Changes
Everything for Biosimilars

LCMS-8060
LIQUID CHROMATOGRAPH MASS SPECTROMETER
Why does this take so long?  
Why wait for affinity results?

I need my results no

Wish I could get protein interaction data on my own.

BLItz™
www.blitzmenow.com

Octet® K2 System

Get them now with the BLItz system!

Do-it-yourself protein analysis. Simple-to-use, affordable, and label-free, the BLItz™ personal assay system lets you characterize protein interactions right at your own bench.

Now you can get rich insight into real-time antibody and protein binding interactions without the wait.

- Entry-level system opens new doors for measuring real-time biomolecular interactions
- Easy to use Bio-Layer Interferometry (BLI) for reliable and versatile label-free assays
- No instrument fluidics means less downtime

For more information, please contact:

Anand +91 97399 77876 | Umamahesh +91 72594 44949
Email: pallindiamarketing@ap.pall.com
When *Nature* published a cover article last year on the human proteome – more than a decade after publication of the draft human genome sequence – it was a moment of joy and pride for proteomics scientists in India. The country had missed the genomics bus earlier but a Bangalore-based group more than made up for the missed opportunity by identifying 17,294 protein-coding genes and providing evidence of tissue- and cell-restricted proteins through expression profiling.

The same issue of *Nature* carried another important paper which gave assembled protein evidence for 18,097 genes in ProteomicsDB and highlighted the utility of the data.

Proteomics has witnessed a boom globally in the last decade, but the India story is especially stunning. “The success in Indian proteomics is mixed,” says John Yates, American chemist and professor of chemical biology at The Scripps Research Institute in La Jolla, California. “Some are doing very well, but others are struggling. I think success revolves around people who have come back to India after working in major proteomics laboratories in the West,” says Yates, best known for the development of the SEQUEST algorithm for automated peptide sequencing and Multidimensional Protein Identification Technology (MudPIT).

According to Pierre Legrain, past President of the Human Proteome Organisation (HUPO), the Indian proteomics community will continue to contribute more in the future, through a network of “talented postdocs and PhD students sent worldwide and coming back to their country developing their own teams and projects.” “The Indian community was very well represented in Human Proteome Project from the inception. We now see many more, younger scientists playing an important role,” he adds.

One challenge, Yates points out, that India needs to take care of is patchy infrastructure. “Funding agencies are providing money to buy the necessary mass spectrometers but having the skill sets and experience in the
methods and protocols is also important. The spotty infrastructure in some places is a problem as mass spectrometers are sophisticated electronic equipment and they do not like dirty or spotty electrical power.” It will be helpful if people trained in this start new labs in other places in India.

“If funding agencies in India are serious about proteomics they should provide fellowships for research fellows to train in high profile labs in the West to learn with the provision they come back to India,” he remarks.

India is perched on the edge of a remarkable evolution in proteomic science. “India’s proteomic scientists are first rate. The national growth in therapeutics, especially therapeutic proteins, is stimulating the growth of proteomic skills and applications,” says Catherine Fenselau, founding president of the US-HUPO and a professor in department of chemistry and biochemistry at the University of Maryland.

“My only concern is about the selection and admission of graduate students – several young people have told me that they had to work as low-paid technicians for four or five years before they could hope to be admitted to a Ph.D. program,” she says.

There are challenges galore in a country always trying to make ends meet with its shoestring research and development funding. Nature India takes this opportunity to capture India’s big bang achievements in global proteomics research following the draft of the human proteome maps. This special issue seeks to analyse the trends and roadblocks in India’s research scene, the problems scientists face in translating research from bench to the bedside and some key lessons this country has learnt while looking at proteomics in the context of social innovation.

The Nature India Special Issue on “Proteomics Research in India” also aims to be a compendium for researchers anywhere with its listing of e-learning initiatives, next generation proteomics tools and tips on how to analyse large datasets to detect scientifically significant events. The issue also talks about the proteomics databases and repositories across the world and looks closely at the trends in cancer, malaria and plants proteomics.

The issue is being brought out with support from the Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay (IITB).

References

Subhra Priyadarshini
Editor
IN THIS ISSUE

EDITORIAL
3 India making a mark in global proteomics research
   Subhra Priyadarshini

6 Trends and roadblocks in proteomics research in India
   Sandipan Ray & Sanjeeva Srivastava

COMMENTARIES
9 Draft of human proteome maps: Significant milestone from India
   Sanjay Navani, Akhilesh Pandey & Harsha Gowda

11 Bench to Bedside: Still a pipedream?
   Ravi Sirdeshmukh, Surekha M. Zingde, K. Dharmalingam & Moorkambeswaran A. Vijayalakshmi

13 E-learning resources and virtual labs
   Sandipan Ray, Sanjeeva Srivastava, Bipin Nair and Shyam Diwakar

16 Social innovation with proteomics technology
   Vural Özdemir, Edward S. Dove & Sanjeeva Srivastava

29 Cancer proteomics in India
   Ravi Sirdeshmukh, Kumar Somasundaram & Surekha M. Zingde

31 Proteomics in malaria research
   Pushkar Sharma, Inderjeet Kaur, Pawan Malhotra & Sanjeeva Srivastava

36 Plant proteomics: An essential tool for food security
   Subhra Chakraborty, Renu Deswal & Niranjan Chakraborty

ON THE COVER

India is emerging as a strong player in the proteomics research arena.
© Ivary Inc./Alamy
Illustration: Ajay Kumar Bhatt
Generic pictures for editorial and commentaries shot at NIPGR, New Delhi & IIT Delhi © S.Priyadarshini

18 Next generation proteomics tools
   Srikanth Rapole, Mahesh J. Kulkarni, Sunan S. Thakur, Shantanu Sengupta

22 Analysing large datasets in the omics era
   Sarath Chandra Janga & Debasis Dash

26 Protein databases from India
   Harsha Gowda, Aditi Chatterjee & T. S. Keshava Prasad

ARTICLE
40 A draft map of the human proteome
   (In the print version)
At the turn of this century, the newly emerging field of ‘proteomics’ began showing promise in various aspects of clinical and industrial research. While India was not able to play a vital role in genome sequencing projects, in the post-genomic era the country is playing an increasingly significant role in global proteomics research.

In 2005, eminent Indian scientist and late President A P J Abdul Kalam noted that India has the “potential to tap research opportunities in proteomics and biochips to help understand the biological processes and treat diseases. This is possible even though the country has missed the opportunity to partner in the human genome project.”

Trends and roadblocks in proteomics research in India

Sandipan Ray1, 2 & Sanjeeva Srivastava1

1Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India (sanjeeva@iith.ac.in). 2Wellcome Trust-MRC Institute of Metabolic Science, University of Cambridge, United Kingdom (sr738@medschl.cam.ac.uk).
The booming Indian proteomics scene

In India, proteomics research was initiated over a decade ago. Research groups in premier institutes started adopting proteomics technologies in biological research projects and the emerging field got considerable support from central research agencies. In 2009, the Proteomics Society, India (PSI) was established as a platform to foster interactions within the Indian proteomics community and to encourage exchange of ideas, enhance collaborations and boost innovations at the national and international level.

Although the development of proteomics research in India was rather slow in the beginning, the last few years have seen a significant expansion in the proteomics community. Presently, there are over a hundred research laboratories in 76 academic or research institutes across India involved in proteome-level research investigations (Figure 1).

Several research groups from India are actively involved in world-class research on proteomics of different human cancers and infectious diseases, and are also effectively contributing towards diverse aspects of bacterial, plant and animal proteomics at the global level.

Notable achievements

High quality data repositories are indispensable to the globalisation of proteomics research. Researchers from the Institute of Bioinformatics (IOB), Bengaluru have developed the Human Protein Reference Database (HPRD) and Human Proteinpedia (www.humanproteinpedia.org/), while important contribution in the Human Protein Atlas (http://www.proteinatlas.org/) has come from researchers at Lab SurgPath, Mumbai.

Creation of the ‘Human Proteome Map’ has been one of the most remarkable achievements in proteomics research in recent times. Pandey and Kuster labs have independently drafted the ‘Human Proteome Maps’ using high-resolution mass spectrometry. More recently, a comprehensive tissue-based map of the Human Proteome using antibody-based microarrays was reported from Uhlen’s group. Indian researchers played a significant role in two of these three important projects contributing towards the characterisation of each and every protein present in the human body.

The Indian proteomics community has been on top of the learning curve, being exposed to international proteomics conferences, meetings and workshops from the very beginning of the proteomics boom. Besides, Indian researchers have developed various e-learning...

Figure 1. Laboratories across India involved in proteome-level research investigations.
resources on proteomics, such as one of the first virtual lab projects dedicated to proteomics (http://iitb.vlab.co.in/?sub=41&brch=118) at the Indian Institute of Technology Bombay, Mumbai. The effort is now recognised internationally and is being incorporated as a part of the International Proteomics Tutorial Programme (IPTP) conducted by the Human Proteome Organization (HUPO) and the European Proteomics Association (EuPA).

Keeping pace with the growing proteomics research efforts, India is actively participating in global proteomics organisational activities and initiatives including the Human Proteome Organization (HUPO), Chromosome centric Human Proteome Project (C-HPP), Asia Oceania Human Proteome Organization (AOHUPO), International Plant Proteomics Organization (IPPO) and Asia Oceania Agricultural Proteomics Organization (AOAPO). India's involvement in cutting-edge proteomics research is receiving worldwide attention. Consequently, the 6th Annual Meeting of PSI – International Proteomics Conference on 'Proteomics from Discovery to Function' (December 2014) was attended by eminent scientists involved in path-breaking proteomics research and the pioneers of the Human Proteome Organization (HUPO). This year the Journal of Proteomics, which serves as an official journal of the EuPA, is also publishing a special issue on 'Proteomics in India' to highlight the recent growth of proteomics research in India.

Hurdles and the way ahead

Despite some success stories, India is still a long way off from successful translation of promising laboratory findings into practical applications. However, armed with technology and expertise India is capable of this translation through long-term multi-disciplinary and multi-institutional research programmes. Such translational research requires advanced infrastructure and substantial enduring financial support, which isn't available to most research laboratories in low and middle-income countries such as India.

Lack of adequate and long-term funds is one of the prime reasons behind the failure of many promising research ventures. These limitations can be overcome with pre-competitive data sharing of existing resources and data repositories, collaborations, joint grant applications and linkages with relevant industries. India needs focussed policies to promote translational research through specialised mega projects. This would ensure that the benefits of proteomics technologies reach one and all.

References

Draft of human proteome maps: Significant milestones from India

Sanjay Navani1, Akhilesh Pandey2,3 & Harsha Gowda†

Human genome sequencing was a very important milestone that transformed biological research. The human proteome was the next frontier in the post-genomics era. Three different research groups have now completed this milestone independently1,2,3. Of these, two groups employed mass spectrometry-derived proteomics while the third used protein specific antibodies. Interestingly, two of these three studies involved two independent research teams from India.

Scientists at the Institute of Bioinformatics (IOB) in Bangalore played a major role in mapping the human proteome by employing Fourier transform mass spectrometry. They carried out an unbiased survey of human proteins across 17 adult tissues, 7 fetal tissues and 6 purified primary hematopoietic cells (Figure 1). The expression pattern of the identified proteins across these 30 human tissues/...
For example, a unique proteogenomics strategy developed by IOB for hundreds of proteins that were designated “missing proteins” in the human genome. They also provided first protein level evidence of protein expression for nearly 90% of the annotated genes in the human genome. They also provided first protein level evidence of protein expression for nearly 90% of the annotated genes in the human genome. They also provided first protein level evidence of protein expression for nearly 90% of the annotated genes in the human genome.

