

# Chapter 10

## Abiotic Stress Resistance in Tobacco: Advances and Strategies



**K. Sarala, K. Prabhakara Rao, C. Nanda, K. Baghyalakshmi,  
Reza Darvishzadeh, K. Gangadhara, and J. J. Rajappa**

**Abstract** Tobacco is a commercial crop cultivated globally in about 3.62 million hectares mainly in semi-arid and rain-fed areas. The crop is often confronted with various forms of abiotic stresses viz., drought, flood, high temperature (heat), cold, salinity, ozone, low and high light intensity, chlorides, heavy metals, ultraviolet radiation etc. Further, climate change in terms of higher temperature, and changing rainfall patterns is going to have remarkable effect on tobacco productivity and quality. The abiotic stresses usually play a negative role in the growth and development of tobacco plant. Tobacco plant respond to stresses in various ways with alterations at morphological, molecular, physiological and cellular levels involving switching on or off stress responsive genes. Inherent capacity of tobacco genotypes need to be improved for achieving higher and stable yields under different abiotic stresses. Degree of abiotic stress and crop stage of stress occurrence vary from year to year and place to place making it difficult to breed resistant varieties. Currently, very fewer number of abiotic stress resistant tobacco varieties are developed through conventional breeding. The Complicated nature of abiotic stresses, lack of suitable morphology based screening techniques, polygenic nature of resistance mechanisms, lack of sources of resistance etc. are limiting the progress that could be made in abiotic stress breeding. These limitations can be successfully overcome through molecular breeding and genome designing strategies. Hence, in this chapter, an attempt made to summarize the available knowledge about tobacco genetic resources, germplasm

---

K. Sarala (✉) · K. P. Rao  
ICAR-Central Tobacco Research Institute (CTRI), Rajahmundry, India  
e-mail: [K.Sarala@icar.gov.in](mailto:K.Sarala@icar.gov.in)

C. Nanda · J. J. Rajappa  
ICAR-CTRI Research Station, Hunsur, India

K. Baghyalakshmi  
ICAR-Central Institute for Cotton Research, Research Station, Coimbatore, India

R. Darvishzadeh  
Department of Plant Production and Genetics, Faculty of Agriculture, Urmia University, Urmia, Iran

K. Gangadhara  
ICAR-CTRI Research Station, Kandukur, India

characterization through molecular markers, molecular maps, markers and QTLs linked to various abiotic stresses, omics resources and databases, abiotic resistant genes studied, etc. for their utilization in designing tobacco genomes for abiotic stress resistance. Further, recent technological advances in marker-assisted breeding, gene editing and genome designing technologies were discussed for effectively utilizing them in developing abiotic stress resistant tobacco genotypes.

**Keywords** Tobacco · *Nicotiana* · Abiotic stress · Molecular · Markers · Maps · Genome sequencing · Linked markers · QTL · Genome designing · MAS · Gene editing · Cloning · Databases

## 10.1 Introduction

### 10.1.1 Economic Importance of the Crop

Tobacco belongs to family Solanaceae with more than 75 species, among which *Nicotiana tabacum* and *N. rustica* are the cultivated species (Goodspeed 1954; Chase et al. 2003; Sierro et al. 2014). Tobacco is a most important non-food crops cultivated in more than 120 countries (FAO 2019). Tobacco is defined by different criteria such as region of production, intended use (i.e. cigar filler, binder and wrapper, bidi, chewing, hookah and cigarette manufacturing), methods of curing (flue-, air-, sun-, smoke-, pit- and fire-cured tobacco) as well as morphological and biochemical characteristics (i.e., aromatic fire-cured, bright leaf tobacco, Burley tobacco, Turkish or oriental tobacco, etc. (Ren and Timko 2001). It is grown on wide variety of soils and climate in less than one percent of the world's agricultural land. Currently, tobacco is being grown in an area of 3.62 million hectares with the global tobacco production of 6.69 million tons (FAO 2019). Some major countries growing tobacco are China, India, Brazil, United Republic of Tanzania, Indonesia, Zimbabwe, Malawi, USA, Zambia, Mozambique, Turkey, Democratic People's Republic of Korea, Bangladesh, Argentina and Pakistan. China is the largest producer of tobacco (2.61 million tons/year) in the world followed by India (0.8 million tons/year), both accounting for more than 50% of World's total. Harvested tobacco leaves are cured and used for smoking in the form of cigarettes, cigars, pipe tobacco, and flavored shisha tobacco. Some tobacco are consumed in the form of snuff, chewing tobacco, dipping tobacco and snus. Tobacco is instrumental in generating enormous revenue to the national governments and providing employment to billions of people.

### 10.1.2 Reduction in Yield and Quality Due to Abiotic Stresses

Land plants are surviving in inherently harsh environment ever since their emergence. A large variety of physical or chemical factors are found to be hostile to them,

including low or high temperature, deficient or excessive water, high salinity, heavy metals, and ultraviolet (UV) radiation, among others (He et al. 2018). These stresses, collectively termed as abiotic stresses, are posing problem to agriculture and the ecosystem and accounting for significant crop yield loss (Wang et al. 2003; Wania et al. 2016). In the field, plants are usually exposed to an unpredictable combination of various stresses rather than a single one (Wania et al. 2016), which is even worse in the context of environmental pollution, soil salinization and climate change.

Globally, tobacco crop is cultivated largely in semi-arid and rain-fed areas and is often confronted with various forms of abiotic stresses viz., excess or deficient water, high salinity, high (heat) or low (cold) temperature, heavy metals, salinity, ozone, low and high light intensity, chlorides, heavy metals, ultraviolet radiation etc. The abiotic stresses usually play a negative role in the growth rate of tobacco plant due to different molecular, physiological and cellular response of plant.

Drought is considered the most destructive condition influencing the growth of crop plants consequently leading to decreased yield (Lambers et al. 2008). The amount of damage to yield depends on the severity and duration of stress, plant resistance and plant growth stage (Robertson et al. 2004). Drought stress is a major constraint to tobacco production and yield stability in many rainfed regions of tobacco cultivation.

In addition to its vulnerability to moisture scarcity, tobacco is susceptible to injury from flooding or saturation of the soil with water. Waterlogging lowers the growth and productivity of tobacco. Rainfall accompanied by high winds may cause lodging along with the root damage causing soiling of leaves and make it difficult to manage plants. Flooding of the field for less than 24-h may not significantly affect the growth and yield of the plants whereas flooding for more than 48 h significantly reduced yield even up to 80% than that of the unflooded plants (Campbell 1973; Anuradha et al. 2013) and soil saturation for several days may lead to serious permanent injury or death of the effected plants (Nurhidayati et al. 2017). Flooding also changes leaf chemical quality parameters (Campbell 1973; Anuradha et al. 2013) and found to accelerate flowering time (Higase 1959) which in turn may reduce the yield due to reduction in period of vegetative growth.

Tobacco is a well-known thermophilic crop plant growing under tropical and subtropical region which is highly sensitive to the change of temperature and doesn't require either very high temperature or very low temperature during its growing period (Yamori et al. 2010; Popov et al. 2013; Zhang et al. 2013). Tobacco requires 3–4 months of frost-free period from transplanting to harvesting of the crop ([www.fao.org](http://www.fao.org)). A slight change in the temperature will make the leaves brittle and affect making it less acceptable. High temperature in tobacco usually disrupts the production of nicotine and other associated pyridine alkaloids (Oeung et al. 2017). The growth of tobacco plants is restricted at the temperature lower than 10–13 °C, affects the morphogenesis of tobacco plants, delays flower initiation and senescence and plants even die when the temperature drops to 2–3 °C (Zhang et al. 2013; Yang et al. 2018). Exposing to 18.5 °C promoted the elongation of petiole and stem, reduced leaf area and increased the angle between leaf and stem of tobacco plants (Yang et al.

2018). Cold stress at harvest period affects agronomic characteristics, leaf quality and curing characteristics (Li et al. 2021).

Higher temperature inhibits the growth and development of tobacco, accelerates flowering and leaf senescence triggering death of plants (Belknap and Garbarino 1996; Yoshida 2003; Lim et al. 2007; Djanaguiraman and Prasad 2010; Gill and Tuteja 2010; Suzuki et al. 2012; Yang et al. 2018) affecting yield and quality. The senescence caused by raise in temperature reduces both the yield and quality of the crop (Chéour et al. 1992; Navabpour et al. 2003; Kim et al. 2011) as early senescing leaves may not grow to their full potential and accumulate the necessary phyto-hormones required for quality (Liu et al. 2015; Nisar et al. 2015).

Light serves as primary energy source in the phototrophic lifestyle of plants. Moderate heat stress and fluctuating light are typical conditions in summer in tropical and sub-tropical regions. The low intensity light deprives the photosynthetic activity and high intensity light damages the photosynthetic apparatus (Tan et al. 2020) leading to lower yields. Exposure to UV-B may lead to either damaging DNA, with subsequent heritable mutations, or by eliciting various regulatory effects that are injurious to plant physiological functions (Lidon 2012).

Weather fleck in tobacco is a leaf spot syndrome induced by air-polluting ozone (Heggestad and Middleton 1959). Weather fleck has caused extensive losses to tobacco growers in the US and Canada since 1955 (Heggestad 1966). The production areas most seriously affected were cigar wrapper in Connecticut and Florida and flue-cured in southern Ontario, Canada. Weather fleck also occurs sporadically in burley, flue-cured, Wisconsin, and Maryland tobacco-producing areas. A number of air pollutants can cause injury to crop plants; evidence, however, indicates that ozone causes the flecking observed on tobacco leaves. Ozone appears to be the primary injury causing pollutant in the Maryland tobacco-producing area as a result of the region's proximity to Washington DC and the high concentration of automobiles, which are considered to be the major source of ozone pollution.

Chlorine in small amounts promotes growth and leaf expansion (Franco-Navarro et al. 2016, 2019) and so improves yield and certain quality factors such as color, moisture content, elasticity, burning and keeping quality of tobacco leaves (McEvoy 1957). However, larger amount of chloride has many adverse effects on the quality of tobacco, so much so that the chloride content in tobacco leaves is considered as a major factor determining the quality of tobacco (Akehurst 1981; Juan and del 1986; Guardiola et al. 1987; King 1990; Chari 1995). An excess level of chlorine produces leaves with poor burning capacity, muddy appearance and undesirable odor as well as highly hygroscopic nature causing discoloration during storage (Karaivazoglou et al. 2005).

Salinity is termed as the total amount of mineral salts dissolved in water and soil (Grattan 2012). Salt stress is the most stubborn one magnified by ever-increasing salinization of arable land worldwide (Munns and Tester 2008; Yuan et al. 2015). More than 20% of cultivated land worldwide is affected by salt stress and is increasing over the time. Salt stress will result in water stress, affecting the plant growth and development, leading to reduced leaf yield (Flowers 2004). Prolonged stress condition leads to death of leaves (Cramer and Nowak 1992). Higher salinity impinges on

plant lifecycle affecting seed germination, seedling establishment, vegetative growth, and flower fertility (Flowers and Colmer 2008; Guo et al. 2012, 2015, 2018).

