11

OMIC Technologies in Bioethanol Production: An Indian Context

Pulkit A. Srivastava and Ragothaman M. Yennamalli*

Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Waknaghat, Himachal Pradesh, India

Abstract

Swift depletion coupled with hazardous effects of nonrenewable energy sources (such as petrol, coal, and diesel) on the environment has inspired all governments around the globe to participate in identification, development, and commercialization of the technology to produce environment-friendly renewable fuels. In this context, recent studies on bioethanol have reflected its viability to be one such energy source due to its production by renewable and sustainable biomass with comparatively less emissions than fossil fuels. Meanwhile, advancement in next-generation sequencing technology and genetic engineering has provided an opportunity to not only inspect into omics data but also to produce hyper-productive strains for efficient degradation of lignocelluloses material at the industry level. This chapter talks about the availability of omics data for microorganism strains and biomass required for bioethanol production in India (third largest energy consumer, second most populated, and counted among the fastest developing countries).

Keywords: Agricultural bioinformatics, biofuel, genome and proteome data, future trends

11.1 Introduction

Natural fossil fuels like natural gas, coal, petroleum, and diesel have been a primary source for the world's economy. However, the over consumption of fossil fuels has not only resulted in depletion of these nonrenewable

**Corresponding author*: ym.ragothaman@juit.ac.in; ragothaman@gmail.com

Rintu Banerjee, Garlapati Vijay Kumar, and S.P. Jeevan Kumar (eds.) OMICS-Based Approaches in Plant Biotechnology, (217–244) © 2019 Scrivener Publishing LLC

resources but also played a major role in the increase of greenhouse gases (GHGs), global warming, and climate change. These limitations accompanied by an increase in the world's energy requirement have made an emphasis to search for alternate energy sources with the following characteristics: renewable, efficient, sustainable, and cost effective with negative or less emission of GHGs [1]. Fuels generated from plant biomass can be considered as substitute of fossil fuels, termed as biofuel. Although, biofuel is similar to hydrocarbons with respect to emission, it can be more environment-friendly if the production and distribution of biofuel are taken care of.

The very first advantage that biofuels hold over conventional fuels is that they are biological molecules and hence biodegradable. This means that they will not persist in the environment for long durations and would not make an area uninhabitable. Secondly, biofuel production can be manipulated to eliminate sulfur to reduce the net impact of acid rain. Further, the limitations imposed by GHG on use of fossil fuels can be easily overcome by the use of biofuel. Lastly, biofuel production would decrease the energy dependence, i.e., countries having land resources to grow biofuel feedstock could produce their own energy and would no longer be dependent on fossil fuels that are limited to few geographical locations in the world.

The production of biofuel has changed over the course of time and biofuel can be broadly classified into three generations, namely, the first, second, and third generations. First-generation biofuel is made using feedstock that is also consumed as human food like sugar, vegetable oil, and/or starch. However, the food vs. fuel debate paved the path for second-generation biofuel, where only sustainable feedstock is considered for biofuel productions, which are either a non-food crop or plant products that are no longer useful for consumption. Since the second-generation biofuel feedstock covers a wider range from forestry to agriculture and waste material enriched in lignocellulosic content that is difficult to extract, it involves advanced conversion technologies. Therefore, the quality of second-generation biofuel is majorly driven by the production routes opted by the industry generally having the following sequence: pretreatment, hydrolysis, fermentation, distillation, and purification of bioethanol.

Moreover, the outcomes of recent advancement and experimentation on algae-based biofuel production in terms of diversity and quantity with lower resource inputs as compared to that of second-generation biofuel production have given more impetus toward the usage of algae for biofuel production. Besides the diversity of carbon sources they can use and quantity they can produce, the simplified path to genetically manipulate them to produce any kind of biofuel is most remarkable. Hence, algae-based biofuel production marks the third generation. However, they require more understanding and further scientific research.

Bioethanol is a biofuel that can be easily produced by renewable and sustainable plant resources that it can be extensively used for transportation with a possibility of minimized particulate production in compression–ignition engines [47]. In Europe and in the United States, bioethanol has been used since the 1900s but was hampered due to its high production cost. However, after the oil crisis in the 1970s, production of bioethanol was resumed with renewed interest. Although bioethanol has lower energy content than petrol (68%), its cleaner combustion and reduction of CO_2 emission level (by 80%) enable it to be a viable alternate source of energy. Also, studies conducted in the past couple of decades have also shown that replacement of fossil fuels by ethanol-based fuels has resulted a drastic reduction in net average of GHGs by 71% for cellulosic ethanol and 31% for bioethanol [2].

Therefore, the scenario presented by second- and third-generation biofuel (mainly bioethanol) production provides an ample area where even a small improvement or discovery could give a huge impetus toward the efficient production of bioethanol. The major areas that need scientific attention are raw materials, organisms that can produce enzymes being used in pretreatment process, and availability of omics data for further research. Hence, the chapter focuses on the availability of omics data and biomass found specifically in India for bioethanol production and the bacterial organism indigenous to India having cellulolytic enzymes.

11.2 Indian Scenario

Bioethanol in India is generally derived from molasses (by-product of sugar production process). However, with the increase in the energy demand of India and its population, it is impossible that molasses alone would be sustainable for bioethanol production. Besides, the debate on first-generation biofuels as they are derived from food crops has raised questions about net energy, water utilization, and net GHG balance. These criticisms have led to increased interest in second-generation biofuels [3]. In India, more than half of its land is used for agriculture. Also, it is the world's first largest producer of rice (https://www.worldatlas.com/articles/the-countries-producing-the-most-rice-in-the-world.html) and the second largest producer of sugarcane (https://www.worldatlas.com/articles/top-sugarcane-producing-countries.html), wheat (https://www.worldatlas.com/articles/top-wheat-producing-countries.html), and sorghum (https://www.worldatlas.com/articles/top-sorghum-producing-countries-in-the-world.html). This leads to production of massive amounts of crop residues that could be used for second-generation biofuel

production. Most of these crop residues are burned by farmers that leads to air pollution. Specifically, this event has been seen in the northern states of India. Agricultural waste has not been given much attention and usually ends being composted or grazed by animals. Agricultural waste, such as rice husk, has been traditionally being burned in the states of Haryana, Punjab, UP, and Rajasthan. This burning is not only an inefficient method of eliminating biomass; it also leads to increased levels of particulate matter in the northern parts of India. In November 2015, the state of Delhi reported the highest particulate matter between the ranges 368–122, much higher than the safe limits (http://www.dailypioneer.com/todays-newspaper/delhi-gasps-for-fresh-air-as-particulate-matter-soars-by-350.html and https://www.nature.com/news/new-delhi-car-ban-yields-trove-of-pollution-data-1.19385). The lack of a supply-chain mechanism of utilizing the biomass left after harvesting is a plausible reason as to why the farmers resort to burning. Thus, industries that can utilize the agricultural waste are a potential area of development that needs attention. Moreover, sustainably produced second-generation biofuel can potentially promote rural development and improve economic conditions in developing regions.

