NEXT GENERATION SEQUENCING TECHNOLOGIES TOWARDS EXPLORATION OF MEDICINAL PLANTS

Sunanya Das, Rukmini Mishra*

Department of Botany, School of Applied Sciences, Centurion University of Technology and Management, Odisha, India.

Received – May 30, 2021; Revision – July 12, 2021; Accepted – August 19, 2021
Available Online – August 30, 2021

DOI: http://dx.doi.org/10.18006/2021.9(4).507.516

ABSTRACT

With the ever-increasing population, the plant cover is decreasing at an alarming rate. The medicinal plants are most affected by this because they are present in the last tier of cultivation. Let it be pharmaceutical companies or people using it for herbalism, medicinal plants have been exploited without getting a chance to flourish in their natural environment. Modern biotechnology acts as a bridge between the cultivation and utilization of medicinal plants. Next Generation Sequencing (NGS) technology which is a decade old but emerging field helps to unveil the importance of medicinal plants. Thus, it paves the way for sustenance of medicinal plants by molecular breeding, micropropagation, large-scale tissue culture, and other methods to conserve the plants with great medicinal value. Various NGS technologies can be found in the market like Ilumina, PacBio, Ion Torrent, and others. The present review will summarize the NGS technologies and their potential use to study the genomes, transcriptome, epigenome, and interactome of medicinal plants towards the identification of bioactive compounds.

KEYWORDS

Medicinal Plants
Next Generation Sequencing
Transcriptomics
Genomics
Ilumina
PacBio
Ion Torrent

* Corresponding author
E-mail: rukmini.mishra@cutm.ac.in (Rukmini Mishra)

Peer review under responsibility of Journal of Experimental Biology and Agricultural Sciences.
1 Introduction

The medicinal plants have proven useful since the ages of nomadic humans when they were transitioning from hunting to scavenging and cultivating. One or more parts or organs have therapeutic properties or contain the precursors of the bioactive compounds. The word herb is derived from the Latin word “herbe”. During the nomadic period, the man had their eyes fixed only on the major food supplier crop species like rice, wheat, maize, soybean and potato, etc., and medicinal plants were forgotten for decades (Siahsar et al., 2011). Later on, when human beings faced the brutal force of nature they learned to protect themselves and combat diseases with herbs and plants. Human being explored nature for two major purposes i.e. food for survival and herbs to treat different ailments (Morgan, 2002). The use of plants or plant products for treating various diseases and ailments has been mentioned in the different historical literature of countries like India, Egypt, China, and Japan, etc.

Even though we are aware of the advantages and diverse applications of medicinal plants, they are still placed at the bottom wave of domestication (Nogueira et al., 2018). With advances in time and technologies, researchers became aware of the importance and use of medicinal plants. The existing population of medicinal plants is reducing at an alarming rate due to the unsustainable and unethical harvesting from the wild and destruction of its natural habitat (Niazian, 2019). With the increase in population, the exploitation of these plants has also increased. Without the help of humans, these plants are not capable of reproducing on their own at a faster rate to meet our demands. Therefore the need to conserve and careful domestication of the medicinal plants is the need of the hour.

The development of the medicinal plant sector will lead to decreases in the dependency of the developing countries on highly expensive synthetic drugs. To produce high-quality plant based bioactive compounds, profound research on medicinal plants is the need of the hour. High scale genome/transcriptome exploration of the medicinal plants will contribute to gaining more insight into biosynthetic processes. NGS are offering us the platform to access the genomic, transcriptomic, and epigenomic aspects of these non-model plants or the medicinal plants at comparatively low cost and efforts (Gandhi et al., 2015).

