Chapter 1

Introduction

1.1 Sweet potato

Sweet potato [*Ipomoea batatas* L. (Lam.)], is a dicotyledonous plant belongs to the family *Convolvulaceae* (morning glory). It is a perennial plant which serves as one of a major food and animal feed, besides industrial raw materials with a significant contribution as energy supplement and phytochemical source of nutrition. It is widely cultivated in the tropics, sub tropics and even in some temperate zones of the developing world [1]. Sweet potato is the seventh most important crop in terms of global production for food and other industrial materials after wheat, rice, maize, potato, barley, and cassava. In developing countries, it ranks third in value of production and fifth in caloric contribution to human diet. Based on the variability in flesh color, white-fleshed to cream-fleshed sweet potato is widespread in the Pacific whereas yellow- to orange-fleshed sweet potato is predominantly found in the United States [2, 3]. It is herbaceous perennial crop having alternate heart-shaped or palmately lobed leaves and intermediate sized sympetalous flowers. It is the only economically important plant of the family *Convolvulaceae*. Climbing stems, branches, spirally arranged leaves on stems and roots are the unique characteristic features of sweet potato plant(s). In addition, single plant produces one to quite a few tuberous roots at maturity.

Sweet potato contains a significant level of starch, soluble sugars, vitamins, minerals, and other nutrients. It has 30% more starch than that of rice and corn and 49% than wheat and is a storehouse of numerous vital pigments like β-carotene, anthocyanin, among others, which act as excellent antioxidants. Developing countries produce more than 95% of the global sweet potato crop and majority of the population in these countries highly dependent on this tuberous root crop for food, nutrition, and income. In addition to nutritional benefits, its adaptability to broad range of agroecological conditions and minimal growth requirements, favors it as a crop of high commercial significance [4, 5]. It has recently been found out that

Taxonomic classification of Sweet potato	
Kingdom:	Plantae
Division:	Magnoliophyta
Class:	Magnoliopsida
Order:	Polemoniales
Family:	Convolvulaceae
Genus:	*Ipomoea*
Species:	*Ipomoea batata*

sweet potato can control blood sugar levels and insulin resistance and it is labeled as anti-diabetic food by World Health Organization. A daily intake of 100 g of orange fleshed sweet potato is anticipated to prevent vitamin A deficiency in children and lactating mothers, according to recent studies. Contribution of sweet potato as an important food for health is acknowledged due to the high nutrient contents and its anti-carcinogenic and cardiovascular disease-preventing properties. Almost all cultivars of sweet potato are excellent source of vitamins C, B_2, B_6 and E as well as dietary fiber, potassium, copper, manganese and iron, and are low in fat and cholesterol [4, 6, 7].

1.1.1 Sweet potato: Origin and importance

Sweet potato [*Ipomoea batatas* (L.) Lam] is originated in Central America and Northern South America or Mexico and has been cultivated initially at least 5000 years ago. It was subsequently dispersed to North America, Europe and Pacific. Columbus brought this crop to Spain and was introduced to Asia and Africa in due course. Family *Convolvulaceae* includes approximately 45 genera and 1000 species; however, only sweet potato is of economic significance as food. The cultivated sweet potato is hexaploid (2n=6x=90) with high level of variations within the species having a DNA content of 1050 Mbp. The complex polyploidy of sweet potato may have been a derivative of the ancient polyploidyzation of *I. triloba* L. (diploid, x=15) (A genome type) and two other B genome types, namely, *I. trifida* (diploid, x=15) and *I. tabascana* (tetraploid) [3, 6].

Overnutrition (excessive intake of nutrients) rather than undernutrition presents a major health challenge at present, in many developed countries. Nonetheless, undernutrition, food insecurity issues, droughts, and inadequate agricultural technologies are major problems from a global perspective. Most of the cultivation in developing countries is highly dependent on root and tuber crops, as contributing, either major sources of food and nutrition or a source of income [8]. From this point of view, it's imperative to critically reevaluate versatile, easily available root and tuber crops with wide ecological adaptability for their efficacy in human nutrition. Sweet potato is one such crop due to its high harvest index and drought tolerance, with a broad range of adaptability to diverse climates and farming systems [9, 10]. Furthermore, it has been emphasized that the sweet potato roots and leaves can support more people per unit hectare than

any other food. The leaves of the sweet potato are dark green and are expected to have nutritive values comparable to common dark green leafy vegetables [11, 12].

The nutritive value of sweet potato is very high as it ranks as one of the healthiest vegetables, because of high levels of vitamins A and C, iron, potassium, and fiber. They are also an excellent source of the vitamin A precursor, β-carotene. As described earlier, the importance of sweet potato is increased significantly due to its easily cultivable nature, low requirements of fertilizer, high adaptability, drought tolerance and the multiple uses of the crop as animal feed and varied processed forms for human consumption (**Fig 1.1 A**). Sweet potato is considered as a security crop and a major staple food crop in subsistence and rural economies. In addition to its direct use as source of food and feed-stock for animals, sweet potato is also a candidate for the production of renewable plant products such as ethanol, high-grade starches, stable natural dyes and vitamin precursors [11].

1.1.1.1 Sweet potato: An efficient source of energy

An adequate, healthy diet must satisfy human needs for energy and all essential nutrients and energy is required to sustain the cellular metabolism. Furthermore, dietary energy needs and recommendations cannot be considered in isolation of other nutrients in the diet, as the lack of one will influence the others. Sweet potato in particular, not only the tuberous roots but vines is also a good source of energy. Excluding cassava (590 kJ), sweet potato (465kJ) is a superior source of energy than other root and tuber crops including potato (335 kJ), yam (434 kJ) and taro (432 kJ). Sweet potato starch is easily digestible and consequently a useful constituent in the preparation of excellent weaning meals [13]. A wide range of nutrients including carbohydrates, vitamins, minerals and proteins are present in varying levels in the different sweet potato genotypes and plant parts. In comparison to cereals sweet potato is a very efficient food crop and produces more dry matter, protein and minerals per unit area. It has been reported that sweet potatoes being the staple food in the developed countries account for 130 kcal of energy per person per day against 41 kcal in the developing countries where it is still considered as vegetable. Despite of being a rich source of starch, sweet potatoes contain an excellent quantity of secondary metabolites and small molecules which play a crucial role in various processes. Many of the compounds present in sweet potato are imperative because of their beneficial effects on health, therefore, are highly desirable in the human diet and functions as an efficient food

[14]. Sweet potato can, and does, play a diverse roles in the human diets being either supplemental or an extravagance food besides being a staple crop for some parts of the world. In Asia, sweet potato uses range from supplementary food of diminutive status to a very vital supplementary food to rice and/or other root and tuber crops [15]. In United States and other developed countries, the use of sweet potato is strictly as a luxury food and in other parts of the world such as in Japan, where it plays its role as novel plant products and/or nutraceuticals [16].

1.1.1.2 Biochemical and nutritional composition of sweet potato

Due to its nutritional composition and exceptional agronomic features sweet potato has immense prospects to facilitate and/or prevent and reduce food insecurity and mal-, under-, and overnutrition in developing and developed countries. However, dearth of information regarding the nutritional composition of sweet potato to a great extent limits its exploitation. Improved knowledge of the nutritional quality, utilization, and future economic importance of the crop has important implications for human food systems, nationally and internationally. The sweet potato contains many nutrients including carbohydrates, protein, minerals (calcium, iron, and potassium), vitamins, carotenoids, dietary fiber, very little fat, and sodium [17]. As described in the concomitant sections, the nutrient composition of the sweet potato varies greatly according to genetic and environmental factors. Compositional differences with reference to proximate composition, starch, total sugars, and minerals possibly reflect the genetic effects rather than environmental effects. While differences exist even when mostly samples were subjected to similar agronomic practices and cultivated precisely in the same environment. However, it should be noted that environmental factors are also known to contribute significantly to the variability in nutrient composition [11].

1.1.1.2.1 Carbohydrate

In most sweet potato cultivars, the percentage of dry matter is reported up to 44%, out of which approximately 90% is carbohydrate [11]. The major carbohydrate components in sweet potatoes are starch having 30-40% amylose and 60-70% amylopectin [18-21]. In raw uncooked roots sucrose is found to be the main sugar although, glucose and fructose are also present. On the other hand in cooked roots maltose, a major product of starch conversion is usually found

