The major castor-producing countries are India, China and Brazil, accounting for 90% of castor produced in the world. India is the largest producer of castor, contributing to almost 65% of the world share (Figure 2). Castor seeds are oval in shape and light to deep chocolate in colour with different molting or beautifications on the seed coat (Figure 1b). Castor seeds contain the finest oil available from a plant source. Approximately 90% of the oil from castor bean seeds is composed of the unusual hydroxylated fatty acid, ricinoleic acid. This makes the

**Figure 1.** a, Castor bean plant. b, Castor seeds.

**Figure 2.** Castor-producing countries and their production volume (source: FAOSTAT, http://faostat.fao.org).
RESEARCH NEWS

Figure 3. Ricin protein domains are represented at the top by blue boxes, and grey boxes represent protein sequences from this gene family aligned to the ricin precursor protein sequence used as a reference. The ruler indicates the amino acid coordinates. The ricin and RCA genes are indicated and the amino acid sequence length for each gene model is shown in parenthesis. Pairs of adjacent gene models that could belong to a single pseudogene are shown in grey (adapted from Chan et al.).

Texas, USA, and is the major domestically grown variety in USA. From cytological studies, it has been well-established that the chromosome number of castor genome is 10 (2n = 20) and the genome size is ~320 Mb. The draft sequence of the castor genome was generated by producing ~2.1 million high-quality sequence reads from plasmid and fosmid libraries, and then using the Celera assembler to build consensus sequences or contigs that were linked to form 25,800 scaffolds using two ends sequences from individual clones. The assembly covered the genome ~4.6x, spanning 350 Mb, which was consistent with the previous genome size. Protein-coding genes were annotated using multiple gene-prediction programs, homology searches against sequence databases and the cDNA spliced-alignment tool PASA (Program to Assemble Spliced Alignments). To aid genome annotation, 52,165 expressed sequence tags (ESTs) were also generated from five cDNA non-normalized libraries.

The annotation of the genome sequence focused mainly on three areas, viz. (i) ricin gene family, (ii) oil metabolism genes and (iii) disease resistance genes. The ricin gene family encodes three domains: an N-terminal RIP domain and two C-terminal lectin domains. The draft sequence contains 28 putative genes in the family, whereas it has been reported that it contains only 6–8 members. The members of this lectin gene family tend to be clustered in the castor genome. Out of 28 genes of the lectin family, seven encode proteins that contain RIP and the two lectin domains, nine encode proteins with only RIP domain, and nine encode proteins with one or two lectin domains only (Figure 3). Castor, known for its high oil content, contains 71 genes in its genome involved in biosynthesis of fatty acids and triacylglycerols, mainly ricinoleic acid. There are many genes coding for enzymes which contribute to this oil metabolism that include oleic acid hydroxylase (FAII), glycerol-3-phosphate acyltransferase (G3PAT), lysophosphatidic acid acyltransferase (LPAT), phosphatidic acid phosphatase (PAP), diacylglycerol acyltransferase (DGA1), choline phosphotransferase (CPT), etc. The genome was also analysed for biotic stress resistance and 121 predicted disease resistance proteins were identified. The majority of these predicted proteins belong to the nucleotide-binding leucine-rich repeat class, followed by the less common extracellular leucine-rich repeat containing proteins, and dirigent-like proteins that have been associated with disease resistance. Till date, genetic improvement of castor was based on mass selection and pedigree methods for developing elite genotypes with desirable attributes. India has made significant progress in the development of hybrids during the last three decades. Based on the exotic pistillate line TSP-10R, the first hybrid castor GCH-3 was developed. Subsequently, the development of indigenous pistillate line, VP-1 has given new impetus to hybrid castor development which resulted in the release of three hybrids, GAUCH1, GCH-2 and GCH-4 during 1990s and 10 more high-yielding varieties later on. Though these varieties are high-yielding, the presence of ricin, still poses a major drawback in their extensive cultivation. Moreover, they are vulnerable to a number of diseases which include wilt (Fusarium oxysporum spp. ricini), root rot (Macroloba phaseolina), bacterial blight (Xanthomonas oryzae), and seedling blight (Phytophthora cactorum), etc. Disease resistance in castor plant has been imparted by producing transgenic lines through Agrobacterium-mediated transformation.

Earlier attempts to produce ricin-free castor using traditional breeding methods have been partially successful with reduction of ricin and RCA up to 75%. However, complete elimination of ricin from castor has not been achieved. Now, with the whole genome sequence available, it is possible to produce ricin-free castor plants by silencing or knocking out the genes responsible for the production of ricin. Gene silencing can be done by RNAi or epigenetic methods (DNA methylation). Alternatively, the genes responsible for oil metabolism can be introduced in an established ricin-free crop. Castor seed triacylglycerols contain 90% ricinoleic acid. This unusual fatty acid possesses a hydroxyl group responsible for the unique chemical and physi-
RESEARCH NEWS

cal properties. Recent efforts to transfer genes encoding the proteins responsible for unusual fatty acid biosynthesis to higher-yielding plants have generally produced limited success\(^1\). The yields of oil from these transgenic plants are considerably low. The main reason for such low yields is the poor understanding of the triacylglycerol biosynthetic pathway in castor plant, i.e. not all the regulatory and metabolic networks involved in the pathway are characterized. The castor genome sequence and its annotation provide us an insight for the identification of all the genes involved in castor oil biosynthesis and thus paves the way for production of ricin-free crops with high yields of castor oil. Thus, the availability of the whole genome draft sequence has formed the foundation for extensive study on this plant to derive maximum benefits.


R. Ramprasad and R. Bandopadhyay* are in the Department of Biotechnology, Birla Institute of Technology, Mesra, Ranchi 835 215, India.

*e-mail: rajib_bandopadhyay@bitmesra.ac.in

---

**Smile with Science**

By - Sumanta Baruah

"Hello captain, Mr. Selkirk refuses to be rescued. He just wants to get his internet connection repaired."