

polyynes with 4, 6, 8, 10, 12, 14, 16, 18, 20 and 22 triple bonds. The significant outcome of the study is that the step-wise synthetic procedure developed for such polyynes can be scaled up to make higher polyynes and even carbyne.

These polyynes are not particularly sensitive to light, moisture or oxygen, and they can be handled and characterized under normal laboratory conditions. Their X-ray crystallographic analysis, ^{13}C NMR and UV-visible spectral studies shed much light on their properties.

It is evident from X-ray crystallographic analysis that the structure shown in Figure 3 *d* (containing Tr^* and TIPS as end-capping moieties) is slightly bent and that of Figure 3 *g* (with Tr^* as end-capping groups) is almost linear (see also the corresponding space-filling model and ORTEP representations in Figure 3). These observations reveal that the Tr^* group extends well outside the sp-carbon framework and offers significantly more steric shielding against unwanted intermolecular interactions that reportedly lead to decomposition (schematically shown in red in Figure 3 *b* and *e*)¹². The bond lengths disclose the presence of alternating carbon-carbon single and triple bonds.

^{13}C NMR spectra of the synthesized polyynes provide ample evidence for the presence of sp-hybridized carbon. For the longest polyyne with 22 triple bonds,

21 NMR signals are discernable (two overlapping signals are observed at 64.27 ppm). The signals appear in the narrow range of 62.10–64.60 ppm, with a median value of 63.7 ppm, except for the two end carbons attached to Tr^* .

As expected, in the UV-visible spectrum of the polyynes, the lowest energy band experiences a bathochromic shift with increasing length. This causes lowering of the energy between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). A plot of λ_{max} values of the polyynes versus *n* (the number of triple bonds) reveals a convergence limit (Figure 3 *a*). From the convergence limit of λ_{max} , it has been estimated that the HOMO-LUMO energy gap for carbyne is 2.54 eV.

Thus the search for new molecular and regular polymeric allotropes of carbon has stimulated the preparation and study of π -conjugated acetylenic macrocycles, which are likely to exhibit exciting properties like the other carbon nanomaterials.

1. Kroto, H. W., Heath, J. R., O'Brien, S., Curl, R. F. and Smalley, R. F., *Nature*, 1985, **318**, 162–163.
2. Iijima, S., *Nature*, 1991, **354**, 56–58.
3. Iijima, S. and Ichihashi, T., *Nature*, 1993, **363**, 603–605.
4. Bethune, D. S., Kiang, C. H., de Vries, M. S., Gorman, G., Savoy, R. and

Beyers, R., *Nature*, 1993, **363**, 605–607.

5. Webster, A., *Mon. Not. R. Astron. Soc.*, 1980, **192**, 7–9.
6. Hayatsu, R., Scott, R. G., Studier, M. H., Lewis, R. S. and Anders, E., *Science*, 1980, **209**, 1515–1518.
7. El Goresy, A. and Donnay, G., *Science*, 1968, **161**, 363–364.
8. Cataldo, F., *Carbon*, 2005, **43**, 2792–2800.
9. Brady, M., Weng, W. and Gladysz, I. A., *J. Chem. Soc. Chem. Commun.*, 1994, 2655–2657; Bartik, T., Bartik, B., Brady, M., Dembinski, R. and Gladysz, J. A., *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 414–417; Dembinski, R., Bartik, T., Bartik, B., Jaeger, M. and Gladysz, J., *J. Am. Chem. Soc.*, 2000, **122**, 810–822.
10. Gbttner, T., Hampel, F., Gisselbrecht, J.-P. and Hirsch, A., *Chem. Eur. J.*, 2002, **8**, 408–432.
11. Eisler, S., Aaron, S. E., Slepko, D., Elliott, E., Luu, T., McDonald, R., Hegmann, F. A. and Tykwinski, R. R., *J. Am. Chem. Soc.*, 2005, **127**, 2666–2676.
12. Chalifoux, W. A. and Tykwinski, R. R., *Nature Chem.*, 2010, **2**, 967–971.

C. Srinivasan* lives at 1, Subash Street, Kalvinagar, Theni Main Road, Nagamalai, Madurai 625 019, India; R. Saraswathi is in the Department of Materials Science, Madurai Kamaraj University, Madurai 625 021, India.