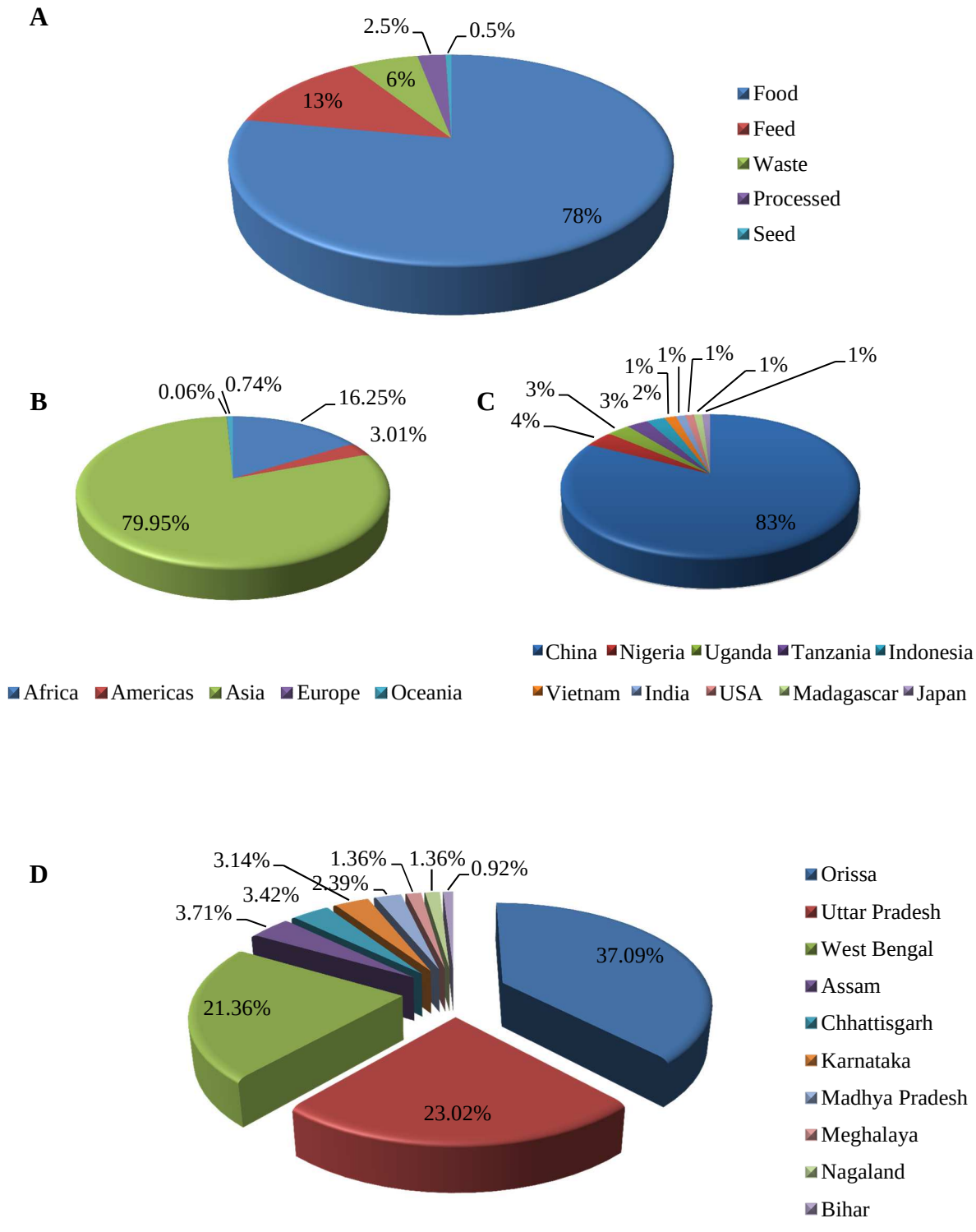


Fig. 1.1 Production and uses of sweet potato. Different uses of sweet potato (**A**), Production of sweet potato in different continents (**B**) and Top ten sweet potato producing countries (**C**) and Production of sweet potato in India (D). (Source FAOSTAT, 2013).

[22]. The rest of carbohydrates (principally cellulose, hemicelluloses and pectins) are known as fibers collectively.

1.1.1.2.2 Protein

A variable range of protein content is reported in sweet potato by several workers. The variations in protein content in sweet potato have been attributed to their genetic background, environmental niche [23]. In some part of the world for instances Africa and Japan, the leaves of the sweet potato are also used, and the protein content has been reported to be as high as 27% protein on dry weight basis (dwb). However, 18.4% protein content in sweet potato leaves was also reported [24, 25]. In general, sweet potato is considered as a low protein crop ranging from 1 to >10%, and mostly ranging between 1.0 and 8.5% in roots [17, 26]. **Ravindran, et al. (1995)** had reported a variable range of protein (2.95-7.1985 g/100g) in different sweet potato cultivars [27]. However, Sosinski (2002) reported an exceptionally high protein content of 14.2% [15]. With reference to FAO reference protein, sweet potato protein contains a good quality of essential amino acids except tryptophan and sulphur amino acids [28].

Recently sweet potato became an apprehension in the applications of food processing, and has received much attention from breeders for the possibilities of its genetic improvement(s) due to its lower protein content.

1.1.1.2.3 Fiber and ash content

Dietary fibers are extremely important for noncommunicable disease prevention. Dietary fiber contains all indigestible polysaccharides and lignin of the diet by human enzymes. It includes resistant starches, cellulose, hemicelluloses, ß-glucans, pectins, nonstarch polysaccharides and lignin. For populations that consume sweet potato as a staple food, its dietary fiber contribution could be vital [29, 30]. An extensive disparity in fiber and ash contents in different roots and tuber crops including sweet potato is evident from various reports. The fiber content in particular varies to a great extent depending on varietal differences and age of the crop, wherein the fiber content increases with the maturity. The total dietary fiber content of orange fleshed sweet potato (OFSP) cultivars has been reported in a range from 2.0-3.2 g/100 g fwb, which is higher than of 0.7 g/100 g listed in the United States Department of Agriculture (USDA) database. Whereas, 2.3-3.9 g/100 g and 2.3-3.3 g/100g is reported, respectively for

purple-fleshed and yellow/white-fleshed sweet potato cultivars [31]. Oboh et al. (1989) had reported a range of 3.8-5.9% of crude fibers in 49 varieties of sweet potatoes analyzed [32].

The crude ash content basically represents the mineral content of a food. Ash content of sweet potato is reported usually between 3- 4% and 1.55-2.04% (dwb). Moreover, ash content of the sweet potato peel is found to be higher than that of the flesh [11, 33]. The average ash content in the peel and the flesh is reported as 4.1% and 4.6% (dwb) respectively [34].

1.1.1.2.4 Vitamins and minerals

Sweet potato is a concentrated source of many essential vitamins and minerals. It is reported to contain significant amount of ascorbic acid (vitamin C), moderate quantities of thiamin (vitamin B_1), riboflavin (B_2) and niacin. Besides this it contains a little amount of pantothenic acid (B_5), pyridoxine and its derivatives (B_6), folic acid and has also been reported to contain reasonable quantities of tocopherol (vitamin E). It has been argued that the mineral content of agricultural crops including sweet potato varies with geographic location. In most of the cases potassium are considered to be the most abundant mineral in sweet potato tubers followed by magnesium and calcium [35]. Nevertheless, **Makki et al. (1986)** reported a higher concentration of calcium than that of potassium [36]. Above and beyond this manganese, iron, copper, and zinc are also present in low amounts [17].

1.1.1.2.5 Antioxidants: phytophenols and carotenoids

Sweet potatoes are rich in antioxidants, such as phytophenolics and carotenoids, which also provide distinct flesh colors. The phytophenols and carotenoids often vary across the genotypes and are thought to be associated with genetic factors, which play a vital role in the formation of secondary metabolites [7]. In recent times, phenolic compounds and their succeeding involvement to the antioxidant capacity of sweet potato have been receiving more attention as they append functional characteristics to it and are considered positive attributes. Whereas the focus is usually on the storage roots as an important food crop, in many countries the aerial portion of sweet potatoes are also used as leafy vegetables. Total phytophenolic compounds present in the leaves were also found to be greater than that of the concentrations in the stems and storage roots [37]. Anthocyanins are a collection of water- soluble pigments accountable for the attractive colors of fruits and vegetables. Anthocyanins obtained from plants

are frequently used in soft drinks, jams, confectioneries, and bakery products as a natural colorant. Deep purple sweet potato flour (SPF) and paste are also used as coloring materials for bread, snacks, and noodles [38, 39]. Due to much higher color retention and more stability than commercially available colorant, sweet potato colorant are more prevalently used [40]. Sweet potato accumulates anthocyanins in the storage root and much interest has been given to categorize the types and structure of those responsible for imparting color. Several workers have isolated acylated as well as diacylated anthocyanins from sweet potato [41-44]. The acylation found in sweet potato anthocyanins is an important characteristic as it adds heat and light stability to the pigment, which has applications in food processing [41]. Sweet potato is also rich in carotenoid contents and more specifically in β-carotene. While orange-fleshed cultivars are putatively known for their high β-carotene content, high genetic variation has also been observed [11, 45-47]. Moreover, carotenoids, such as β-carotene are also used as food colorants due to their natural origin, lack of toxicity, and flexibility of providing both lipo- and hydrosoluble colorants [48].

Table 1.1: Nutritional composition of raw sweet potato (per 100 g).

Nutrient	**Value (per 100 g)**	**Nutrient**	**Value (per 100 g)**	**Nutrient**	**Value (per 100 g)**	**Nutrient**	**Value (per 100 g)**
Water	77.28 g	Ash	0.99 g	Potassium	337 mg	Vitamin B2	0.1 mg
Energy	360.00 kJ	Crude fiber	3.0 g	Phosphorus	47 mg	Vitamin B3	0.61 mg
Carbohydrate	20.12 g	Calcium	30.00 g	Sodium	55 mg	Vitamin B5	0.8 mg
Starch	12.7 g	Iron	0.61 mg	Vitamin C	2.4 mg	Vitamin B6	0.21 mg
Protein	1.6 g	Magnesium	25 mg	Vitamin A	14187 IU	Vitamin B9	11 µg
Total lipid (fat)	0.05 g	Phosphorus	47 mg	Vitamin B1	0.1 mg	ß-carotene	8509 µg

Source: USDA National Nutrient Database (2009)

1.1.1.3 Health benefits of sweet potato

In recent years, quite a few reports suggested that the phytochemicals in sweet potato display antioxidative or radical-scavenging activity and exerted several health-promoting functions. Besides providing distinctive flesh colors to sweet potatoes, phytophenols and carotenoids also provide potential antioxidant activities. [7, 50, 51]. Use of the sweet potato has always been related with excellent health and improved nutrition. However, inadequate research has addressed the role of sweet potato in human nutrition and health. Cellular damage from reactive oxygen species (ROS) is ubiquitous. ROS are known to have carcinogenic potential and natural antioxidants such as phytophenols and carotenoids are protective against ROS by neutralizing their effects [52, 53]. It has been reported that white, yellow, orange, and purple-fleshed sweet potato cv. have antioxidant and radical-scavenging activities. Interestingly, the antioxidant activity in sweet potato skin was found to be more or less three times higher with respect to the other parts. Purple-fleshed sweet potato extracts are known to inhibit reverse mutation in some pathogens which was conjectured to be due to the anthocyanin pigments [54, 55]. Antimutagenicity of anthocyanins isolated from some sweet potato cultivars having high instances of inhibition against some common mutagens. The high concentration of anthocyanin and β-carotene in sweet potato and other dietary carotenoids, are putatively known to account for anticarcinogenic/antitumorigenic properties. Anthocyanins isolated from some cultivars of sweet potato have been evaluated for their ability to suppress glucose metabolism via inhibition of action of α-glucosidase. It provides a means to treat with noninsulin-dependent diabetes by delaying glucose absorption in the small intestine thereby preventing an excessive rise in blood glucose levels or hyperglycemia [56]. An antidiabetic effect in diabetes mellitus and antibacterial efficacy of sweet potato is also well documented [57]. Sweet potato also contains the compounds like coumarins scopoletin (3, R-H), aesculetin (3, R-OH), and umbelliferone (3, R-OMe), which have anticoagulation properties and are predicated to restrain HIV replication [58].