There are two essential requirements for successful production of bioethanol in a developing country like India: availability of second-generation plants and crop residues that are sustainable, and identification of native strains of microorganisms (preferably bacteria) that produce cellulases. The latter is most important as characterization and further manipulation of catalysts via protein engineering could lead to preparing and commercialization of enzymatic cocktails. While enzymatic cocktails are available from companies such as Novozymes, there is a lack of enzymatic cocktail of Indian origin that in our opinion has the ability to revolutionize bioethanol production in India.

11.3 Cellulolytic Enzymes Producing Bacterial Strains Isolated from India

An impediment toward the conversion of plant biomass into bioethanol is the enzymatic hydrolysis of lignocellulosic material. It is a process whereby lignocellulosic content from plant biomass is converted into fermentable sugars (Figure 11.1) through the use of enzymes like cellulases and xylanases. However, production of these enzymes at industrial scale by bacterial and fungal strains involves relatively high cost, and hence

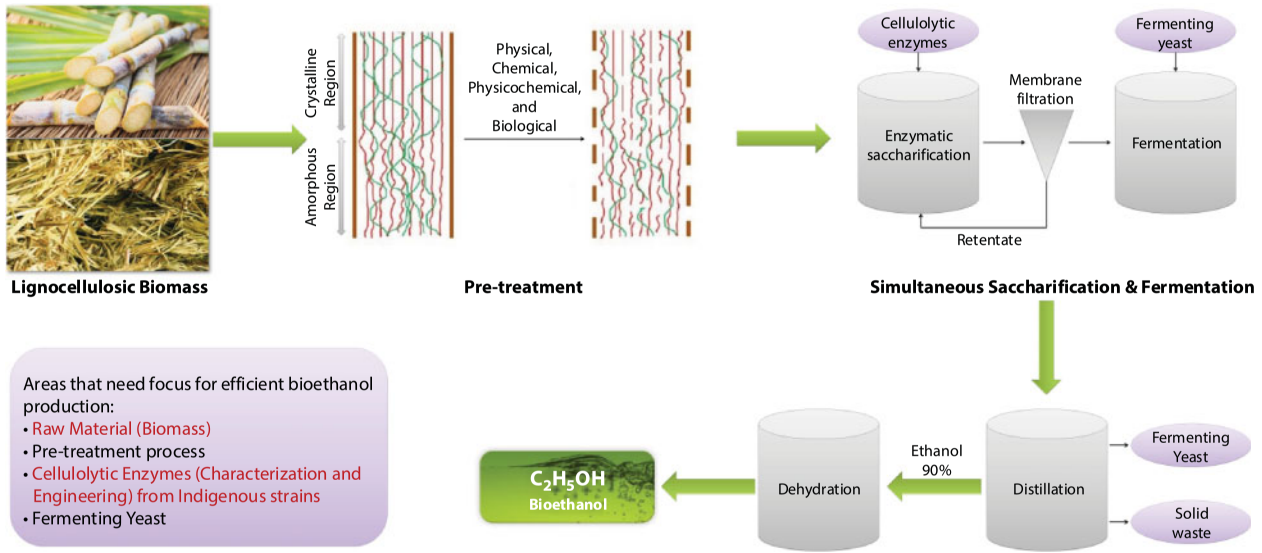

Figure 11.1 Schematic diagrams explaining the bioethanol production from second-generation biomass. Conversion of second-generation sustainable biomass (described in the text) to bioethanol involves multiple steps, where pretreatment renders the cellulose fibers to be exposed and to detach from other recalcitrant polymers (hemicelluloses and lignin). Further saccharification and fermentation followed by distillation delivers the bioethanol that can be blended with gasoline/petrol. However, certain areas require focused attention, with respect to India, such as procurement of biomass and preparing an enzymatic cocktail, which involves complete characterization and protein engineering of cellulases identified from strains isolated in India.

identification of new and active enzymes besides improvement in the efficiency of already known enzymes is one of the major areas of research in bioethanol production.

Cellulases working synergistically with xylanases are the most common enzymes produced by fungal and bacterial strains. But enzymes like amylase, holocellulase, and pectinase not only increase the range of plant biomass that can be degraded but also reduce the cost if they can be produced by only one organism. The greater interest lies in the enzymes produced from thermophiles ranging from 50°C to 90°C as it can be used for harsh industrial process. Moreover, at the industrial scale, it is important to produce copious amount of cellulases, and hence omics study of organisms that can be easily engineered genetically besides having higher growth rate plays a crucial role (bacteria). Here, we give an overview of various promising microorganism strains indigenous to India for bioethanol production. Table 11.1 provides a list of bacterial organisms that can produce the plant biomass-degrading cellulolytic enzymes (such as endo/exoglucanases, cellobiohydrolases, β-glucosidases, and others) with its genome availability.

11.3.1 *Bacillus* Genus of Lignocellulolytic Degrading Enzymes

Bacillus, a genus of gram-positive and rod-shaped bacteria, has around 20 different species (Table 11.1) capable of producing lignocellulolytic degrading enzymes isolated from India. Many of these have the ability to grow at varying pH and temperature making them attractive for industrial purpose of cellulase production.

Among the 20 strains listed, two strains of *Bacillus cereus* and *Bacillus amyloliquefaciens*, respectively, can produce a variety of enzymes like β-glucosidase, cellulase, pectinase, xylanase, and β-1,4-endoglucanase for hydrolysis of plant biomass and have been extensively studied.

11.3.2 *Bhargavaea cecembensis*

Gram-positive, nonmotile, nonspore forming, and rod-shaped bacteria isolated from the Indian Ocean are reported about their cellulase-producing capability [4]. Named after the late Indian scientist, PM Bhargava, the strain of *Bhargavaea cecembensis* was isolated from the hot spring of Manikaran (India). This thermophillic organism is potentially a source of cellulase that can be scaled up for industrial-scale production.

Table 11.1 List of microorganism strains and enzymes isolated reported to be indigenous to India and are proposed to be industrially important for bioethanol production.