*e-mail: ceesri@yahoo.com

Wheat genome sequence: challenges and success

Satish Kumar, Chandan Kumar Singh and Rajib Bandopadhyay

Wheat is one of the most important staple food crops of the world, occupying 17% (one-sixth) of crop acreage worldwide, feeding about 40% (nearly half) of the world population and providing 20% (one-fifth) of total food calories and protein in human nutrition¹. Global wheat production has been under threat in recent years from increasing demand and climate change. China, the world's largest wheat producer and consumer, is facing its own grain security problem and is eager to increase crop yields amid shrinking farmland and a rising population. Recently, Russia, one of the biggest

producers of wheat, banned all export of wheat after severe drought and wildfire charred crops around the country. Major floods in Pakistan and mudslides in China made wheat prices go up even further. Canada and several other countries also expect their wheat harvest to be much lower compared to the previous year due to weather conditions. The move raised worldwide concerns about possible wheat shortages and has sent wheat prices soaring. Sequencing the wheat genome could help identify and manipulate specific genes for useful traits, such as tolerance to drought, resis-

tance to crop diseases, or better grain quality. We can expect that improved crop management will provide at least 50% of the solution. Further, because of its recent history, hexaploid wheat is a good model to study polyploidy, a driving force for plant genome evolution.

Wheat is adapted to temperate regions of the world and was one of the first crops to be domesticated some 10,000 years ago. At the cytogenetic level, common wheat is known to have three sub-genomes (each sub-genome has seven chromosomes, making $n = 21$) that are organized in seven homoeologous

groups; each group has three closely related chromosomes, one from each of the three related sub-genomes. The diploid progenitors of the A, B and D sub-genomes have been identified, although there has always been a debate regarding the progenitor of the B genome. It has also been found that common wheat behaves much like a diploid organism during meiosis, but its genome can tolerate aneuploidy because of the presence of triplicate genes. These features along with the availability of a large number of aneuploids (particularly including a complete set of monosomics, a set of 42 compensating nullisomic-tetrasomics and a complete set of 42 ditelocentrics developed by Sears²) and more than 400 segmental deletion lines (developed later by Endo and Gill³) greatly facilitated wheat genomics research.

The wheat genome is the largest genome decoded to date. It is five times larger than the human genome and is known to be a complex structure, comprised of three independent genomes. Its large genome (>16 GB) and the high proportion of repetitive DNA (>80%) make this crop less attractive for genomics research. Nevertheless, a variety of genomic resources, including the following have become available for this crop, which should facilitate further research: (i) genome-wide genetic and physical maps (including expressed sequence tag (EST) maps and molecular function maps); (ii) a large number of known quantitative trait loci (some of them even cloned) for a variety of agronomic traits¹; (iii) >1.07 million wheat ESTs (dbEST, 1 December 2010; http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html); (iv) BAC libraries for each of the three sub-genomes (A, B and D) and for specific chromosomes or chromosome arms^{4,5}; (v) identified gene-rich regions (GRRs) on all wheat chromosomes⁶, which will be the target for genome sequencing under the aegis of the recently launched 'Wheat Genome Sequencing Consortium' (WGSC; <http://www.wheatgenome.org>), and (vi) availability of BLAST-searchable wheat genomic sequences that yielded ~1× genome coverage of Chinese Spring genome (http://www.cerealsdb.uk.net/search_reads.htm).

The wheat genome project, funded by the Biotechnology and Biological Sciences Research Council (BBSRC, UK), was undertaken at the Centre for Genomic Research, The University of

Liverpool, UK. Neil Hall and Anthony Hall (The University of Liverpool), in collaboration with Keith Edwards and Gary Barker (University of Bristol) and Mike Bevan (John Innes Centre) have sequenced the entire wheat genome and now the DNA data are available via EMBL/GeneBank and CerealsDB to crop breeders to help them select key agricultural traits for breeding. These data are expected to be published soon. The decoding was done with the help of advanced sequencing technology developed by 454 Life Sciences, USA, which reads the DNA sequence hundred times faster than the system used for human genome sequencing. The genote sequence machine used enabled researcher to analyse a million strands of DNA at a time. The process used in this case is called pyrosequencing, a technique which involves extracting DNA, suspending it in fluid, breaking it apart with bursts of gas and using chemical reactions and a high-resolution camera to infer its make-up. This is a classical example of world-class skills in sequencing and wheat genetics to deal with large genome sequence, as mentioned by Richard Summers, Vice-Chairman of the British Society of Plant Breeders. This technique enables scientists to complete the wheat genome sequencing in around a year compared to 13 years in case of the human genome sequencing which has five times shorter genome sequence.