1.1.2 Sweet potato: Production and distribution across the globe

Sweet potatoes are the seventh most important food crop grown in around 111 countries. Nowadays, the main commercial contributors of sweet potatoes are China, Indonesia, Vietnam, Japan, India, and Uganda [5]. It is commercially grown as a high value vegetable crop under intensively managed production systems in most of the developed and developing countries.

Sweet potato is playing a major part as a food resource in Asia and Pacific islands. This region contributes more than 90% of the global sweet potato production. China is the major contributor which accounts for more than 80% of the world's total production alone. India, Japan, Indonesia, Vietnam, Republic of Korea and Philippines have the biggest area of landscape under sweet potato cultivation (**Fig. 1.1 B, C**). Africa's largest producing countries are Rwanda and Uganda. Sweet potato production in Latin America and in Caribbean is comparatively small [59].

India produces about one million tonnes of sweet potato (**Fig. 1.1 D**) and different varieties are grown in various parts of the country. In India Orissa is the largest sweet potato producing state followed by Uttar Pradesh, West Bengal, Assam and Chhattisgarh.

1.1.3 Sweet potato: Botany and physiology

Tuberous root is the edible portion of the crop though young leaves and shoots are occasionally used ornamental vines or leafy vegetables. Depending upon the type of cultivars the root color varies from red to purple, brown and white whereas the root flesh color ranges from beige through white, yellow, red, violet, orange and purple [60]. However, storage root production is inconsistent from plant to plant. The fundamental genetic mechanisms and the factor(s) that promote storage root development, one of the most significant physiological processes in sweet potato, are not clearly understood. Storage roots *i.e.* tubers usually arise from the underground part of a vine cutting that is being used as planting material. Primarily, white adventitious roots develop, and some of these roots subsequently undergo changes rapidly in their growth pattern and expand into storage roots. The tuberization is a complex process involving various metabolic cues, leading to massive accumulation of starch and proteins. Tuber as a storage organ acts as a sink and competes for the available photo-assimilates. The allocation of photoassimilates is an important criterion for plant productivity based on the harvest index. In tuberous crops, a higher harvest index indicates an efficient diversion of photoassimilates sink [6, 61, 62]. The inconsistency in tuber yield among and within the sweet potato genotypes has been attributed to genotypes, propagation material and agroecological factors [63]. The composite interaction between environmental and genetic factors, source and sink relationship, photosynthesis, translocation and respiration in sweet potatoes is well documented. Besides their direct impact on tuberization, genetic and environmental factors also influence leaf area, leaf

production and abscission, photosynthesis, total dry matter production and dry matter partitioning [64, 65].

A tuber of sweet potato is developmentally and anatomically a true root, lacking nodes and related preformed meristematic tissue. Quite the opposite, potato (*Solanum tuberosum* L.) tubers are thickened underground stems, and therefore having meristematic buds found at vestigial nodes. Sweet potato tubers can instigate adventitious meristematic activity and produce sprouts and/or roots from the tuber. In the primary stage of root development, sweet potato initially forms colorless fibrous roots. Some of these fibrous roots become pigmented and begin to swell eventually developing into storage roots as the root development proceeds further. The storage root or tuber has been defined as the root in which there is anomalous secondary cambial activity inside the primary cambium [2, 66]. A straightforward link between lignification and storage root initiation, *i.e.*, lignification "prevents" storage root development [67]. Promotion of linear growth as well as the bulking of the storage root and/or yield have been revealed to be affected by ecological factors, as well as soil, light, photoperiod, temperature, humidity, carbon dioxide, and drought [2, 68-71].

Growth hormones play a pivotal role in the tuber development. Auxins and cytokinin levels have been found to be elevated throughout the initial stages of tuberization, suggesting the involvement of these hormones for the inception and concomitant primary thickening growth of storage roots [72, 73]. Induction of storage root formation in the presence of high sucrose concentrations by the exogenous application of cytokinin further corroborated this finding [74]. Tuber yield is mostly dependent on secondary growth that is the phenomenon of later stage of tuber development. Secondary thickening growth is positively correlated with the concentrations of abscisic acid (ABA) and cytokinin, on the other hand, a progressive decrease was observe with auxin levels [73, 75]. This led to the proposition that auxin, cytokinin and ABA possibly have different roles in the induction and/or thickening growth of storage roots. While, auxin is involved in the initial stages of formation of the storage root, ABA is in the later secondary thickening growth and cytokinin is active throughout the developmental stages. Whereas, treatment of GA3 had a negative impact on storage root development [76]. However, gibberelllin is traditionally associated with cell division and elongation. The upregulation of gibberellin-responsive *GASA* genes in storage root in storage roots in comparison to fibrous roots had

strengthened the possible role of gibberelllin in association with auxins in storage root development. Nevertheless, distinct role of each hormone has thus so far not been directly elucidated [6].

1.1.4 Cultivation

Sweet potato can be cultivated even in the adverse climatic conditions. Nonetheless, it can be grown in many agricultural environments and having quite a few natural enemies. It is a perennial crop but can be cultivated even as an annual in the tropics and subtropics. Sweet potato needs moderate temperature of 21-26°C and a well dispersed rainfall of 75-150 cm is adequate for its cultivation. Drought as well as water logging both is unfavorable for sweet potato yield. Drought stress is detrimental during the crop establishment and development of storage root [77]. It alters the source-sink liaison by affecting the production, translocation and partitioning of photosynthates. Extended periods of drought minimize the yield, quality of roots, and in severe cases causes overall crop failure [78]. At the initial stages of tuberization, drought in general, results in lower numbers of smaller sized storage roots which ultimately reduces sweet potato yield. However, to some extent it can tolerate drought but cannot survive in water logging. It requires ample of sunlight and shade aggravates yield reduction. Nonetheless, with the aim of crop intensification and profit maximization sweet potato can be intercropped with other seasonal crops like pigeonpea, maize, etc. [79]. Although, well drained loam and clay loam soils are suitable, sandy loam with clay subsoil are ideal for sweet potato farming. Due to its firmness, heavy clayey soils limit the storage root development whereas, sandy soils favour long cylindrical pencil like roots. Nevertheless, sweet potato is grown on a broad range of soils, well sapped light and average textured soil with a pH range of 4.5-7.0 is relatively favorable for it. Since, it is mostly grown in acidic soils an optimum soil pH range of 5.5-6.5 is found to be more suitable for growth and yield [79, 80]. Alkaline soil pH encourages pox and scurf diseases in sweet potato, whereas sweet potato suffers from aluminium toxicity in extremely acidic soil. Sweet potato is also susceptible to saline and alkaline conditions [81, 82]. It is mostly cultivated by vine cuttings and it is relatively easy to propagate [11, 79].

1.1.5 Sweet potato: Cultivars, varieties and hybrids

An improved understanding of germplasm diversity is critical for developing new varieties and helpful for conducting basic research into the biology of a crop. In sweet potato, cultivars play an important role in yield improvement. Many organizations working on sweet potato have the most important aim of developing location specific sweet potato cultivars. Different methods of breeding were used to develop the elite clones *viz.* clonal selection, open pollinated selection, hybridization, mutation and genetic engineering were mostly used universal methods [83]. More than 6000 accessions of landraces, breeding lines, and advanced cultivars of sweet potato were held by the International Potato Center (CIP), in its genebank, seems just a small fraction of the available varieties in the world. Nevertheless, many wild varieties have not been domesticated hitherto [84]. While, numerous sweet potato genotypes are maintained globally by national gene banks, few cultivars have predominated sweet potato cultivation in each major sweet potato producing country. The preference of cultivars depends largely on the yield and utility. As the cultivars considered vital, that is, extensively produced, differ from country to country, the dominant cultivars in some of the major sweet potato growing countries are given in **Table 1.2**.

Table 1.2 Sweet potato varieties grown in different countries.