Enzyme	Name	Genome	Strain	Reference
Cellulase	*Agrobacterium larrymoorei*	Present	A1	[24]
Amylase, cellulase, xylanse	*Alkalibacterium* sp.	Absent	JN645863	[11]
Endoglucanase, exoglucanase, xylanase	*Amycolatopsis* sp.	Absent	GDS	[25]
Cellulase, xylanase	*Aneurinibacillus aneurinilyticus*	Present	DBT87, DBT15	[26]
Cellulase, β-glucosidase, pectinase, xylanase, β-1,4-endoglucanase	*Bacillus amyloliquefaciens*	Present	IARI-SP-2 KF240848, IARI-R-25	[27, 28]
Cellulase	*Bacillus aquimaris*	Present	NT1	[29]
Cellulase, β-glucosidase, xylanase	*Bacillus baekryungensis*	Absent	IARI-ABL-36	[28]
Cellulase	*Bacillus brevis*	Present	VS-1	[30]
Cellulase	*Bacillus cellulasensis*	Present	NIO-1130T HQ858013	[31]
Amylase, cellulase, pectinase, xylanase, β-glucosidase, β-1,4-endoglucanase	*Bacillus cereus*	Present	JX312575, JN411376, IARI-SP-7 KF240853, DBT14, IARI-L-73	[26–28, 32]

(*Continued*)

Table 11.1 List of microorganism strains and enzymes isolated reported to be indigenous to India and are proposed to be industrially important for bioethanol production. (*Continued*)

Enzyme	Name	Genome	Strain	Reference
Amylase, β-glucosidase, pectinase	*Bacillus firmus*	Present	JN411422, IARI-L-21	[28, 32]
Amylase, cellulase, pectinase, xylanase, β-glucosidase	*Bacillus flexus*	Present	JN645852, IARI-ABL-37	[11, 28]
Amylase, β-glucosidase, pectinase, β-1,4-endoglucanase	*Bacillus licheniformis*	Present	JX312587, IARI-SP-3 KF240849, IARI-ABL-38	[27, 28, 32]
Amylase, cellulase, pectinase, xylanase, β-glucosidase	*Bacillus marisflavi*	Present	IARI-ABL-39	[28]
β-1,4-Endoglucanase	*Bacillus megaterium*	Present	JF343146, IARI-SP-9 KF240855	[27]
Amylase, β-glucosidase, pectinase	*Bacillus mojavensis*	Present	IARI-ABL-40	[28]
β-Glucosidase, pectinase	*Bacillus muralis*	Present	IARI-L-74	[28]
β-1,4-Endoglucanase	*Bacillus mycoides*	Present	JN411283, IARI-SP-8 KF240854	[27, 32]
β-1,4-Endoglucanase	*Bacillus oceanisediminis*	Present	IARI-SP-10 KF240856	[32]

(*Continued*)

Table 11.1 List of microorganism strains and enzymes isolated reported to be indigenous to India and are proposed to be industrially important for bioethanol production. (*Continued*)

Enzyme	Name	Genome	Strain	Reference
Cellulase, β-glucosidase	*Bacillus pumilus*	Present	JN411426, NAIMCC-B01415, IARI-SP-5 KF240851, IARI-L-54	[28, 32, 33]
Amylase, cellulase, xylanase	*Bacillus selenatarsenatis*	Present	JN645858	[11]
β-1,4-Endoglucanase	*Bacillus* sp.	Absent	SV1, 313SI, IARI-SP-4 KF240850	[27, 34, 35]
Amylase, xylanase, β-glucosidase, pectinase	*Bacillus subtilis*	Present	JF343233, JN411353, LFS3, IARI-SP-1 KF240847, IARI-L-69	[27, 28, 32, 36]
β-1,4-Endoglucanase	*Bacillus thuringiensis*	Present	IARI-SP-6 KF240852	[27]
Cellulase	*Bhargavaea cecembensis*	Present	JX312640	[4]
Cellulase	*Cellulomonas cellulans*	Present		[37]
Xylanase	*Cellulosimicrobium cellulans*	Present	CKMX1	[38]
Cellulase	*Citrobacter freundii*	Present		[39]

(*Continued*)

Table 11.1 List of microorganism strains and enzymes isolated reported to be indigenous to India and are proposed to be industrially important for bioethanol production. (*Continued*)

Enzyme	Name	Genome	Strain	Reference
Amylase, cellulase, β-glucosidase, pectinase	*Desemzia incerta*	Present	IARI-L-46	[28]
Amylase, β-glucosidase, pectinase, xylanase	*Exiguobacterium antarcticum*	Present	IARI-L-70	[28]
Amylase, cellulase, β-glucosidase, pectinase, xylanase	*Exiguobacterium indicum*	Present	IARI-R-137	[28]
Amylase, β-glucosidase, pectinase	*Exiguobacterium undae*	Present	IARI-L-116	[28]
Amylase, cellulase, xylanase	*Halobacillus trueperi*	Present	JN645845	[11]
Amylase, cellulase, xylanase	*Halomonas caseinilytica*	Present	JN645870	[11]
Amylase, cellulase, xylanase	*Halomonas muralis*	Present	JN645872	[11]
Amylase, β-glucosidase	*Jeotgalicoccus halotolerans*	Absent	IARI-ABR-5	[28]
Cellulase	*Marinobacter* sp.	Absent	MSI032	[40]
Cellulase	*Paenibacillus barcinonensis*	Absent	MG7	[41]
Cellulase, endoglucanase	*Paenibacillus pabuli*	Present	JF343230	[32]

(*Continued*)

Table 11.1 List of microorganism strains and enzymes isolated reported to be indigenous to India and are proposed to be industrially important for bioethanol production. (*Continued*)

Enzyme	Name	Genome	Strain	Reference
Amylase, cellulase, β-glucosidase	*Paenibacillus* sp.	Present	IHBB 10380, IHB B 3415	[42, 43]
Amylase, cellulase, β-glucosidase, pectinase, xylanase	*Paenibacillus terrae*	Present	JF343212, IARI-L-127	[28, 32]
Amylase, β-glucosidase, pectinase	*Paenibacillus xylanexedens*	Present	IARI-L-76	[28]
Amylase, cellulase, xylanase	*Planococcus antarcticus*	Present	IARI-ABL-9	[28]
β-Glucosidase, pectinase	*Planococcus donghaensis*	Present	IARI-L-39	[28]
Amylase, β-glucosidase, pectinase	*Planococcus kocurii*	Present	IARI-ABL-3	[28]
Amylase, cellulase, β-glucosidase, pectinase, xylanase	*Pontibacillus* sp.	Absent	IARI-ABL-41	[28]
Cellulase	*Pseudomonas fluorescens*	Present		[44]
Amylase, cellulase, xylanase	*Salicola marasensis*	Absent	JN645869	[11]
Amylase, cellulase, xylanase	*Sediminibacillus halophilus*	Present	JN645856	[11]
Cellulase, xylanase	*Serratia rubidaea*	Present	DBT4	[26]
Amylase, β-glucosidase, pectinase	*Sinobaca beijingensis*	Absent	IARI-ABL-18	[28]