Soil and water contaminated with heavy metals have now become one of the major constraints to the crop productivity and quality. From the past few decades we are witnessing rapid growth in industrialization and modern agricultural practices which have led to the environmental contamination (Miransari 2011). The increasing population and the continuing food demand add much more to the contamination. Lands are mostly being contaminated due to the use of fertilizers, pesticides, municipal and compost wastes, and also due to release of heavy metal from metalliferous mines and smelting industries (Yang et al. 2005b). Tobacco leaves found to naturally accumulate relatively high levels of heavy metals and particularly cadmium in leaves (Lugon-Moulin et al. 2008; Kaličanin and Velimirović 2012; Ajab et al. 2014; Regassa and Chandravanshi 2016). Cadmium is toxic and non-essential to both plant and humans. The accumulated heavy metals get transferred to human being through cigarette smoking (Jarup et al. 1998; Nordberg et al. 2007; Verma et al. 2010) causing a significantly serious damage on human health (Stojanovic et al. 2004; Norom et al. 2005; Sharma and Dubey 2005; Lugon-Moulin et al. 2006). Heavy metal toxicity symptoms are usually associated with stunted stem and root growth, leaf chlorosis of younger leaves (extends to the older leaves after longer exposure), disturbs phytohormone levels in the leaves etc. (Fontes and Cox 1998; Reddy et al. 2005; Gangwar and Singh 2011; Srivastava et al. 2012). Excess Zn can also give rise to copper (Cu) and manganese (Mn) deficiencies in plant shoots. Mercury toxicity induces physiological disorder in tobacco plants (Zhou et al. 2007). Excess amount of Chromium (Cr) in the soil results in reticence of plant growth, nutrient imbalance, chlorosis in young leaves, root injury, wilting of tops (Scoccianti et al. 2006) along with inhibition in chlorophyll biosynthesis (Vajpayee et al. 2000). Arsenate (As) acts as an analogue to phosphate and competes in the root zone of plants (Meharg and Macnair 1992).

### ***10.1.3 Growing Importance in the Face of Climate Change and Increasing Population***

Climate change is resulting in altering weather patterns, rising sea levels, and weather events are becoming more extremes. It is affecting lives in country on every continent and disrupting national economies. The year 2019 found to be the second warmest year on record and the period 2010–2019 was the warmest decade ever recorded. carbon dioxide (CO<sub>2</sub>) levels and other greenhouse gases in the atmosphere rose to new records in 2019 (United Nations 2019) Worldwide there is a steady increase in temperature from 1.5 to 2 °C for the past 60 years (IPCC 2019). Warming has resulted in an increased frequency, intensity and duration of heat-related events, including heatwaves in most land regions. Frequency and intensity of droughts has increased in some regions (including the Mediterranean, west Asia, many parts of

South America, much of Africa, and north-eastern Asia and there has been an increase in the intensity of heavy precipitation events at a global scale.

Global warming has led to shifts of climate zones in many world regions, including expansion of arid climate zones and contraction of polar climate zones. Raising temperatures are changing rainfall intensity, flooding, drought frequency and severity, heat stress, dry spells, wind, sea-level rise and wave action, and permafrost thaw with outcomes being modulated by land management. Climate change has been affecting the food security due to warming, changing precipitation patterns, and greater frequency of some extreme events.

Climate change resulting in unpredictable rainfall and increasing temperature with heat waves is going to have remarkable effect on tobacco productivity and quality in view of its sensitivity to these events. Drastic shifting of rainfall pattern and frequent dry spells cause moisture stress especially in critical periods of crop growth, significantly affecting growth, yield and quality of tobacco. Leaf is the important economic product in the tobacco plant. Leaves are harvested when they mature and before they reach senescence. The raise in temperature usually results in earlier senescence of leaves there-by reducing the quality harvest due to insufficient accumulation of quality related phytochemicals. Climatic variability was found to decrease tobacco productivity in Indonesia in 2013 and 2016 (Muttaqin et al. 2019). Both the premature and late-matured tobacco leaves will seriously affect the yield and quality of tobacco crop (Chéour et al. 1992; Navabpour et al. 2003; Kim et al. 2011).

Rapid growth in industrialization and modern agricultural practices in the past few decades has led to the environmental contamination (Miransari 2011). The increasing population and the continuing food demand add much more to the contamination. Most of the lands have been contaminated due to the use of pesticides, fertilizers, municipal and compost wastes, and heavy metals released from smelting industries and metalliferous mines (Yang et al. 2005b). Increased in salinization and heavy metals concentrations affects yield and quality of tobacco. Accumulated heavy metals makes tobacco consumption further harmful.

#### ***10.1.4 Limitations of Traditional Breeding and Rational of Genome Designing***

Conventional breeding continues to play an important role in improving tobacco productivity under different climatic situations including abiotic stress conditions. Abiotic stresses are complex characters and the success of breeding program primarily depends on the existence of variability for characters that contribute to stress mechanism (Fita et al. 2015). The achievements that can be realized through conventional breeding are limited by non-availability of sources for resistance to abiotic stresses and yield contributing traits, narrow genetic variability, natural barriers of crossing among existing species, longer period for developing stable homogenous lines, undesirable associations between resistant genes and desirable traits either due

to pleiotropic effects of the resistance genes or due to linkage drag effects caused by the presence of deleterious genes linked to gene of interest (Legg et al. 1981; Friebe et al. 1996; Brown 2002; Chaplin et al. 1966; Chaplin and Mann 1978). Recombination suppression within introgressed chromatin (Paterson et al. 1990; Liharska et al. 1996) may interfere in alleviating linkage drag effects through back crossing (Stam and Zeven 1981; Young and Tanksely 1989) and also complicate the efforts to distinguish between pleiotropic and linkage drag effects (Purrington 2000; Brown 2002).

Often the abiotic stresses are controlled by polygenes or many genes with smaller effects and modifier genes with pleiotropic effects. Many drought-inducible genes are also induced by salt stress and cold, which suggests the existence of similar mechanisms of stress responses. Hundreds of genes are thought to be involved in abiotic stress responses (Seki et al. 2003; Baloglu et al. 2012). Undesirable linkage of such genes with other deleterious genes makes pooling the resistance genes into a cultivar difficult through classical breeding. The appearance of various abiotic stresses in the crop growth period may vary from year to year and place to place. As the response and vulnerability of various crop stages (seedling, growth, maturity, flowering etc.) to stresses vary it may be difficult to breed lines uniformly resistant for different stages.

Other limitations in conventional breeding are the relatively longer time periods required to combine different target genes and laborious methods of screening/phenotyping segregating generations for abiotic stresses. The success of the abiotic stress resistance breeding depends upon the efficiency of screening techniques for abiotic stresses. The importance of developing reliable screening techniques has been realized very early (Levitt 1972). Plants exhibiting resistance to abiotic stresses can be identified based on their performance under different abiotic stresses after completion of their life cycle under field condition making it difficult to select plants in early stage of their life. Appearance of other stresses during the growth of the plants interfere in clearly assessing the effect of target stress. Such screening related issues are important limitations in achieving progress in resistance breeding.

Various limitations of traditional breeding mentioned can be overcome through the genome designing strategies (Kole 2017). The advancements in the field of genomic designing strategies including molecular breeding, transgenics, genomic-assisted breeding, and the recently emerging genome editing tools are providing a greater promise for improving tobacco for abiotic stress resistance. Whole-genome sequencing and genotyping-by-sequencing methods adopted in tobacco for mapping and trait discovery in recent years may pave the way for obtaining precise information about the genes conferring abiotic stress resistance. The polygenes identified for abiotic stress resistance can be effectively transferred through various molecular breeding methods. Handling of target genes in these methods overcome the issues of undesirable linkages and avoids the transfer of non-target genes, thereby reduce the time taken for elimination of non-target traits. Screening of target traits with tightly linked markers overcome the phenotyping requirements under stress environments and allows early generation screening. The gene editing tools and transgenic



approaches can be of great help in cases where genetic sources of resistance are not available. The emerging genomics-aided techniques including genomic selection, allele mining, gene discovery, and gene pyramiding for developing adaptive varieties holds a great promise for improving tobacco cultivars in terms of abiotic stress resistance in near future.

## 10.2 Description on Different Abiotic Stresses

### 10.2.1 Root Characters

Understanding various root characteristics of tobacco viz., manner of branching, the depth of penetration and the lateral spread of the root systems, and of absorbing areas of root under different soil conditions makes it possible to clearly interpret the responses of the plant to the various factors of its environment.

The tobacco has a taproot system that consists of primary, lateral and adventitious roots (Xi et al. 2011). Transplanted tobacco plant possesses an extensive but comparatively shallow fibrous root system ([https://ctri.icar.gov.in/for\\_morphology.php](https://ctri.icar.gov.in/for_morphology.php)). Most of these roots develop adventitiously from the portion of the main stem buried during transplanting. In general, most of the root system (72% of which were adventitious) of a mature plant filled all the cultivated layer of the A horizon (Gier 1940). The average total length of a mature root system was 260 m with a maximum length of 432 m. The shoot–root ratio ranged from 4.95:1 to 13.0:1 with an average of 10:1. Genotypes differ in length of roots and branching pattern (Jones and Shew 1995).

Bruner (1931) made detail study on the root development in tobacco. He reported that the first structure of the seedling is its main or taproot. The root system of the one month old tobacco plant is succulent and covered with root hairs. Absorption begins at first through the epidermis and soon increases rapidly with the appearance of root hairs even when the root is only a fraction of an inch long. Later, branches appear while the root is still only a few inches long and are soon covered with root hairs, and absorption is greatly increased. The main or taproot and its branches develop in the form of a more or less symmetrical cone which increases in size as the roots develop. As the roots continue to grow the majority of the older absorbing rootlets die and absorption is carried on by the younger rootlets. Those which do not die increase in length and usually in diameter and become permanent roots. The immediate environment determines just which of the rootlets will become permanent roots. The older portions of the roots do not absorb directly and frequently bear no absorbing rootlets. They may be reinvested with absorbing rootlets if the soil moisture is replenished. The absorbing portions of the root system is larger during moist periods owing to an increase in the number and length of absorbing rootlets. The absorbing portion of the root system decreases during periods of drought owing to the fact that the temporary or deciduous absorbing rootlets die much more rapidly



than they are produced. Competition with roots of plants in adjacent areas check the lateral spread of the horizontal roots. This is due, in part at least, to the drying of the soil. This causes most of the main roots to develop in a plane perpendicular to the row. There is a greater tendency for the roots to intermingle if the soil is kept moist as in periods of frequent precipitation.

However, root systems developing from the transplanted tobacco had no taproot and lacks the symmetry (Bruner 1931). Root branches developed from transplanted portion of the original root system and grew at all angles from the base of the plant. Some follows a horizontal course in the moist surface loam (soil) but some grew directly downward or at more or less of an angle. Competition among the horizontal roots caused them to develop less strongly at the extremities as the season progressed. The competition stimulated the development of their longer branches some of which penetrated downward and not infrequently the branch became the main root later in the season. Thus, many roots which appeared to turn rather sharply downward about 2 feet from the plant were not the result of curvatures of the original root but were developed from a portion of a horizontal root and one of its lateral branches. This occurred. Where the distal portion of the horizontal root ceased to function or became unimportant as an absorbing structure.

### ***10.2.2 Drought Tolerance***

Drought may be defined as the inadequacy of water availability, including precipitation and soil moisture storage capacity, in quantity and distribution during life cycle of crop plant restricting the expression of genetic potential of the plant. Drought can be the result of an overall decline in rainfall in wet or dry season, a shift in the timing of the wet season, or a strong local warming that exhausts water bodies and soils through evaporation. Drought is considered the most destructive condition influencing the growth of crop plants consequently leading to decreased yield (Lambers et al. 2008). The amount of damage to yield depends on the severity and duration of stress, plant resistance and plant growth stage (Robertson et al. 2004). Drought stress is a major constraint to tobacco production and yield stability in many rainfed regions of tobacco cultivation. Drastic shifting of rainfall pattern and frequent dry spells cause moisture stress especially in critical periods of crop growth, significantly affects growth, yield and quality of tobacco. Drought stress limits the growth and economic yield of tobacco through reduction in leaf growth, chlorophyll concentration, soluble protein concentration, stomatal conductance, accelerating senescence of leaves and reducing the rate of photosynthesis etc.