Next Generation Sequencing proved to be a powerful tool in the discovery of important traits and genes involved in various biochemical pathways of medicinal plants. The release of high throughputs sequencing platforms led to the emergence of Next Generation Sequencing (Edwards & Batley, 2010). The term next generation sequencing which is more than a decade old referring to the methods or platforms that produce data at or beyond the genomic scale (Metzker, 2010). Next Generation Sequencing allows analysis of whole-genome to find and determine the genotype that leads to the difference observed in phenotype. This helps us to identify and combine the traits which are of greater importance. The use of NGS is not only limited to this, but it also helps in whole-genome sequencing of wild related species and provides species-specific data that can use it in generating species-specific PCR primers. By sequencing the products of PCR amplification of target genes, the genetic variation in large numbers of genes can be explored. With the help of NGS, novel allelic variations can be characterized in close or distant relatives. Figure 1 summarized the role of biotechnology in the improvement of medicinal plants.

Figure 1 Role of Biotechnology in the improvement of medicinal plant
Next generation sequencing has played a pivotal role in revolutionizing research in the field of genomics. The human genome project, which took thirteen years to sequence and decipher the entire human genome, used Sanger sequencing strategies. This entire feat can now be achieved in a single day using NGS technologies. NGS platforms work on the basic principle of sequencing millions of small fragments of DNA in parallel. The output is analyzed using various bioinformatics strategies. The individual reads are then mapped and arranged according to the reference genome available. NGS technologies are not only limited to sequence the whole genome but can also be used to sequence special areas of interest of a small number of individual genes or including all of the 22,000 coding genes (Behjati & Tarpey 2013). NGS platforms are now highly used to decipher the scattered pieces of the normal functioning of medicinal plants.

2 First-generation sequencing technology

Maxam-Gilbert and Sanger sequencing technologies are considered as the first generation sequencing technology (Sanger & Coulson, 1975). Sanger sequencing is also known as the dideoxynucleotide method or chain termination method. Although Sanger sequencing is widely used in the last few decades and is still used for low throughput DNA sequencing, it is time-consuming and extremely expensive. These limitations of the first generation sequencing led the way for the development of next generation or second generation sequencing. Table 1 shows the comparison between Sanger Sequencing and NGS Technologies.

3 Next Generation Sequencing Technologies

The availability of NGS as commercial platforms was not possible until 2005. The first platform in this segment was introduced by Solexa Technologies. Since then, various sequencing technologies have been developed and major improvements have been observed. Among various NGS technologies, 454 Genome Sequencer (Roche), Illumina Sequencing (Illumina), and Ion Torrent Sequencing (Thermo Fisher Scientific) are generally used (Cheng et al., 2014; Hu et al., 2015). In 2013, 454 Genome Sequencer was shut down by Roche as this technology became non-competitive as compared to other NGS technologies. Thus this technology is considered obsolete and no longer used for NGS analysis. Nanopore Sequencing (Oxford Nanopore Technologies) is also gaining popularity and is actively used for both amplicon sequencing and genome assembly with ultralong reads (Liu et al., 2012; Li & Weng, 2017). To date, Illumina and Ion Torrent Sequencing remain as the most preferred NGS techniques for improvement in medicinal plants because of their short read-length.

3.1 Illumina Sequencing

Illumina Genome Analyser was initially developed by Solexa which works on the principle of the sequence by synthesis (Figure 2). It uses solid-phase bridge amplification where 5' and 3' adapters are ligated to each end of the inserted DNA template (Liu et al., 2012). The library containing fixed adapters is denatured to form single strands and added to the flow cell. Clusters of DNA strands are generated by bridge amplification which takes place in a flow cell. Subsequently, these strands are denatured which results in single stranded forward and reverse strands of DNA. Four kinds of nucleotides (ddATP, ddGTP, ddTTP, ddCTP) containing different cleavable fluorescent dyes are then added to these strands for sequencing (Egan et al., 2012). The process of elongation continues after the fluorescent dye moiety is cleaved. This cycle is repeated until the DNA fragment has been synthesized to its target length. The mean read length ranges from 50 SE to 101 PE.