(*Continued*)

Table 11.1 List of microorganism strains and enzymes isolated reported to be indigenous to India and are proposed to be industrially important for bioethanol production. (*Continued*)

Enzyme	Name	Genome	Strain	Reference
Amylase, β-glucosidase, pectinase	*Sporosarcina aquimarina*	Absent	IARI-L-77	[28]
Amylase, cellulase, β-glucosidase, pectinase	*Sporosarcina globispora*	Present	IARI-R-111	[28]
Amylase, cellulase, pectinase	*Sporosarcina psychrophila*	Present	IARI-R-110	[28]
Amylase, β-glucosidase, pectinase	*Staphylococcus arlettae*	Present	IARI-ABL-33	[28]
Amylase, β-glucosidase, pectinase	*Staphylococcus xylosus*	Present	IARI-ABR-1	[28]
Xylanase	*Stenotrophomonas maltophilia*	Present	X6 JX220726	[45]
Endoglucanase, cellulase	*Streptomyces albogriseoulus*	Present	C2-221	[7]
Cellulase	*Streptomyces albospinus*	Absent	MTCC 8768	[8]
Cellulase	*Streptomyces griseorubens*	Present	AB184139	[9]
Cellulase	*Streptomyces matensis*	Absent	KF553639	[10]
Endoglucanase, cellulase	*Streptomyces nitrosporeus*	Absent	N2-14	[7]

(*Continued*)

Table 11.1 List of microorganism strains and enzymes isolated reported to be indigenous to India and are proposed to be industrially important for bioethanol production. (*Continued*)

Enzyme	Name	Genome	Strain	Reference
Holocellulase, cellulase, endoglucanase	*Streptomyces* sp.	Absent	ssr-198, S7, M7a, M7b	[5–7]
Cellulase, xylanase	*Streptomyces viridiviolaceus*	Absent	JN645842	[11]
Amylase, cellulase	*Thalassobacillus* sp.	Absent	JN645857	[11]
Amylase, cellulase, β-glucosidase, xylanase	*Virgibacillus halodenitrificans*	Present	IARI-ABR-18	[28]

11.3.3 *Streptomyces* Genus for Hydrolytic Enzymes

Besides being the largest genus for antibiotic production, several species of *Streptomyces* isolated from India have shown to have cellulase-producing ability [5–11]. *S. albogriseoulus*, *S. albospinus*, *S. griseorubens*, *S. matensis*, *S. nitrosporeus*, and *S. viridiviolaceus* have been reported in various reports exhibiting cellulase-producing capability.

Although a significant effort has been made to reduce the production cost of cellulase at industrial scale, it still equals twice the cost of producing ethanol from starch. Hence, realizing the paramount importance of cellulolytic enzymes producing organism, it is necessary to explore new organism and cocktails of enzymes that can be used for production at industrial scale. Besides, modulating transcriptional regulators to produce the desired cocktail of lignocellulolytic degrading enzymes with high yield and productivity can also be an alternative for cost reduction. However, to attain this, one must have in-depth knowledge of the organism's genome, and hence its public availability plays a crucial role.

11.4 Biomass Sources Native to India

In other countries, where bioethanol blending has been successfully implemented (such as the USA and Brazil), second-generation biomass has been actively researched upon and new technologies are developed and simultaneously commercialized for a sustainable bioethanol production. One of the successful crops that have garnered massive attention is switchgrass, a perennial prairie plant mostly found in Midwestern United States. In the case of India, multiple labs have focused on other perennial plants that are not cultivated for food. Here, we list some of those grasses/plants that are found geographically in India and highlight the advances as well.

11.4.1 *Albizia lucida* (Moj)

Albizia lucida, known as Moj, is native to the northeastern part of India. It is a fast-growing tree that is majorly used as fuel by local tribal people. Physiochemical analysis of the plant is carried out to check its use in biofuel production, and it was found that it can be a good alternate for bioethanol production. Specifically, its high cellulose with lower lignin makes it an attractive biomass. Further, the crystallinity content of Moj analyzed using the Ash test was found to be 46.43% and degrades at temperature of 355°C [12].

11.4.2 *Areca catechu* (Betel Nut)

Areca catechu is a species of palm that majorly grows in Asia, parts of east Africa, and tropical Pacific. In India, it is used extensively for cooking and heating. Interestingly, the husk of the betel nut palm is proposed to be used as biomass source. Compared to Moj, the betel nut husk contains higher crystallinity of cellulose and hemicelluloses of 63.84% [12].

11.4.3 *Arundo donax* (Giant Reed)

Arundo donax is a perennial dry energy grass native to Middle East Asia, Indian subcontinent, and Mediterranean Basin. Similar to *Miscanthus*, it has a rhizoidal root system that is tough and forms an extensive knotty network through which it propagates further [13]. It has an ability to grow in erosive and marginal, degraded soils with remarkable increase in biomass in response to nitrogen fertilizers. Additionally, it can improve soil fertility and has higher yield of bioethanol, making it a prominent candidate for bioethanol production [14].

11.4.4 *Pennisetum purpureum* (Napier Grass)

The hybrid variety of Napier grass developed by Tamilnadu Agricultural University, Coimbatore, has a higher biomass yield compared to other perennial grasses. While it has been bred for feedstock purposes, it is highly tolerant to salinity and reduces the pH of the saline soil [15].

11.4.5 *Brassica* Family of Biomass Crops

Brassica carinata (Ethiopian mustard) is a newly introduced semiarid crop in India that is agriculturally significant as it is used as a source for jet engine aviation biofuel [16, 17]. The environmental performance of the biofuel in comparison to conventional gasoline was estimated to be better for ethanol-based application in passenger cars [17].

Another member of *Brassica*, *Brassica napus* (rapeseed), a popular crop for edible oil production, has also been proposed to be a sustainable biomass source for bioethanol production. The investigation of the various factors that influence the saccharification of rapeseed stalks concluded that the lower the quantity of pectin and arabinogalactans, the better the saccharification [18].

Multiple studies have been reported on the common mustard plant, a widely grown agricultural crop in India and South Asia. Specifically, the structural features of dilute acid, steam exploded, and alkali-pretreated mustard stalks and their impacts on enzymatic hydrolysis are investigated [19].