Country	Variety
Bangladesh	BARI SP-6 (Lalkothi), BARI SP-7 (Kalmegh), BARI SP-8 and BARI SP-9
China	Xuzhou 18
Ethiopia	Balella and Bareda
India	Pusa Safed, Co-1, VL Sakarkand-6, Sree Nandini, Sree Vardhini, Co-2, Co-3, Samrat, Sree Bhadra, Sree Arun, Sree Varun, Sankar, Gouri, Kalinga, Gautam, Sourin, Kishan, H-41, H-42, H-268, Rajendra Sakarakand-5, Sree Rethna, Sree Kanaka, Kamala Sundari, S-1221, WBSP-4, Tripti, Bidhan Jagannath, BCSP-5, Birsa Sakarakand-1 and Indira Sakarakand-1
Japan	Quick Sweet, Kokei No.14, Beniazuma, Norin No.2, Koganesengan and Satsumahikari
Korea	Mokpo 32, Mokpo 34, Hongmi, Hwangmi, Sinjami, Shinmi, Wonmi, Poongmi, Borami and Enumi
Malaysia	Jalomas, Minamiyutaka, Pisang Kapas, Madu, Bawang, Gendut, Telong, Kangkung Cina, Ikan Selayang, Kangkung Kampung, Bukit Naga, Taiwan and Pasar Borong-1, Serdang, Suberang Perai, Kundang, Bidor, Pontian, Rhu Tapai, Sungai Baging and Kuala Linggi
New Zealand	Toka Toka Gold, Owairaka Red and Beauregard
Papua New Guinea (PNG)	Koitaki 2, K9, K42, UIB016, Wanmun murua, Wanmun and Large
Taiwan	Taoyuan No.1, Taoyuan No.2 and Tainung No.71
Thailand	PIS 205, PIS 65-16, PIS 166-5, Maejo and Taiwan
United States of America	Goldrush, Redgold, Centennial, Beauregard and Jewel
Vietnam	H12, K51 and TV1

Source: Nedunchezhiyan et al., 2012

1.1.6 Sweet potato: Problems, challenges and future perspectives

The agronomical development in sweet potato has elevated the tuber yield by 10-30% in most of the sweet potato growing countries. Nonetheless, sweet potato is known as a robust crop which is competent to withstand many adversities; numerous pre and/or post harvest abiotic as well as biotic constrains have been reported. Besides drought and water logging stress as described earlier, moisture stress and storage stress are few of the major abiotic stress reducing sweet potato yield. Moisture stress has become a major abiotic constrain to crop productivity worsened by climatic changes [85]. Soil moisture accessibility determines the external water status available for the plant and the internal plant water status as well inside the tissue of the plants. Plants encounter the moisture stress as the readily available soil water nearby the root zone is exhausted [86]. Photosynthetic performance and translocation of assimilates are significantly reduced by the drought stress thereby affecting the yield [87, 88]. The high moisture content is an indicator of lower storage quality in tubers. During postharvest storage, tuberous roots of sweet potato generally undergo a biotic and abiotic stress influencing protein expression pattern and substance contents. Therefore, postharvest storage of sweet potato is the bottleneck for industrial applications and the problems that should be addressed are rottenness and high rate of the loss of carbohydrates [89]. Besides few of the abiotic constrains biotic constraints such as weeds, insect pests and diseases have also been proved to be unfavorable for overall productivity of sweet potato. Sweet potato weevil is one of the major biotic impediments of sweet potato in the tropics [11, 90, 91]. Whereas, alternaria blight, sweet potato virus disease (SPVD) and root-knot nematodes (*Meloidogyne* spp) are mostly found in the temperate zones. Furthermore, SPVD is a combination of sweet potato chlorotic stunt virus (SPCSV) and sweet potato feathery mosaic virus (SPFMV) [92, 93]. In a nutshell, there are approximately 20 viruses or virus like diseases, roughly 35 bacterial and fungal diseases, about 20 nematodes, and about 20 insect pests that influence sweet potato [94-96]. However, some socio-economic constraints such as deprived agronomic varieties, poor post-harvest handling and storage facilities, lack of processing skills and clean seed materials, poor seed distribution system, and inadequate capacity also limits the optimum exploitation of sweet potato.

To overcome the aforesaid limitations, global development of elite cultivars is the major challenge for researchers working on sweet potato. Therefore, there is an urgent need to establish

a breeding program for developing drought and disease resistant sweet potato varieties to mitigate its effects. To embark on the regional task, breeders are required to adapt the approach for breeding narrowly and broadly adaptable varieties. Since genotype x environment interaction (G X E) is an important factor involved therein, therefore, in order to categorize the clones with broad and/or narrow adaptation, breeding strategies are required that enable assessment of genotype(s) by environmental effects, to identify the genotypes with excellent performance and yield across the ecological zones or location. However, the breeders should ensure that the variety released satisfies the need of farmers. It is also important to understand the function and inheritance of gene(s) responsible for higher productivity and/or tolerance so that it can be introduced into the suitable broadly adaptable varieties. Due to numerous factors including recurrent cross incompatibilities, the complex inheritance of traits in sweet potato due to polyploidy (2n=6x=90), and the need to sustain heterozygosity to circumvent inbreeding depression in F1 breeding populations, it is impossible to introduce the gene(s) into high yielding, indigenous varieties, using a conventional breeding strategy. Therefore, it is pertinent to use the genetic engineering approach which is a better alternative to guarantee a sufficient supply of quality protein in sweet potato. This intervention will help in meeting the ever-increasing demand for sweet potato, as a healthy food and for revenue generation. Sweet potato is a good source of nutrients except for proteins [17, 26]. Therefore, its nutritional improvement in terms of protein is essential since it is a staple food crop and among the few sources of food in many developing countries. In the period of environmental changes and soil deprivation, generation of improved eco-friendly, location specific sustainable techniques through well designed research are indispensable to improve sweet potato productivity. Well planed and sustainable investments into sweet potato research will play a major role to the food security and improvement of the subsistence of underprivileged people.

1.2 Nutrition and its 'omic' studies

Nutritional requirements are not typically generalized to a population as a whole; rather, they are adapted to population subgroups. Nutrition is an integrative discipline. From its inception, nutritional science has been distinguished by its integration of knowledge, and technology derived from the biological and physical sciences. This assimilation is in essence to understand the role of nutrients and other dietary components in human health and disease

throughout the life cycle, and the translation of that information for the improvement of public health [97]. Nutritional quality and agricultural productivity are the two key questions raised several times in the context of an ever-growing population where human health largely depends on plants pertaining to sustainable food production globally. Humans require a diverse and nutritionally well-balanced diet to maintain optimal health and depend largely on plants for their daily nutritional requirements. Moreover, a large section of the world's population is undernourished. Thus, nutritional improvement of crop plants is an urgent worldwide health issue as basic nutritional requirements for much of the world's population are still not met.

1.2.1 Nutritional genomics or nutrigenomics

The diverse tissue and organ-specific effects of nutrients comprise of gene-expression (transcriptome); chromatin organization (epigenome); protein-profiling (proteome) and metabolite profiles (metabolome). The current technologies have focused on the completion of the sequenced genomes of the staple crops. The enormous information originated from this achievement has not only impelled the expansion of analytical platforms and data mining software, but have also revolutionized the concept of Greek suffix 'ome', meaning "complete" or "all," describing the global analysis of genes (genomics), mRNA (transcriptomics), proteins (proteomics), and metabolite (metabolomics) [98-100]. **Fig. 1.2 A** gives a brief account of various omics technologies into nutritional research.

Nutritional genomics is a science studying the association between human genome, nutrition and health and is a term referring to a combination of biochemistry, genetics, molecular biology and genome-based technologies to investigate and manipulate plant compounds with nutritional value [101]. It can further be subdivided into two disciplines *i.e.* (I) Nutrigenomics and (II) Nutrigenetics. The prospective analysis of differences among nutrients in the regulation of gene expression *i.e.* studies of the effect(s) of nutrients on the genome, proteome, and metabolome comes under the realm of nutrigenomics. Whereas, nutrigenetics is the retrospective analysis of genetic diversity among individuals with respect to the interactions between diet and disease. Nonetheless, nutrigenomics and nutrigenetics, two fields with distinct approaches to elucidate the interaction between diet and genes but with a common ultimate aim to optimize health via the personalization of diet, offer robust approaches to unravel the multifaceted

relationship between nutritional molecules, genetic polymorphisms, and the biological system as a whole [98].

1.2.2 Nutritional epigenomics

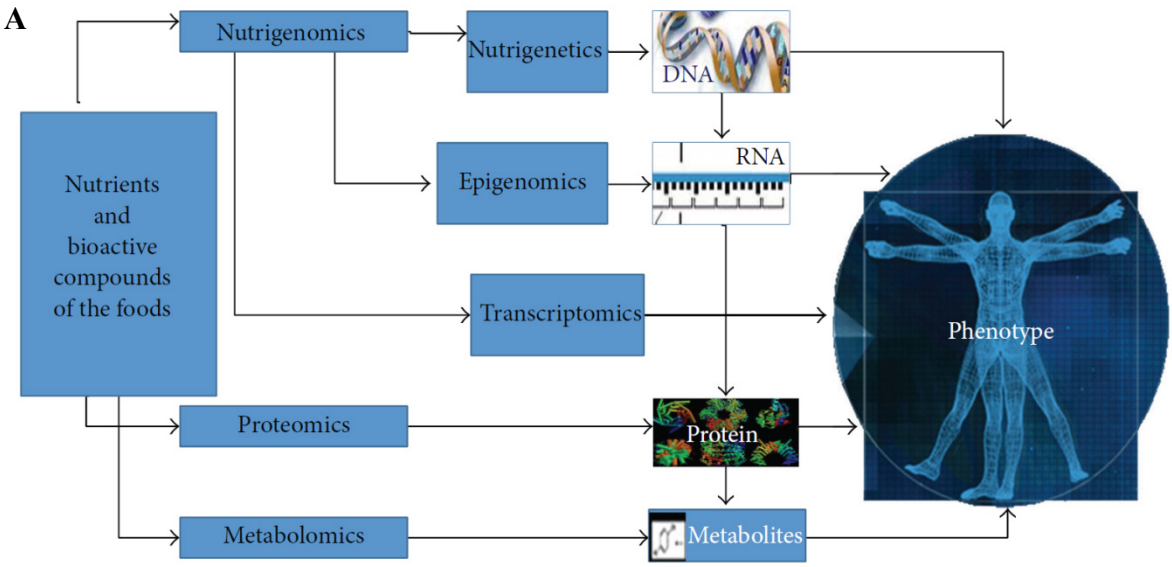
The study of the complete set of epigenetic modifications in a cell or in a tissue at a given time can be defined as epigenomics. The epigenome consists of chemical compounds that alter or mark the genome and these marks are called epigenetic marks or epigenetic signature. These epigenetic characters are inherited from one cell to another and thus will be passed from generation to generation. These signatures are influenced by genotypes and the nutrients available therein and eventually determine the phenotype. Bioactive food components can modify epigenetic events and other critical factors to selectively activate the gene functions. An array of regulatory proteins including histone-modifying enzymes, chromatin remodeling factors, DNA methyltransferases, methyl-cytosine guanine dinucleotide binding proteins, and their multimolecular complexes are involved in the overall epigenetic process [99, 100].