A research group at Lab Surgpath in Mumbai played a major role in mapping the human proteome using an antibody-based approach. The tissue-based analysis detected more than 90% of the putative protein-coding genes. This approach was used to explore the human secretome, the membrane proteome, the druggable proteome, the cancer proteome, and the metabolic functions in 32 different tissues and organs. All the data were integrated in an interactive web-based database that allows exploration of individual proteins, as well as navigation of global expression patterns, in all major tissues and organs in the human body.

The Lab Surgpath pathology group has annotated more than 13 million images following immunohistochemical labeling of tissue sections of all major organs in the human body, at the rate of 8,000 images being evaluated every day (Figure 2). This includes 44 normal human tissues, 20 cancer types and 46 human cell lines. All images with their detailed annotations are freely accessible at ‘The Human Protein Atlas’ portal (www.proteinatlas.org). This interactive portal is aimed at researchers interested in human biology, translational medicine and surgical pathologists with interest in immunohistochemistry.

Together, these efforts to map the human proteome, which included a significant contribution from India, provided evidence of protein expression for nearly 90% of the annotated genes in the human genome. They also provided first protein level evidence for hundreds of proteins that were designated “missing proteins” by the research community due to lack of protein level evidence. For example, a unique proteogenomics strategy developed by IOB led to identification of 193 novel protein coding regions in the human genome that were not reported earlier. A large majority of these novel regions are annotated as pseudogenes or regions that code for non-coding RNAs.

In addition, this study also reported identification of more than 100 novel coding exons and novel open reading frames for some of the annotated genes in the human genome. The Human Protein Atlas portal allows exploration of tissue-elevated proteomes in specific tissues and organs. In addition, it allows analysis of tissue expression profiles for specific protein classes, including proteins involved in housekeeping functions in the human body, such as cell growth, energy generation, and metabolic pathways; groups of proteins involved in diseases; and proteins targeted by pharmaceutical drugs. The impact of these studies on basic research as well as biomedical research will become evident in the coming years. Methods used in these studies will also be embraced by the scientific community for further exploration of the human proteome.

The contribution of Indian scientific groups in these seminal studies to characterize the human proteome holds enormous significance for Indian science. These studies have clearly demonstrated that India has the talent, infrastructure and capability to be a major player in global science. It also highlights the ability of researchers from India in carrying out large-scale collaborative international research projects that have global impact.

Both research groups from India, which participated in this global effort, are from private institutions. Institute of Bioinformatics is a 13-year-old private non-profit organization engaged in biological research. The institute, besides publishing over 200 papers in international journals, has developed several world-class biological databases that are extensively used by the global scientific community. Lab Surgpath is the first private surgical pathology laboratory in India to have participated as a major collaborator in an international project with the annotated images resulting in more than 300 publications. Lab Surgpath offers both research and diagnostic services.

Given that strategic initiatives are being enacted in most countries, this makes a strong case for enhanced support by Indian funding agencies to institutions based on their scientific productivity and excellence regardless of their governmental affiliation. Such an approach is likely to foster parallel development of novel research ecosystems in India in addition to government research institutions.

The United States of America has successfully nurtured such an ecosystem for several decades with the Broad Institute being one such example. To keep the momentum going in proteomics, India needs bold initiatives with funding that supports scientific projects where India can make a global impact. This will not only help the country build human resource required for successful execution of such large-scale projects but will also make Indian scientists a scientific force to contend with in the international arena.

**References**

Bench to bedside: Still a pipedream?

Ravi Sirdeshmukh1,2, Surekha M. Zingde3, K. Dharmalingam4 & Mookambeswaran A. Vijayalakshmi5

The proteomics approach has taken centre stage in biology research. However, scientists are yet to fully explore its clinical relevance. Every proteomics researcher is confronted with the question – can proteomics deliver the expected promises? The answer isn’t very encouraging right now – not because of the lack of potential or efforts, but perhaps because of the complexity of proteomics science.

Heterogeneity in disease phenotypes poses a big challenge and the application of proteomics techniques to individual samples is not yet robust or cost effective. While these are issues for proteomics researchers everywhere, Indian scientists face additional challenges that need particular attention.

What’s ailing translation?

Among the two million proteoforms in the human proteome, one or more represent a specific disease condition. These disease specific proteoforms could be used in many ways. They could help predict, prevent and better manage diseases. Indian researchers are pursuing the discovery and functional evaluation of the protein biomarkers vigorously and with considerable success. However, the clinical validation phase is yet to take off in India.

There is a great need for knowledge exchange between the scientist and the clinician so that each understands mutual strengths and needs. The clinical queries of relevance can then be addressed with appropriate technologies and specimen cohorts so that the journey from the bench to bedside becomes more achievable.

Most hospitals do not have an institutionalised clinical record management system, nor is there a system at the national level to accommodate these important aspects of translation, including patient follow-ups. Further, there is need to catalogue protein biomarkers and their variants for diseases prevalent in the country and use them intelligently to apply for specific clinical questions. In the present scenario, most proteomic studies are in the domain of cancer biomarkers, neglecting metabolic disorders and other diseases. For example, the recognition of other glycated proteins in addition to haemoglobin can be important for clinical applications. Similarly, the detection of protein variants in cardiovascular disorders can be interesting to pursue.

Publishing discoveries or filing patents is not adequate. Engaging industry for licencing discoveries and an active effort for product development is required with greater intensity. However, the current environment in India is not very conducive for strong large-scale interactions between academia and industry for translation of technologies and concepts. Companies, especially multinationals, are also bound to their headquarters for R&D. The lukewarm interest shown by industry partners in taking discoveries to the clinic is a big limitation for translation.

Suboptimal infrastructure, lack of financial support and trained manpower to handle data in an integrated manner are some other limiting factors.

The sporadic progress in translational proteomics research in India can also, in part, be attributed to the reluctance of young investigators in taking up this field of research. Established senior investigators are slow to appreciate the advantages of omics platforms in general and proteomics in particular and venture into these new technologies. Rapid changes in instrumentation platforms and the prohibitive costs for front line equipment prevents less-endowed labs and educational institutions from undertaking this area of research.

There are also challenges on the technology front. Advances in quantitative proteomics have made it possible to pinpoint even minor differences in the protein levels between normal and pathological samples. However, more innovative methods to mark even structural differences in proteins introduced by mutations or structural variations induced by post-translational modifications or protein truncation that are associated with pathologies would be useful and an important value addition.

The road ahead

Not many studies in India have addressed the clinical course of diseases or defined the source material for targeted proteomic inquiries. They have focussed on differential expression of proteins between normal and diseased samples. Some questions begging primary attention pertain to clinical subtypes, sub-sites, subcellular and subcellular number of samples, when to use tissue or body fluids, time and mode of collection, storage, transport, likely concentration of the marker, pre-fractionation of the protein mixture and technology platform to be used.

1Mazumdar Shaw Center for Translational Research, Bangalore, India. 2Institute of Bioinformatics, Bangalore, India (ravisirdeshmukh@gmail.com). 3CH3-53, Kendriya Vihar, Khaghar, Navi Mumbai, India (surekha.zingde@gmail.com). 4Aravind Medical Research Foundation, Dr. G. Venkataswamy Eye Research Institute, Aravind Eye Care systems, Madurai, India (dharmalingam.k@gmail.com). 5Centre for Bio-Separation Technology (CBST), VIT University, Vellore, India (indviji@yahoo.com).
At this important crossroad in proteomics research, scientists must go beyond generic research and utilise discovery data to address unmet clinical needs. It is essential to move beyond studies of protein expression and define the intended use of such data in clinics. Focusing on specific areas such as cancers and infectious diseases prevalent in the country would be of utmost importance. Study designs should direct specific /discrete action to identify markers for risk assessment, early detection, diagnosis, prognosis, prediction or potential targets for therapy so that the outcome of the investigations have clinical relevance.

Organised tissue repositories and hospital information systems for clinical data are the primary need for translation. Only then, the biomarkers emerging from discovery research can be taken forward for multi-centre validation in larger cohorts. These issues are now receiving some attention and need more intense effort.

For clinical applications, body fluid proteomics occupies a key position. To overcome the limitations of the present techniques to achieve the depth of the proteome in body fluids or to detect protein variants associated with clinical conditions, researchers need simple technologies which can be tailored for on-line pre-fractionation or identification of the variants based on specific recognition sites like reporter amino acids. The immobilised metal-ion affinity (IMA) concept, based on the recognition of accessible histidine residues on a protein by divalent transition metal ions, has excellent potential for the simple detection of variants. Such newer approaches may be explored. Finally, special attention for the development of user friendly, point-of-care devices is particularly important in the Indian system, given large number of centres with limited technical expertise.

India with its diverse demography is a great resource of clinical material. With scientific expertise, optimal funding and increased communication between the scientist and the clinician, India can contribute effectively to the bench to bedside translation. CME programmes, workshops and free interactions between the two groups is the way to develop an integrated discipline. While this might take some time, the volume of omics information publicly available today has opened the door to an era of integrative and hypothesis-driven science, connecting with even metabolomics in the downstream. The largest international proteomics forum, the Human Proteome Organization (HUPO), has taken up several proteomics initiatives with implicit translation goals. However, exploration of human biology is an integral element that cannot be side-lined. So, it is imperative to continue scientific efforts in acquiring new methods and trends while simultaneously pursuing improvements in clinical paradigms to enable more effective fusion of the two segments. There are some publications from Indian groups now which show the promise of translation both in thinking and in effect.

India needs to strengthen the public-private-partnership (PPP) model for the industry partners to shape and develop technologies with government funding. Research-oriented hospitals, who could join hands in clinical proteomics, should be encouraged by government agencies. It is time that funding agencies considered proteomics as truly translational just like genomics. Establishment of tripartite partnerships involving clinicians, academia and industry will be the way forward for clinical proteomics in India so that the bench to bedside dream moves from being a pipedream to a reality.

References
5. https://www.hupo.org/hupo-sponsored-initiatives
E-learning resources and virtual labs

Sandipan Ray1,2, Sanjeeva Srivastava1, Bipin Nair3 and Shyam Diwakar3

India’s recent strides in information technology have propelled the growth of web-based digital learning in most disciplines of science and engineering education. Distance education and open learning endeavours offer many advantages in resource-limited developing countries, where the number of potential learners is much higher than the number of experienced teachers or advanced educational institutes1.

However, these endeavours alone have proved insufficient in providing practical skills for science experiments or analysis of scientific data. Virtual laboratories, which act as free, round-the-clock replicas of actual laboratories, could be an effective alternative. Learners in a virtual laboratory can understand scientific theories and also experience practical experimental procedures2,3. As educational budgets in developing and under-developed countries continue to shrink, e-learning and open-learning programmes are gaining popularity4.

E-learning proteomics resources

E-learning and virtual labs are rapidly changing the culture of education in developing countries5. India is playing an imperative role in the development of diverse e-learning resources and virtual labs in proteomics and other disciplines of biotechnology (see box on right). In recent years, proteomics and related disciplines have been incorporated into academic curricula across the globe due to their increasing impact on clinical and industrial research.

The Indian Institute of Technology Bombay has developed pioneering proteomics learning resources such as the Virtual Proteomics Lab, Clinical Proteomics Remote Triggering Virtual Laboratories, and other related e-learning initiatives supported by India’s ministry of human resources and development (MHRD) with a goal to disseminate high-quality educational content exclusively in proteomics6. The resource contains modules on gel-based proteomics, mass spectrometry-based proteomics and bioinformatics, each with a set of experiments (http://iitb.vlab.co.in/?sub=41&brch=118). The course contents of Virtual Proteomics Lab have now been incorporated as a tutorial article under the International Proteomics Tutorial Programme (IPTP 14) developed by the Human Proteome Organization (HUPO) and the European Proteomics Association (EuPA)7.