### ***10.2.3 Flooding and Submergence Tolerance***

Tobacco plants are mostly cultivated in a dry climate. Tobacco plants require dry land conditions for 2–3 months after planting to harvest leaves and for the ripening process (Muttaqin et al. 2019). One risk of tobacco cultivation in general is the high rainfall causing waterlogging. This type of environmental stress might happen because of unpredictable season occurring in tropical region and global climate change as a consequence of rapid growing industries all over the world.

Tobacco is the most susceptible crop to injury from flooding or saturation of the soil with water. Waterlogging lowers the growth and productivity of tobacco, which is very sensitive to an excess of water. It exhibits two types of reaction to flooding, immediate but temporary wilting accompanying temporary flooding, and severe permanent injury caused by longer periods of flooding (Kramer 1951; Kramer and Jackson 1954; Campbell 1973). If the soil is suddenly saturated by a downpour of rain and the sun later shines bright and hot, sudden wilting of the leaves, often termed “flopping” by tobacco growers, sometimes occurs. This sudden wilting occurs where drainage is slow and the soil remains saturated for at least a few hours after a rain. High air temperatures and bright sun accentuate this sudden wilting and its occurrence may depend also on the condition of the plants, apparently being much more severe if the tobacco has been growing rapidly and therefore is somewhat soft and succulent. If high winds accompany the rainfall, blow over may occur along with the root damage. Lodged or blown-down tobacco can be difficult to manage. If the excess soil moisture drains away within a few hours the plants usually recover from this type of wilting with little or no permanent injury. The growth and yield of the plants may not be significantly affected if plants are flooded for less than 24-h, whereas flooding for longer than 48 h significantly reduced yield even up to 80% than that of the unflooded plants (Campbell 1973; Anuradha et al. 2013).

### ***10.2.4 Light Stress***

Light serves as primary energy source in the phototrophic lifestyle of plants. The photoreceptors of plants perceive the light and regulate different metabolic processes through gene expression (Gyula et al. 2003; Kami et al. 2010; Jenkins 2014). The changes in light conditions directly affect the photosynthetic reactions within chloroplasts. The detrimental effect of high or low intensity light on biological and metabolic processes of plant is denoted as light stress. The low intensity light deprive the photosynthetic activity and high intensity light damages the photosynthetic apparatus. The decline of photosynthetic activity due to intense incoming light is known as photoinhibition. The shuttle changes in light quantity and quality primarily cause imbalances in the light reactions of photosynthesis and the carbon fixation reactions. However, light stress is not a major issue of concern in tobacco cultivation and currently researchers are not seriously attempting to breed tobacco for light stress.

### 10.2.5 UV Stress

Plants are exposed to ultraviolet-B (UV-B, 280–320 nm) at varied intensities based on the solar angle and spread of stratospheric ozone layer in the specified region. Although UV-B is only a minor component of the total solar radiation (0.5%), increase in its intensity has devastating effects on biological systems. In the experiments conducted with tobacco seedlings (*N. tabacum* L. cv. K326), exposure to UV-B stress increased the carotenoid synthesis capability of plants (Shen et al. 2017). The plants could deplete the carotenoids to scavenge excess reactive oxygen species (ROS) at high UV-B radiation levels, which protects the tobacco plant from oxidative damage caused by UV-B stress. While increasing the photosynthetic efficiency it was found that expression of the carrot lycopene  $\beta$ -cyclase (DcLCYB1) in *N. tabacum* cv. Xanthi resulted in increased carotenoid accumulation, faster plant growth, early flowering and increased biomass there-by higher yields in constant and fluctuating light conditions (Juan et al. 2020). Further, UV stress also induces physiochemical changes in tobacco leaf, reduces the amount of wax deposited on the adaxial leaf surface and also alters the density of trichomes in tobacco leaf (Barnes et al. 1996). Breeding tobacco for UV stress is not a priority to tobacco researchers as UV stress is not a major limiting factor in tobacco cultivation.

### 10.2.6 Weather Fleck (Ozone Pollution)

Weather fleck in tobacco is a leaf spot syndrome induced by air-polluting ozone (Heggestad and Middleton 1959). Weather fleck has resulted in extensive losses to tobacco growers in the US and Canada since 1955 (Heggestad 1966). The two production areas most seriously affected have been cigar wrapper in Connecticut and Florida and flue-cured in southern Ontario, Canada. Weather fleck also occurs sporadically in burley, flue-cured, Wisconsin, and Maryland tobacco-producing areas. A number of air pollutants can cause injury to crop plants; evidence, however, indicates that ozone causes the flecking observed on tobacco leaves. Ozone appears to be the primary injury causing pollutant in the Maryland tobacco-producing area as a result of the region's proximity to Washington, DC and the high concentration of automobiles, which are considered to be the major source of ozone pollution.

### 10.2.7 Chloride Stress

Among mineral nutrients, chlorine is recognized as an essential micronutrient in tobacco cultivation. Tobacco is known to accumulate chloride very rapidly in considerable amounts, and an amount up to 100 g Cl kg<sup>-1</sup> leaf dry matter have been observed. Chlorine in small amounts results in promoting growth, leaf expansion, a

better hydration state, reduced transpiration, higher water use efficiency (WUE), and water saving (Franco-Navarro et al. 2016, 2019) and so improve yield and certain quality factors such as color, moisture content, elasticity, burning and keeping quality of tobacco leaves (McEvoy 1957). However, larger amount of chloride has many adverse effects on the quality of tobacco, so much so that the chloride content in tobacco leaves is considered as a major factor determining the quality of tobacco. An excess level of chlorine produces leaves with poor burning capacity, muddy appearance and undesirable odor as well as highly hygroscopic nature causing discoloration during storage (Karaivazoglou et al. 2005). The threshold value for chloride in a good and acceptable tobacco leaf is usually set at below 1.5% (Chari 1995), the values greater than 2% inhibit the burning properties of tobacco (Akehurst 1981; Juan and del 1986; Guardiola et al. 1987; King 1990). Tso (1990) reported that various soil and fertilization conditions, as well as tobacco type, variety and methods of harvesting may contribute to the differences in the absorption, distribution of chloride with respect to stalk positions and the total leaf chloride content.

### ***10.2.8 Salinity Stress***

Salinity is termed as the total amount of mineral salts dissolved in water and soil (Grattan 2002). More than 20% of cultivated land worldwide is affected by salt stress and is increasing over the time. It is mainly related to increase in  $\text{Na}^+$  and  $\text{Cl}^-$  ions and decrease in  $\text{K}^+$  and  $\text{Ca}^+$  ions in plants (Perez-Alfocea et al. 1996; Shilpim and Narendra 2005). Salt stress will lead to water stress there-by affecting the leaf growth and development. Salinity stress negatively influences the cell division and expansion as well as stomatal opening and closing (Flowers 2004). In tobacco, soil salinity is known to reduce the plant growth through osmotic stress followed by ion toxicity. The salt stress in tobacco can be divided into ion toxicity (such as destroying plasma membrane structure, hindering the absorption of mineral elements, etc.) and the secondary stress effect (oxidative stress, drought stress, etc.) (Sharma et al. 2019).

### ***10.2.9 Heavy Metal Stress***

The tobacco farmers are bound to use huge amount of fertilizers and pesticides which contain high levels of metals (Karaivazoglou et al. 2007; Lecours et al. 2012). The main reason of heavy metals is the phosphate fertilizers utilized in the tobacco cultivation. The levels of metal accumulation in the leaves are found to vary (Lugon-Moulin et al. 2006) based on the area in which tobacco is cultivating. Vardi and Venkatrayulu (2019) reported the contamination of water samples with heavy metals namely, Arsenic, Lead, Cadmium, Mercury, Iron, Manganese, Copper, and Zinc which was above the WHO standards. Heavy metal toxicity symptoms are usually associated with stunted stem and root growth, chlorosis in younger leaves which

extends to the older leaves after long term exposure (Gangwar and Singh 2011; Srivastava et al. 2012).

### ***10.2.10 Traditional Breeding Methods Addressing Abiotic Stresses***

Tobacco is a self-pollinated crop with 5–10% out-crossing. Hence, all the breeding methods such as introduction, mass selection, pure line breeding, pedigree method, back cross breeding, mutation breeding, interspecific hybridization etc. that are commonly used in self-pollinated crops are being used in tobacco breeding (Bowman and Sisson 2000; Sarala et al. 2012). However, development and release of resistant cultivars for specific abiotic stresses is not there in tobacco. In view of selection of plants in low moisture regimes and having chloride contents below 1 ppm in breeding programs may be ensuring in developing cultivars that can adopt to water and chloride stress to certain extent (Sarala et al. 2012).

In order to get higher and stable yields under different abiotic stresses, the inherent capacity of tobacco genotypes need to be improved. However, development of stress tolerant tobacco cultivars requires thorough understanding of plant responses to stress environment. Information on availability of stress resistance sources, understanding the genetics and inheritance pattern of genes involved in stress resistance, molecular mechanisms conferring resistance and genome sequences associated with abiotic stress resistance are essential inputs in the development of abiotic stress resistant tobacco cultivar.

The first major category of breeding for abiotic stress environments is the indirect method, this approach attempts to breed for high yield and quality under several environments ranging from optimum conditions to stress conditions. The genotypes selected with high yield and quality under optimum conditions will also excel under drought conditions. As such, in this approach selections are not based directly on stress factors. There is evidence of existence of high positive correlation between performance of genotypes in optimum and stress conditions (Johnson and Frey 1967). The selections having high yield and quality over environments are further evaluated under a range of environments and those selections showing high performance with stability can be released for its cultivation in stress prone areas.

In order to overcome the problems of indirect breeding, second category for breeding approach for abiotic stress resistance (drought) has been advocated (Hurd 1971) in which test materials are evaluated in deliberately chosen testing sites that represent drought conditions reliably and uniformly. It is obvious that uniformity of drought conditions in field trials cannot be imposed while making selection among the genetic materials as drought is highly unpredictable and varies over the years and locations, thus resulting in reduced effectiveness of selection especially for yield which has low heritability.

The third category of abiotic stress resistance breeding is through incorporation of characters that contribute to drought mechanism into a high yielding variety. The choice of the character/characters to be incorporated will depend mainly on the importance of the character in enhancing abiotic stress resistance without much compromise on yield; the ease, speed, inexpensive and accuracy in measurement of the characters; targeted trait should have high heritability than yield and positively correlated with yield and the character should be reasonably stable over the time and should withstand minor environmental fluctuations.

Methodology of the incorporation of such characters depends upon the gene controlling the character and heritability. Characters controlled by oligogenes that are simply inherited, and if the aim is to transfer few characters backcross breeding is the usual choice. Characters controlled by polygenes that are considerably affected by environmental factors pedigree method are followed. A combination of drought contributing characters rather than a single character is much more useful selection criteria for drought resistance (Lundlow and Muchow 1990). When such is the aim for simultaneous selection of drought contributing characters having different heritability's, as is usually the case in many breeding programs, modification of bulk method called single seed descent is advocated. Jinks et al. (1977) while evaluating the random sample of 59 F<sub>1</sub> of *N. rustica* lines through single seed descent method had isolated superior recombinant lines for flowering date and plant height and demonstrated that the mean performance and environmental sensitivity were largely under control.