1.2.3 Nutritranscriptomics

At present, the most extensively used tool for transcriptomics is DNA microarrays, which simultaneously allows evaluation of the expression level of transcripts in bulk. Now a day, DNA microarray technology is used in cell culture systems or laboratory animals to elucidate the cellular responses to dietary constituents and their molecular targets. Several studies revealed that the DNA microarray technology is also an efficient approach for evaluation of the safety of foods and their components. However, despite of few limitations such as requirement of significant quantities of tissues material, expense of analysis and lack of highly efficient and reliable algorithm to facilitate analysis and interpretation of the experimental results, nutritranscriptomics is still a promising area [102].

1.2.4 Nutriproteomics

Nutriproteomics is a naive area of research exploiting the multifaceted dynamics of proteomic tools to distinguish the molecular and cellular changes in protein expression and function globally as well as to elucidate the interaction of proteins with food nutrients and the mechanism(s) underlies. Since nutrients are present in composite forms, its bioavailability and



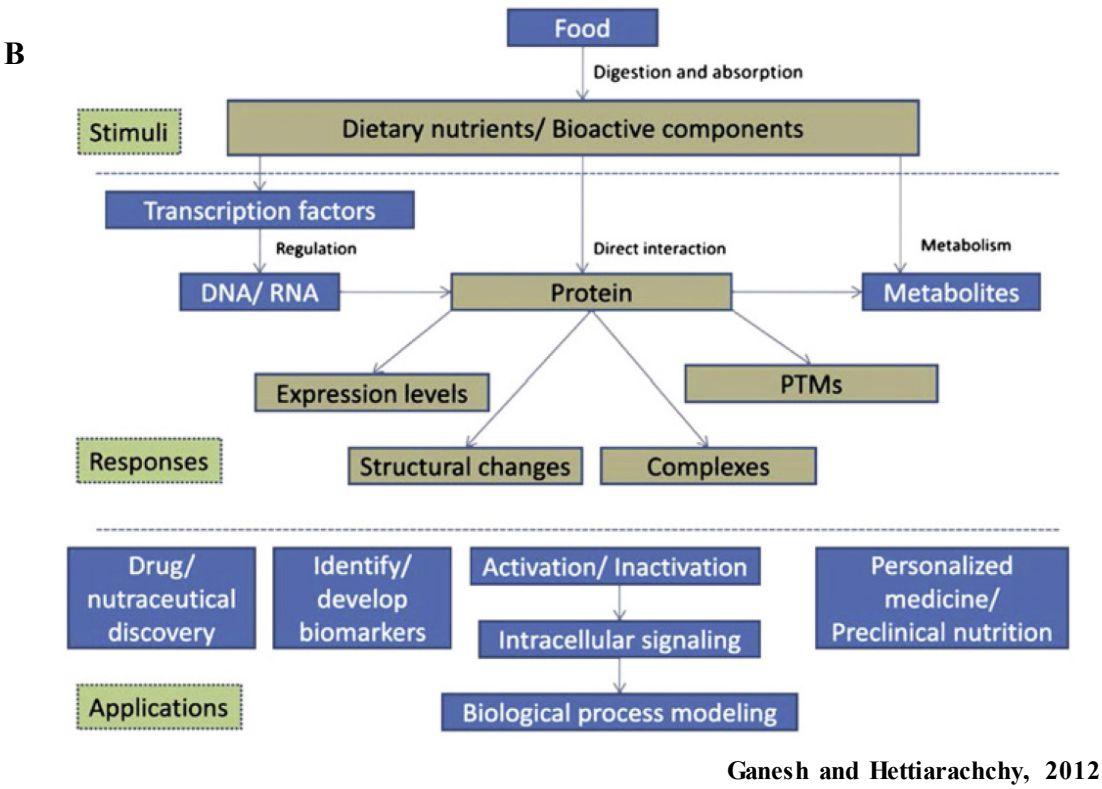


Fig. 1.2. Analysis of transgenic plants at different integration levels of evaluation. "omics" technologies into nutritional research (**A**) and Nutrients and their effects on the major applications of nutriproteomics (**B**).

functions may possibly be influenced by the coexistence of other nutrients and their interactions. The nutriproteomic tools for identification, characterization, quantification and analyses of proteins includes, 2-DE, mass spectrometry, chromatography, protein microarray and other promising technologies involving visual proteomics. The prospects of nutriproteomics in nutritional research for transforming biomarker development, nutraceutical innovations, biological process modeling, nutrition linking diet and diseases and restructuring ways to a personalized nutrition. Nevertheless, numerous challenges such as protein dynamics, analytical complexity, cost and resolution still exist, the extent of the application of proteomics to nutrition is rapidly expanding and more holistic as well as promising strategies are emerging [103]. A comprehensive depiction of how nutrients can affect proteins and the major applications of nutriproteomics is shown in **Fig. 1.2 B**.

1.2.5 Nutritional metabolomics

Metabolome consists of a set of primary and/or secondary metabolites of an organism or species. Metabolomics is the area of functional genomics that deals with the changes in metabolites, with the ultimate goal of isolation and characterization. Metabolomics enables the perception of metabolic arrangements and instabilities that is especially due to the intervention from the diet. In the area of nutrition, metabolomics plays a vital role to understand the role of dietary constituents and their effects on cellular metabolome. This contributes to the understanding of how the surplus or lack of some nutrients or compounds (secondary metabolites) present in food can affect the health/illness or overall development of an individual [99, 100].

1.3 Enhancement of protein quality in crops

Protein is one of the most essential nutrients for human and animals, as evinced by the origin of its name from Greek word *proteios* for primary. The protein is made up of 20 amino acids out of which 10 are essential because animals including humans are incapable of synthesizing them; consequently these 10 amino acids have to be supplemented from diet. Nevertheless, for proficient protein synthesis, amino acids derived from ingested protein must be present in balanced amounts. Lysine, tryptophan, and methionine have received the most attention among the essential amino acids, as they are most limiting in cereals (predominantly

Lys and Trp) and legume crops (particularly Met). Lysine is one of the most limiting amino acids in cereal grains, which are commonly used as a principal energy source in humans and livestock. On the other hand, legume seeds, which are important sources of protein in human and animal's diet, generally contain an adequate supply of lysine, but are deficient in the sulfur containing amino acids, methionine and cysteine. Conversion of methionine to cysteine by animals including human is possible but they cannot convert cysteine to methionine. Therefore, methionine can provide the complete requirement for sulfur containing amino acids but cysteine cannot. The sulfur containing amino acids manifested intake of animal must consequently contain a minimum amount of methionine [104, 105]. Therefore, dependency on vegetarian diet of restricted diversity has detrimental consequences on developmental aspects. Amino acid imbalance in diet may have some adverse effects on the physical and mental development of children and can irreversibly retard overall growth and development for instances kwashiorkor is a disorder caused by protein deficiency in infants. Nutritional improvement of crop plants in essential amino acids has economical as well as humanitarian interest.

1.3.1 Improvement of protein quality in crops: Strategies and approaches

The protein of the majority of the important crops should be modified as per the specific dietary requirements of the humans and livestock [106]. Suitable applications of biotechnology can help address problems by improving cereals and other crop species in terms of their nutritive values. Three fundamental approaches are being applied to engineer enhanced content of storage proteins in association with balanced essential amino acid content in storage organs such as seeds or roots in plants.

The **first approach** includes the engineering of the amino acid metabolism to facilitate the increase in respective essential amino acid both qualitatively as well as quantitatively. The transgenic plants have been generated expressing the aspartate kinase isoenzymes and/or aspartate kinase dihydrodipicolinate dehydrogenase (DHDPS), the key regulatory enzymes in the pathway of isoleucine, lysine, threonine, and methionine biosynthesis. Increase in the amount of free Thr and Lys were observed in these transgenic plants [107, 108]. Furthermore, more than 100-fold increase in free lysine in transgenic soybean (*Glycine max*) and canola (*Brassica napus*) by the seed specific expression of aspartate kinase and dihydrodipicolinate was also reported,

equating to nutritionally significant increase in total seed lysine by 25% and 100%, respectively [109].

The **second approach** includes either *in vitro* mutagenesis to engineer the endogenous storage proteins by mutating suitable amino acid codons into essential amino acid codons such as Lys and Met or insertion of stretches of additional codons of these amino acids [110]. For instances, *β-phaseolin* from *Phaseolin vulgaris* has been modified by the addition of a 45 bp nucleotide sequence encoding a methionine rich region from maize and 15 kDa zein storage protein. The modified phaseolin protein had 9 methionine residues instead of 3 in wild type. Additionally, mRNA abundance in tobacco revealed that the expression of modified *phaseolin* gene was comparable to the wild type gene in seeds of transgenic tobacco. However, this modification led to the unstability of phaseolin in the developing seed [111].