The Clinical Proteomics Remote Triggering Virtual Laboratory (http://iitb.vlab.co.in/?sub=41&brch=237) creates a realistic virtual environment for learners to get a first-hand experience of performing different proteomic technology experiments commonly used in clinical proteomics research. Additionally, 40 hours of a web-based video lecture course from National Programme on Technology Enhanced Learning (NPTEL) and Open Source Courseware Animations Repository (OSCAR) provides basic working principles and comprehensive details of advanced proteomics technologies using videos, animations and interactive simulations. These new e-learning resources in proteomics serve as valuable global platforms for students and researchers from different disciplines of proteomics.

Virtual labs for rural and urban India

A study of online statistics indicates that virtual lab users have been increasing in India. The study also suggests increasing usage trends in times to come8.

The researchers tried to figure out the impact and penetration of virtual labs through hands-on workshops in rural South Indian biotechnology and engineering institutes and compared them with

Key biotechnology virtual lab and e-learning initiatives in India

- “Sakshat” Virtual Biotechnology Engineering Labs: http://www.vlab.co.in/
- Technology Enhanced Learning (NPTEL): http://nptel.iitm.ac.in/
- Open Source Courseware Animations Repository (OSCAR): http://oscar.iitb.ac.in/oscarHome.do
- National Knowledge Network (NKN): http://www.nkn.in/
- Amrita Virtual Interactive E-Learning World (A-VIEW): http://aview.in/
- Online Labs for schools: http://www.olabs.co.in/

1Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India (sanjeeva@iitb.ac.in). 2Wellcome Trust-MRC Institute of Metabolic Science, University of Cambridge, United Kingdom (sandipan@iitbomab@gmail.com). 3Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham (Amrita University), India (shyam@amrita.edu, bipin@amrita.edu).
data from urban areas9. The feedback showed that 60% of students rated virtual labs as user-friendly tools that made their biotechnology courses interesting and easier to comprehend; 65% found virtual labs to be good online material for better understanding of the basic concepts in biotechnology and about 10% reported difficulty using them due to computer illiteracy or network connectivity issues.

Among 250 teachers surveyed, around 85% suggested that virtual labs could be used as an autonomous, supplementary learning and teaching material for enhancing laboratory education. 67% of the teachers from rural areas adopted virtual labs in their teaching while only 33% from urban areas opted for them.

Virtual labs are a technological innovation providing new learning environments for proteomics and biotechnology users. Simulation-based virtual labs can now train a huge cluster of potential researchers, who could play an important role in bioinformatics analysis of big data sets generated by scientific research labs across the world and effectively accelerate high-throughput translational research.

Virtual and open learning initiatives are poised to bring a dramatic change in science education but cannot completely substitute existing educational institutes or hands-on practical laboratory courses. Effective expansion of science education is possible by taking benefits of the affordances of both the approaches10. Innovative and forward-looking initiatives for distributed learning practices through e-learning and virtual labs would certainly enrich the global community of students, scientists and citizens.

References
Proteomics – Indian scientists working with GE Life Sciences!!!

A list of publications from some renowned scientists in India collaborating with GE Healthcare Life Sciences, India.

Proteomic Investigation of Falciparum and Vivax Malaria for Identification of Surrogate Protein Markers
Published: August 9, 2012, DOI: 10.1371/journal.pone.0041751

2D and DIGE technology was used to analyze alternations in the human serum proteome as a consequence of infection by malaria parasites. Dr. Chennareddy, GE Healthcare and Dr. Srivastava’s lab at IIT, Mumbai identified malaria specific proteins differentially expressed in serum due to parasite infection.

Single cell-level detection and quantitation of leaky protein expression from any strongly regulated bacterial system
Analytical Biochemistry, Volume 484, 1 September 2015, Pages 180-182
Kanika Arora, Sachin S. Mangale, Purnananda Guptasarma

Deconvolution microscopy was used to study fluorescently tagged protein localization. Dr. Guptasarma’s lab at IISER Mohali with Dr. Sachin Mangale, GE Healthcare used the Deltavision™ microscope to examine low levels of bacterial protein localized to nucleoids at the single cell level.

Interaction of ATP with a Small Heat Shock Protein from Mycobacterium leprae: Effect on Its Structure and Function
Sandip Kumar Nandi, Ayon Chakraborty, Alok Kumar Panda, Sougata Sinha Ray, Rajiv Kumar Kar, Anirban Bhunia, Ashish Biswas
PLOS, Published: March 26, 2015, DOI: 10.1371/journal.pntd.0003661

Surface Plasmon Resonance technology was used to study interaction between HSP18 and ATP molecules. In this and other papers, Dr. Biswas’ lab at IIT Bhubaneswar with Dr. Sougata Ray, GE Healthcare used Biacore™ and for the first time reported kinetics of interactions between ATP and HSP18.

Evaluation of crocin and curcumin affinity on mushroom tyrosinase using surface plasmon resonance
International Journal of Biological Macromolecules, Volume 65, April 2014, Pages 163-166
Sushama Patil, Sistla Srinivas, Jyoti Jadhav

Surface Plasmon Resonance technology was used to study binding kinetics of crocin and curcumin with mushroom tyrosinase. Dr. Jadhav’s lab at Shivaji University, Kolhapur and Dr. Srinivas Sistla, GE Healthcare used Biacore to compare binding constants and study the role of these molecules as inhibitors.

Investigation of serum proteome alterations in human endometriosis
Journal of Proteomics, Volume 114, 30 January 2015, Pages 182-196
Mainak Dutta, Elavarasan Subramani, Khushman Taunk, Akshada Gajbhiye, Shubhendu Seal, Namita Pendharkar, Snigdha Dhali, Chaitali Datta Ray, Indrani Lodh, Baidyanath Chakravarty, Swagata Dosgupta, Srikanth Rapole, Koel Chaudhury

2DE and DIGE technology was used to compare serum proteome of endometriosis patients and healthy subjects. Dr. Shubhendu Seal, GE Healthcare and Dr. Rapole’s lab at the National Centre for Cell Science, Pune were able to identify three novel proteins as promising candidate markers for diagnosis of endometriosis.

gelifesciences.com

For any further queries, please write to us at supportdesk.india@ge.com
© 2015. All Rights Reserved.
Social innovation with proteomics technology

Vural Özdemir1,2, Edward S. Dove3 & Sanjeeva Srivastava4

Countries around the globe have embraced knowledge-based innovation as a strategy for smart growth in the past couple of decades. However, India has achieved something that many countries are still struggling to – sustained investments in national technology policy and a sharp focus on biotech and proteomics1,2,3,4.

This has paid off handsomely, as evident through landmark studies by a Bangalore-based team, which was among three global groups involved in the “Human Proteome Draft”, part of the Human Proteome Project (HPP) that debuted in 20105. Indian scientists have also made significant contribution to the recently published tissue-based map of the human proteome1.

These scientific victories – a culmination of years of painstaking work – deserve unmitigated celebration. However, it would be a mistake to rest on the laurels since the past is not a guarantor of future success when it comes to disruptive innovation in life sciences and medicine6,7. The challenge is to lay out science and technology innovation strategies to ensure India capitalises on the current opportunities emergent from the HPP.

Steering proteomics towards innovation

The ‘push’ strategy of innovation

How does one steer a technology such as proteomics into innovative applications? Vannevar Bush’s much-cited report ‘Science, The Endless Frontier’ has come to define not only past but also current perceptions of translational research8. This policy suggests, however, a linear, narrow and one-way technology transfer from the laboratory to field applications. It also embodies a romanticised vision of science as pure and apolitical, primarily occurring within the confines of the laboratory isolated from society, and having invariably benevolent outcomes.

At its core, Bush’s innovation model proposed ‘science push’ and grossly overlooked the needs, values and expertise of user communities to innovate (‘science pull’). However, regular day-to-day science is, and always has been, inherently political. It faces multiple possible future(s), encompassing both benevolent and uncertain societal outcomes9,10. Unfortunately, user communities such as patients, clinicians and citizens continue to be neglected when biomedical innovation is designed in the laboratory isolated from society. This results in significant research waste that can otherwise be avoided. For example, out of nearly US$ 200 billion spent annually on biomedical research globally, up to 85% is estimated as inefficient11,12. The key reasons for research waste have been poor targeting, i.e. finding the right answers for the wrong questions, where research findings have little or no relevance for the communities meant to benefit from it.

The ‘pull’ strategy of innovation

The social and economic benefits of science, especially biotechnology, demand more than a marriage of biology and technology. Innovations in part emerge out of the needs of the users. As such, users should be a part of the development process. Notably, ‘lead-users’, whose needs are well beyond those of the average users, have a distinct role to play. This elite group pushes the boundaries of available products, status quo technology design and manufacturing processes, as they begin innovating on their own. Notable examples include mountain bicycles pioneered by cycling enthusiasts, online mobile banking in Africa and emerging markets, and patients with rare diseases who learned to come up with new or improved treatments for their conditions13,14.

FIGURE 1. Building social innovation in India based on innovation push, pull innovation and nested technology governance systems informed by technology ethics and ethics-of-ethics.

1Faculty of Communications and the Department of Industrial Engineering, Office of the President, International Technology and Innovation Policy, Gaziantep University, Gaziantep, Turkey (vural.ozdemir@alumni.utoronto.ca). 2Amrita School of Biotechnology, Amrita Vishwa Vidyapeetham, Kollam, Kerala, India. 3Proteomics Laboratory, Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India (sanjeeva@iith.ac.in).
Understanding the ‘social life’ of proteins

To sustain its current global leadership in proteomics, India must rethink the existing translational research model that rests primarily on a ‘science push’ factor inefficiently delivering users’ needs. India needs a new strategy to cultivate disruptive innovations centered on social innovation as proteomics technology begins to grow from discovery science to societal applications.

Social innovation is a relatively new concept. It has roots in, and yet is distinct from, a variety of emergent scientific practices such as open innovation, lead-user innovation, nested knowledge co-production, and distributed innovation. Social innovation brings together the previously isolated strategies and factors of science push and science pull. However, social innovation is more than joining forces of these two innovation strategies – the innovation actors and day-to-day tactics of creative co-production are fluid and interchangeable in the case of social innovation. The sharp demarcation between designers and users does not exist anymore. Users can become designers and designers can empathise with users’ priorities to intelligently and responsibly steer the research and product development trajectory between lab and society. This trajectory permits a two-way exchange of expertise. Innovators must ask both ‘on-frame’ and ‘in-frame’ questions at the intersection of technology and society. In-frame questions are utilitarian and enable a technology and its transfer to products. Though important, they are only a small subset of the numerous societal and policy relevant questions worth contemplating at this early stage. Yet, contracting out these questions entirely to ethicists, social scientists or humanists will neither suffice nor serve the concept of social innovation well.

On-frame questions, on the other hand, would address more deeply rooted and fundamental sociological questions such as: Are there alternatives to proteomics technology? What are the opportunity costs of investing in technology A versus technology B? Should we single-mindedly approve the innovation acceleration discourses a priori and endorse new technologies without thinking of their broader impacts? Should we also not pay due attention to decelerate the innovation trajectory when/if the broader technology impacts suggest unsustainable futures and adverse societal impacts? Importantly, who should govern and regulate the conduct of ethicists, social scientists and those who are entrusted to look after societal development of proteomics applications?