11.4.6 *Cajanus cajan* (Pigeon Pea)/*Cenchrus americanus* (Pearl Millet)/*Corchorus capsularis* (Jute)/*Lens culinaris* (Lentil)/*Saccharum officinarum* (Sugarcane)/*Triticum* sp. (Wheat)/*Zea mays* (Maize)

These crops are grown in huge quantity in India covering almost all parts of the country. Various parts of plants like straw, stalk, waste, and husk are utilized for the production of bioethanol. The roadmap for biofuel production in India [3] shows the total net ethanol availability in 2010/2011 as 27.9 billion liters and estimates an increase in the availability (37.3% by 2020/2021) with the increase in production of these plants. As stated earlier, among the world's largest producer of rice, pigeon pea, pearl millet, jute, sugarcane, wheat, and maize, there are high rates of having biomass resource from various parts that are not edible or form waste.

11.4.7 *Medicago sativa* (Alfalfa)

Alfalfa is a perennial flowering plant of pea family grown almost all around the world. Due to its rich lignocellulosic content and being herbaceous non-edible energy crop, it is considered as a potential source of biomass for bioethanol production. Though not produced in large quantity, West and South-Central India is an ideal location for the production of alfalfa.

11.4.8 *Manihot esculenta* (Cassava)/*Salix viminalis* (Basket Willow)/*Setaria italica* (Foxtail Millet)/*Setaria viridis* (Green Foxtail)

With the efficient and cost-effective lignocellulosic bioethanol production, bioethanol industry would need a more reliable plant biomass that can be produced using minimum water, fertilizer, and agricultural land. To address this issue, various C4 energy plants could be used for bioethanol production as they use an efficient system when compared to C3 plants, hence yielding high amount of biomass. In a recent article [20], the authors investigated the optimization techniques for biomass production and feedstock quality with respect to bioethanol production in C4 plants.

11.4.9 *Vetiveria zizanioides* (Vetiver or Khas)

In many parts of India, Khas grass grows wild and is an important plant for industrial and medicinal uses. Khas/Vetiver's roots are used for cooling purposes, and oil is extracted as well, which is prized for its cooling effect (http://www.vetiver.org/).

Harvesting of Khas/Vetiver's roots is done by uprooting the entire plant and removing the roots after cleaning. Interestingly, systematically cultivated Khas/Vetiver is harvested by removing the roots after 10–12 months, mostly for medicinal purposes. However, the pulp of Khas/Vetiver goes into paper and straw board production.

11.4.10 Millets and *Sorghum bicolor* (Sorghum)

Although there are plants and crop wastes that are being utilized for biofuel production in India (like wheat bran, wheat straw, sugarcane bagasse, rice straw, and rice husk), there are plants that have been investigated by few labs for their utilization in biofuel production. However, the lack of omics (genome, transcriptome, and proteome) data for these plants is acting as a roadblock in the efficient production of bioethanol. Table 11.1 lists a total of 39 Indian plants that have been investigated for their bioethanol production capability and availability of the omics data on the National Center for Biotechnology Information (NCBI) portal. As discussed in Ref. [3], wastes from sugarcane, wheat, maize, jute, pearl millet, pigeon pea, and lentil are being extensively used for bioethanol production. However, plants like Moj, Bonbogori and German grass, foxtail millet, basket willow, and green foxtail can be investigated in depth to evaluate their efficiency in the production of bioethanol. Specifically, sorghum has received a lot of attention recently [21, 22]. Therefore, availability of genome, transcriptome, and proteome of plants not only would pave a better way to understand and overcome the complexity involved in bioethanol production at the industrial level but also would provide various reliable resources from all over the country to meet the challenge of fuels. Moreover, in a country like India with vast land resources to grow biofuel feedstock, it has a potential to not only produce the energy for itself but also distribute the bioethanol to other countries.

11.5 Omics Data and Its Application to Bioethanol Production

"Omics" is a general term applied to biological data that result after sequencing the genome of a particular organism. Derived from the suffix "ome," omics has multiple flavors, such as transcriptomics, metabolomics, etc. With the advent of next-generation sequencing techniques, it has been a matter of hours for any organism's genome to be elucidated. The implication of genome sequence availability is quite long reaching. For example, once the genome of an organism is known, the functional genomics

(process of identifying all the genes and annotating their functions) step enables to study specific genes that influence in a specific pathway.

In the case of bioethanol production, the availability of omics data is a major limitation. Among the ~65 strains isolated from India, there are only 49 species' genome available (73% of strains) in the NCBI. While published reports emphasize on the specific enzyme that was isolated and characterized from the strains, they do not report the genome availability or sequencing efforts. The genome data available for most of the ~65 strains of Indian origin can be analyzed with standard organism's genome available in the NCBI using BLAST and other genome comparative tools. As a general rule of thumb, these genome data can be used as the strain's genome, with the caveat that there might be some unique genes present in the Indian strain that are missing in the standard genome's data. To overcome this limitation, it would be ideal to have the strain's genome sequenced by NGS and deposited in the NCBI. Also, many of the metagenome efforts have led to identification of indigenous strains that have possible implications in bioethanol production in India.

The other way around would be a metagenomic study of microorganisms with characteristics like low pH and sustainability at higher temperature. The strains that produce cellulases and possess any one of the above characteristic can be further explored, i.e., what regulatory factors play a critical role in production of cellulase and other plant-biomass degradation. However, all research to manipulate the production of cellulase at industrial scale requires the presence of omics data. Additionally, availability of genomic information would also increase the computational studies so that it can be well integrated with experimental studies to not only increase the production of cellulase but also at the same time decrease the experimental cost.

Similarly, among the biomass sources that are native to India, there are only 15 genome data available [such as *Brassica juncea* (mustard), *Brassica napus* (rapeseed), *Brassica rapa* (field mustard), *Cajanus cajan* (pigeon pea), *Cenchrus americanus* (pearl millet), *Corchorus capsularis* (jute), *Gossypium barbadense* (Brazilian cotton), *Gossypium hirsutum* (upland cotton), *Manihot esculenta* (cassava), *Populus euphratica*, *Populus trichocarpa* (black cottonwood), *Setaria italica* (foxtail millet), *Sorghum bicolor* (sorghum), *Triticum* sp. (wheat), and *Zea mays* (maize)]. Among these 15, only 10 organisms' proteome data are available. However, there are 21 transcriptome data available for the biomass sources. For strain engineering and development, it is essential that the organism's genome, transcriptome, and proteome are all available. Thus, the availability of all three data for sorghum, maize, upland cotton, and Brassica species will enable for further research (Table 11.2).

Table 11.2 Availability of genome, proteome, and transcriptome data for biomass sourced from India.