The success of any of the above breeding approaches and methods depends upon the efficiency of screening techniques of drought resistance. The importance of developing reliable screening techniques has been realized very early (Levitt 1972). Numerous screening techniques for drought under laboratory and field conditions have been developed and breeder can adopt any of the techniques depending upon the situation. Since increasing yield is the ultimate goal of any plant breeding program, breeders emphasize on higher yields under moisture stress conditions. Therefore selections based on drought index, which provides a measure of drought based on loss of yield under drought condition verses optimum/stress condition is advocated for drought screening (Clark et al. 1984; Ndunguru et al. 1995). At ICAR-Central Tobacco Research Institute (CTRI), breeding program for drought has been initiated utilizing drought resistant entry, MRS 3 as one of the parent. The material is being screened under optimum conditions coupled with in vitro screening for germination under high molecular weight poly-ethylene glycol (PEG).

### ***10.2.11 Use of Morphological Markers***

Morphological markers in tobacco are related to easily identifiable variation in different plant, leaf, flower, capsule and seed characters (Sarala et al. 2018a, b). These markers can be easily scored and doesn't require sophisticated equipment or

preparatory procedures. Breeding for abiotic stresses aims to incorporate morphological characters that directly or indirectly contribute to resistance/tolerance mechanism into a high yielding variety. The choice of the character/characters to be incorporated will depend mainly on the importance of character in enhancing abiotic stress resistance without much compromise on yield; the ease, speed, inexpensive and accuracy in measurement of the characters; targeted trait should have high heritability than yield and positively correlated with yield and the character should be reasonably stable over the time and should withstand minor environmental fluctuations.

Morphological, anatomical and compositional characters identified to be associated with abiotic stress tolerance/resistance (especially drought) in tobacco that can be used as morphological markers to develop a model plant conferring abiotic stress resistance are discussed below.

Increased rooting depth and rooting density have been found to correlate positively with resistance to water deficit (Tuberosa et al. 2002). The length, weight, volume, penetrative ability, density of plant roots were reported to be associated with drought resistance (Tuberosa 2012). When water reserve exists at depth, a decrease in the cytokinin (CK) level or a reduction in CK signaling can lead to an enlarged root system (Macková et al. 2013) that reaches the water resources. When moisture reserves are confined to the upper layer of soil, number of lateral roots gain importance than deep root system. Breeding for small xylem vessels in the seminal roots has been suggested as a means for increasing the resistance (Passioura 1983). Increased root/shoot ratio is also associated with drought resistance (Bliss et al. 1957).

As tobacco varieties are bred for higher leaf yields, mechanisms that regulate transpiration improves WUE and confers drought resistance in plants. Morphological responses like, narrow and thick leaves (Nobel 1980; Abrams et al. 1990), leaf area (Turner 1986; Pereira and Chaves 1993), surface leaf rolling, increasing cuticular waxes deposition (Cameron et al. 2006), hairs on leaf surface, covering with trichomes on leaves, mid-veins, stalks, and floral parts of many *Nicotiana* species (Goodspeed 1954) aid in reduction of transpiration. Low stomatal frequency can reduce potential transpiration and improve WUE (Wilson 1975; Wang et al. 2012; Dias de Oliveira et al. 2013). Stomata size, distribution and sunkenness plays a critical role in regulation of transpiration.

Trichomes present in tobacco play several roles in the defense against abiotic stresses such as moderation of leaf temperature or water loss through increased light reflectance, sequestration and compartmentalization of heavy metals etc. (Wagner 1991; Harada et al. 2010). Trichomes also play an important role in ion and metal homeostasis of plants. Hence, recording observations on trichome densities is important in identifying resistant genotypes in breeding programs aiming at multi stress-tolerant genotypes.

Altered anatomical properties like thicker palisade tissue, a higher ratio of palisade to spongy parenchyma thickness and a more developed vascular bundle sheet reduce excess water loss and enhance water holding ability there-by improves drought tolerance (Esau 1960; Guha et al. 2010). The fortified sclerenchyma can reduce the damage from wilting and protect plants from direct light radiation (Terashima 1992).



The phenomenon of the accumulation of various organic and inorganic substances such as sugars, polyols, amino acids, alkaloids and inorganic ions in the cytochylema reduces the osmotic potential, increases cell water retention and elasticity of cell (Morgan 1984; Rhodes and Samaras 1994). Such osmotic adjustment, sustains cell structure and photosynthesis, delays leaf senescence and improves root growth under stress conditions.

Tobacco has evolved a C3 path way for carbon dioxide and carbohydrate fixation which is not as efficient as other two path ways viz., crassulacean acid metabolism (CAM) and C4, because of photorespiration. Photorespiration increases with the temperature and under moisture conditions (Rivero et al. 2009; Huang et al. 2016). A balance between growth and carbon supply is achieved through a complex regulatory network in which sugars (e.g., glucose, sucrose and starch) and phytohormones, mainly abscisic acid (ABA) and cytokinins (CK) perform central roles (Shinozaki and Yamaguchi-Shinozaki 1996; Rolland et al. 2006; Havlov et al. 2008; Nishiyama et al. 2011). Under moisture stress conditions roots of plants triggers a huge increase in *de nova* synthesis of ABA (Sauter et al. 2001) and transported mainly to leaves as an intercellular messenger and recognized by guard cells which trigger stomatal closer via intercellular single transduction and weakening the metabolic activities related to plant growth (Boursiac et al. 2013). The influence of ABA has multiple effects on drought response encompassing the regulation of stomatal closure, channel activities in guard cells, transcriptional levels of calmodulin protein and the expression of some ABA responsive genes (Cocucci and Negrini 1988; Rabbani et al. 2003). Researchers are paying attention to improve photosynthesis efficiency for increasing yields under different conditions including stresses (Zhu et al. 2010; Long et al. 2015; Ort et al. 2015).

Majority of the stresses affect plant growth, development and morphogenesis. Recording observations on relative plant growth rates and morphogenesis patterns are essential in identifying the effect of various stresses and resistant genotypes. Time taken for stress symptom occurrence after the stress incidence, its relative severities, symptom progression and recovery patterns after stress amelioration are important morphological parameters in identifying abiotic stress resistant genotypes.

### ***10.2.12 Limitations and Prospect of Genomic Designing***

The genome designing strategies overcome the limitations of conventional breeding as they deal at the level of genomes and manipulate the gene sequence to achieve desired phenotypic traits. Transfer of desired traits from tertiary gene pools and other unrelated sources to cultivated tobacco can also be successfully achieved through trans- and cis-genesis approaches involving gene mapping, identification, gene transfer, gene editing etc. With the rapidly evolving technological advancements, marker and genome assisted breeding approach is going to accelerate the progress made in breeding programs. Targeted modification or designing of plant

genome including addition of alien genes will accelerate the tobacco varietal developmental process through precise manipulation of gene functions for higher yields and stress resistance.

Published draft genome of *N. tabaccum* and a few wild species and data sharing and analysis platforms (databases) available in recent times, made it possible to use innovative bioinformatics tools for the in depth study of genomes and their comparative genomic analysis. Such studies are helping to understand genes, their sequences and linked molecular markers for target traits. This information can successfully be utilized to edit the genome sequences with rapidly evolving precise gene editing tools viz. meganucleases, zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), homing endonucleases, CRISPR/Cas 9 etc. Gene editing technologies at present do suffer from lower specificity due to their off-targets side effects (Khan 2019). High density of molecular maps and genome information are now offering scope for providing the knowledge of linked molecular markers and quantitative trait loci (QTLs) that are either tightly linked or present within the target gene (s) and also allow map-based cloning of desirable traits. Linked markers and QTLs identified in tobacco for various abiotic stress responsive genes are going to pave the way for marker assisted breeding for resistance to abiotic stresses. The available information on linked markers and traits can be effectively used in estimating the breeding value of individuals in genomics aided breeding and accordingly desired plants can be selected.

Though genetic engineering (GE) tools offer a number of advantages, they do have certain limitations. Time-consuming and complicated protocols, potential tissue damage, incorporation of DNA of selection marker in the host genome, and low transformation efficiency etc. are some of the limitations of GE technologies. Compared to tradition breeding, genome designing techniques are resource intensive and require technology expertise for handling the protocols and processes.

### 10.3 Genetic Resources of Resistance Genes

Availability of genetic resources with stable and heritable resistance factors for abiotic stresses can facilitate breeding resistant varieties. The gene pools consist of easily crossable tobacco lines and *Nicotiana* species are to be explored for such variability. In case of non-availability of sources of resistance in any of these gene pools, variability need to be created through mutations or incorporated through genome designing approaches.

Currently, large number of cultivated tobacco varieties and around 83 *Nicotiana* species are available (Lewis 2011; Berbeć and Doroszewska 2020). Taxonomy Browser of National Centre for Biotechnology Information (NCBI) lists around

92 *Nicotiana* species and varieties (<https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi>). 307 records of tobacco varieties and *Nicotiana* species and available in The Plant List database (<http://www.theplantlist.org/tpl1.1/search?q=Nicotiana>). The International Plant Name Index providing 440 records with the keyword 'Nicotiana' through its database (<http://www.ipni.org/ipni/plantnamesearchpage.do>). Large number of these species are reliable sources of resistance to various stresses (Lewis 2011). Wild *Nicotiana* species are good sources of cytoplasmic male sterility (CMS) for developing male-sterile isolines of inbred lines and cultivars. CMS is a prerequisite in tobacco for technically feasible and economically viable seed production of hybrid varieties. Various available sources of resistance in tobacco genetic resources are discussed below.

### 10.3.1 Primary Gene Pool

The primary gene pool includes genotypes that are crossable and produce fertile offspring with the cultivated tobacco. They may be cultivated species and in wild gene pools. The cultivated gene pool comprises of commercial varieties, as well as landraces. Wild gene pool comprises closely related species and putative ancestors that have fair degree of crossability with the cultivated tobacco. Large number of varieties are developed by breeders in different countries and fairly large collections of germplasm are available in *N. tabacum* and *N. rustica* that can be explored for the identification of abiotic stress resistant factors as gene transfers from such sources is easy.

### 10.3.2 Secondary Gene Pool

The secondary gene pool refers to crop wild relatives that can cross with the cultivated tobacco at least to some extent to produce some fertile offsprings. These sources are distinct from the cultivated species and include closely related species, primitive cultivars, old land races evolved and adopted to different environments and hence, valuable source for resistance to abiotic stresses. *N. tabacum* is found to hybridize with majority of *Nicotiana* species (58 no.) directly or through sister *Nicotiana* species (Berbeć and Doroszevska 2020). *N. tabacum* is found to yield inviable hybrids with *N. africana*, *N. excelsior*, *N. goodspeedii*, *N. gossei*, *N. maritima*, *N. megalosiphon* and *N. velutina* when crossing at 28 °C (Tezuka et al. 2010). However, Type II hybrid lethality showing the characteristic symptoms of browning of hypocotyls and roots observed in these crosses can be suppressed at higher temperatures (34–36 °C). Utilization of genes from these materials is tedious due to incompatibility and undesirable linkages. Genomic tools can assist in overcoming such difficulties.

### 10.3.3 Tertiary Gene Pool

More distantly related crop wild relative species are included in this pool. *N. tabacum* had no hybridization records with 14 *Nicotiana* species (*N. azambujae*, *N. acaulis*, *N. ameghinoi*, *N. paa*, *N. cutleri*, *N. longibracteata*, *N. spagazzini*, *N. faucicola*, *N. fatuhivensis*, *N. heterantha*, *N. monoschizocarpa*, *N. stenocarpa*, *N. truncata*, and *N. symonii*) (Berbeć and Doroszewska 2020). Hybrid lethality in crosses with incompatible *Nicotiana* species may be due to the genes in both the S and T sub-genomes of *N. tabacum* (Tezuka and Marubashi 2012). Specific techniques such as bridge crossing, ovary/ovule culture, embryo rescue, various sorts of treatments of male and/or female flower parts, partial genome transfer (chromosome addition and/or substitution lines, translocation breeding, mutagenesis, cell fusion, etc.) chromosome and genome manipulation (polyploidization or haploidization), exchange of nuclear and cytoplasmic genomes (mitochondrial and/or chloroplastic), grafting, marker-assisted breeding (MAB), tissue culture and genetic engineering are needed to transfer genes from such pools (Weil et al. 2010).