Proteins deserve a social life. This can be achieved by in-frame and on-frame societal questions posed in tandem that can best serve the proteomics innovation ecosystem by creating nested governance structures. Proteomics is an ideal case for social innovation given India’s current worldwide lead, large population with diverse cultures and needs, and the infrastructure in proteomics technology built over the past decade by a consistent and supportive national innovation policy.

References

Next generation proteomics tools

Srikanth Rapole1, Mahesh J. Kulkarni2, Suman S. Thakur3 & Shantanu Sengupta4

In the early era of proteomics, conventional tools such as two dimensional gel electrophoresis to separate a large number of proteins, spot picking, tryptic digestion of individual spots and MALDI-TOF-MS based identification were quite popular. However, they were tedious, time consuming, lacked throughput and quantitative ability.

The quantitative ability of mass spectrometry was realised with the development of nano liquid chromatography-based separation and high resolution mass spectrometry. Subsequently, shot gun proteomics and semi quantitative labeling technologies such as isobaric tags for relative and absolute quantitation (iTRAQ) and stable isotope labeling by amino acids in cell culture (SILAC) were developed.

The multiplexing ability of iTRAQ allowed discovery of candidate biomarkers including tissue biomarkers, serum biomarkers and drug resistance markers despite the limitations of variability in labelling efficiencies and quantitation at MS/MS level. SILAC is another mass spectrometry based quantitative approach predominantly compatible with cell culture system, which eventually became a powerful tool in quantitative biology. Label free approaches like MS6 (MS at elevated energy) facilitate untargeted quantitation.

Quantitation methods

Historically, triple-quadrupole based mass spectrometers were used for quantitation because of their high sensitivity and scan speed. These instruments facilitated development of quantitative approaches such as multiple reaction monitoring (MRM) or selected reaction monitoring (SRM), which recently have made a mark in the area of proteomics for biomarker validation. In MRM, a specific precursor and fragment ion are monitored for quantitation. MRM is highly reproducible and provides absolute concentration if stable isotope-labeled internal standards are included in the workflows. MRM based targeted quantification is becoming quite popular in the proteomics community, as this approach is able to replace expensive antibody-based quantification like Western blotting and ELISA.

1Proteomics Lab, National Centre for Cell Science, Pune, India (rsrikanth@nccs.res.in). 2Proteomics Facility, Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Pune (mj.kulkarni@ncl.res.in). 3Proteomics and Cell Signaling, CSIR-Centre for Cellular & Molecular Biology Hyderabad, Andhra Pradesh, India (sst@ccmb.res.in). 4CSIR-Institute of Genomics and Integrative Biology, New Delhi (shantanus@igib.res.in).
Several recent studies have used this approach for targeted quantitation. One such example is quantification of cancer related proteins in body fluids using targeted proteomics. The sensitivity, selectivity and scan speed of triple quadrupole mass spectrometers have been incorporated into high resolution mass spectrometers such as QTOF and Orbitraps, for identification and quantification of the proteome. ABI5600, Q-Exactive HF and Xevo G2 XS QTOF are some of the instruments that can perform the dual functions of identification and quantitation.

MRM performed on QTOFs and Orbitraps are called pseudo MRM or high resolution MRM (HR-MRM). Sometimes it is also referred to as parallel reaction monitoring (PRM). Unlike MRM, in PRM it is not possible to monitor the specific fragment ion during acquisition. Post mass spectral acquisition, extracted ion chromatograms (XIC) for selected ions are used for quantitation. The high scan speed facilitates development of sequential window acquisition of all theoretical mass spectra (SWATH).

In this approach, a spectral library is created by information dependent acquisition (IDA), later the instrument is specifically tuned for the selection of precursor ions from an overlapping window of 25 m/z spread over a precursor mass range of 400-1250 m/z window 25 m/z wide. Peptides are quantitated by targeted data extraction of SWATH-MS data.

**Post-translational modifications**

Post-translational modifications are vital for regulating a number of cellular processes and cellular control mechanism. High resolution accurate mass spectrometers like TOFs and Orbitraps also facilitate better characterisation of PTMs. Various mass spectrometric strategies like MS/MS, neutral loss and electron transfer dissociation (ETD) are being used for precise characterisation of PTMs. However, the quantitation of PTMs heavily relies on the fragment ion library.

Thus construction of fragment ion library for synthetically modified peptides becomes a prerequisite for quantification of PTMs. Once the library for the modified peptide is established, the PTMs can be quantified by either MRM, PRM or SWATH. One of the inevitable consequences of post-translational modifications is generation of various ‘proteoforms’. They also arise as a result of genetic variations, mutations and splicing.

Proteoforms are important since they can be differentially expressed in disease conditions and activate different pathways leading to completely different disease physiology. Proteoforms play an important role in biological processes and could be potentially used as biomarkers. One of the popular examples of such proteoforms is HbA1c, or glycated haemoglobin, used as a diagnostic marker to assess glycaemic status over preceding 100-120 days.

There is an increasing need to discriminate and quantify such isoforms with accuracy and efficiency. But highly homologous amino acid sequences and a great variation in their cellular concentrations pose a major challenge. However, advancements in mass spectrometry-based techniques have now enabled the identification and quantification of these protein isoforms.

**Bottom-up or top-down?**

A major roadblock in identification of proteoforms using the popular bottom-up approach is that very few peptides of an intact protein can be detected unambiguously. In most cases, the unique peptide of an isoform may be missed or may remain undetected. Thus, the top-down approach can be useful as it provides more information about the single intact protein.

Top-down proteomics is one of the classical techniques in mass spectrometry with great potential. It is not as popular as the bottom-up approach due to its complexity – it is still in the developing phase.
In top-down proteomics, the focus is to get an intact protein mass and characterise it by fragmentation of different charge states. Top-down proteomics has great benefits, especially to detect degradation products, sequence variants and simultaneous detection of multiple post-translational modifications.

Top-down proteomics experiments have been performed using different modes CID, HCD, ECD, ETD for different proteins, non-covalent complexes and subunit complexes using a variety of new generation instruments such as modified FT-ICR MS, Orbitrap, Q-TOF with ion mobility and triple quadruple. Top-down proteomics will play an important role in clinical and translational research and also in identifying unique protein forms or proteoforms. However, this method is less widely used due to the lack of MS compatible methods and instruments.

Proteogenomics

Proteogenomics is another upcoming discipline that utilises proteomics, genomics and transcriptomics information to find, assign and confirm the functional gene, as individually they don't provide wholesome information about a system. Genomics and transcriptomics experiments help make customised protein sequence database.

De novo sequencing is also very helpful to find novel peptide sequences and novel genes by searching against databases. Notably novel peptides are identified using shotgun proteomics through customised databases that are absent in reference protein sequence databases. Proteomics is playing an important role in providing protein level evidence on the basis of expressed proteins and finally enriching gene information.

Proteogenomics has great importance in clinical research and biomarker discovery. Onco-proteogenomics is becoming especially popular in cancer research to find tumour specific signature peptides. As normal shot gun proteomics does not provide information regarding point mutation, unusual splice variants and gene fusion, proteogenomics will be helpful in this regard. Proteogenomics has the potential to link DNA, RNA and protein expression information in the perspective of central dogma.

MALDI imaging

Matrix assisted laser desorption ionisation imaging mass spectrometry (MALDI-IMS) is emerging as a powerful tool to explore the molecular content of tissues within their morphological context. It allows direct measurement of proteins, peptides, metabolites, lipids and drugs from tissue sections.

Distribution of detected compounds can be seen as an image. Being a label free approach, MALDI-IMS is increasingly being recognised in the field of biomarker discovery, especially tissue-based research. The technical developments in MALDI imaging acquisition and data analysis are facilitating this approach for better understanding of molecular changes associated with the progression of disease.

Protein microarrays

The protein microarray platform is another gel free approach emerging as a powerful tool to study thousands of proteins simultaneously. Protein arrays are miniaturised 2D arrays generally printed on functionalised glass slides comprising immobilised proteins of interest which can be analysed in a high throughput manner.

There are several protein microarray formats including tissue arrays, reverse-phase arrays, capture arrays and lectin arrays that have advanced in recent years. These are being successfully applied in various fields including protein-protein interaction studies, immunological profiling, biomarker discovery and vaccine development. Rapid advances in next generation proteomics tools are delivering meaningful biological insights in modern biology.

References

nSMOL: Simplifying Biosimilar Quantitation

When it comes to addressing biosimilar and mAb quantitation in biological matrix, research groups and pharmaceutical organization face two glaring challenges: First, poorly developed pretreatment process in terms of selectivity and repeatability. Secondly, compromising on speed and/or sensitivity of LC-MS/MS machines. One might consider ELISA to be an alternate solution, but lag in development time, accrued cost, high failure rate and cross reactivity compel scientist to look at mass spectrometry based solutions.

Shimadzu Life Science Research Center has been relentlessly working on establishing universal bioanalytical pretreatment method for IgG derived mAb's which is easy and more selective. Current proteolysis methods can make identifying the signature peptide amongst gamut of peptides very difficult, thereby decreasing the quantitative limits. To simplify this process, Shimadzu has devised a novel techniquest SMOL (nano-surface and molecular-orientation limited proteolysis, Analyst. 2014 139(3):576) which can be applied to all mAbs.

nSMOL works on selective proteolysis of Fab by making use of the difference in size of the protease nanoparticle diameter (200 nm) and the antibody resin pore size (100 nm). With the use of nSMOL one can maintain the specificity of the antibody sequences while minimizing the sample complexity as well as the elimination of extra protease. This approach leads to the shortening of analytical time, LCMS robustness, wide dynamic range, and considerable improvement in sensitivity. This technique can already boast of ready to use completely validated methodology for bioanalysis of Trastuzumab in human plasma in accordance with the Japan Guideline on Bioanalytical Method Validation in Pharmaceutical Development from Notification 0711-1 of the Evaluation and Licensing Division, Pharmaceutical and Food Safety Bureau, the Ministry of Health, Labour and Welfare, dated July 11, 2013.

For addressing the second challenge, Shimadzu has introduced new LCMS-8060 triple quadrupole mass spectrometer which is a part of the ultra-fast mass spectrometry platform of MS/MS systems. With new UF Qarray ion guide technology increasing ion production and signal intensity, the LCMS-8060 brings you a new distinct vision of sensitivity that makes a real difference to working better and faster. Unparalleled data acquisition scan speed of 30,000 u/sec and polarity switching time of 5 msec allows LCMS-8060 to bring new levels of data quality and confidence at the highest sensitivity with maximal throughput. In a nutshell, LCMS-8060 is designed to push the limits of quantitation for complex bioanalysis experiments by providing highest sensitivity, robustness and short analysis time.

When powerful techniques complement each other, only then can scientist find a wholesome solution to their riddles. nSMOL and LCMS-8060 perfectly helps fix biosimilar and mAb quantitation puzzle and brings it in right perspective both during pre-clinical and clinical phases of development.
Analysing large datasets in the omics era

Sarath Chandra Janga1,2 & Debasis Dash3

Advances in high-throughput technologies in a number of omics-related fields in the post-genomic era have revolutionised the approach to understanding biomedicine. Reductionism, which has been the standard paradigm in biological research for more than a century, has armed researchers with immense knowledge of individual cellular components, their functions and mechanisms. Despite its huge success over the years, post-omics biology has increasingly made it clear that discrete biological functions can only rarely be attributed to an individual molecule. Instead, most biological outcomes in a cell arise from a complex interplay of different cellular entities such as proteins, DNA, RNA and metabolites. This has brought forth the notion of using multi-dimensional data-driven approaches in a number of biomedical settings. Genomics and proteomics have been particularly influenced by an avalanche of datasets originating from a number of laboratories and large-scale consortium funded projects. For instance, the rate of growth of Genbank has been exponential, doubling every 18 months1 with specific genomic surveys such as the Global Ocean Sampling (GOS) expedition alone contributing to more than 6 million proteins2. In fact, the more recent next-generation sequencing technologies show a declining trend in the cost of sequencing as prices go down by half every 5 months3. These massive increases in omics data, often referred to as big data, naturally bring in a new set of challenges for the scientific community.