The **third approach** involves transformation of genes encoding storage proteins with a high content of essential amino acids. The gene *AmA1* from *Amaranthus hypochondriacus*, Met-rich 2S albumins from seeds of Brazil nut, sunflower and *Arabidopsis*, Met-rich zeins from maize and high-molecular-weight glutelin subunits from wheat, have been successfully integrated and expressed in transgenic plants [112-115]. However, the allergic response of some proteins, such as Brazil nut 2S albumin limits their usefulness [116]. The advancement of sequence-based approaches and functional prediction of protein structure, folding and topology have made it promising to artificially design storage proteins, which are rich in essential amino acids and do not trigger allergic response in humans. Overexpression of a synthetic storage protein gene (HEAAE-DNA) in potato was reported earlier [117]. However, this gene could not be expressed and translated enviably and consequently this gene was modified (named as *asp1*) and overexpressed in tobacco and sweet potato, coding for a storage protein rich in quite a few essential amino acids (80%) as well as 13% methionine residues [118, 119]. This showed high protein and essential amino acid levels in leaves and root, respectively than that of wild type. Based on the structurally well characterized storage zein proteins from maize (Z19 and Z22), and the ASP1 (Artificial Storage Protein 1) were designed to have a stable storage protein like structure in plants [120].

1.3.2 Transgenic approach vs. conventional breeding

Conventional plant breeding involves the interbreeding (crossing) of closely or distantly allied species or individuals in order to generate novel crop varieties or lines with expedient characteristics. Plants are crossbred to introduce desirable traits/genes from one variety or line into a new genetic background. However, introduction of new traits/genes in conventional plant breeding transpires only between the individuals of the same species, or, in few of the cases, between closely related species. Nonetheless, relatively limited success was procured by genetic approaches when tried in different crop species. This is most likely because of restricted availability of genetic resources for plant breeding. While, genetic traits for high contents of Lys, Trp, or Met do not work in a seed-specific manner they are normally related with anomalous plant growth. Contrastingly, genetic engineering appears to be more promising, principally because this method enables seed-specific expression of desired traits, by means of seed-specific promoters. Genetically engineered traits can be introduced into several plant species and genotypes and can function synergistically with many other agronomically significant traits is an added advantage. These genetic engineering approaches were usually designed for precise improvements of essential amino acid metabolic pathways and expressing native and genetically engineered proteins augmented with essential amino acid contents. However, a detailed understanding of how these pathways interact with regulatory networks and accustomed with plant development is required for the improvements of metabolic pathways by genetic engineering. Modern systems biology approaches, including transcriptomics, proteomics, and metabolomics are used to elucidate the underlying mechanism.

1.4 Seed storage proteins and its role in nutritional improvement

Storage organs exhibit assorted nutritional excellence and complex multistep development and act as sinks in plants. The composition of nutrients in the storage organs, carbon (C) and nitrogen (N) in particular, greatly influence the organ development and determine the nutritive quality. Both C and N metabolites are known to act as signals that influence many cellular processes, for example, development, metabolism related to N assimilation, and amino acid synthesis, in addition to their essential role as macronutrients in living organisms, including plants. Seed storage proteins in plants, rich in essential amino acids, are thought to serve as C and N source for the growing seedling and meet the chief dietary protein requirement of over

half of the world population. Thus, they presumably play an indispensable role in productivity and protein quality and have received considerable attention owing to their postulated dual role in growth and importance as a component of the human diet [121].

Seed protein can be categorized broadly into two major categories, *viz.* housekeeping protein and seed storage protein. The former is accountable for maintaining normal cellular metabolism. A moderately recent classification categorizes these proteins into storage, structural and biologically active proteins [104, 122]. Lectins, enzymes and enzyme inhibitors are the major biologically active proteins. Though, these proteins may be considered as inconsequential in terms of their role in growth and metabolism but they may have nutritionally more balanced amino acid composition than storage proteins. On the contrary, seed storage proteins are non-enzymatic and have the exclusive function of providing proteins to fulfill the nitrogen and sulfur source requirement during germination and establishment of a new seedling. Based on their solubility, seed proteins were classified by Osborne (1924) [123] as follows:

Albumin: Soluble in water or dilute buffer at neutral pH; (1.6S-2S).

Globulin: Soluble in dilute alkali or acid solution; (7S-13S).

Glutelin: Soluble in aqueous alcohols.

Prolamin: Soluble in 70-90% aqueous alcohols and is a major storage protein of the cereals.

Other protein storage bodies includes

Lectin: Capable of binding sugars and of agglutinating red blood cells; may have defensive role.

Late embryogenesis abundant (LEA): Highly hydrophilic proteins that accumulate late in seed development.

Albumins and globulins include the storage proteins of dicots, while prolamin and glutelin are most important proteins in monocots. Table 1.3 depicts the amino acids composition of some representative seed storage protein in comparison to WHO recommended value of

storage protein for crop improvement. The seed storage proteins can be distinguished from other proteins by some of their characteristics.

1. These accumulate in mid-maturation stage of developing seed and are consumed during germination.
2. These are exclusively synthesized in the seeds (in cotyledon or in endosperm).
3. They do not have any other functional activity besides storage.
4. They are stored typically in special storage organelles called protein bodies.

Table 1.3 Amino acid composition of some seed proteins (mole %) and comparison with WHO recommendation for essential amino acids

Amino Acid	WHO	*Zea mays* 2S	*Brassica* 2S	*Chenopodium* 2S Protein	*Amaranthus* AmA1	*Pisum* Legumin
Ala (A)	-	13.7	4.2	4.3	5.3	6.0
Arg (R)	-	1.2	4.2	7.5	5.3	10.0
Asx(B)	-	4.7	1.6	5.0	16.6	13.1
Cys(C)	-	Trace	6.6	5.8	0.7	1.2
Gly(G)	-	2.1	6.6	5.8	12.0	6.9
His(H)	-	0.9	3.3	5.7	3.3	1.8
Ile(I)	4.0	4.1	4.2	3.8	5.0	4.0
Leu(L)	7.0	20.0	6.6	3.0	7.6	7.6
Lys(K)	5.5	0.2	6.6	4.9	6.6	4.2
Met(M)	3.5*	Trace	3.3	2.8	1.6	0.7
Phe(F)	6.0**	6.3	1.6	5.4	5.6	2.3
Pro(P)	-	9.7	8.2	6.1	3.6	5.5
Ser(S)	-	6.6	4.9	5.6	6.6	5.8
Thr(T)	4.0	2.4	3.3	5.1	5.3	3.1
Trp(W)	1.0	N.D.	0.82	N.D.	2.3	-
Tyr(Y)	**	3.1	0.82	3.0	4.8	3.6
Val(V)	5.0	3.6	4.9	3.4	5.6	5.0

†, Derived from DNA sequence; N.D., not determined; *, Met plus Cys; **, Phe plus Tyr. Source: Mandal and Mandal, 2000.

1.4.1 Seed storage proteins: A vista for nutritional improvement

Seed storage proteins, anticipated as a resource of nitrogen for germinating seedlings, form an essential source of dietary protein for human beings. Humans require a diet with a balanced amino acid composition, but in many cases seeds are deficient in some of these essential amino acids. Over the years improvement of the balance of essential amino acids has been attempted by several plant breeders in few important crop plants. For improving the nutritional value of seed proteins molecular approaches are the better alternatives to the conventional approaches. In certain seed proteins *in vitro* mutagenesis in coding sequence has been strived to enhance the contents of essential amino acids [124, 125]. Introduction of heterologous storage protein genes that encode proteins with higher levels of limiting amino acids is an alternative approach [126]. However, expression of a particular amino acid at higher level by heterologous gene transfer or by mutation may be unfavorable to the normal physiology of seed development. This may also produce seeds with a biased amino acid composition. Therefore, to express a gene with a balanced amino acid composition may be the solution of this problem.

1.4.2 Seed storage protein AmA1: A promising candidate for nutritional improvement

The grain of *Amaranthus hypochondriacus*, a pseudocereal contains a high protein content (17-19% of seed dry weight) when compared to other conventional crops which have an average of about 10% protein. Its protein is rich in essential amino acids like lysine (5.0%, more than twice that of wheat flour), threonine (2.9%), tyrosine (3.4%), tryptophan (2.3%) and sulfur-containing amino acids (4.4%). Earlier, a nutritionally balanced 35-kDa seed albumin protein (with 304 amino acids and 7.4 pI value), amaranth seed albumin (AmA1) from *Amaranthus hypochondriacus* was identified. It shows a number of characteristics that distinguish it from other seed storage proteins. AmAl is synthesized during early embryogenesis and stored until maturation. The AmA1 protein is hydrophilic in nature with small stretch of hydrophobic amino acids region at N terminus [127]. Further, the full-length cDNA clone encoding *AmA1* was isolated and successfully introduced in potato under the control of constitutive (CaMV 35S) as well as tuber specific (GBSS) promoter. The transgenic potato showed an increase in all essential amino acids with respect to wild-type. A significant 2.5- to 4-fold increase in lysine, methionine, cysteine, and tyrosine contents was observed in constitutive transgenic lines whereas, tuber

specific lines showed 4- to 8-fold increase [112]. Furthermore, AmA1 was overexpressed in seven genotypic backgrounds suitable for cultivation in different agroclimatic regions and transgenic tubers revealed up to 60% increase in total protein content. Different amino acids revealed a significant increase in amino acids, particularly lysine, tyrosine, and sulfur amino acids, which are otherwise limited in potato tubers. In addition, there was an increase in aspartic acid, glutamic acid, arginine, leucine, and isoleucine levels [113].