### 10.3.4 Artificially Induced/Incorporated Traits/Genes

Creation of mutations (physical and chemical mutagens), genetic engineering for transfer of alien genes, gene manipulation and genome editing technologies are to be adopted in developing resistant cultivars when source of resistance is not available in any of the above pools.

## 10.4 Glimpses on Classical Genetics and Traditional Breeding

### 10.4.1 Classical Mapping Efforts

Very few classical studies were reported in tobacco. Clausen and Goodspeed (1926) established that one of the two types of monosomics (haplo-C, then called “corrugated”), involved the chromosome in which the basic color factor, *Wh*, is located. Anderson and Dorothea (1931), East (1932), and Brieger (1935) reported linkage between a pollen color factor and the sterility factors. Later, Brieger (1935) established the first two linkage groups 1. self-sterility allele (*S*) and lethality (*I*) 2. *C* is the basic gene for anthocyanin color and a recessive gene causing a peculiar type of growth *cr* (*crassa*) based on the linkage data on in *N. langsdorfii* and *N. sanderae*. Smith (1937) confirmed the existence of linkage between self-sterility and pollen anthocyanin color in tobacco.

Later, Clausen and Cameron (1944) established location of 18 genes on nine chromosomes through the transmission studies between monosomics and mendelian characters using complete set of 24 monosomics. However, due to its allopolyploid nature (Suen et al. 1997; Narayanan et al. 2003) genetic linkage maps are not fully developed in tobacco.

### ***10.4.2 Limitations of Classical Endeavors and Utility of Molecular Mapping***

Mapping based on morphological markers is tedious and time taking and genes governing quantitative traits cannot be mapped (Worland et al. 1987). To make gene maps more comprehensive it would be necessary to find characteristics that were more distinctive and less complex than visual ones. But, only a fraction of the total number of genes in tobacco exist in allelic forms that can be distinguished conveniently making it difficult to construct classical maps. One of the reasons why our knowledge of the details of inheritance in tobacco was so meager, is because of the prevalingly quantitative or semi-quantitative nature of majority of characters including flower color in tobacco (Clausen and Cameron 1944). Abiotic stress responses is the result of action of numerous genes with major and minor effects with low heritability and are influenced by environmental factors of the gene which adds to the genotyping woes.

Recent enormous progress in the field of biotechnology, especially with the advent of DNA markers, QTL mapping techniques, genome sequencing techniques, gene/genome editing techniques and genome wide association mapping techniques, identification and mapping of candidate genes/markers conferring abiotic stress resistance/tolerance is becoming more feasible.

However, in comparison to the other Solanaceae crops such as the tomato, potato, and pepper plants, molecular marker development and genetic map construction in tobacco have lagged behind (Tanksley et al. 1992; Barchi et al. 2007). The molecular marker based maps can be effective anchoring points for identification of linked traits for their isolation, cloning and also for use in marker-assisted breeding.

### ***10.4.3 Breeding Objectives***

The tobacco breeding mainly aims at enhancing leaf yield potential of the cultivar in addition to maintaining leaf quality, and resistance to biotic and abiotic stresses. Numerous studies have identified plant characters that are associated with various abiotic stress responses. For example, increased root depth and root density has been found to correlate positively with resistance to water deficit (Tuberosa et al. 2002). Increased root/shoot ratio is also associated with drought resistance (Champoux

et al. 1995; Tavakol and Pakniyat 2007; Ali et al. 2009; Pallardy 2010). Morphological responses like, narrow and thick leaves (Nobel 1980; Abrams et al. 1990), leaf area (Turner 1986; Pereira and Chaves 1993), surface leaf rolling, increasing cuticular waxes deposition (Cameron et al. 2006), hairs on leaf surface, covering with trichomes on leaves, mid-veins, stalks, and floral parts of many *Nicotiana* species (Goodspeed 1954) aids in reduction of transpiration. Trichomes are involved in the moderation of leaf temperature or water loss through increased light reflectance (Wagner 1991). Stomata size, distribution and sunkenness plays a critical role in regulation of transpiration. Low stomatal frequency can reduce potential transpiration and improve WUE (Wilson 1975; Wang et al. 2012; Dias de Oliveira et al. 2013). Altered anatomical properties like thicker palisade tissue, a higher ratio of palisade to spongy parenchyma thickness and a more developed vascular bundle sheet reduce excess water loss and enhance water holding ability there-by improves drought tolerance (Esau 1960; Guha et al. 2010). The fortified sclerenchyma can reduce the damage from wilting and protect plants from direct light radiation (Terashima 1992). Menser and Street (1962) showed that N nutrition critically affected the weather fleck susceptibility of Catterton tobacco. Fleck and N supply were related inversely. Selection for these associated traits in the desired direction, positive or negative, would yield desired results.

#### 10.4.4 Classical Breeding Achievements

Traditional tobacco breeding aimed at developing improved tobacco varieties with higher yield, better leaf quality, resistance to biotic and abiotic stresses. Significant progress has been made over the years in enhancing the tobacco leaf yield through both varietal and hybrid development, in addition to improving disease and insect resistance without significantly sacrificing in ease of curing. However, success in case of abiotic stress resistance/tolerance is relatively low.

Janardhan et al. (1994) identified Bell No. 10, Bigorinico, Cocker 128, F. 207 and F. 212 as tolerant to drought using sprinkler line-source technique under field conditions during rain-free post monsoon season. Sarala et al. (1998) identified tobacco genotypes, Cy 113, Cy 118, Kanchan, L 621, VA 21 and CM 12 as drought tolerance lines under cyclic water stress (Sarala et al. 1998). Cultivated Varieties viz., Zhubo-1, G 80, K346 Sahyadri, N-98, Tugabhadra, Anand-119, GT-4 CTRI Special, Jayashree, 16/103, Godavari Special, Hema, VT1158, Rathna etc. are found to be drought resistant/tolerant (ICAR-CTRI 2021). FCV tobacco entries, FCR-23 and FCR-15 recorded higher pollen and seed germination under higher PEG concentration which indicates their drought tolerance capacity (ICAR-CTRI 2016).

Povilaitis and White (1966) used a segregating population of flue cured 'Delcrest' to develop fleck tolerant 'Delcrest 66'. McKee (1968) developed Maryland 64 by crossing Catterton X Wilson and selected for an intermediate, high yielding type. His efforts led to the most fleck-resistant of the Maryland cultivars currently grown although obtaining higher fleck resistance was not his main objective.

Nurhidayati et al. (2017) identified tobacco varieties, Kemloko 3 (index value of 0.03), Paiton 2 (index value of 0.18), and Kemloko 2 (index value of 0.42) as resistant to water logging stress based on the sensitivity index. FCV cultivars viz., FCJ-11 and FCR-15 found to withstand wet foot to certain extent (Sarala et al. 2020).

#### ***10.4.5 Limitations of Traditional Breeding and Rationale for Molecular Breeding***

Most of the stress tolerance traits are mainly quantitative trait loci and greatly influenced by environment, thus making selections difficult (Anderson et al. 2014). The mechanisms of abiotic stresses like drought tolerance are highly complex and recent advances have provided insight into plant gene regulatory network system, which is mainly composed of inducible-genes (environmental factors and developmental cues), expression programming and regulatory elements (cis-element and trans-element), corresponding biochemical pathways and diverse signal factors (Tang et al. 2003; Wang et al. 2003; Zhu 2003; Munns 2005). Many drought-inducible genes are also induced by salt stress and cold, which suggests the existence of similar mechanisms of stress responses. Hundreds of genes are thought to be involved in abiotic stress responses (Seki et al. 2003; Baloglu et al. 2012). The biggest challenge in traditional breeding is the environmental interaction and low heritability of the genes involved in regulating resistance/tolerance mechanism which drastically hinders the progress of incorporation of characters of abiotic stress resistance and may not reflect the desired yield coupled with resistance due to difficulties in selections owing to considerable variation in the imposition of stresses in field conditions.

The major limitation in traditional breeding for abiotic stress response in tobacco is undesirable gene association is polygenic inheritance coupled with low heritability of the genes involved in abiotic resistance/tolerance reaction.

With the advent of genome designing techniques like, marker-assisted selection (MAS), plant transformation, various gene editing tools etc. and identification of several candidate genes with major effects, is possible to develop tobacco cultivars having resistance/tolerance reaction to various abiotic stresses.

### **10.5 Brief on Diversity Analysis**

Lack of diversity in cultivated crops can lead to crop losses due to reduced flexibility of varieties to adapt to changing environmental conditions such as increasing temperatures or salinity and to combat infestations of new strains of biotic stresses (Moon et al. 2009a). Hence, genetic diversity analysis of germplasm is essential for identifying sources for economically important traits and diverse parentals to



create maximum genetic variability in the breeding populations for effecting selection in breeding (Barrett and Kidwell 1998). Deploying tobacco varieties developed from diverse genetic backgrounds insulate the crops from changing environmental stresses. Thus, genetic diversity analysis is an essential step for continued progress in breeding as well as for adaptation to future environmental challenges.

### ***10.5.1 Phenotype-Based Diversity Analysis***

Phenotypic diversity in terms of morphological, karyotypical and physiological characters have been regularly studied in tobacco germplasm (Goodspeed 1954; Zhang 1994; Lu 1997). Diversity is found to exist in tobacco germplasm for several agro-morphological traits (Zhang 1994; Wenping et al. 2009; Zeba and Isbat 2011; Baghyalakshmi et al. 2018; Sarala et al. 2018), chemical and cytological traits (Tso et al. 1983; Okumus and Gulumser 2001; El-Morsy et al. 2009; Darvishzadeh et al. 2011). Agro-morphological traits are found to vary with environment and their diversity estimates are affected under different environments (Lu 1997). Studying the diversity of morphological characters for importing abiotic stress resistance and yield contributing traits is essential for breeding abiotic stress tolerance in tobacco.

The tolerance/resistance mechanism to different abiotic stresses is now being extensively studied through high throughput phenotyping where the system quantifies a number of traits in a population with automated image collection and analysis. This technology can effectively be utilized in breeding genotypes for abiotic stresses to its non-destructive sampling methods, rapid screening of larger population under artificially created abiotic stress conditions (Buschmann and Lichtenthaler 1998; Goggin et al. 2015). Such high throughput phenotyping systems could possibly reduce the amount of labor and screening time for identifying plants that are tolerant and have desirable traits.

### ***10.5.2 Genotype-Based Diversity Analysis***

Limited information has been available confirming the relationship between morphological variability and genome diversity in cultured tobacco. In view of this, attempts made to examine the degree of relatedness, among tobacco cultivars and diversity of germplasm, based on variability at DNA level. As about 77% of the total genomic DNA content is composed of repetitive sequences in tobacco, the remaining non-repetitive sequences part is responsible for variability in morphological and quality traits (Narayan 1987).

The advent of different types of molecular markers over the last two decades has revolutionized the entire scenario of biological sciences including tobacco (Liu and Zhang 2008). These markers are abundantly available throughout the genome and offer advantages such as highly polymorphic nature, codominant inheritance,

easy access, easy and fast assay, high reproducibility and easy exchange of data between laboratories. Molecular markers provide a relatively unbiased estimation of genetic diversity in plants. DNA-based molecular markers have acted as versatile tools and have found their own position in various fields like characterization of genetic variability, genome fingerprinting, genome mapping, gene localization, analysis of genome evolution, population genetics, taxonomy, genome comparisons, gene mapping, quantitative trait loci analysis, marker-assisted breeding diagnostics, etc.