3.2 Ion Torrent Sequencing

Ion torrent is one of the unique NGS technologies as it is not based upon fluorescent dyes (Figure 3). The working principle of this system is based on measuring the pH change as a result of the release of an H+ ion during the incorporation of nucleotide using semiconductor technology (Liu et al., 2012). The machine detects the nucleotides which are added sequentially to the growing strand.

| Sample | Sanger Sequencing | NGS Technologies |
|--------|-------------------|------------------|
| Sequence chemistry | PCR products or plasmid after purification | NGS library made from genomic DNA |
| Data capture | Only homogeneous DNA sequence can be identified | Numerous DNA sequences are determined simultaneously |
| Advantage | High precision and simple data analysis | Efficient for massive sequence determination |
| Disadvantage | Analyzing heterogeneous amplified products in sample is difficult | For data interpretation, advanced bioinformatics analysis is required |
3.3 Pacific Biosciences

Pacific biosciences released its RS SMS platform in 2010 (Figure 4). In this platform, both the ends of target DNA molecules are ligated with hairpin adaptors to form an SMRT bell. DNA polymerase molecules are immobilized in Zero Mode Wavelength (ZMW) that restricts the volume of nucleotides and light emission to a small area surrounding the DNA polymerase (Liu et al., 2012). The polymerase binds to the hairpin adaptor and replications begin where four fluorescent labeled nucleotides are incorporated with the release of fluorescent signals. The light pulses in ZMW chambers are recorded and interpreted into base sequence results. PacBio is best known for sequencing unknown samples and library samples with co-existing variants.

3.4 Oxford Nanopore Technologies

Oxford nanopore technologies are one of the most advanced third generations sequencing technologies. GridION and PromethION are developed by this. A strand of DNA or RNA can be directly analyzed on this platform (Liu et al., 2012). The DNA strand to be sequenced is ligated with adapter and processing enzymes using library preparation kits. When the DNA-enzyme complex reaches the nanopore, double stranded DNA is unwound by helicase to create single stranded DNA and this will pass through the narrow aperture of the nanopore. A constant electric field is configured across the membrane and when DNA strands pass through the nanopore disruption of electric current occurs. This change in electric signal is recorded and interpreted into a base sequence of
DNA nucleotide (Schneider & Dekker, 2012). The schematic diagram for the workflow of Oxford Nanopore Technologies is represented in Figure 5. A comparison between all the four sequencing platforms viz., Illumina, Ion Torrent, PacBio, and Nanopore sequencing technologies has been represented in Table 2.

### 4 Flow of NGS analysis

Before sequencing the obtained samples, the fragmented DNA is first end-repaired and then ligated with specific adaptors to construct the library. Approaches like AFLP, RFLP, and target enrichment can be used to select target sequences for the preparation of the NGS library. As the size distribution and DNA concentration play a vital role in library construction they should be checked for quality control. If there are multiple samples involved then the number of samples that could be pooled depends on various factors such as (i) availability of unique molecular identifier UMIs number (ii) sequencing capacity of the NGS platform or technology chosen (iii) number of amplicons produced per sample and (iv) the required or predetermined sequencing depth (Hamady et al., 2008; Smith et al., 2010). When the raw NGS data is acquired from sequencers they are sent for data analysis. From the results of the analysis, it can be identified that

---

**Figure 5** The schematic diagram for the workflow of Oxford Nanopore sequencing

| Sample Preparation | Nanopore Sequencing | Signal Detection | Raw Signal Processing and Base Calling | Post Sequencing Processing |
|--------------------|---------------------|------------------|---------------------------------------|-----------------------------|
| Insolation and purification of DNA or RNA. | Disruption in ionic current | Change in ionic current is detected |
| Ligation of adapters. | |
| Attachment of tethers. | |