Challenges in big data analysis

One of the most crucial challenges in analysing big data is acquisition and pre-processing. Some data sources, such as mass spectrometers and DNA sequencing facilities can produce staggering amounts of raw data. Much of this data is of no interest, and can be filtered and compressed by many orders of magnitude. For instance, quality scores of reads from next generation sequencing data may not be of much use to most downstream analytical pipelines once high quality reads are identified. Likewise, spectra once mapped to peptides and their abundance estimated may not be of much use for downstream analysis. So a natural challenge is to define filters in the pipelines in such a way that they do not discard useful information.

While these big datasets hold great promise for discovering patterns despite heterogeneities in the data, their massive sample sizes and high dimensionality introduce unique computational and statistical challenges, including scalability and storage bottleneck, noise accumulation, spurious correlation as well as measurement errors4. Broadly, any analyses of large-scale omics datasets can be divided into three major steps – data acquisition and pre-processing, data analysis and interpretation or visualisation.

1Department of Biohealth Informatics, School of Informatics and Computing, Indiana University Purdue University, Indianapolis, Indiana, USA (scjanga@iupui.edu). 2Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Health Information and Translational Sciences (HITS), Indianapolis, Indiana, USA. 3Institute of Genomics and Integrative Biology, New Delhi (ddash@igib.res.in).
to know the software and corresponding parameters used to generate a particular file format, which encodes the compressed information, so that the end user knows how the data is organised and stored based on the raw data. Frequently, the information collected is not in a format ready for analysis. Therefore, these two steps involve an information extraction process that pulls out required information from the underlying sources and expresses it in a structured form suitable for analysis.

Data analysis is also becoming considerably more challenging not only because of the volume of data but also due to the heterogeneity of the embedded datatypes which need to be integrated into computing frameworks. For instance, when computing on large datasets from diverse sources, data analytic frameworks need to consider the available computing infrastructure, scalability of the algorithmic implementation and level of automation that is desirable and possible for the problem being addressed. The last of these factors requires differences in data structure to be expressed in formats that can be automatically resolved for building efficient workflows for high-throughput data processing.

Mining data also assumes that the data is clean and in a format ready for analysis and that there are algorithms available to process the data on computing clusters. Both of these assumptions may not always be true. For instance, most current algorithms in omics cannot be readily deployed in Hadoop clusters and hence new code needs to be written to make them usable in cloud environments which can significantly speed up running times as compared to traditional multi-node computing clusters where there is no interaction between the compute nodes. Also since data stored in Hadoop clusters is typically replicated, computing resources and infrastructure have to be taken into consideration.

Likewise, if a noSQL framework is used as the underlying database, data needs to be imported to the data server to facilitate such efforts. These challenges also provide unique opportunities for exciting inter-disciplinary collaborations with experts from biomedical data science and engineering.

The most important step from an end user’s perspective is data interpretation. Unless the results of an analysis and its process are well documented and visualised, the analysis is of little value to the user. So it is essential to document the various steps of the implemented framework along with user-friendly visualisations which can enhance usability of the software. Providing a workflow would allow users to not only vary the choice of the parameters to study the impact on their results but also help understand the causes of noise in the dataset. Such an effort from the developer can also help in iteratively improving the software in the long run, especially if a feedback system or a listserv is maintained. Such level of information would also enable the user to realise the potential and utility of the processed data in making interpretations or in using it to integrate with other in-house datasets.

Scientific research has been revolutionised by big data. Several resources such as Gene Expression Omnibus (GEO) from National Center for Biotechnology Information (NCBI) and the PRoteomics IDEntifications (PRIDE) database from European Bioinformatics Institute (EBI) have become the central resource for omics researchers. Omics fields are being transformed from one where investigators measured individual genes or proteins of interest to one where the levels of all genes or proteins across a number of conditions and contexts are already in a database and the investigator’s task is to mine for interesting genes and phenomena. In most omics settings, there is a well-established tradition of depositing scientific data into a public repository to create public databases that can be used by all.

Data sharing in proteomics (e.g. PRIDE, Peptide Atlas, Massive and ProteomeXchange consortium) have led to free availability of data in the public domain. This has enabled researchers to develop new algorithms, reannotate and reinterpret, thereby providing deeper insight. Many Indian laboratories have contributed to and benefitted from this global sharing of high quality proteomics data and added value to the field through their participation.

References
In recent years, rapid technological advancements in LC-MS methodologies have brought state-of-the-art liquid chromatography and high-resolution accurate mass spectrometry (HR/AMS) together. The duo have gained immense popularity and have become a preferred analytical choice for accurate quantification of protein therapeutics from diverse biological matrices. Historically, ligand-binding assays (LBA) served as gold standard for protein bioanalysis. Recently, LC-MS based approaches have become a focal point of analytical research interest owing to an amalgam of factors including high sensitivity, throughput, and robust nature of these assays, with additional advantages of multiplexing capabilities and also alleviate the need to raise antibodies against target analytes, thereby complementing LBA.

India’s opportunity from Biological Patent Expiration: Twenty one biologics with a market value of about $54 billion will be losing patent protection by 2019 in the U.S. The bulk of global revenues for the Indian generic drug come from the U.S. market, and now a similar opportunity knocks on the door of the Indian pharma sector in the form of biosimilars. Globally, biosimilars are expected to generate a market revenue of about $3.987 billion by 2017. From 2012 to 2017, global pharmaceutical sales are expected to rise 13% in the “pharmaking markets” compared to 2% for the top mature markets, according to the IMS Institute for Healthcare Informatics.

Biosimilars are biologics containing the same active ingredient as the originator, they have the exact amino acid sequence composition and highly similar glycosylation patterns overlapping with the originator reference product. Biosimilar protein characterization studies involve extensive structural characterization and include an impressive arsenal of MS based techniques needed as per various regulatory requirements, e.g. intact antibody analysis, peptide mapping, multiple post translational modifications analysis, glycosylation patterns, disulfide bond mapping, oxidation, deamidation and sequence truncation, etc. Analytical characterization provides unique insights into the development potential of biopharmaceuticals.

Monoclonal antibodies (mAbs) are one of the fastest growing class of pharmaceutical products, these are considerably large, highly complex and heterogeneous glycoproteins with more than 20,000 atoms and molecular weight approximately 150 kDa. Post translational modifications on mAbs are critical in elucidating immune response and can significantly influence the overall efficacy and stability of the final product, detection and quantification of these subtle modifications can be achieved with unprecedented accuracy by deploying high resolution mass spectrometers either at intact protein level or separated light and heavy chains of the antibody, usually referred as ‘top down’ proteomics approach or by an alternate approach which involves enzymatic digestion usually referred as bottom-up proteomics approach.

**Top Down mass spectrometry based proteomics**

Top Down mass spectrometry based proteomics approach embraces highly simplified sample preparation, intact protein of interest is separated on LC, sequentially introduced into the mass spectrometer by electro spray ionization (ESI-MS) forming multiple charged species and are further subjected to fragmentation for both identification and in-depth characterization. High resolution mass spectrometer efficiently resolves co-eluting intact proteins as well as isotopic peaks of highly charged proteins for determination of charge states and accurate mass. Amgen and Amylin used Top Down for characterization (Kelleher, N. L. et al) of recombinant antibodies (Gadgil, H. S. et al) and endogenous secretory peptides (Ghosh, S. S. et al), respectively. Intact analysis of immunoaffinity enriched and reduced Adalimumab was performed (Peterman, S. et al) from plasma with a detection limit of 1.8ug/mL using Q Exactive mass spectrometer, both heavy and light chains were analysed with a full scan taken from m/z 900-4500 at a resolving power of 17,500 (FWHM).

Supercharging of proteins and peptides in electrospray ionization mass spectrometry (ESI-MS) enhances electro spray response, and sensitivity of proteomics experiments (Kuster, B. et al) by reducing overall ion mass-to-charge (m/z) ratio resulting in improved tandem mass spectrometry efficiency (Tsibin et al), more recently (Loo, J. A. et al) demonstrated efficient improved disulfide bond cleavage efficiency with various proteins including bovine β-lactoglobulin, soybean trypsin inhibitor, chicken lysozyme, etc. Also our preliminary results from intact protein analysis carried with a customized Agilent HPLC-Chip/MS enabled with post column mixing of a supercharging reagent resulted in enhanced fragmentation efficiency and enhanced MS response of standard proteins cytochrome c and myoglobin.

Biopharma clinical trials routinely include pharmacokinetics (PK) based studies to understand the drug efficacy and safety aspects. Targeted proteomics is a focused-driven strategy; primary aim here is to monitor a select few proteins/peptides of interest with high sensitivity, reproducibility and quantitative accuracy. It is a hypothesis-driven approach and
is emerging as an assay technology capable of selective and sensitive detection and quantification of potentially any protein of interest in the proteome. Pre-selected peptide fragments are analyzed using a triple quadrupole mass spectrometer using selected reaction monitoring or multiple reaction monitoring (SRM/MRM) process. The development of a SRM technique began with the development of triple quadrupole mass spectrometer (QQQ) 1970 for small-molecule analysis (Enke CG et al). Ion filtering enables the selection of predetermined masses for fragmentation followed by their subsequent detection. Q1 section of MS acts as a mass filter, selects specific ‘precursor ions’ and these are subjected to fragmentation by collision induced dissociation (CID) in Q2 and the resulting ‘product ions’ are selected in Q3 and further guided to the detector for quantification, resulting in a trace of signal intensity versus retention time for each precursor ion-product ion pair. MRM has a greater sensitivity towards low abundance peptides and relatively good quantitative precision compared to other methods. It is capable of detecting attomole concentrations of peptides across a dynamic range of up to 10^3. MRM-MS-based assay relies on efficient selection of a subset of peptides used as surrogate -signature peptide candidates for accurate quantification of a selected protein candidate. Signature peptides are uniquely sequence specific and are highest responding peptides (Proteotypic peptides) for each protein. Proteotypic peptides are referred to those class of peptides which are observed by mass spectrometer and can easily be identified from either databases of MS experimental data including Peptide Atlas and heterogeneous glycoproteins with approximately >20,000 atoms and molecular weight app150 kDa. Post translational modifications on Monoclonal antibodies (mAbs) are one of the fastest growing class of pharmaceutical products, these are considerably large, highly complex disulfide bond mapping, oxidation, deamidation and sequence truncation, etc. Analytical characterization provides unique insights into the highly similar glycosylation patterns overlapping with the originator reference product. Biosimilar protein characterization studies involve extensive structural characterization and include an impressive arsenal of MS based techniques needed as per various regulatory services through a competent force of human skills and advanced infrastructure.

**References:**

In recent years, there has been an exponential surge in biological data. The advent of high-resolution instruments with better sensitivities and new-age software suites have made it easier for data analysts and researchers to use omics approaches in biomedical investigations. Using such approaches, researchers are generating vast amounts of molecular data that provide a platform for several novel hypotheses.