S. No.	Name (common name)	Genome	Transcriptome	Proteome	References
1	*Albizia lucidia* (Moj)	Absent	Absent	Absent	[12]
2	*Areca catechu* (betel nut)	Absent	Present	Absent	[12]
3	*Arundo donax* (giant reed)	Absent	Absent	Absent	[13, 46]
4	*Bambusoideae* (bamboo)	Absent	Present	Absent	–
5	*Brassica carinata*	Absent	Present	Absent	[17]
6	*Brassica juncea* (Mustard)	Present	Present	Absent	[18]
7	*Brassica napus* (rapeseed)	Present	Present	Present	[18]
8	*Brassica rapa* (field mustard)	Present	Present	Present	[19]
9	*Cajanus cajan* (pigeon pea)	Present	Absent	Absent	[3]
10	*Cenchrus americanus* (pearl millet)	Present	Present	Absent	[3]
11	*Chrysopogon zizanioides* (Vetiver)	Absent	Absent	Absent	–
12	*Corchorus capsularis* (jute)	Present	Absent	Present	[3]
13	*Cynodon dactylon* (Bermuda grass)	Absent	Absent	Absent	[46]

(*Continued*)

Table 11.2 Availability of genome, proteome, and transcriptome data for biomass sourced from India. (*Continued*)

S. No.	Name (common name)	Genome	Transcriptome	Proteome	References
14	*Echinochloa polystachya* (German grass)	Absent	Absent	Absent	[3]
15	*Gossypium barbadense* (Brazilian cotton)	Present	Present	Absent	–
16	*Gossypium hirsutum* (upland cotton)	Present	Present	Present	[48, 49]
17	*Helianthus tuberosus* (sun root)	Absent	Absent	Absent	[50]
18	*Hibiscus cannabinus* (Kenaf)	Absent	Present	Absent	[51]
19	*Lantana camara* (red sage)	Absent	Absent	Absent	[52]
20	*Lens culinaris* (lentil)	Absent	Present	Absent	[3]
21	*Linum perenne* (blue flax)	Absent	Absent	Absent	–
22	*Manihot esculenta* (cassava)	Present	Present	Present	[20]
23	*Medicago sativa* (alfalfa)	Absent	Present	Absent	[20]

(*Continued*)

Table 11.2 Availability of genome, proteome, and transcriptome data for biomass sourced from India. (*Continued*)

S. No.	Name (common name)	Genome	Transcriptome	Proteome	References
24	*Miscanthus fuscus* (elephant grass)	Absent	Absent	Absent	[20]
25	*Paspalum notatum* (Bahia grass)	Absent	Absent	Absent	[46]
26	*Pennisetum purpureum* (Napier grass)	Absent	Absent	Absent	[46]
27	*Phalaris arundinacea* (reed canary grass)	Absent	Absent	Absent	[13, 53]
28	*Phragmites australis* (common reed)	Absent	Absent	Absent	[54]
29	*Populus euphratica*	Present	Present	Present	[3]
30	*Populus trichocarpa* (black cottonwood)	Present	Present	Present	[3]
31	*Saccharum officinarum* (sugarcane)	Absent	Present	Absent	[3]
32	*Salix viminalis* (basket willow)	Absent	Present	Absent	[3]
33	*Sesbania bispinosa* (prickly Sesban)	Absent	Absent	Absent	–
34	*Setaria italica* (foxtail millet)	Present	Present	Present	[3]

(*Continued*)

Table 11.2 Availability of genome, proteome, and transcriptome data for biomass sourced from India. (*Continued*)

S. No.	Name (common name)	Genome	Transcriptome	Proteome	References
35	*Setaria viridis* (green foxtail)	Absent	Absent	Absent	[3]
36	*Sorghum bicolor* (sorghum)	Present	Present	Present	[21, 22]
37	*Triticum* sp. (wheat)	Present	Present	Absent	[3]
38	*Zea mays* (maize)	Present	Present	Present	[3]
39	*Ziziphus rugosa* (Bonbogori)	Absent	Absent	Absent	[12]

Besides, the higher content of lignin hinders the production of bioethanol and the reason can be imparted to complexity involved in separating lignin from other polymers for efficient enzymatic hydrolysis [23]. Hence, one can think of genetically engineering fuel plants to generate less amount of lignin as compared to other biopolymers and would require information of regulatory factors involved in the synthesis of lignin. Though the biosynthetic pathway of lignin has been deciphered in great depth, regulating its production in various plants would surely involve the knowledge of genome of that particular plant. In addition, one can go for transcriptomics study to know what all genes and regulatory factors are involved in higher production of cellulose and xylose in a given plant so as to increase the content of bioethanol in the long term.

11.6 Conclusion

The success of bioethanol blending as proposed by the Government of India is dependent on two things: (1) availability of biomass sources that are sustainable and (2) characterization of catalysts. For the former, the major limitation is incentive to the farmers to provide the industry enough biomass (such as second-generation biomass) to make it sustainable. In this regard, this chapter summarizes not only agricultural crop residues but also plants that are used majorly as feed or for other purposes. With regards to catalysts, it has been shown that for a developing country like India, there are abundant strains available that produce cellulases and other associated enzymes. Some of these enzymes are attractive for industry in terms of thermostability and other properties that are deemed desirable. At the same time, some of these organisms' genome has been sequenced, and thus there is availability of omics data that can be tapped for further characterization and development of strains that are highly suitable for industrial purposes. In the future, there is a need to generate more omics data (genome, transcriptome, and proteome) to be generated for both the biomass source and strains native to India.

References

1. Nigam P.S., Singh A., Production of liquid biofuels from renewable resources, *Prog. Energy Combust. Sci.*, 37, 52–68, 2011.
2. Koh, L., Ghazoul, J., Biofuels, biodiversity, and people: Understanding the conflicts and finding opportunities. *Biol. Conserv.*, 141, 10, 2450–2460, 2008.
3. Dhar, S., Shukla, P., Low carbon scenarios for transport in India: Co-benefits analysis. *Energy Policy*, 81, 186–198, 2015.