Molecular markers such as restricted fragment length polymorphism (RFLP), randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), microsatellites or simple sequence repeat (SSR), single-nucleotide polymorphism (SNP), inter simple sequence repeats (ISSR) etc. have been employed in studying genetic diversity, gene mapping and marker-assisted breeding of tobacco. RFLP markers were the first molecular markers used in tobacco research specially to study the function of few cloned genes (Bretting and Widrechner 1995). Invention of polymerase chain reaction (PCR) technology and PCR-based markers such as RAPD and AFLP emerged in the beginning of nineties and later microsatellite markers were used by different workers to study genetic diversity in tobacco. Compared to RFLP, these PCR based markers are preferred because of the relative ease with which PCR assays can be carried out. Both RAPD (Xu et al. 1998; Del Piano et al. 2000; Evanno et al. 2005; Zhang et al. 2005, 2008; Arslan and Okumus 2006; Sarala and Rao 2008; Sivaraju et al. 2008; Denduangboripant et al. 2010; D'hoop et al. 2010) and AFLP (Huang et al. 2008; Zhang et al. 2008; Chuanyin et al. 2009; Liu et al. 2009) were used to analyze the genetic diversity and varietal identification in tobacco.

Soon after the discovery of simple sequence repeat (SSR) markers in late 90s and the beginning of twenty-first century, they became markers of choice as they could be able to eliminate all drawbacks of earlier DNA marker technologies (Jafar et al. 2012). Considering that the genetic diversity between tobacco cultivars (particularly between those of the same type) is very limited (Del Piano et al. 2000; Rossi et al. 2001; Julio et al. 2006) and that cultivated tobacco is a tetraploid species with a very large genome (Livingstone et al. 1999; Ren and Timko 2001; Doganlar et al. 2002), making the development of PCR-based molecular markers generally inefficient. However, Bindler et al. (2007) for the first time employed around 637 functional SSR markers (out of which 282 were highly polymorphic) for variety identification. Since then, SSR markers are being regularly used in estimating the diversity in tobacco. An additional set of 5,119 new and functional SSR markers were developed for mapping and diversity studies by Bindler et al. (2011). Later, Tong et al. (2012) developed another set of SSR markers [including 1365 genomic SSRs and 3521 expressed sequence tag (EST)-SSRs] that slightly overlapped the set published by Bindler et al. (2007, 2011). Madhav et al. (2015) developed and validated a new set of microsatellite markers for their applicability in differentiating different types of tobacco, diverse cultivars of flue-cured virginia (FCV) tobacco, and the transferability of these markers in a wide range of *Nicotiana* species. Cai et al. (2015) utilized the database of tobacco EST for the development of EST-SSR markers and validated them in studying the genetic differentiation among tobacco accessions.

Wang et al. (2018) detected a total of 1,224,048 non-redundant NIX (*Nicotiana* multiple (X) genome) markers (SSRs) through comparative genome wide characterization of ~20 Gb sequences from seven species viz. *N. benthamiana*, *N. sylvestris*, *N. tomentosiformis*, and *N. otophora*, and three *N. tabacum* cultivars (TN90, K326, and BX) (Wang et al. 2018). Such large scale development of SSR markers in tobacco has led to the analysis of molecular diversity of genetic resources (Moon et al. 2009b; Davaliev et al. 2010; Fricano et al. 2012; Gholizadeh et al. 2012; Prabhakararao et al. 2012; Xiang et al. 2017), distinctiveness uniformity stability (DUS) testing (Binbin et al. 2020), genetic relatedness of cultivated varieties (Moon et al. 2008), estimating the changes in diversity due to breeding interventions (Moon et al. 2009a) and also for the identification of linked markers and QTLs to abiotic stresses (Hatami et al. 2013).

Markers such as ISSRs (Yang et al. 2005a, 2007; Qi et al. 2006) and inter-retrotransposon amplification polymorphism (IRAP) markers (Yang et al. 2007) have also been employed to assess genetic diversity in tobacco.

Even though the application of SNPs in tobacco is complicated and challenging due to its tetraploid nature and complex genetic architecture (Ganal et al. 2009), recent studies identified number of SNPs in tobacco (Xiao et al. 2015; Thim-megowda et al. 2018; Tong et al. 2020). These SNPs are being used in characterizing of germplasm for markers linked to economically important traits including abiotic stresses, development of molecular maps, and studying genome structure and organization.

Wang et al. (2021) identified 47 core Kompetitive allele specific PCR (KASP) and 24 candidate core markers based on SNP data. KASP markers can able to discriminate between two alleles of a SNP using a common reverse primer paired with two forward primers, one specific to each allele. These core markers were utilized for the identification of tobacco varieties and fingerprinting of 216 cigar germplasm accessions.

### 10.5.3 Relationship with Other Cultivated Species and Wild Relatives

Cultivated tobacco belongs to the genus *Nicotiana* and family Solanaceae. Evolution and the genetic diversity in genus *Nicotiana* was studied through comparison of morphological, cytological and biochemical traits, organellar (plastid and mitochondrial) genome organization and analysis of molecular features, such as repetitive DNA sequences and the structure of various nuclear gene families (Kostoff 1943; Goodspeed 1954; Komarnitsky et al. 1998; Lim et al. 2000; Liu and Zhang 2008).

In habit and habitat, the genus *Nicotiana* resembles the two genera, *Cestrum* (8 pairs of chromosomes) and *Petunia* (7 pairs of chromosomes) (Darlington and Janaki Ammal 1945). The genus is envisaged as derived from a pre-generic reservoir of two related genera and evolving into three complexes, at the 12-paired level,

that are hypothetical precursors of the three modern sub-genera. Although 6-paired species of *Nicotiana* is not known, the predominance of 12-paired species and their compound morphological character, along with a frequency of 4–8 pairing with a mode of 6 pairs in large number of F<sub>1</sub> hybrids combining 12-paired species, indicates that 6 is the basic chromosome number for *Nicotiana* and both 12 and 24 are derived numbers. The 24-paired species including *N. tabacum* and *N. rustica* are modern descendants of the 12-paired progenitors entered into amphiploid origin.

Goodspeed (1954) and Goodspeed and Thompson (1959) presented the systematic classification of the genus based mainly on cytogenetic studies involving chromosome morphology, behavior in interspecific hybrids and amphiploids and aneuploids. Subsequently, additions to this classification were made by Burbidge (1960). However, in the revised systematic classification based on molecular research, subgenera were dropped retaining the division into sections (Chase et al. 2003; Clarkson et al. 2004; Knapp et al. 2004). As per the new classification, *N. trigonophylla* Dun. was renamed as *N. obtusifolia* Martens et Galeotti, *N. affinis* Hort is considered synonymous with *N. alata* Link et Otto, and *N. bigelovii* (Torrey) Watson with *N. quadrivalvis* Pursh. *N. sanderae* Hort. is considered to be hybrid between *N. alata* and *N. forgetiana* Hemsl. (*Nicotiana* x *sanderae*) and *N. eastii* Kostoff as an autotetraploid variant of *N. suaveolens* Lehm. (Chase et al. 2003; Knapp et al. 2004).

Among the 83 identified wild species of *Nicotiana* in the genera, *Nicotiana tabacum* and *N. rustica* are the cultivated species (Goodspeed 1954; Chase et al. 2003; Lewis 2011; Sierro et al. 2014; Berbeć and Doroszewska 2020). Both the cultivated tobacco are allopolyploid species ( $2n = 4x = 48$ ) with basic chromosome number of  $x = 12$  (Gopalachari 1984; Knapp et al. 2004). Tobacco stands out as a complex allotetraploid with a large 4.5 Gb genome with significant proportion (>70%) of repeats (Zimmerman and Goldberg 1977; Renny-Byfield et al. 2011). *N. tabacum* constitutes wide range of morphological types having diversified utilities viz., smoking, chewing, snuff, etc. Other *Nicotiana* species cultivated in smaller scale are *N. repanda* Willd ex Lehm., *N. attenuata* Torrey ex S. Watson and *N. quadrivalvis* Pursh are for smoking, *N. sylvestris* Spegazzini & Comes, *N. alata* Link and Otto, *N. langsdorffii* Weinmannm, *N. forgetiana* Hemsley, and *N. sanderae* (Hort) for ornamental and *N. glauca* Graham for industrial purpose (Lester and Hawkes 2001).

*N. tabacum* is natural amphidiploid (allopolyploid,  $2n = 4x = 48$ ) ascended by hybridization of wild progenitor species, *N. sylvestris* (S-genome) x *N. tomentosiformis* (T-genome) and *N. rustica* L. from *N. paniculata*/*N. knightiana* (P/K-genome) x *N. undulata* Ruiz & Pav. (U-genome) (Goodspeed 1954; Clarkson et al. 2005; Lim et al. 2005; Leitch et al. 2008; Edwards et al. 2017; Sierro et al. 2018). Whole-genomic sequence studies indicated that the genome of *N. sylvestris* contributes 53% and *N. tomentosiformis* 47% to *N. tabacum* (Sierro et al. 2014). The comparative mapping studies suggested that the tetraploid tobacco genome has undergone a number of chromosomal rearrangements after the polyploidization (Wu et al. 2009; Gong et al. 2016). Number of reciprocal translocations and inversions (>10) has been found to differentiate the ancestral tobacco genomes from the tomato genome (Wu et al. 2009). Using cpSSRs and MtSSRs, Murad et al. (2002) concluded that the S genome

in tobacco was originated from *N. Sylvestris* ancestor. Chloroplast genome studies indicated that *N. otophora* is a sister species to *N. tomentosiformis* and *Atropa belladonna* and *Datura stramonium* are the closest relatives (Asaf et al. 2016).

Sierro et al. (2018) reported that 59% *N. rustica* genome originated from the maternal donor (*N. paniculata*/*N. knightiana*) and 41% from the paternal donor (*N. undulata*). Comparison of families of repetitive sequences proved that P- and U-genomes of *N. rustica* was similar to the putative parents, *N. paniculata* and *N. undulata*, respectively (Lim et al. 2005). Genomic in situ hybridization studies confirmed that *N. rustica* is an allotetraploid between *N. paniculata* (maternal P-genome donor) and *N. undulata* (paternal U-genome donor) and interlocus sequence homogenization has resulted in the replacement of *N. paniculata*-type intergenic spacer (IGS) of rDNA in *N. rustica* with *N. undulata*-type of sequence (Matyasek et al. 2003). However, analysis of nuclear genome, chloroplast genome and functional genes indicated that *N. knightiana* is more closely related to *N. rustica* than *N. paniculata*. Gene clustering revealed 14,623 ortholog groups common to other *Nicotiana* species and 207 of them are unique to *N. rustica* (Sierro et al. 2018).

Around 40% of *Nicotiana* species are allopolyploids and considered to be generated independently in six polyploidy events several million years ago (Clarkson et al. 2004; Leitch et al. 2008). Many of the diploid genome donors that make up various allopolyploid species are closely related and others are members of distantly related taxonomic sections (Clarkson et al. 2004; Leitch et al. 2008).

*N. tabacum* and *N. rustica* shares their basic chromosome number ( $n = 12$ ) with other Solanaceous species such as tomato, potato, pepper and eggplant (Lim et al. 2004; Clarkson et al. 2005). Microsynteny observed at the protein level between the genomes of *N. tabacum* cv. TN90, K326 and BX and those of tomato and potato (Sierro et al. 2014).