Biological databases play an important role in assimilating large amounts of data and enabling users to access this information using dynamic query systems. These diverse datasets also serve as a medium to identify enzyme-substrate- and metabolic- networks involved in various biological processes, which can be helpful in molecular diagnostics and therapeutics for human diseases; and genomic selection of better biological traits in crop plants and dairy animals. Databases also provide a scaffold for the execution of regular updates, international data exchange and global meta-analysis.

Databases of protein features, resources for signaling and metabolic pathways that drive specific biological processes, and repositories of disease-specific molecular level alterations will help researchers apply systems biology approaches to unearth mechanisms or leads with greater biological significance.

Large protein databases

Leading the Indian proteomic database revolution is Bangalore-based Institute of Bioinformatics (IOB; http://ibioinformatics.org) with its flagship Human Protein Reference Database (HPRD) and other noteworthy platforms – the Human Proteinpedia, Human Proteome Map, a database on molecular alterations reported in pancreatic cancers and several human signaling pathways.

The HPRD (http://www.hprd.org/) has highly curated information on a non-redundant set of 30,047 human proteins, which include 40,042 protein-protein interactions and 1,09,518 post-translational modifications. A number of biomedical scientists use HPRD either directly by downloading data or indirectly by using the information available in RefSeq and Entrez Gene databases of the National Center for Biotechnology Information (NCBI) and University of California Santa Cruz (UCSC) genome browser.

The Human Proteinpedia (http://www.humanproteinpedia.org) provided a platform for over 200 proteomic laboratories for storing, sharing and dissemination of multidimensional proteomic datasets, even before publication. Data from Human Proteinpedia is also made available to the larger scientific community through HPRD.

Human Proteome Map (http://humanproteomemap.org/) is another interactive web portal, which represents the largest mass spectrometry-derived label-free quantitative proteomic data from 30 different human tissues. The Plasma Proteome Database (PPD, http://plasmaproteomedatabase.org/), initiated as a part of the Human Plasma Proteome Project of Human Proteome Organization, contains information on 10,546 proteins detected in human serum/plasma.

The institute has also created a highly curated database on molecular alterations reported in pancreatic cancers, initially published as a compendium of overexpressed proteins, and followed up with the database of molecular alteration at mRNA, protein and miRNA.

IOB’s centralised resource for human signaling pathways is called NetPath (http://www.netpath.org). NetPath contains manually curated data for 36 signaling pathways including prolactin, gastrin, corticotropin-releasing hormone (CRH), fibroblast growth factor-1 (FGF1), interleukins (IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, and IL-11), brain-derived neurotrophic factor (BDNF), leptin, oncostatin-M and RANKL, among others. IOB scientists have curated several signaling pathways such as Delta-Notch, EGFR1, Hedgehog, TNF-alpha and Wnt for Cancer Cell Map (http://cancer.cellmap.org/cellmap/), which is a database of human cancer focused pathways developed by Memorial Sloan-Kettering Cancer Center in New York, USA.

A slimmer version of signaling pathways annotated in NetPath were gathered to form NetSlim (http://www.netpath.org/netslim/), which comprises a graphical network of core signaling reactions.

The Bioinformatics Centre at Institute of Microbial Technology in Chandigarh has created many web servers and protein databases to perform structure and function of proteins based on their amino acid sequences, potential MHC class I and II binding regions in antigens, subcellular localisation and classification of eukaryotic and prokaryotic proteins, identification of bacterial toxins and several analytical tools for pattern finding in genome annotation. The IMT group has created a curated database of proteins associated with cervix cancer – CCDB; a database of anticancer peptides and proteins called CancerPPD, and a database of hemolytic and non-hemolytic peptides - Hemolytik.

The National Centre of Biological Sciences (NCBS) in Bangalore has created a number of publicly available databases of protein...
Family trees by analysis of protein motifs and domains. These resources include SMOS.2, LenVarDB, 3PFDB, SUPFAM; GenDiS18, MegaMotifbase, iMOTdb and PASS29. A database on disulphide bonds (DSDBASE)20, A database of olfactory receptors (DOR)21 and a resource for transcription factors responding to stress in Arabidopsis thaliana (STIFDB)22 were also developed by the team. NCBS has also developed the Database of Quantitative Cellular Signaling (DOQCS), which represents the collection of basic models of signaling pathways23, 24, 25, 26.

Researchers at the Indian Institute of Science have created a number of protein databases related to structure and function of protein kinases. These include 'KinG', a database of kinases27, NrichD28, PALI, PRODOC and MulPSSM29, resources of protein domains and alignments.

Databases for analysis of secondary structures of proteins, which include Conformation Angles DataBase of proteins (CADB), a web-based database of Transmembrane Helices in Genome Sequences (THGS)30 and Secondary Structural Elements of Proteins (SSEP)31 have also been received well by researchers worldwide.

At the Center of Bioinformatics in Pondicherry University, researchers have developed a number of protein databases including Peptide Binding Protein Database, Immune Epitope Prediction Database & Tools, Structural Epitope Database (SEDB)32, Clostridium-DT(DB)33, a comprehensive database for potential drug targets of Clostridium difficile and Viral Protein Structural Database (VPDB)34.

A manually curated database of rice proteins (http://www.genomeindia.org/biocuration) was developed by a team from the University of Delhi South Campus and is an important plant protein database from India35.

Future outlook

India is uniquely positioned to take a lead in developing and maintaining world class biological databases. A large base of human resource in biological sciences as well as software technology serves as an advantage. Although there are independent efforts from different research labs in India to build and maintain biological databases, there hasn’t been a dedicated effort to do it at a scale that is required to build and maintain world class databases much like how NCBI and EBI are doing for several years.

There are several companies in India that are offering annotation and database services to industries, which clearly demonstrates existence of such capability in India. Much of the data that is being generated within India is also not organised for the immediate use by fellow researchers.

Dedicated funding from government and corporates for academic research groups with demonstrated capability in developing and maintaining world class databases could be a good starting point. Such resources will become a necessity in the future as the amount of data being generated continues to grow.

References

Enormous efforts have gone into understanding the mechanism behind cancer pathogenesis using cell and molecular biology approaches, cell lines and animal models. Several investigators from India are actively contributing to the understanding of oncogenic processes using these approaches.

Cancer proteomics in India started with work on gliomas – the most common and aggressive adult primary tumours of the brain and oral cancers, which form a large part of the total head and neck cancer burden in the country. This multi-institutional programme by the Council of Scientific and Industrial Research included gene expression analysis at the transcript level.

The first publication on cancer proteomics from India was on gliomas in 2005. In the last decade, over 50 other publications have appeared on oral cancers, gastro-intestinal tract cancers and some others.

Analysis methods

Early cancer proteomics efforts relied upon available technology and infrastructure – the 2DE-MS approach and its advanced variation 2D-DIGE. These approaches offered limited proteome depth leading to application of alternative strategies that enabled deeper protein analysis. These included shotgun proteomics through the liquid chromatography–mass spectrometry (LC-MS or MS) approach for the first time in oesophageal cancers and gliomas and generated larger and deeper protein datasets.

Scientists prefer the antibody based approaches (ELISAs, tissue microarrays or TMAs) for confirmation and validation of markers. TMAs have been used for the validation of candidate proteins for oesophageal tumours and glioblastomas and in a variety of cancers in the Human Tissue Proteome Map.

Direct profiling has remained a major technical challenge for body fluid proteomics. Indian researchers have used tumour cell secretome analysis as an important alternative. Exosomes, which carry circulatory nucleic acids and proteins, are also emerging as an attractive choice for the analysis of circulatory biomarkers.

Indian studies

Indian proteomics studies have largely remained at the first stage of biomarker development – the discovery stage – though recently there have been some efforts to translate discovery leads to clinical applications.

A glioma research group has used clinical tissues, plasma samples and cell lines to get large data using gel-based and gel-free proteomics approaches. The data is extensively annotated for biology and tumour related processes. Candidate proteins from this dataset are being evaluated for molecular typing of tumours, post-treatment surveillance and therapeutic applications.

Indian scientists have also conducted extensive gene expression studies on gliomas and extrapolated them to the protein level. These studies have identified potential markers, undertaken molecular typing of glioma grades and analysed candidate regulatory molecules using cell lines and some groups have used a different approach to screen gliomas for autoantibodies on human protein arrays. Their investigations are targeted towards identifying proteins specific to the grade, aggressiveness and invasiveness of these tumours.

Oral cancers form a major proportion of the cancer burden in the country. Using 2DE-MS based investigations and quantitative LC-MS/MS methods with oral cancer tissues, differentially expressed proteins have been identified. Candidate markers are being evaluated in different cohorts and for more defined clinical queries. Some of them are also being assessed to distinguish histologically normal surgical margins for tumour areas. Some investigators have also identified predictive markers in pre-malignant lesions. Specific proteins eliciting an autoantibody response in oral cancer patients have been identified using immunoproteomics. Secretomes of cell lines from head and neck cancers have been analysed. Cell line models chronically exposed to tobacco extracts or smoke have been developed to create risk prediction markers for tobacco induced oral cancers, and differentially altered proteins have been identified. Studies are on to distinguish dysplastic leukoplakia, early stage and late stage oral tumours using salivary proteins.

Oesophageal squamous cell carcinoma and gastric cancer are significantly common malignancies in Asia meriting many investigations by Indian groups. Indian scientists have also conducted extensive gene expression studies at mRNA and protein level. Using phosphoproteomics and functional assays, signalling molecules such as cell surface receptors have been identified. A key molecule possibly involved in tumour development has been identified and validated in clinical specimens.

India falls under the endemic region for gall bladder cancer and efforts are on to understand the disease at the molecular level through proteomic and phosphoproteomic profiling of gall bladder cancer.
cell lines. Some other cancers studied with proteomic approaches are urinary bladder cancer\(^2\) and ovarian tumours\(^9\).

**On the way to translation**

Some of the leads on gliomas are showing translational promises\(^{15, 16, 17}\). Glioma proteomics data is well on its ways for use in tumour typing and post treatment surveillance\(^3\).

Autoantibody response to specific tumour antigens have shown merit in oral cancer prognosis\(^{11}\). Indian scientists are also exploring the potential utility of candidate salivary proteins to define dysplastic leukoplasia. A biomarker-based, ultrasensitive chip consisting of arrayed immunosensors is currently being examined as a point-of-care screening device.

A recent report\(^{10}\) describes analysis of predicting treatment response in patients with head and neck cancers – a step towards personal medicine. Proteomics-based clinical applications backed by state-of-the-art experimental and analytical methodologies are thus yielding some encouraging results for Indian proteomics researchers. However, these efforts need intensification. Specific clinical questions – big and small – have to be asked for the benefit of the diverse Indian patient population. Appropriate patient cohorts and accessibility of large number of clinical specimens are going to be a key requirement.

In recent years, many medical centres have set up big and small sample repositories. The bigger challenges to India's health care system, however, are lack of physical samples, scarce clinical annotation and poor follow-up patient data. The country’s medical record informatics system needs a big overhaul with long term support from both public and private sector so as to facilitate more effective translation in this multi-disciplinary, multi-centric field.

**References**

Proteomics in malaria research

Pushkar Sharma¹, Inderjeet Kaur², Pawan Malhotra² & Sanjeeva Srivastava³

Genome sequence and proteome analysis of many parasitic organisms have provided new hope for the identification of new vaccine/drug targets and their corresponding inhibitors or drugs. Among these, the most significant progress has been made with the malaria parasite.

Malaria is the most prevalent tropical parasitic disease killing at least a million people annually. Genome sequences of different species of *Plasmodium* and its insect as well as vertebrate hosts have propelled the growth of integrated omics approaches in different spheres of malaria research, including understanding of the host-pathogen interactions, disease etiology and pathogenic mechanism, characterisation of stage-specific parasite proteome as well as post-translational modifications and elucidation of mechanisms of antimalarial drug action.