---

**Table 2** Comparison between NGS Platforms.

|                      | Illumina Sequencing | Ion Torrent Sequencing | PacBio Sequencing | Nanopore Sequencing |
|----------------------|---------------------|------------------------|------------------|---------------------|
| Major sequencing platform | MiSeq and HiSeq | Ion Proton and Ion PGM | Sequel and RSL | GridION and PromethION |
| Sequencing chemistry | Reversible dye terminators | pH value change to detect H+ ion | Single Molecule Real Time Sequencing | Changing magnitude of electrical conductivity across nanopore surfaces |
| Library preparation | Bridge amplification | Emulsion PCR | Blunt-end ligation with hairpin adaptors to form SMRT bell | Adaptor ligation and motor protein attachment |
| Read length | 100–150 bp (HiSeq), 400–500 bp (MiSeq) | 200–600 bp (Ion S5), ~200 bp (Ion Proton), 200–500 bp (Ion PGM) | 2000–5000 bp on average | Variable up to 900 kb |
| Sequencing time | 1–6 days | 2–4 hrs | 0.5–4 hrs | 1 min–72 hrs |
| Output data per run | 500–1000 GB | Up to 10 GB | 0.1–1 GB | 15 GB–8.6 TB |
| Advantages | High throughput and low cost with respect to per base | Low cost of the instrument and user friendly when compared with others | Long sequencing read length (~20 kb) | Long read DNA molecule can be sequenced without PCR amplification or chemical labeling, |
| Limitations | Short sequence reads around (~250 bp) are obtained | When met with homopolymer region sequencing error may occur | Lower throughput, higher error rate and higher cost per base | Error rate of using change in electric current to determine nucleotide sequences can be as high as 15% |

---

Journal of Experimental Biology and Agricultural Sciences http://www.jebas.org
the particular genes are involved or responsible for which particular trait, function, or pathway (Lo & Shaw 2019). Figure 6 represented the general workflow NGS technologies used for the improvement of medicinal plants.

5 Advancement in medicinal plant research with NGS technologies

This section of the review deals with the reported studies for the advancement of medicinal plants using NGS technologies. With the ever-increasing population and demands of medicinal plants, the need to look further into their genomes, transcriptome, epigenome, and interactome is very high (Hamilton & Robin Buell, 2012). NGS is used for dissecting the complexity of plant genomes. It was observed that medicinal plants were much more complex than model plants. Thus more study into the complexity would result in a better understanding of the plant. NGS is mainly used for studying comparative genomics (Hirsch & Robin Buell, 2013). Comparative genomics help to identify the novel chemical entities in the particular plant, by comparing its genome with its closely related species. *Rehmannia glutinosa*’s root is known to possess various bioactive molecules mainly iridoids. A reference genome of *R. glutinosa* was reported using Nanopore technology, Illumina, and Hi-C sequencing. The assembly genome is 2.49 GB long with a scaffold N50 length of 70 Mb and high heterozygosity (2%) (Ma et al., 2021). NGS is also used to reveal unidentified and novel bioactive principles (Sharma & Shrivastava, 2016). The chloroplast genome of *Ocimum gratissimum* which is 152,469 bp in length was sequenced by using Illumina paired-end sequencing data. It was found that the genome contains 138 unique genes, including 85 protein-coding, 45 tRNA, and eight rRNA genes (Balaji et al., 2021). Identification of the candidate genes expressed in a particular secondary metabolite pathway is a pivotal part of the whole transcriptomic study (Luo et al., 2012). *Panax notoginseng* is particularly valued for containing Triterpenoid saponins in their roots. A combination of single molecular real-time (SMRT) and Illumina Sequencing (SGS) technologies were used to map 2,383 full-length non-chimeric (FLNC) reads. 1,925 mis-annotated genes were corrected and merged into 927 new genes (Zhang et al., 2020). NGS is used to discover small and long non-coding RNA species. Non-coding RNAs are known to play an important role in the regulation of

