hedgegaard J, Thorsen K, Lund MK, Blin A, Leroy Q, et al. Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. Mol Oncol 2015;9:783–90. http://dx.doi.org/10.1016/j.molonc.2014.12.004.

Hedgegaard J, Thorsen K, Lund MK, Blin A, Leroy Q, et al. Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. Mol Oncol 2015;9:783–90. http://dx.doi.org/10.1016/j.molonc.2014.12.004.

Hedgegaard J, Thorsen K, Lund MK, Blin A, Leroy Q, et al. Circulating tumor DNA as a non-invasive substitute to metastasis biopsy for tumor genotyping and personalized medicine in a prospective trial across all tumor types. Mol Oncol 2015;9:783–90. http://dx.doi.org/10.1016/j.molonc.2014.12.004.