Proteomics is an effective tool for the identification of next-generation biomarkers and potential drug/vaccine targets. Nonetheless, this emerging field is also fraught with various challenges, such as in vivo proteomic profiling of human malaria parasites, expression and purification of proteins in large quantities, difficulties in targeting low-abundance target analytes within complex biological samples, and analysis and interpretation of huge multi-omics datasets.

Research leads

*Plasmodium* genomes encode about 5,300 proteins, more than half of which are hypothetical proteins since they do not show sufficient similarity with proteins from other organisms¹. In the last decade, a series of proteomics studies have tried to understand the expression of *Plasmodium* proteins at different parasite stages²,³ and also illustrated post-translational modifications that many of these proteins undergo⁴,⁵,⁶,⁷,⁸.

These modifications have been crucial for protein functions such as haemoglobin degradation, host invasion and merozoite egress⁵. Proteome analyses have further led to the identification of a library of cell surface and secreted proteins that probably are responsible for host cell invasion and immune modulations⁹, organelle specific proteins and drug sensitive proteins¹⁰. A significant number of proteomics studies have also been performed to understand the development, pathogenesis and drug resistance in apicomplexan parasites¹²–¹⁹. Additionally, chemical proteomics approaches have provided new chemistries to develop new anti-malarials²⁰. A number of novel *Plasmodium* secretory proteins at asexual blood stages have been identified alongside a haemoglobin degradation-hemozoin formation complex²¹,²².

Proteome analysis studies have led to the identification of potential new targets such as haemoglobin degradation enzymes²², enzymes/proteins of purine salvage pathway and of protein and polyamine metabolism²³, proteins associated with parasite specific trafficking/transport pathways²⁴,²⁵, GPI anchored proteins²⁶, proteins associated with proteasome machinery and proteins linked with spread of drug resistance²⁷. These global proteomic studies have provided researchers enough arsenal to develop novel anti-parasitic strategies both for new drugs and vaccine development.

Signaling studies

Researchers have also noted the presence of several putative effectors of cell signaling in the parasite genome²⁸.

The post-genome era saw a flurry of activity in the use of reverse genetics to understand the function of enzymes like protein kinases and phosphatases⁴,²⁹,³⁰,³¹—major modulators of protein phosphorylation. While these and other efforts revealed that signaling may regulate most stages of parasite development, the underlying mechanisms remained unclear and the signaling map of the parasite remains ambiguous.

Efforts to understand the role of second messengers like calcium and phosphoinositides (PIPs) in the parasite have resulted in explaining novel signaling and trafficking pathways. For instance, researchers have shown that phospholipase C-mediated calcium release may regulate protein kinases like PIPCDPKs and PIPKB, which in turn may regulate key processes like host erythrocyte invasion and sexual differentiation²².

Using mass spectrometry, several regulatory phosphorylation sites on PIPCDPK1³¹ and its substrates like PIPAP45³² have been identified. Indian teams have also identified several novel substrates for *Plasmodium* kinases. Some of these substrates have been validated by performing *in vitro* kinase assays with recombinant substrate proteins and LC-MS/MS analysis.

A combination of traditional approaches with the omics approach will help understand the mechanism through which signaling pathways regulate the development of malaria parasite.

¹National Institute of Immunology, New Delhi, India (pushkar@nii.ac.in). ²International Centre for Genetic Engineering & Biotechnology, New Delhi, India (pawanm@icgeb.res.in; inderjeet@icgeb.res.in). ³Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India (sanjeeva@iith.ac.in).
Identifying diagnostic and prognostic markers

Most of the attention on infectious diseases of the developing world has focused on the development of rapid diagnostic tests and novel therapeutics to ensure timely treatment and improved survival rates. Existing diagnostic tests are either too expensive or time consuming or difficult to implement in developing countries due to the lack of resources and expertise. The inability to predict disease severity is also a major challenge to effective clinical management and prevention of long-term malaria complications. These limitations have spurred the search for better diagnostic and prognostic markers in malaria that can be easily measured in body fluids.

For over a decade now, several attempts to discover novel biomarkers in human bio-fluids such as serum, plasma and urine have been made by various research groups. Such studies have involved the use of proteomic technologies to profile host responses to infectious diseases. The high-throughput proteomic technology platforms not only investigate the systemic alterations of protein expression in response to diseases but also enable visualisation of the underlying interconnecting protein networks and signaling pathways, facilitating the discovery of unique markers of infection. Proteomic technologies have also been used to discover biomarkers that demonstrate the presence of the infecting organisms. One of the first attempts to unravel the proteome of the malaria parasite, Plasmodium vivax from clinical samples provided new leads towards the identification of diagnostic markers, novel therapeutic targets and an enhanced understanding of malaria pathogenesis.

Efforts to decipher host responses to malaria infection have revealed a panel of proteins with a distinct pattern of differential abundance that can discriminate malaria patients from healthy subjects and patients with other infectious diseases. A comparative serum proteomic analysis of severe and non-severe malaria in search of prognostic markers using quantitative proteomics has highlighted the presence of muscular, cytoskeletal and anti-oxidant proteins in patient sera revealing extensive oxidative stress and cellular damage in severe malaria. These findings are currently being validated in a larger cohort of patients using immunoassays.

The application of proteomic technologies has shown promising leads. However, early disease detection, measurement of therapeutic efficacy, prediction of disease severity and tailored patient therapy are still some distance away.

References

OMICS Conferences

6th International Conference on Bioinformatics
Mar 29-30, 2016 Valencia, Spain

6th International Conference and Expo on Proteomics
Mar 31-Apr 2, 2016 Atlanta, USA

5th International Conference and Exhibition on Metabolomics
May 16-18, 2016 Osaka, Japan

International Conference on Structural Biology
Jun 23-24, 2016 New Orleans, USA

International Conference on Next Generation Sequencing
Jul 21-22, 2016 Berlin, Germany

International Conference on Nucleic Acids
Aug 04-06, 2016 Seattle, USA

2nd International Conference on Geneexpression & Transcriptomics
Aug 18-20, 2016 Portland, Oregon

Biotechnology World Convention
Aug 18-20, 2016 Sao Paulo, Brazil

2nd Glycobiology World Congress
Aug 29-31, 2016 Atlanta, USA

Indian Conferences

Indo-Global Summit and Expo on Health Care
Oct 05-07, 2015 New Delhi, India

6th World Congress on Biotechnology
Oct 05-07, 2015 New Delhi, India

7th Indo- Global Summit and Expo on Food & Beverages
Oct 06-10, 2015 New Delhi, India

Pharmaceutical Summit and Expo
Oct 08-10, 2015 New Delhi, India

6th Indo Congress and Expo on Dental & Oral Health
Oct 19-21, 2015 Bangalore, India

7th Indo-Global Diabetes Summit & Medicare Expo
Oct 19-21, 2015 Bangalore, India

International Conference on Lipid Science
Oct 06-08, 2016 Miami, USA

International Conference on Biochemistry
Oct 13-15, 2016 Kuala Lumpur, Malaysia

7th International Conference & Expo on Proteomics
Oct 24-26, 2016 Rome, Italy

7th International Conference on Bioinformatics
Oct 27-28, 2016 Chicago, USA

2nd International Conference on Genetic and Protein Engineering
Nov 14-16, 2016 Atlanta, USA

6th International Conference and Exhibition on Metabolomics
Nov 28-30, 2016 Chicago, USA

International Conference on Glycomics
Dec 01-03, 2016 Chicago, USA

World Congress on Amino Acids and Proteins
Dec 08-09, 2016 Baltimore, USA

Contact us
America: 2360 Corporate Circle, Suite 400 Henderson NV 89074-7722, USA   Tel: +1-650-268-9744,
Europe: Traian Str., No. 61, Constanta 900720, Romania
Asia-Pacific: 6th Floor, North Block, Divyasree Building, Raidurg, Hyderabad, India 500032, Tel: 040-33432300
contact@conferenceseries.com
OMICS International believes in the fact that education is the base for progress of any society and emphasizes its ideology through knowledge dissemination using its worldwide Medical and Healthcare conferences and Journals as tools to connect the global educational, scientific, technological and entrepreneurial communities.

Information is power and OMICS International has been serving as a power capacitor to rejuvenate collaborations across the globe in various fields of study and research through its vast number of Medical, pharmaceutical, Clinical, Management and Life science Journals, Conferences, Workshops and Symposium.

OMICS International took a quantum leap by acquiring New York based Private Equity Forums, and Allied Academies of USA, Insight Medical Publications (ImedPub), UK and Research & Review journals of India. With 700 + peer reviewed, Open Access Journals on Clinical, Medical, Pharmaceutical, Management, Engineering & Technology and Life sciences, OMICS International has been enhancing its portfolio in Science, Technology, Business, Management and Academic fields.

Our journals are double peer reviewed, and strictly adhere to the publication standards. They are contributing immensely to the academic and scholarly research by offering unrestricted flow of well researched information. This information, free from all subscription hurdles is well archived and indexed in reputed database including Scopus, CHORUS, PubMed, ISI, Chemical Abstracts Service, ProQuest, and Gale EBSCO.

OMICS International is equally instrumental in spreading its presence to great global destinations through its series of ‘Science Events’ on diversified and unique topics through International science conferences, Seminars, Symposia and Workshop on Environmental, Chemical, Engineering, Healthcare, Business and Pharmacy themes across USA, Europe, Asia, Middle East, Australia, Germany, Spain, China, Canada, Italy, South Korea, South Africa, Japan, and Turkey.

OMICS International could carve a niche of its own through its 700+ science journals and 1000+ International Science events bringing together academia, scientific, research and business communities through its panel of Editors and Conference Organizing Committees. Presence of 75 000+ reputed scientists, scholars and academicians with hands on expertise in their respective fields in its editorial panel is adding value to our journals and conferences.

Noble laureates, Reputed Scientists for excellence, outstanding researchers with several patents on their kitty, veteran academicians, CEOs and Entrepreneur of top management cadre that are rich in theory and practice grace our conferences to debate the novice and the latest happenings in the field of science and technology.

B2B & Global Networking:
OMICS International B2B meetings help the participants in maximizing the benefits as they provide a forum to meet the successful business and trade gurus of the globe. Our interactive sessions with participants from across the knowledge society including government, local bodies, NGOs and corporate, would enable the participants to pick relevant contacts and build association over it to further their business and academic interests. Companies connect with each other in an effort to build up common strategies in this fast globalizing world with a motto of ‘Earning while Learning’ through these B2B sessions. B2B meetings are pre-arranged and targeted appointments. This would help buyers and exhibitors to meet each other in a simple, efficient and pro-active manner to work together.

OMICS International Conferences on Successful path:
OMICS International’s world class science events on Cardiology, Dental, Biotechnology, Food Technology, Dementia, Nutritional Science, Vaccines, Immunology, Materials Sciences, Virology, Endocrinology, Toxicology, Cancer Science, Cell Therapy, STD &AIDS, Petro Chemistry, Diabetes, Nursing, Nanotech, Obesity, and Clinical Pharmacy have registered a great success during 2015-16. Its business conferences have witnessed the presence of a cross section of business and corporate, denoting a steady progress of participants including 22% of Investors, 11% of Corporate Business Executives, 23% of C-Level Executives, 21% of Academy and Industry, 11% of Clinicians/Physicians, 2% of Service Providers, and 10% of Manufacturers and R&D. OMICS international Academic conferences could attract 12% of Executives, 11% of Directors/CEO’s, 21% of Research Analysts, 23% of Scientists, 8% of Managers, 8% of Associates, 12% of Students and 5% of participants constitutes other categories.