### 10.5.4 Relationship with Geographical Distribution

The genus, *Nicotiana* is presumed to have had its original habitat in and around the Andes region in South America and Central America, possibly from the mild to low altitude forest margin (Goodspeed 1954). While, occurring naturally as a perennial plant, tobacco is evolved as an annual crop. Twenty of the *Nicotiana* species are native to Australia, one to Africa and 54 to North/South America (Goodspeed 1954; Burbridge 1960; Clarkson et al. 2004). *N. benthamiana*, a species indigenous to Australia, is being used extensively as a model system to study various biological processes. The cultivated species, *N. tabacum* L. (common tobacco) and *N. rustica* L are native to America and several commercial varieties of them are cultivated extensively throughout the world.

Darvishzadeh et al. (2011) reported clustering of oriental-type tobacco genotypes based on morphological traits was in agreement with their geographical distribution. The genetic diversity studies with SSR markers in oriental (Darvishzadeh et al. 2011), and RAPD and AFLP markers in flue-cured tobaccos (Zhang et al. 2008) could not

indicate any such clear pattern based on their geographical origins. However, the clustering of tobacco genotypes based on molecular diversity found to correspond to commercial classes (Flue-Cured, Burley, etc.), manufacturing trait and parentage (Sivaraju et al. 2008; Fricano et al. 2012).

### 10.5.5 Extent of Genetic Diversity

Existence of morphological diversity is observed in the tobacco germplasm collections maintained at tobacco gene bank in India (Baghyalakshmi et al. 2018; Sarala et al. 2018). Similarly, Moon et al. (2009b) observed large average genetic diversity among *N. tabacum* accessions from the U.S. *Nicotiana* Germplasm Collection compared to FCV tobacco accessions. While, lower SSR diversity per locus reported in similar investigations carried out on TI accessions of tobacco (Fricano et al. 2012). Low degree of genetic polymorphism was observed among tobacco cultivars by different workers (Xu et al. 1998; Del Piano et al. 2000; Rossi et al. 2001; Yang et al. 2005a; Zhang et al. 2005; Julio et al. 2006). In contrast, Xiang et al. (2017) reported the richest genetic diversity for local group of tobacco varieties and lower diversity for introductions, and higher genetic similarity values between introductions and breeding group. While, the variation among the tobacco lines for chemical traits was found to be higher (Tso et al. 1983; Darvishzadeh et al. 2011) The relatively low levels of diversity in tobacco cultivars may be due to the utilization of only small proportion of the variability of the gene pools of the progenitor species in breeding programs (Ren and Timko 2001; Lewis et al. 2007). Among the species, the level of polymorphism among the varieties of *N. tabacum* was reported to be higher compared to *N. rustica* (Sivaraju et al. 2008). However, genetic diversity among wild tobacco accessions was found to be higher (Chuanyin et al. 2009).

## 10.6 Association Mapping Studies

Association analysis, also known as linkage disequilibrium (LD) mapping or association mapping is based on linkage disequilibrium, which detects the frequency of significant association between the genetic variation of markers or candidate genes and target traits in the natural population (Bradbury et al. 2007; Pritchard et al. 2000). Linkage disequilibrium (LD) is defined as the non-random association of alleles at two or more loci (Fricano et al. 2012). Association analysis does not require the construction of specific mapping populations and genetic maps, thereby, considerably reduce the workload. Further, it uncovers (explores) elite genes from a certain scale (quantity) of germplasm resources in a single instance providing evidence for genetic diversity. LD mapping has advantage over traditional mapping because in a random-mating population over several generations, only close linkage between markers and traits remains, thus facilitating fine mapping.

### ***10.6.1 Extent of Linkage Disequilibrium***

Extent of linkage disequilibrium in crop plants is influenced by mutation rate, genetic drift, selection, mating system, recombination rate, gene conversion, and size and structure of the population (Flint-Garcia et al. 2003). Long- and short-range LD could be identified through high-density genome fingerprinting. In species with large genomes, a lower number of molecular markers can be tested for the identification of large LD, although this will result in a lower mapping resolution (Waugh et al. 2009). Conversely, if large panels of markers are available, short-range LD enables the fine mapping of causal polymorphisms (Myles et al. 2009).

When conducting genome-wide association studies (GWAS), the knowledge on the extent of LD is essential to estimate the minimum distance required between markers for effective coverage of the genome. Fricano et al. (2012) identified 89 tobacco genotypes that captured the whole-genetic diversity at 49 SSR loci and evaluated LD using 422 SSR markers mapped on seven linkage groups. This study clearly indicated that LD in tobacco was dependent on the population structure and extended up to a distances of 75 cM with  $r^2 > 0.05$  or up to 1 cM with  $r^2 > 0.2$ .

### ***10.6.2 Target Gene Based LD Studies***

LD has been used locate QTLs or major genes, based on the co-segregation of specific marker alleles and traits in tobacco (Zhu et al. 2008; Rafalski 2010). Zhang et al. (2012) conducted association analysis and detected 18 sequence-related amplified polymorphism markers significantly associated with six agronomic traits in 258 flue-cured tobaccos. One SSR marker and six microsatellite-anchored fragment length polymorphism markers found to be associated with the levels of tobacco-specific nitrosamines (Yu et al. 2014). Twenty four SSR loci associated with aroma substances in tobacco (Ren et al. 2014) and one SSR locus from linkage group 13 with low chloride accumulation rate in 70 oriental-type tobaccos (Basirnia et al. 2014). Fan et al. (2015) performed a marker–trait association analysis and obtained 11 SSR markers associated with potassium content in tobacco; five among the 11 SSR markers were selected to validate the stability of the associated markers by scanning 130 other tobacco germplasms. Tong et al. (2020) made association analysis of leaf chemistry traits in natural populations using a large number of tobacco germplasms based on genome-wide SNPs.



### ***10.6.3 Genome Wide LD Studies***

Genome wide LD studies are not reported in tobacco as on date. Association mapping studies largely depends on population genetic structure. Based on the existing molecular diversity in germplasm collections, the population structure could be reconstructed in tobacco for association studies (Moon et al. 2009b). Ganesh et al. (2014) observed that 25 unlinked SSR markers delineated genetic structure of 135 FCV (flue-cured virginia) tobacco genotypes revealing a total of 85 alleles with an average of 3.4 alleles per locus.

### ***10.6.4 Future Potential for the Application of Association Studies for Germplasm Enhancement***

The population-based association studies utilizes the available broader genetic variations in wider background and detects marker-trait correlations. Hence, they can lead to construction of higher resolution maps with broader allele coverage because of the utilization of majority recombination events from a large number of meiosis throughout the germplasm development and exploits historically measured trait data without the development of expensive and tedious biparental populations saving time and costs involved (Abdurakhmonov and Abdugarimov 2008). LD based association studies provide an opportunity to dissect and exploit existing natural variations in tobacco germplasm resources for tobacco improvement. Availability of large collection of germplasm resources in tobacco over worldwide facilitate to detect neutrally inherited markers linked to genetic causatives or genes controlling the complex quantitative target traits including resistance to abiotic stresses.

## **10.7 Brief Account of Molecular Mapping of Resistance Genes and QTLs**

### ***10.7.1 Brief History of Mapping Efforts***

The identification of linkage between pollen color factor and the sterility factors in tobacco, initiated gene mapping research in tobacco (Anderson and Dorothea 1931; East 1932; Brieger 1935; Smith 1937). Establishing two types of monosomics, Clausen and Goodspeed (1926) demonstrated haplo-C (then called “corrugated”) is involved in the chromosome in which the basic color factor, Wh, is located. Later, the association in transmission between 24 monosomics developed by Clausen and Cameron (1944) led to location of 18 genes in nine chromosomes. Though genes regulating various traits and their linkages with other genes are identified, detailed

map based entirely on genes is not available in tobacco (Suen et al. 1997; Narayanan et al. 2003). Efforts were initiated during early nineties to map and tag resistant genes linked to various stresses with DNA markers such as RFLP, RAPD and AFLPs. Initially, RFLP and RAPD markers were used to map *Nicotiana* spp. (Lin et al. 2001). Later, RAPD, AFLPs and ISSR were used in construction of genetic maps (Lin et al. 2001; Nishi et al. 2003; Julio et al. 2006; Xiao et al. 2006). With the sighting of SSR markers in late 1990s, SSR based molecular map showing 24 linkage groups was developed in *N. tabacum* (Bindler et al. 2007). This SSR map was improved further with identification of a more number of SSRs (Bindler et al. 2011; Tong et al. 2012). With the identification of SNPs in recent years, high density SNP based tobacco genetic map has been developed with 24 linkage groups (Tong et al. 2020). Currently, maps are available for selected *Nicotiana* spp. FCV and burley tobacco types (see Sect. 10.7.5).

### 10.7.2 Evolution of Marker Types: RFLPs to SNPs

Variations existing among individuals for specific regions of DNA are deduced by molecular markers and hence serves as useful tools in mapping of genetic material. Molecular genetic markers, such as RFLP, RAPD, AFLP, microsatellites or SSRs, and SNPs have been used in genetic linkage mapping and QTL mapping in tobacco (Liu and Zhang 2008). In the initial stage, PCR-based RAPD markers were used by different researchers to map and tag resistant genes linked to abiotic stresses due to their relative ease in spite of the reproducibility issues. Although, reproducibility and sensitivity of AFLPs markers is higher, they were used in a limited degree in mapping of resistance genes due to their lengthy and laborious detection method, low reproducibility and non-suitability to automation. In late 1990s with the discovery of SSR markers, SSRs and EST-SSRs became markers of choice for mapping in tobacco (Bindler et al. 2007, 2011; Tong et al. 2012). Currently, more than 10,000 SSR markers are available in tobacco for their use in QTL/gene mapping studies (Bindler et al. 2007, 2011; Tong et al. 2012; Cai et al. 2015; Madhav et al. 2015). In addition, Wang et al. (2018) identified a huge number of about 1,200,000 non-redundant and novel NIX (*Nicotiana* multiple (X) genome) markers (SSRs) for use in tobacco.

With the advent of SNPs in recent past, mapping of tobacco genome using these markers proved an easy platform for mapping. Xiao et al. (2015) developed SNPs using two different methods (with and without a reference genome) based on restriction-set associated DNA sequencing (RAD-seq). Through whole-genome resequencing of 18 FCV tobacco genotypes, Thimmegowda et al. (2018) identified and positioned SNPs into linkage groups. Using *N. tabacum* (K326 cultivar) as a genome reference, Tong et al. (2020) identified and mapped 45,081 SNPs to 24 linkage groups on the tobacco genetic map. Adding advantage to the above markers others such as; sequence-specific amplification polymorphism (SSAP), sequence-related amplified polymorphism (SRAP), cleaved amplified polymorphic sequence

(CAPS) and diversity arrays technology (DarT) were also used in molecular mapping in tobacco.

### ***10.7.3 Mapping Populations Used***

Diverse populations viz.,  $F_2$  populations, doubled haploid (DH) lines, recombinant inbred lines (RILs),  $BC_1$  progenies,  $BC_1F_1$ ,  $BC_4F_3$  populations etc. have been used as the for molecular mapping in tobacco (Table 10.1). Majority of the maps developed were based on  $F_2$  and DH populations and other maps were developed based on next-generation sequencing (NGS) technologies. Practically, the population size need to be around 99–381 individuals in a mapping population for higher resolution and fine mapping.

### ***10.7.4 Mapping Software Used***

In developing molecular maps of tobacco, softwares such as Mapmaker program (Lander et al. 1987; Lin et al. 2001; Wu et al. 2010), JoinMap<sup>®</sup> 3.0 program (Van Ooijen and Voorrips 2001; Bindler et al. 2007, 2011), Map Manager QTXb20 (Manly et al. 2001; Bindler et al. 2011; <http://www.mapmanager.org>) and JoinMap 4.0 (Van Ooijen 2006; Lu et al. 2012; Tong et al. 2016) and LepMap3 software (Rastas 2017; Tong et al. 2020, 2021) were used. Among the softwares, JoinMap 3.0/4.0 program is the widely used one for construction of molecular maps in tobacco. Similarly, LepMap3 software was used in the building of maps using NGS data.