During the past decade, several potential candidate genes have been targeted for the nutritional improvement of protein content in crops, namely: Brazil nut 2S albumin, *AmA1* (Amaranth Albumin1), β-phaseoline, HS-7 zein, cruciferin, sunflower seed albumin, and S-rich zein. However, except *AmA1*, introduction of these genes in target plants has often resulted in an increase in one of the amino acids at the expense of others, leading to an imbalance of the amino acid profile in transgenic crops [113]. Interestingly, AmA1 has great agricultural importance because it is a well balanced protein in terms of amino acid composition, possessing even better values than recommended by the World Health Organization for a nutritionally rich protein. It is acknowledged that most allergenic proteins tend to have characteristic sequence stretches. The amino acid sequence of AmA1, however, did not show any homology with known allergenic candidates when searched against the allergen online database (www. allergenonline.com). **Chakraborty et al. (2010)** had established that AmA1 potato tubers are nontoxic, nonallergenic, and safe for consumption by using two different animal models (rat and rabbit) [113]. Therefore, AmA1 protein has a great potential as a protein donor for the following salient features: (1) unlike the majority of seed proteins, it is a well balanced protein in terms of amino acid composition and even better than the values recommended by the WHO for a nutritionally rich protein; (2) it is a nonallergenic protein in its purified form; (3) Unlike many seed storage proteins, AmA1 is encoded by a single gene and thus would facilitate gene transfer into target plants with less difficulty; (4) An increased nutritive value of transgenic potato by expressing a seed albumin gene *AmA1* from *A. hypochondriacus* is proved earlier; and (V) *AmA1* gene is obtained from a natural source of edible vegetable plant *A. hypochondriacus*. Therefore, it may prove to be more acceptable to the general public than currently used genetically modified crops.

Phylogenetic study of AmA1 reveals the presence of three major clades which is based on the presence of agglutinin domain in association with other domains. As per Pfam database 5,

agglutinin family members have varied architecture. Amaranthus AmA1 protein, encompassing a homodimer, shows homologous relationship with peach agglutinin. Grape, castor and foxtail were present in the same group but having different architecture with aerolysin. *Brachypodium, Aegilops* and wheat were present in different clades. *Selaginella* and Maize were grouped together as they were distinct from other proteins thereby strengthening the possibility that these proteins evolved separately from other proteins. InterProScan 5 analysis of AmA1 suggests the presence of two agglutinin domain. Agglutinins are sugar-specific lectins that can agglutinate erythrocytes. Amaranthus agglutinin contains two identical subunits and each subunit contains two homologous globular domains. Each homodimer consists of two beta-trefoil structure (**Fig. 1.3 A-C**). Moreover, lectins are known as recognition molecules involved in cell–molecule and cell–cell interactions in a variety of biological systems. However, the biological function of agglutinin is still not well characterized particularly in plants but expression and characterization of two domains of agglutinin from *Pinellia ternate,* its overexpression in tobacco chloroplast ssuggests its role in the tolerance against a varieties of pathogens [128, 129]. Recently, by a multivariate comparative proteomic approach was used to elucidate the AmA1-regulated molecular mechanism affecting increased protein synthesis, reserve accumulation, and enhanced growth [121]. A comparative proteomic as well as metabolomics analysis of AmA1-potato and their wild type counterparts revealed its role in protein enhancement, cell growth and organ development tuberization in particular.

1.5 Assessment of genetically modified crops

Genetic modifications offer important prospective in mounting crop production and diversification of the nutritional base. However, risk of inadvertent effects due to transgene integration is one of the major concerns. Due to the random insertion of specific DNA sequences into the plant genome (intended effect), the disruption, alteration(s) or silencing of active genes or the activation of silent genes may occur, which may result in overall alteration in metabolic pathway by the formation of either new metabolites or altered levels of existing metabolites. Moreover, safety could also be compromised due to formation of novel fusion proteins, or other pleiotropic effects, that may generate new allergens or toxins. These effects may be predictable to a certain extent on the basis of understanding of the site of transgene insertion, the characteristics of the transgene, or its involvement in metabolic pathways. While due to the

inadequate knowledge of gene regulation and gene-gene interactions other effects (pleiotropic effects) are unpredictable. It should be emphasized that the occurrence of these effects is not specific for genome modification through rDNA technology, it also occurs frequently in conventional breeding [130, 131]. Scrutiny of the agronomical/morphological traits and an extensive chemical analysis of essential nutrients, anti-nutrients and toxicants could be helpful to identify these unintended effects. The possible occurrence of anonymous toxicants and anti-nutrients, predominantly in edible plant species with no history of safe use and the unavailability of adequate detection methods are the major limitations of this analysis. Therefore, it is pertinent to develop a high throughput approach to identify the secondary effects of genetic modifications. Different approaches may be applied in order to categorize potential secondary effects of the genetic modification, which would result in alterations in the composition of genetically modified crops. These approaches may broadly be categorized as the targeted (compound-specific) approach, or the non-targeted (profiling/ fingerprinting) approach.

1.5.1 Targeted (compound-specific) approach using single compound analysis

Targeted or compound-specific investigation should comprise of baseline analyses of a number of key nutrients such as proteins, carbohydrates, fats, vitamins and other nutritional and/or anti-nutritional compounds which, if inadvertently modified, might affect nutritional value and safety. Choice of the key nutrients and toxicants needs to take into account the target species, structural and functional aspects of the transformed gene(s), their possible interferences in metabolic pathways at different integrated level (**Fig. 1.4**). Selection of compounds may be restricted which represents crucial biochemical/physiological pathways in the organism. It is conceivable, that the alterations anticipated in the metabolism as a possible result of the genetic modification will be recognized by the analysis of a huge number of components, but unexpected changes are purely identified inadvertently. The targeted approach has several limitations pertaining to unknown anti-nutrients and natural toxins, particularly in less recognized crops [130].

1.5.2 Non-targeted (profiling/ fingerprinting) approach using profiling methods

An unconventional (non-targeted) approach to reveal the transgene-mediated effect is the use of profiling techniques. Quite a few novel methods have been developed so far which allow

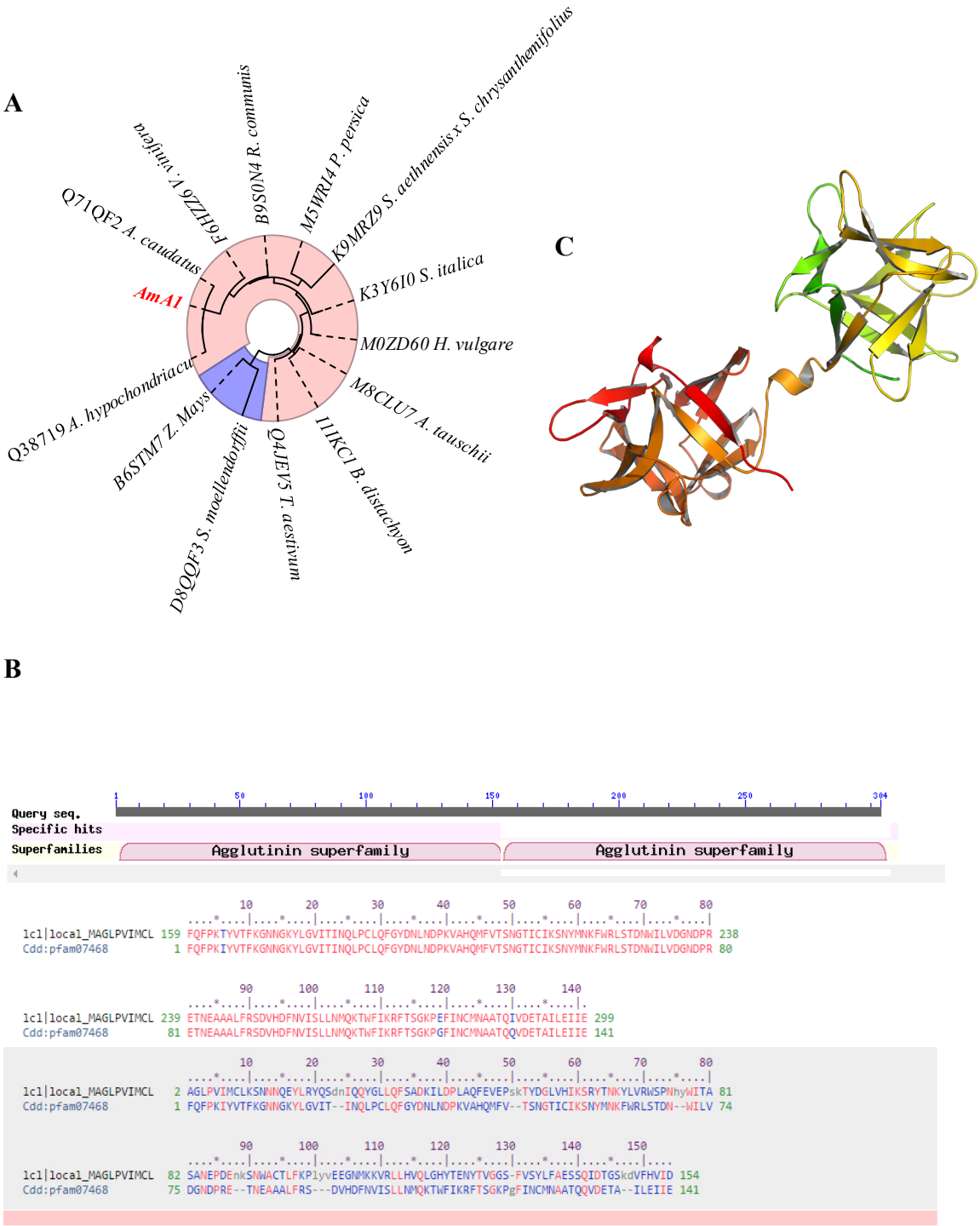


Fig. 1.3. *In silico* analysis of AmA1. Phylogenetic study of *AmA1* (**A**), Functional domain of AmA1 (**B**) and Secondary structure prediction of AmA1 (**C**).