---

**Figure 6** The schematic diagram for the workflow NGS technologies for improvement of medicinal plants.
Das & Mishra  protein synthesis. If these RNAs are discovered in medicinal plants, then it would be helpful to discover more about the transcriptional regulation of the proteins of biosynthetic pathways (Huang et al., 2013). The berberine biosynthesis pathway was investigated in *Berberis koreana* using the PacBio sequencer platform. Sequencing of the cDNA library along with isoform detection of expressed genes led to the identification of 23,246 full-length unigenes. (Roy et al., 2021). The involvement of NGS technologies provided the platform for large-scale analysis of transcriptomes from medicinal plants to identify the key genes associated with the biosynthesis of bioactive compounds (Chen et al., 2013). Table 3 comprises recent work done on medicinal plants using NGS technologies.

### 6 Limitations of Next-Generation Sequencing Technologies

Various strategies including template preparation, sequencing, imaging, and sequence assembly constitute NGS technologies. The main goals of NGS platforms are to achieve longer read lengths, higher sequence accuracy, and greater throughput. Although SGS produces high accuracy and higher throughput at the cost of read length but the currently achievable read lengths are low. The third generation of sequencing platform offers a very long read length and low GC bias but has modest throughput and low accuracy. Short read sequence data is a limitation of *de novo* sequence assembly. The library construction involves multiple steps and is very time-consuming and the gel-based size selection limits the sequence-ready library preparation significantly. During PCR amplification errors in the template sequence and amplification bias can be introduced. The challenges faced from massive high-throughput sequence data generated from NGS technologies are storage and quality control, data transfer, informatics operations to align or assembly read data, and laboratory information systems. Recently, the hybrid-sequencing strategy has been developed which uses the high-throughput and high-accuracy short reads data yielded from SGS platforms to correct errors in the long reads of PacBio. Hybrid sequencing approaches are used for understanding the complete expression of gene isoforms (Luo & Chen, 2006; Gaquere et al., 2013; Techen et al., 2014; Morree et al., 2014; Huang et al., 2015; Niazian et al., 2017; Rai et al., 2017a; Scossa et al., 2018).

### 7 Conclusion and Future Prospects

The potential of discovering new drugs is very high in medicinal plants but the research in this field continues to lag as compared to research and production of synthetic drugs for commercial use. Another reason for facing this backlash is that it requires more investment of time, money, and effort as compared to others. The computational approach towards these problems is the only viable solution. These computational approaches are used to analyze the diversity and importance of phytochemicals or other bioactive compounds present in the plant. It can be concluded that NGS has gained more attention because of its high throughput and low sequencing cost per base. So far, NGS has been used mainly for research purposes in medicinal plants. The ground-level steps such as...
as DNA extraction, sterile techniques, library preparation, reference DNA sequences, and bioinformatics analysis must be maintained and kept under strict supervision regularly.

The population on earth is increasing at such a rate that nature is not able to cope up and sustain itself on its own. The need and effort to feed and maintaining good health has become a challenge, especially in developing countries. The long-term consumption of synthetic medicines has caused serious health problems in both developing and developed countries. The need for the incorporation of plant-based products is very high. The phytochemicals are a promising material to provide a continuous and high quality of health. Most plants grow in the wild thus are exploited without being able to procreate in a short period. Thus they are getting threatened and their population is decreasing gradually placing them under the endangered plant category. Cultivation and other traditional breeding practices have become obsolete and are not useful to improve the neglected condition of these medicinal plants. Thus we can rely on biotechnology for help and NGS to identify and classify the wild population of medicinal plants.

Our ancestors have been using plants as remedies for several illnesses. According to the survey conducted by various researchers and others, a large part of the world population still relies on herbal medicines as a major source of health care. The biopharmaceutical companies exploit these plants and use them either as main active ingredients or as nutritional supplements. Integration of genomics and metabolomics has begun to revolutionize our perspective on herbal medicines. Generally, there are two aspects involved such as (i) the biosynthesis and pathway regulation of bioactive compounds of importance and (ii) the mechanism of action of the phytochemicals on human health and metabolism. The need of the hour is to develop a fast, more informative, and comprehensive analytical approach. These challenges can be overcome by developing computational metabolomics strategies, reconstructing the De Novo pathway, and by analysis of natural variation.