OMICS International is keen in delivering every innovation over the years to come and its success can be attributed to its readers, researchers and participants that have been extending relentless support.
Plant proteomics: A potential approach for food security

Subhra Chakraborty¹, Renu Deswal² & Niranjan Chakraborty¹

Genome sizes are increasing but gene numbers are not going up proportionately. This interesting nugget highlights the fact that higher organisms, including plants, might be engaging in combinatorial interactions between genes and gene-products, the proteins, to achieve cellular complexity.

Living cells are exceptionally complex and may consist of more than 100,000 protein species at any given time with different physical and functional properties. The bridge between proteome and phenome is at the core of biological processes. Proteomics involves the determination of the proteome using large scale protein identification and elucidation of their function. Proteins are required not only in metabolic activities in cell, but also play a key role in integration of internal and external signals. They are subjected to a constant turnover maintaining cellular homeostasis, an essential feature of their regulation¹. The misregulation of protein expression causes an imbalance in cell milieu, which eventually affects plant growth and development, and reduces crop yield.

Plants contribute to economic sustainability and security, being the source of food, feed, fiber, and fuel. India possesses substantial plant biodiversity. The country's agriculture contributes 8% of the world's agricultural gross domestic product and supports 18% of global population. Nevertheless, about 80% of India's land mass is highly vulnerable to external threats². Therefore, conservation of biodiversity, acceleration of plant productivity and nutritional security are of paramount importance.

Designer crops

Currently plant breeders and plant biologists are focusing on the development of designer crops that are better equipped to withstand a wider range of climatic variability and have better nutrient availability. Conventional breeding approaches are handicapped because breeders require precise gene modifications with targeted traits. In the post-genomic era, the integration of proteomics into plant science will help accelerate the development of new generation of food crops.

Despite having scientific infrastructure and potential, India was not a part of the international human genome sequencing project. The country got involved in genome sequencing in June 2000 with the International Rice Genome Sequencing Project (IRGSP), and chose to sequence a part of chromosome 11. Since then plant

¹National Institute of Plant Genome Research, New Delhi, India (schakraborty@nipgr.ac.in; nchakraborty@nipgr.ac.in). ²Department of Botany, University of Delhi, Delhi, India (rdeswal@botany.du.ac.in).
genomics research in India has continued to advance. In contrast, plant proteomics has taken a back seat although it is not just complementary to the genomic information generated by genome research, but also indispensable to the country's science.

Until now, plant proteomic researchers were studying complex biological processes, defining localisation of protein species, and elucidating modification and functionality of protein complexes. Many plant proteomics laboratories in India are using high resolution mass spectrometers for systematic analysis of proteomes at the cell, tissue, and organ or organism level. The proteomics community in India has mostly followed a gel-based approach in combination with different versions of mass spectrometers. Stress proteomics is the most investigated area followed by the study of growth and development and post-translational modifications. It is the need of the hour to adopt and include newer gel-free, label-free approaches, advance imaging systems, high-end mass spectrometer instrumentation, data mining through multivariant analysis, proteogenomics and improvised and corrected databases for faster and accurate analysis.

**Targeted proteomics**

‘Targeted proteomics’ must be adopted to validate potentially important biomarkers for plant research using multiple reaction monitoring (MRM), selected reaction monitoring (SRM) or parallel reaction monitoring (PRM). The proteomics community should aim to provide sets of peptide markers, which are unique spatially, or temporally to give a snapshot of spatio-temporal differential proteome. Extensive genome annotations along with new genome sequence information would help tailor diverse marker peptides for broader applicability. Additionally, nano-proteomics, in combination with micro-dissection and single cell system analysis could distil the information tremendously and provide novel insights.

A systems biology approach, where integration of proteomics with genomics and phenomics would enhance the quality and meaning of the derived biological information, needs to be taken. During the past decades, an array of protein chemistry techniques have generated huge amounts of knowledge on the function and molecular properties of individual proteins. However, proteins rarely act alone, they often team up, and function as complex molecular machines. We now understand that the plant proteome is not only a static linear set of proteins, but exist as a dynamic web of informational interactions that sustain the developmental process and allow its evolutionary modification.

Recent developments in gathering large scale proteomic information pose substantial challenges to bioinformatic processing of this data. The most challenging tasks of proteomics research range from sample preparation, separation of proteome complex and database processing to functional interpretation of biological significance. Therefore, inclusion of ‘interaction proteomics’ would help overcome such challenges, which will eventually generate new knowledge. Development of methods to systematically study protein complexes and their functional annotation opens up new avenues of plant research. It is very likely that such studies will unravel new complexes and their functional annotation opens up new avenues of knowledge. Development of methods to systematically study protein complexes and their functional annotation opens up new avenues of biological processes, defining localisation of protein species, and elucidating modification and functionality of protein complexes. Many plant proteomics laboratories in India are using high resolution mass spectrometers for systematic analysis of proteomes at the cell, tissue, and organ or organism level. The proteomics community in India has mostly followed a gel-based approach in combination with different versions of mass spectrometers. Stress proteomics is the most investigated area followed by the study of growth and development and post-translational modifications. It is the need of the hour to adopt and include newer gel-free, label-free approaches, advance imaging systems, high-end mass spectrometer instrumentation, data mining through multivariant analysis, proteogenomics and improvised and corrected databases for faster and accurate analysis.

**Moulding policy**

To stake India’s claim to leadership in plant research, scientists and policy makers must set a goal over the next 10 years by combining traditional know-how and new proteomic technology strengths, focusing attention on sustainable food and nutrition security. It is obvious that no single entity can effectively achieve such a goal without the active participation of academia, industry, and funding agencies. Since the birth of plant proteomics research in India, the country has been positioned to convene the parties and facilitate the strategies and activities crucial for the success of such endeavours.

The challenge ahead is to develop methods that would allow the generation of new testable hypothesis based on pre-existing biological information. The approach of gene discovery through proteomics is currently proving to be an effective way to speed up crop improvement programmes worldwide. India must harness the fruits of proteomics technology in collaboration with the international community, just the way the genomics research community has done. In a not too distant future, proteomics researchers might find proteomic blueprints of most crop species. This may provide plant breeders the resource to provide food security to the growing population in India and the world as a whole.

**References**

Proteomics and Bioinformatics Lab of Jamia Millia Islamia, New Delhi was established in the year 2007 with an objective to create a plant proteomics facility to study basic proteome profiles of different plant organs (N2-fixing nodule, root and leaf) and plastids (chloroplasts and thylakoids). The lab specializes in membrane proteomics and focuses on understanding the modulation of multi-protein complexes in response to iron, sulfur and cadmium stress (studied in Arabidopsis thaliana, Brassica juncea, Parthenium hysterophorus, Spinacia oleracea and Glycine max). The objectives are achieved mainly through gel-based proteomics (2DE: IEF-SDS-PAGE) and Blue Native (BN)-PAGE (for MPC separation) followed by tryptic protein digestion and peptide mass fingerprinting through MALDI/ESI-MS/MS. The proteins are identified through protein databases and analysed using various bioinformatics platforms. The lab is interested in working out the major protein players against individual as well as multiple stresses in plants.

Indian mustard is a major oilseed crop which is heavily dependent on proper Fe, S and N supplies. Lab has worked out impact of Cd under such nutrient deficiencies. Numerous proteins subunits of PSI, PSII, Cytb6f and ATPase have been identified that are compromised by Cd stress under Fe-deficiency [1]. Further, proteomic changes under said conditions are of prime target. Arabidopsis thaliana is considered for studying basic physiochemical [2] and proteomic changes [3, and unpublished data] revealing some interesting facts of metabolic and proteomic shifts as adaptation to stresses. Among proteins, found associated with stress-tolerance, KIN1, HSP23.5 and LCR17 are being further studied.

With exponential population growth, gap between demand and supply is widening and forcing us to use N2-fertilizers, excessive use of which is not good for environment and animal health. N2-fixers are always considered as best substitute of N2-fertilizers being associated with mutual sustenance with higher plants. Our lab utilizes comparative gel-based/shotgun proteomics and real-time qPCR for identification of stress-responsive proteins and transcripts in Soybean nodules. Besides proteome and transcriptome, a huge change in metabolome was observed in nodules in response to abiotic stresses to confer tolerance [3, and unpublished data].

Parthenium hysterophorus is an invasive weed which is used to get metabolomic and proteomic clues [4] regarding its stress-wielding molecular mechanisms and allelopathy and allergies [5, and unpublished data]. A number of important proteins responsible to impart extraordinary stress tolerance to Parthenium have been identified. The next approach is to develop knockout mutants of selective stress-imparting genes employing CRISPR/Cas9 genome editing tool.

As a fact, plant usually faces multiple stresses at a single time. Considering this, multiplexed stress was exposed to Spinach helping in identifying the common minimum proteome playing role under multiple stresses [6]. Study reveals the identity of some such important proteins including some already know but many novel ones supposed to be anti-stress [7].

**References:**

[4]. Bashir et al. 2015, Under revision
[5]. Qureshi MI. Introducing the Weedomics. ICP-2010, IARI, India
[6]. Ahmad et al. 2015. Nat. Prod. Res. Accepted
NIAS Consciousness Studies Programme
NATIONAL INSTITUTE OF ADVANCED STUDIES
Bengaluru, INDIA
presents

CONSCIOUSNESS, COGNITION & CULTURE
Implications for the 21st Century

an international conference
for researchers, scholars and students to discuss latest research
work on brain, self, cognition and consciousness,
and to meet wellknown experts from around the world

9 -11 Dec, 2015
JRD Tata Auditorium, NIAS

REGISTER TODAY
write to <niasconsciousnessprogramme@gmail.com>
website: niasconsciousnesscentre.com
OMICS International Conferences

1000+ Annual Global Events
Live broadcasting of conferences across the globe
Meet inspiring speakers and experts
Panel discussions and interactive sessions
B2B meetings & World class exhibitions
Global networking opportunities
1200+ National Symposia in Reputed Universities
1200+ Workshops
1000$ Travel Grant to each Young Scientist or Student

OMICS International Journals

700+ Open Access Journals
21 Day rapid review process
75000+ Editorial team
10 Million readers
Publication immediately after acceptance
Quality and quick editorial, review processing

1000+ Global Events on Medical, Pharma, Engineering, Science, Technology and Business

2015-16 Major Destinations-25+ Countries

USA Spain UAE Canada Germany Australia India UK
Italy Brazil France Japan Malaysia Singapore South Africa
Sweden Norway Belgium Ireland Russia Turkey Hong Kong

Israel South Korea China Greece

Covering All Major Disciplines
Medical | Chemistry | Mass Media | Pharma | Genetics | Clinical | Materials Science | Chemical Engineering | Agri, Food & Aqua | Healthcare | Molecular Biology
Earth Science | Nursing | Physics | EEE | Microbiology | Immunology | Veterinary | Neuroscience | Engineering | Business & Management

Contact us
America
2360 Corporate Circle. Suite 400 Henderson NV 89074-7722, USA Tel: +1-650-268-9744

Europe
Traian Str., No. 61, Constanta 900720, Romania

Asia-Pacific
6th Floor, North Block, Divyasree Building, Raidurg, Hyderabad, 500032 Tel: 040-33432300

contact@conferenceseries.com

Proceedings of OMICS International Conferences 2014-15

Medical 17800
Pharmaceutical 15580
Clinical 15565
Life Sciences 15575
Engineering 12800
Nursing & Healthcare 4500
Dental & Diabetes 5600