4. Manorama,R., Pindi,P., Reddy,G., Shivaji,S., *Bhargavaea cecembensis* gen. nov., sp. nov., isolated from the Chagos-Laccadive ridge system in the Indian Ocean. *Int. J. Syst. Evol. Microbiol*, 59, 10, 2618–2623, 2009.
5. Singh, S., Tiwari, R., Renuse, S., Pranaw, K., Nain, L., Proteomic analysis of Streptomyces sp. ssr-198 grown on paddy straw. *J. Basic Microbiol.*, 55, 6, 790–797, 2015.
6. Rathnan, R., and Ambili, M., Cellulase Enzyme Production by Streptomyces Sp Using Fruit Waste as Substrate. *Aust. J. Basic & Appl. Sci.*, 5, 12, 1114–1118, 2011.
7. Chellapandi P., Jani H.M., Production of endoglucanase by the native strains of Streptomyces isolates in submerged fermentation., *Braz. J. Microbiol*., 39, 122–127, 2008.
8. Prasad, P., Bedi, S., Singh, T., *In vitro* Cellulose Rich Organic Material Degradation by Cellulolytic *Streptomyces albospinus* (MTCC 8768). *MJM*, 2012.
9. Prasad, P., Bedi, S., Characterization of a novel thermophilic cellulase producing strain *Streptomyces matensis* strain St-5. *IJCMAS*, 3, 3, 74–88, 2014.
10. Prasad P., Singh T., Bedi S., Characterization of the cellulolytic enzyme produced by *Streptomyces griseorubens* (Accession No. AB184139) isolated from Indian soil, *J. King Saud Univ. - Sci*., 25, 245–250, 2013.
11. Sahay, H., Mahfooz, S., Singh, A., Singh, S., Kaushik, R., Saxena, A., Arora, D., Exploration and characterization of agriculturally and industrially important haloalkaliphilic bacteria from environmental samples of hypersaline Sambhar lake, India. *World J. Microbio.l Biotechnol.*, 28, 11, 3207–3217, 2012.
12. Sasmal, S., Goud, V., Mohanty, K., Characterization of biomasses available in the region of North-East India for production of biofuels. *Biomass and Bioenergy*, 45, 212–220, 2012.
13. Odero, D., Gilbert, R., Ferrell, J., Helsel, Z., *Production of Giant Reed for Biofuel*. University of Florida, Florida, 2008.
14. Aliberti, A., Ventorino, V., Robertiello, A., Galasso, M., Blaiotta, G., Comite, E., Faraco, V., Pepe, O., Effect of Cellulase, Substrate Concentrations, and Configuration Processes on Cellulosic Ethanol Production from Pretreated *Arundo donax*. *BioRes*, 12, 3, 2017.
15. Ma, C., Naidu, R., Liu, F., Lin, C., Ming, H., Influence of hybrid giant Napier grass on salt and nutrient distributions with depth in a saline soil. *Biodegradation*, 23, 6, 907–916, 2012.
16. Lal, B., Rana, K., Rana, D., Shivay, Y., Sharma, D., Meena, B., Gautam, P., Biomass, yield, quality and moisture use of *Brassica carinata* as influenced by intercropping with chickpea under semiarid tropics. *JSSAS*, 2017.
17. González-García, S., Gasol, C., Gabarrell, X., Rieradevall, J., Moreira, M., Feijoo, G., Environmental aspects of ethanol-based fuels from *Brassica carinata*: A case study of second generation ethanol. *Renew. Sust. Energ. Rev.*, 13, 9, 2613–2620, 2009.
18. Wood, I., Wellner, N., Elliston, A., Wilson, D., Bancroft, I., Waldron, K., Effect of *Brassica napus* cultivar on cellulosic ethanol yield. *Biotechnol. Biofuels*, 8, 1, 2015.

19. Kapoor, M., Raj, T. , Vijayaraj, M., Chopra, A., Gupta, R., Tuli, D., Kumar, R., Structural features of dilute acid, steam exploded, and alkali pretreated mustard stalk and their impact on enzymatic hydrolysis. *Carbohydr. Polym.*, 124, 265–273, 2015.
20. Byrt, C., Grof, C., Furbank, R., C4 Plants as Biofuel Feedstocks: Optimising Biomass Production and Feedstock Quality from a Lignocellulosic Perspective Free Access. *J. Integr. Plant Biol.*, 53, 2, 120–135, 2011.
21. Lingle, S., Opportunities and Challenges of Sweet Sorghum as a Feedstock for Biofuel. *ACS Symp. Ser.*, 177–188, 2010.
22. Qi, C., Yadama, V., Guo, K., Wolcott, M., Thermal stability evaluation of sweet sorghum fiber and degradation simulation during hot pressing of sweet sorghum–thermoplastic composite panels. *Ind. Crops Prod.*, 69, 335–343, 2015.
23. Welker, C., Balasubramanian, V., Petti, C., Rai, K., DeBolt, S., Mendu, V., Engineering Plant Biomass Lignin Content and Composition for Biofuels and Bioproducts. *Energies*, 8, 8, 7654–7676, 2015.
24. Bera, A., Ghosh, A., Mukhopadhyay, A., Chattopadhyay, D., Chakrabarti, K., Improvement of degummed ramie fiber properties upon treatment with cellulase secreting immobilized *A. larrymoorei* A1. *Bioprocess Biosyst. Eng.*, 38, 2, 341–351, 2014.
25. Kshirsagar, S., Saratale, G., Saratale, R., Govindwar, S., Oh, M., An isolated Amycolatopsis sp. GDS for cellulase and xylanase production using agricultural waste biomass. *J. Appl. Microbiol.*, 120, 1, 112–125, 2015.
26. Asem, D., Leo, V., Passari, A., Tonsing, M., Joshi, J., Uthandi, S., Hashem, A., Abd_Allah, E., Singh, B., Evaluation of gastrointestinal bacterial population for the production of holocellulose enzymes for biomass deconstruction. *PLoS One*, 12, 10, p. e0186355, 2017.
27. Pandey, S., Tiwari, R., Singh, S., Nain, L., Saxena, A., Evaluation of beta-1,4- Endoglucanases Produced by Bacilli Isolated from Paper and Pulp Mill Effluents Irrigated Soil. *J. Microbiol. Biotechnol.*, 24, 8, 1073–1080, 2014.
28. Yadav, A., Sachan, S., Verma, P., Kaushik, R., Saxena, A., Cold active hydrolytic enzymes production by psychrotrophic Bacilli isolated from three sub-glacial lakes of NW Indian Himalayas. *J. Basic Microbiol.*, 56, 3, 294–307, 2015.
29. Trivedi, N., Gupta, V., Kumar, M., Kumari, P., Reddy, C., Jha, B., Solvent tolerant marine bacterium *Bacillus aquimaris* secreting organic solvent stable alkaline cellulose. *Chemosphere*, 83, 5, 706–712, 2011.
30. Singh, V., Kumar, A., Production and purification of an extracellular cellulase from *Bacillus brevis* vs-1. *IUBMB Life*, 45, 3, 443–452, 1998.
31. Mawlankar, R., Thorat, M., Krishnamurthi, S., Dastager, S., *Bacillus cellulasensis* sp. nov., isolated from marine sediment. *Arch. Microbiol*, 198, 1, 83–89, 2015.
32. Pandey, S., Singh, S., Yadav, A., Nain, L., Saxena, A., Phylogenetic Diversity and Characterization of Novel and Efficient Cellulase Producing Bacterial Isolates from Various Extreme Environments. *Biosci. Biotechnol. Biochem.*, 77, 7, 1474–1480, 2013.