**Author contributions:** RM and SD conceived the idea and SD wrote the manuscript. RM critically reviewed the manuscript. All others have approved it.

**Acknowledgments:** Authors are thankful to the President, Centurion University of Technology and Management for his constant help and support. We apologize to the authors for the omission of their work that could not be included in this paper due to space constraints.

**Funding:** None.

**Conflict of interest:** The authors declare no conflict of interest.

**References**

Balaji R, Ravichandiran K, Tanuja, Parani M (2021) The complete chloroplast genome of Ocimum gratissimum from India—a medicinal plant in the Lamiaceae. Mitochondrial DNA Part B, 6(3), 948-950.

Behjati S, Tarpey PS (2013) What is next generation sequencing? Archives of disease in Childhood. Education and Practice Edition 98(6): 236–238.

Chen J, Dong X, Li Q, Zhou X, Gao S, Chen R, et al. (2013) Biosynthesis of the active compounds of Isatisindigotica based on transcriptome sequencing and metabolites profiling. BMC Genomics 14(1): 1-13. https://doi.org/10.1186/1471-2164-14-857.

Cheng X, Su X, Chen X, Zhao H, Bo C, Xu J, et al. (2014) Biological ingredient analysis of traditional Chinese medicine preparation based on high-throughput sequencing: the story for LiuweiDihuang Wan. Scientific Reports 4(1): 1-12. https://doi.org/10.1038/srep05147.

Edwards D, Batley J (2010) Plant genome sequencing: applications for crop improvement. Plant Biotechnology Journal 8(1): 2-9. https://doi.org/10.1111/j.1467-7652.2009.00459.x.

Egan AN, Schlueter J, Spooner DM (2012) Applications of next-generation sequencing in plant biology. American Journal of Botany 99(2): 175-185. https://doi.org/10.3732/ajb.1200020.

Gandhi SG, Mahajan V, Bedi YS (2015) Changing trends in biotechnology of secondary metabolism in medicinal and aromatic plants. Planta 241(2): 303-317. https://doi.org/10.1007/s00425-014-2232-x.

Gaquerel E, Kuhl C, Neumann S (2013) Computational annotation of plant metabolomics profiles via a novel network-assisted approach. Metabolomics 9(4): 904-918. https://doi.org/10.1007/s11306-013-0504-2.

Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nature Methods 5(3): 235-237. https://doi.org/10.1038/nmeth.1184.

Hamilton JP, Robin Buell C (2012) Advances in plant genome sequencing. The Plant Journal 70(1): 177-190. https://doi.org/10.1111/j.1365-313X.2012.04894.x

Hirsch CN, Robin Buell C (2013) Tapping the promise of genomics in species with complex, non-model genomes. Annual Review of Plant Biology 64: 89-110. https://doi.org/10.1146/annurev-arplant-050312-120237.
Hu L, Hao C, Fan R, Wu B, Tan L, Wu H (2015) De novo assembly and characterization of fruit transcriptome in black pepper (Piper nigrum). PLoS one 10(6): e0129822. https://doi.org/10.1371/journal.pone.0129822.

Huang W, Manglik A, Venkatakrishnan AJ, Laeremans T, Feinberg EN, Sanborn AL, et al. (2015) Structural insights into µ-opioid receptor activation. Nature 524(7565): 315-321. https://doi.org/10.1038/nature14886.

Huang Y, Zhang JL, Yu XL, Xu TS, Wang ZB, Cheng XC (2013) Molecular functions of small regulatory noncoding RNA. Biochemistry (Moscow) 78(3): 221-230. https://doi.org/10.1134/S0006297913030024.