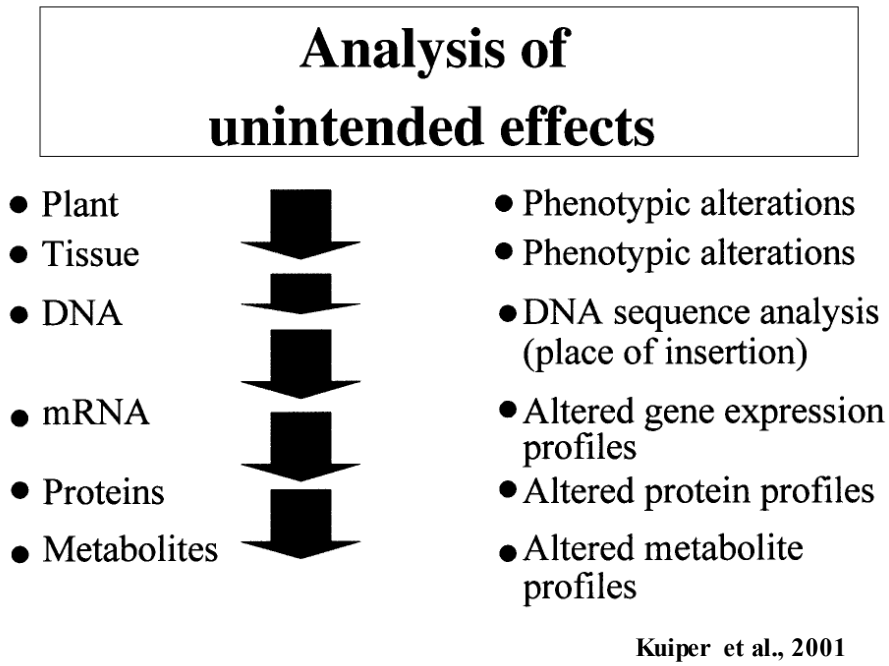


Fig. 1.4. Analysis of unintended effects in transgenic plants.

for the screening of probable changes in the physiology of the modified host organism at different cellular integration levels, at the genome level, transcript level, proteome level, and at the level of metabolome. Several factors, such as genetic uniqueness (cultivar, individual, isogenic lines, heterosis), agronomic factors (soil, fertilizers, plant protection products), environmental influences (location effect, weather, time of day, stress), plant-microbe interactions, maturity stage, and post-harvest effects determine the morphological, agronomic and physiological properties of a food crop. Assessment for prospective changes in these characteristics in genetically modified plants becomes more significant because the recent genetic alterations changing agronomical or nutrition associated properties are more intricate, involving introduction of large DNA fragments or clusters of genes [130].

1.5.2.1 Analysis of genome

The knowledge about the site of transgene insertion and the flanking sequences is one of the most important direct approaches to predict and identify possible occurrence of effects in recipient plant DNA. In case of transgene insertion within or proximity of endogenous gene and the knowledge about the flanking regions will pave the way for further analysis. The chromosomal location and structure of the transgene can be detected by various methods such as *in situ* hybridization, and fluorescence in situ hybridization, and by direct sequencing of flanking DNA [132-135]. The information of the genetic code and the regulation of gene expression with respect to the networks of metabolic activity is increasing. However, knowledge about plant genome is still restricted, including the reliability of annotations in genomic databases. Therefore, the sequencing of the site of insertion(s) will become increasingly informative.

1.5.2.2 Transcript analysis

Transcript analysis is also an important criterion to assess the extent of alteration(s) in the expression of gene(s) of vital metabolic pathways and its regulation. Microarray technology and qRT-PCR are the effective methods to study gene expression. Small scale analysis of expression of a huge number of genes at the same time, in a sensitive and quantitative manner is one of the most important benefits of the microarray technology over conventional gene profiling techniques [136, 137]. In addition, qRT-PCR allows detecting the expression of gene(s) having lower transcript abundance. Furthermore, both the methods allow the comparison of gene-

expression profiles under different conditions. These tools may efficiently be used to screen for altered gene expression, and at the same time provide preliminary information on the nature or type of perceived alterations. These methods may answer the question regarding the safety and nutritional value of the food crop under question. [138-140].

1.5.2.3 Proteomics analysis

In contrast to the genome, which is fairly static, the latter three molecular groups (transcriptome, proteome and metabolome) are highly dynamic and vary greatly according to cell conditions and the life cycle of an organism. The dynamic expression of genes as mRNA (the transcriptome) can be followed in a quantitative and qualitative manner [141]. There are many questions that arise pertaining to the monitoring of changes in the abundance of mRNA: (I) the level of mRNA does not always correlate with the level of protein expression [142, 143]; (II) protein function is controlled by many post-translational modification(s); and (III) protein maturation and degradation are dynamic processes that dramatically alter the final amount of active protein, independent of the mRNA level. Therefore, the proteomics is an indispensible technology to understand the molecular function of cells and consequently understanding of the biological complexities in the plant cell can be expanded by exploiting proteomics, a technique that analyses many proteins simultaneously and will contribute to our understanding of gene function. This process involves the gathering of information about the temporal, spatial, and physiological regulation of proteins, their interacting partners, biochemical activities, and post-translational modifications, which ultimately influence the physiology of the organism. Different technologies such as matrix-assisted laser desorption ionization MALDI, surface-enhanced laser desorption ionization (SELDI-MS), two dimensional gel electrophoresis (2-DE), differential two-dimensional fluorescence gel electrophoresis (2D-DIGE), isotope-coded affinity tag (ICAT), isobaric tags for relative and absolute quantification (iTRAQ), liquid chromatography and mass spectrometry (LC/MS) are in use for protein profiling [144, 145]. Although different approaches for proteome analysis have been developed, a core technology of proteomics is 2-DE because there is no other technique that is capable of simultaneously resolving thousands of proteins in one separation procedure [146, 147]. The proteomics studies revealed the influence and efficacy of such approach to distinguish the landraces, populations, varieties and even species [148, 149]. The use of 2-DE can be extended for the comparison between a genetically modified

organisms/crops with their wild type counterparts. Comparative proteomic analysis of the lines under investigation by 2-DE is the initial step looking for the unintended changes, if any. Normal variations should be evaluated in that case if differences in protein profiles are detected. Furthermore, if the proteomic profile is deviated from the normal variations, identification of the protein must be carried out, which may direct to further toxicological investigations. Additionally, metabolic alterations may be checked if the identified protein has a known enzymatic activity. Protein chip-based approaches are much faster technique to analyze the crops protein profiling [150, 151]. Earlier 2-DE has been used for the identification of African yam bean seed proteins [152]. Prominently resolved polypeptide bands showed sequence homology with a number of known anti-nutrient and inhibitory proteins, which may have implications for the safe use of these seeds as human food.

1.5.2.4 Metabolome analysis

Analyses pertaining to the intended and/or unintended effects as a result of genetic modifications could be indicated by a multi-compositional analysis of biologically active compounds in plants *i.e.* nutrients, antinutritional factors, toxicants and other relevant compounds. The three most significant techniques that have emerged are gas chromatography (GC), high-performance liquid chromatography (HPLC) and nuclear magnetic resonance (NMR). These methods are accomplished of detecting, resolving and quantifying an extensive range of compounds in a particular sample. For example, HPLC method was used in metabolic profiling of isoprenoids with applications to genetically modified tomato and Arabidopsis [153]. Chemical signature of different compounds is determined on the basis of the detection of alterations in NMR spectra obtained from the sample extracts. For instances, analysis of 1H-NMR spectra was used in case of genetically modified tomato varieties, such as the antisense RNA exogalactanase fruit, and from their non-modified counterpart(s) [154].

1.6 Transformation of AmA1 in sweet potato: Needs, challenges and future perspective

Sweet potato is considered as a low protein crop and with reference to FAO reference protein; sweet potato protein contains a good quality as well as the quantity of essential amino acids except tryptophan and total sulfur amino acids (as described earlier). Therefore, it would be beneficial to improve the nutrient value of sweet potato by increasing its essential amino acids

since it is the chief food source for the poorest segment of population in developing countries. To guarantee a sufficient supply of quality protein in sweet potato, specific interventions in genetic engineering are an absolute necessity. On the other hand, the amino acid composition of AmA1 shows a high proportion of essential amino acids such as lysine, leucine, threonine, phenylalanine, valine, tryptophan and sulfur amino acids that are otherwise deficient in the major food crops including sweet potato. Interestingly, high proportion of tryptophan and sulfur amino acids in AmA l is the added advantage which made it a potential candidate of gene transfer in sweet potato. With the increased protein, sweet potato will be balanced staple food for the under developed population. However, sweet potato is reported to be quite recalcitrant to genetic transformation and genotype-dependent [155, 156]. Therefore, development of an ideal and genotype independent transformation system would remains one of the key challenges for sweet potato.

Keeping in mind these facts, the following objectives were undertaken for this research study:

1. *Agrobacterium tumefaciens* mediated transformation of sweet potato by expressing a seed albumin, *AmA1* under the control of constitutive and tuber-specific promoters.
    A. Establishment of transformation and regeneration of selected cultivars of sweet potato.
    B. Screening of putative transformants.
2. Molecular characterization of putative transgenic sweet potato lines.
    A. Determination of transgene introduction.
    B. Expression analysis.
3. Comparative biochemical and agrophysiological characterization of transgenic events.
4. Metabolite profiling of tuber.