33. Padaria, J., Sarkar, K., Lone, S., Srivastava, S., Molecular characterization of cellulose- degrading *Bacillus pumilus* from the soil of tea garden, Darjeeling hills, India. *J. Environ. Biol.*, 35, 3, p. 551, 2014.
34. Thomas, L., Ram, H., Singh, V., Inducible cellulase production from an organic solvent tolerant Bacillus sp. SV1 and evolutionary divergence of endoglucanase in different species of the genus Bacillus. *Braz. J. Microbiol.*, 49, 2, 429–442, 2018.
35. Goyal, V., Mittal, A., Bhuwal, A., Singh, G., Yadav, A., Aggarwal, N., Parametric Optimization of Cultural Conditions for Carboxymethyl Cellulase Production Using Pretreated Rice Straw by Bacillus sp. 313SI under Stationary and Shaking Conditions. *Biotechnol. Res. Int.*, 2014, 1–7, 2014.
36. Rawat, R., Tewari, L., Purification and characterization of an acidothermophilic cellulase enzyme produced by Bacillus subtilis strain LFS3. *Extremophiles*, 16, 4, 637–644, 2012.
37. Premkumar, J., Sudhakar, T., Srikiran, K., Isolation and identification of *Cellulomonas cellulans* from silver fish and characterization of cellulase enzyme. *J. Chem. Pharm. Res.*, 7, 1, 346–349, 2015.
38. Walia, A., Mehta, P., Chauhan, A., Kulshrestha, S., Shirkot, C., Purification and characterization of cellulase-free low molecular weight endo β-1,4 xylanase from an alkalophilic *Cellulosimicrobium cellulans* CKMX1 isolated from mushroom compost. *World J. Microbio.l Biotechnol.*, 30, 10, 2597–2608, 2014.
39. Pawar, K., Dar, M., Rajput, B., Kulkarni, G., Enrichment and Identification of Cellulolytic Bacteria from the Gastrointestinal Tract of Giant African Snail, Achatina fulica. *Appl. Biochem. Biotechnol.*, 175, 4, 1971–1980, 2014.
40. Shanmughapriya, S., Kiran, G., Selvin, J., Thomas, T., Rani, C., Optimization, Purification, and Characterization of Extracellular Mesophilic Alkaline Cellulase from Sponge-Associated Marinobacter sp. MSI032. *Appl. Biochem. Biotechnol.*, 162, 3, 625–640, 2009.
41. Asha, B., Revathi, M., Yadav, A., Sakthivel, N., Purification and Characterization of a Thermophilic Cellulase from a Novel Cellulolytic Strain, *Paenibacillus barcinonensis*. *J. Microbio.l Biotechnol.*, 22, 11, 1501–1509, 2012.
42. Pal, M., Swarnkar, M., Thakur, R., Kiran, S., Chhibber, S., Singh, A., Gulati, A., Complete Genome Sequence of Paenibacillus sp. Strain IHBB 10380 Using PacBio Single-Molecule Real-Time Sequencing Technology. *Genome Announc.*, 3, 2, 2015.
43. Dhar, H., Swarnkar, M., Gulati, A., Singh, A., Kasana, R., Draft Genome Sequence of a Cellulase-Producing Psychrotrophic Paenibacillus Strain, IHB B 3415, Isolated from the Cold Environment of the Western Himalayas, India. *Genome Announc.*, 3, 1, 2015.
44. Sethi, S., Datta, A., Gupta, B., Gupta, S., Optimization of Cellulase Production from Bacteria Isolated from Soil. *ISRN Biotechnol.*, 2013, 1–7, 2013.
45. Raj, A., Kumar, S., Singh, S., A Highly Thermostable Xylanase from *Stenotrophomonas maltophilia*: Purification and Partial Characterization. *Enzyme Res.*, 2013, 1–8, 2013.

46. Hattori, T., Morita, S., Energy Crops for Sustainable Bioethanol Production; Which, Where and How? *Plant Prod. Sci.*, 13, 3, 221–234, 2010.
47. Balat M., Balat H., Recent trends in global production and utilization of bio-ethanol fuel. *Applied Energy*, 86, 11, 2273–2282, 2009.
48. Baig, M., Dharmadhikari, S., Bioethanol Production from Enzymatically Hydrolysed Cotton Stalk: One Approach Towards Sustainable Energy Development. *Curr. World Environ.*, 9, 3, 940–946, 2014.
49. Wang, M., Zhou, D., Wang, Y., Wei, S., Yang, W., Kuang, M., Ma, L., Fang, D., Xu, S., Du, S., Bioethanol production from cotton stalk: A comparative study of various pretreatments. *Fuel*, 184, 527–532, 2016.
50. Song, Y., Wi, S., Kim, H., Bae, H., Cellulosic bioethanol production from Jerusalem artichoke (Helianthus tuberosus L.) using hydrogen peroxide-acetic acid (HPAC) pretreatment. *Bioresour. Technol.*, 214, 30–36, 2016.
51. Saba, N., Jawaid, M., Hakeem, K., Paridah, M., Khalina, A., Alothman, O., Potential of bioenergy production from industrial kenaf (*Hibiscus cannabinus* L.) based on Malaysian perspective. *Renew. Sust. Energ. Rev.*, 42, 446–459, 2015.
52. Kuhad, R., Gupta, R., Khasa, Y., Singh, A., Bioethanol production from Lantana camara (red sage): Pretreatment, saccharification and fermentation. *Bioresour. Technol.*, 101, 21, 8348–8354, 2010.
53. Riffaldi, R., Saviozzi, A., Cardelli, R., Bulleri, F., Angelini, L., Comparison of Soil Organic-Matter Characteristics under the Energy Crop Giant Reed, Cropping Sequence and Natural Grass. *Commun. Soil Sci. Plant Anal.*, 41, 2, 173–180, 2010.
54. Cotana, F., Cavalaglio, G., Pisello, A., Gelosia, M., Ingles, D., Pompili, E., Sustainable Ethanol Production from Common Reed (*Phragmites australis*) through Simultaneous Saccharification and Fermentation. *Sustainability*, 7, 9, 12149–12163, 2015.