Kalmankar NV, Venkatesan R, Balaram P, Sowdhamini R (2020) Transcriptomic profiling of the medicinal plant Clitoria ternatea: identification of potential genes in cyclotide biosynthesis. Scientific Reports 10(1): 1-20. https://doi.org/10.1038/s41598-020-69452-7.

Kotwal S, Kaul S, Sharma P, Gupta M, Shankar R, Jain M, Dhar MK (2016) De novo transcriptome analysis of medicinally important Plantago ovata using RNA-Seq. PLoS one 11(3): e0152073. https://doi.org/10.1371/journal.pone.0152073.

Li FS, Weng JK (2017) Demystifying traditional herbal medicine with modern approach. Nature Plants, 3(8): 1-7. https://doi.org/10.1038/nplants.2017.109.

Li MR, Shi FX, Zhou YX, Li YL, Wang XF, Zhang C, et al. (2015) Genetic and epigenetic diversities shed light on domestication of cultivated ginseng (Panax ginseng). Molecular Plant 8(11): 1612-1622. https://doi.org/10.1016/j.molp.2015.07.011.

Lin W, Sun F, Zhang Y, Xu X, Lu X, Li L, Xu R (2019) Comparative transcriptome and metabolite profiling of four tissues from Alisma orientale (Sam.) Juzepe reveals its inflorescence developmental and medicinal characteristics. Scientific Reports 9(1): 1-12. https://doi.org/10.1038/s41598-019-48806-w.

Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. (2012) Comparison of next-generation sequencing systems. Journal of Biomedicine and Biotechnology http://www.jebas.org

Liu Y, Wang Y, Guo F, Zhan L, Mohr T, Cheng P, et al. (2017) Deep sequencing and transcriptome analyses to identify genes involved in secoiridoid biosynthesis in the Tibetan medicinal plant Swertia mussotii. Scientific Reports 7(1): 1-14. https://doi.org/10.1038/srep43108.

Lo YT, Shaw PC (2019) Application of next-generation sequencing for the identification of herbal products. Biotechnology Advances, 37(8): 107450. https://doi.org/10.1016/j.biotechadv.2019.107450.

Luo H, Chen S (2006) Progress on the Transcriptome Analysis of Medicinal Plants with Next-Generation Sequencing Technologies. Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation 1-21. https://doi.org/10.1002/9780470027318.a9936.pub2.

Luo H, Li Q, Flower A, Lewith G, Liu J (2012) Comparison of effectiveness and safety between granules and decoction of Chinese herbal medicine: a systematic review of randomized clinical trials. Journal of Ethnopharmacology 140(3): 555-567. https://doi.org/10.1016/j.jep.2012.01.031.

Ma L, Dong C, Song C, Wang X, Zheng X, Niu Y, et al. (2021) De novo genome assembly of the potent medicinal plant Rehmannia glutinosa using nanopore technology. Computational and Structural Biotechnology Journal 19: 3954-3963.

Metczer ML (2010) Sequencing technologies—the next generation. Nature Reviews Genetics 11(1): 31-46. https://doi.org/10.1038/nrg2626.

Morgan K (2002) Medicine of the Gods: Basics Principles of Ayurvedic Medicine. Available at http://www.complink.co.uk/~mandrake/Ayurveda.htm access on 29th April 2020.

Moreel K, Saeys Y, Dima O, Lu F, Van de Peer Y, Vanholme R, Boerjan W (2014) Systematic structural characterization of metabolites in Arabidopsis via candidate substrate-product pair networks. The Plant Cell 26(3): 929-945. https://doi.org/10.1105/tpc.113.122242.

Niazian M (2019) Application of genetics and biotechnology for improving medicinal plants. Planta 249(4): 953-973.

Niazian M, Sadat Noori SA, Tohidfar M, Mortazavian SMM (2017) Essential oil yield and agro-morphological traits in some Iranian ecotypes of ajowan (Carum copticum L.). Journal of Essential Oil Bearing Plants 20(4): 1151-1156. https://doi.org/10.1080/0972060X.2017.1326849.