The genome sequences released are fivefold coverage (5×) of the reference bread wheat variety Chinese Spring line (CS 42). This gives the researchers access to an estimated 95% of the genes in this variety of wheat. The genome data released are in a 'raw' format, comprising sequence reads of the wheat genome in the form of letters representing the genetic 'code'. A complete copy of the genome requires further read-through, significant work on annotation and the assembly of the data into chromosomes. By understanding the genetic differences between varieties with different desirable traits, selective breeding can produce new types of wheat that are able to withstand drought and salinity better, and also provide higher yields. The quick release of data into the public domain allows other scientists and wheat-breeding companies to rapidly employ them in practical applications.

The massive genome size of wheat (16 billion base pairs) far larger than

corn (3 billion base pairs), rice (430 million base pairs)⁷ (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC139353/>), is one of the reasons for the delay in analysing the genetic code. This large genome size is the result of repetition of the same gene around six times in the strains of Chinese spring wheat, which is used for sequencing. Also, wheat has a tangled ancestry, tracing its descent from three different species of wild grass.

The future aspect of wheat genome sequencing is the identification of specific genes for useful traits such as tolerance to drought. It is also a source to enable the scientist and breeder worldwide to analyse the sequence and use it in a new breeding method called macro-assisted selection that could dramatically increase the speed and efficiency of plant breeding. Genetic engineers may hope to craft artificial strains of wheat using the code. It will help to overcome the ongoing demand for food, the production of wheat will need to be increased by 50% of the total productions worldwide within the next 40 years (http://esciencenews.com/articles/2010/08/26/decoding_wheat_genome_will_aid_global_food_shortage).

International Wheat Genome Sequence Consortium (IWGSC), an international group working on wheat, compared it with the unordered string of letters from a set of encyclopedia volumes, because this sequence information cannot be considered as a whole genome sequence of wheat, comparable to those produced for crops earlier, like rice and corn, in which the sequence are ordered, annotated, and aligned in such a way that the position of the genes along the chromosomes is well known. This ordering and alignment is essential for linking the genetic information to the agronomically important traits that the breeders are targeting for improving wheat varieties. Thus, significant additional resources, far exceeding those invested to achieve the 5× coverage, are needed over next few years to obtain a wheat genome sequence. IWGSC called this a premature claim jeopardizing the ongoing international efforts to truly achieve a genome sequence with high utility for wheat in the next 5 years.

Whatever be the opinions made by claimants worldwide, this is an impressive first achievement and major breakthrough in the ongoing wheat research and several people around the world are going to be benefited directly or indirectly, although there are still some con-

straints to download the data due to the large size, i.e. (~30 Gb), being stored in a single file initially. To reduce this effort Barker has split these data into 11 smaller files of ~3 Gb each, so that size may not become a hurdle in the way of getting the sequence. Hopefully in the near future, scientists related to this research will publish their full paper with annotation. The whole genome sequence of corn is used for the comparative studies of different varieties of corn. It has been found that more than 100 genes are present in some corn lines, but are missing in others. Heterosis also has been studied through genome sequencing, which is important for crop breeders to produce better hybrid corn (<http://www.sciencedaily.com/releases/2010/11/101123121117.htm>). Rice genome sequence acts as a reference genome, paving the way in understand-

ing the wide range of diversity among cultivated and wild relatives of rice. The centromere and telomere regions have been studied with the help of these genome sequences. Specific characteristics like stress tolerance, disease resistance or high yield are under study, which is not possible without the use of the rice genome sequence⁸. We hope that this wheat genome sequence data may also be used for the comparative genomics study.

1. Gupta, P. K., Mir, R. R., Mohan, A. and Kumar, J., *Int. J. Plant Genomics*, 2008, p. 36; doi: 10.1155/2008/896451
2. Sears, E. R., *Chromosome Manipulation and Plant Genetics* (eds Riley, R. and Lewis, K. R.), Oliver and Boyd, Edinburgh, UK, 1966, pp. 29–45.

3. Endo, T. R. and Gill, B. S., *J. Hered.*, 1996, **87**, 295–307.
4. Paux, E. *et al.*, *Science*, 2008, **322**, 101–104.
5. Charles, M. *et al.*, *Genetics*, 2008, **180**, 1071–1086.
6. Erayman, M. *et al.*, *Nucleic Acids Res.*, 2004, **32**, 3546–3565.
7. Petsko, G. A., *Genome Biol.*, 2002, **3**, PMC139353.
8. Matsumoto, T. *et al.*, *Int. J. Plant Genome* 2008, p. 9; doi: 10.1155/2008/348621.

Satish Kumar, Chandan Kumar Singh and Rajib Bandopadhyay are in the Department of Biotechnology, Birla Institute of Technology, Mesra, Ranchi 835 215, India.*

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