Nogueira M, Enfissi EM, Almeida J, Fraser PD (2018) Creating plant molecular factories for industrial and nutritional isoprenoid production. Current Opinion in Biotechnology 49: 80-87. https://doi.org/10.1016/j.copbio.2017.08.002.

Rai A, Saito K, Yamazaki M (2017a) Integrated omics analysis of specialized metabolism in medicinal plants. Available at https://doi.org/10.1111/tpj.13485 access on 29th April 2020.

Rai M, Rai A, Kawano N, Yoshimatsu K, Takahashi H, Suzuki H, et al (2017a) De novo RNA sequencing and expression analysis of Aconitum carmichaelii to analyze key genes involved in the biosynthesis of diterpene alkaloids. Molecules 22(12): 2155. https://doi.org/10.3390/molecules22122155.
Roy NS, Choi IY, Um T, Jeon MJ, Kim BY, Kim YD, Yu JK, Kim S, Kim NS (2021) Gene Expression and Isoform Identification of PacBio Full-Length cDNA Sequences for Berberine Biosynthesis in Berberis koreana. Plants. 10(7):1314. https://doi.org/10.3390/plants10071314.

Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology 94(3): 441-448. https://doi.org/10.1016/0022-2836(75)90213-2.

Scossa F, Benina M, Alseekh S, Zhang Y, Fernie AR (2018) The integration of metabolomics and next-generation sequencing data to elucidate the pathways of natural product metabolism in medicinal plants. Plantmedica 84(12/13): 855-873.

Sharma S, Shrivastava N (2016) Renaissance in phytomedicines: promising implications of NGS technologies. Planta 244(1): 19-38. https://doi.org/10.1007/s00425-016-2492-8.

Schneider GF, Dekker C (2012) DNA sequencing with nanopores. Nature Biotechnology 30(4): 326-328.https://doi.org/10.1038/nbt.2181.

Siahsar B, Rahimi M, Tavassoli A, Raissi A (2011) Application of biotechnology in production of medicinal plants. American-Eurasian Journal of Agricultural & Environmental Sciences 11(3): 439-444.

Smith AM, Heisler LE, St Onge RP, Farias-Hesson E, Wallace IM, Bodeau J, Nislow C (2010) Highly-multiplexed barcode sequencing: an efficient method for parallel analysis of pooled samples. Nucleic Acids Research 38(13): e142-e142. https://doi.org/10.1093/nar/gkq368.

Techen N, Parveen I, Pan Z, Khan IA (2014) DNA barcoding of medicinal plant material for identification. Current Opinion in Biotechnology 25: 103-110. https://doi.org/10.1016/j.copbio.2013.09.010.

Wang S, Chen J, Ma J, Jin J, Chen L, Yao M (2020) Novel insight into theacrine metabolism revealed by transcriptome analysis in bitter tea (Kucha, Camellia sinensis). Scientific Reports 10(1): 1-11. https://doi.org/10.1038/s41598-020-62859-2.

Yang L, Ding G, Lin H, Cheng H, Kong Y, Wei Y, et al (2013) Transcriptome analysis of medicinal plant Salvia miltiorrhiza and identification of genes related to tanshinonelbiosynthesis. PLoS One 8(11): e80464.

Zhang D, Li W, Chen ZJ, Wei FG, Liu YL, Gao LZ (2020) SMRT- and Illumina-based RNA-seq analyses unveil the ginsenoside biosynthesis and transcriptomic complexity in Panax notoginseng. Scientific Reports 10(1): 1-14.

Zhong F, Huang L, Qi L, Ma Y, Yan Z (2020) Full-length transcriptome analysis of Coptis deltoidea and identification of putative genes involved in benzylisoquinoline alkaloids biosynthesis based on combined sequencing platforms. Plant Molecular Biology 102(4): 477-499. https://doi.org/10.1007/s11103-019-00959-y.