### ***10.7.5 Maps of Different Generations***

For studies of genetics, genomic structure, genomic evolution and for mapping essential traits, genetic linkage maps are vital. In case of tobacco, genetic map construction has lagged behind other *Solanaceae* crops such as the tomato, potato, and pepper plants (Barchi et al. 2007; Jacobs et al. 2004; Tanksley et al. 1992). Till the end of twentieth century, scanty data was available on genetic mapping and molecular marker development in tobacco (Suen et al. 1997). Construction of genetic linkage maps in tobacco was started at the beginning of twenty-first century (Lin et al. 2001). Various maps constructed are briefly discussed here.

**Table 10.1** Molecular linkage maps constructed in Nicotiana

S. No.	Population	Markers used	Linkage groups identified	Total length covered (cM)	References
<b>A Nicotiana species</b>					
1	99 F <sub>2</sub> plants from the cross <i>N. plumbaginifolia</i> × <i>N. longiflora</i>	69 RFLP and 102 RAPD	9	1,062	Lin et al. (2001)
2	<i>N. tomentosiformis</i>	489 SSRs/CAPS	12	~1,000 cM	Wu et al. (2010)
3	<i>N. acuminata</i>	308 SSRs/CAPS	12	~1,000 cM	Wu et al. (2010)
<b>B Burley</b>					
1	125 DH lines, derived from F <sub>1</sub> hybrids between <i>W6</i> and <i>Michinoku I</i>	117 AFLPs	10	383	Nishi et al. (2003)
2	92 DH population derived from Burley 37 and Burley 21	112 AFLP loci and 6 SRAP loci	22	1,953.6 cM	Cai et al. (2009)
<b>C Flue-cured</b>					
1	DH population from a cross between Speight G-28 and NC2326	11 ISSRs and 158 RAPDs	27	2,094.6	Xiao et al. (2006)
2	114 recombinant inbred lines	138 ISSR, AFLP and SSAP markers	18	707.6	Julio et al. (2006)
3	186 F <sub>2</sub> derived from Hicks Broad Leaf and Red Russian	293 SSRs	24	1,920	Bindler et al. (2007)
4	186 F <sub>2</sub> derived from cultivar, Hicks Broad Leaf and Red Russian	2318 SSRs	24	3,270	Bindler et al. (2011)
5	207 DH population derived from FCV cultivars, HHDJY and Hicks Broad Leaf	851 (238 DARt and 613 SSRs)	24	2,291	Lu et al. (2012)

(continued)

**Table 10.1** (continued)

S. No.	Population	Markers used	Linkage groups identified	Total length covered (cM)	References
6	Parents [HD (Honghua Dajinyuan) and RBST (Resistance to Black Shank Tobacco)] and their F <sub>1</sub> (HD × RBST) and 193 BC <sub>1</sub> progenies	Map 1: 4138 SNPs and Map 2: 2162 SNPs	24	Map 1: 1,944.74 (with reference genome); Map 2: 2,000.9 (de novo identification SNPs by RAD-seq)	Xiao et al. (2015)
7	Specific locus-amplified fragment sequencing	4215 SNPs and 194 SSRs	24	2,662.43	Gong et al. (2016)
8	213 BC <sub>1</sub> s derived from Y3 and K326	626 SSRs	24	1,120.45	Tong et al. (2016)
9	Parents (Y3 and K326), their F <sub>1</sub> and 271 RILs	45,081 SNPs	24	3,486.78 cM	Tong et al. (2020)
10	Parents (NC82 and TT8), their F <sub>1</sub> and 200 BC <sub>1</sub> F <sub>1</sub> s	13,273 SNPs	24	3,421.80	Cheng et al. (2019)
11	Parents (Y3 and K326), their F <sub>1</sub> and 381 BC <sub>4</sub> F <sub>3</sub> s	24,142 SNPs	24	2,885.36	Tong et al. (2021)
<b>D</b>	<b>Inter type</b>				
1	187 F <sub>2</sub> derived from the cross Taiyan 7 (FCV) and Bailei 21 (burley)	112 (92 SRAP and 10 ISSR)	26	1,560.2	Ma et al. (2008)

*Note* RFLP—Restriction fragment length polymorphisms; RAPD—Random amplified polymorphic DNA; AFLP—Amplified fragment length polymorphism; ISSR—Inter simple sequence repeat; SSAP—Sequence-specific amplification polymorphisms; SSR—Simple sequence repeats; SRAP—Sequence-related amplified polymorphism; CAPS—Cleaved Amplified Polymorphic Sequences; DArT—Diversity arrays technology; and SNP—single nucleotide polymorphisms; DH—Doubled haploids; RILs—Recombinant inbred lines; FCV—Flue-cured virginia

### 10.7.5.1 Mapping of *Nicotiana* Species

Lin et al. (2001) constructed a genetic linkage map based the F<sub>2</sub> plants (99 individuals) derived from tobacco wild species, *Nicotiana plumbaginifolia* × *N. longiflora*. This map covers 60 RFLP and 59 RAPD loci spread on nine major linkage groups measuring 1,062 cM. The tenth linkage group could not be identified due to unavailability of markers, corresponding to the haploid chromosome number of

*N. plumbaginifolia*. Wu et al. (2010) generated two maps for wild diploid *Nicotiana* species, *N. tomentosiformis* and *N. acuminata* with 12 linkage groups spanning ~1,000 cM. A combination of 489 SSR and cleaved amplified polymorphic sequence (CAPS) markers was used to construct *N. tomentosiformis* map constructed and, while the *N. acuminata* (closely related to *N. sylvestris*) map was generated with a mixture of 308 SSR and CAPS markers (Wu et al. 2010).

### 10.7.5.2 Mapping of Burley Tobacco

A genetic map of the burley tobacco was constructed using AFLP based on DH lines with 10 linkage groups was derived from F<sub>1</sub> hybrids between burley entries, W6 and Michinoku 1 (Nishi et al. 2003). The currently available high density burley linkage map was generated assembling 112 AFLP loci and six SRAP loci into 22 linkage groups (A1-A22) covering ~1,954 cM using a DH population derived from a cross between Burley 37 (high nicotine content) and Burley 21 (low nicotine content) (Cai et al. 2009).

### 10.7.5.3 Mapping of Flue-Cured Tobacco

The first linkage map of flue-cured tobacco based on a DH population was developed from a cross between Speight G-28 and NC2326 (Xiao et al. 2006) using 169 ISSR/RAPD molecular markers covering 27 linkage groups. While a molecular linkage map of flue-cured tobacco with 18 linkage groups covering 138 ISSR, AFLP and SSAP markers based on 114 flue-cured tobacco RILs was constructed by Julio et al. (2006).

Bindler et al. (2007, 2011) constructed an enriched SSR based linkage map with 2,318 SSR markers covering 24 linkage groups with a total length of 3,270 cM using an F<sub>2</sub> population from a cross between Hicks Broadleaf × Red Russian. This map is the most widely used map of tobacco and the average genetic distance between the markers was 1.4 cM. There still exists some gaps of about ~16 cM in this map, in spite of the high-density marker used (Fig. 10.1; Bindler et al. 2011). Tong et al. (2012) used double haploid (DH) lines derived from a cross between ‘Honghua Dajinyuan’ and ‘Hicks Broad Leaf’ with 207 individuals and constructed a genetic map of flue-cured tobacco consisting of 611 SSR loci distributed on 24 tentative linkage groups covering a total length of ~1,882 cM. Tong et al. (2016) constructed a genetic map entailing 626 SSR loci distributed across 24 linkage groups covering a total length of about 1,120 cM utilizing 213 backcross (BC1) individuals derived from an intra-type cross between two flue-cured tobacco varieties, Y3 and K326.

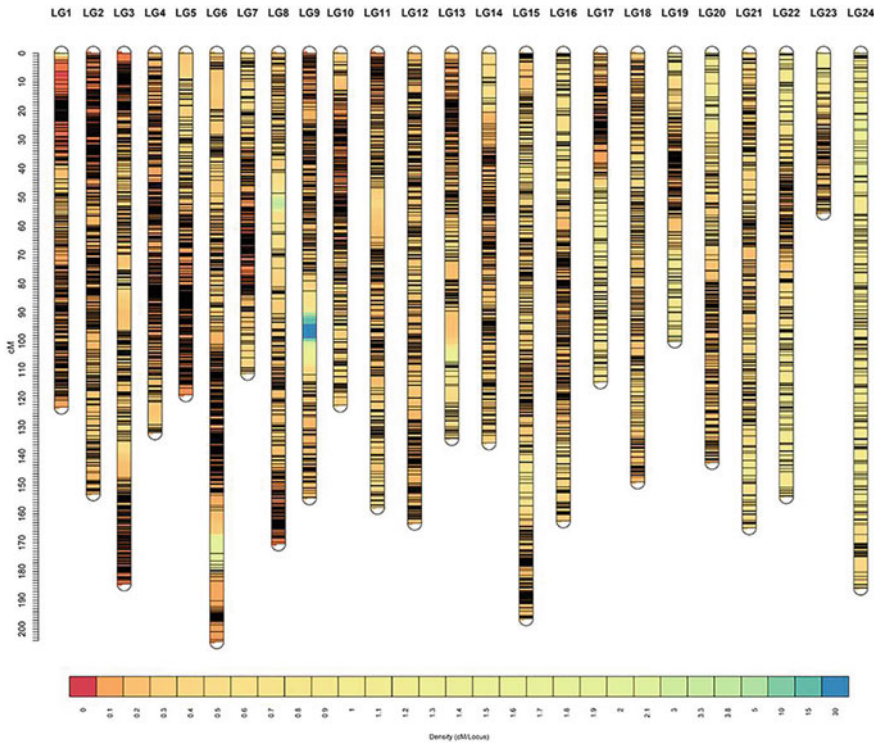
Xiao et al. (2015) constructed two linkage maps with a total of 2,162 and 4,138 SNP markers covering around 2,001 and 1,945 cM, into 24 different linkage groups based on reference genome and without reference, respectively. SOL Genome Network released SNP-based high density genetic map, *N. tabacum* 30k Infinium HD



**Fig. 10.1** Part SSR map (1–6 linkage groups) constructed by Bindler et al. (2011) with 2,318 microsatellite markers covering a total length of 3,270 cM

consensus map 2015 ([https://solgenomics.net/cview/map.pl?map\\_version\\_id=178](https://solgenomics.net/cview/map.pl?map_version_id=178)). With restriction site-associated DNA sequencing, Cheng et al. (2019) constructed a high-density SNP genetic map of flue-cured tobacco using restriction site-associated Illumina DNA sequencing. In this map, a total 13,273 SNP markers were mapped on 24 high-density tobacco genetic linkage groups spanning around 3,422 cM, with a mean distance of 0.26 cM between adjacent markers. Tong et al. (2020) identified a total of 45,081 SNP markers (with 7,038 bin markers) and characterized to construct a high-density SNP genetic map of flue-cured tobacco spanning a genetic distance





**Fig. 10.2** Linkage map constructed by Tong et al. (2020) covering a total length of 3486.78 cM

of 3,487 cM (Fig. 10.2). Tong et al. (2021) successfully constructed another high-density genetic map with 24,142 SNP markers using a  $BC_4F_3$  population derived from inbred of flue-cured tobacco lines Y3 (recurrent parent) and K326 (donor parent). This map included 4,895 bin markers with a genetic distance of  $\sim 2,886$  cM and an average genetic distance of 0.59 cM.

Lu et al. (2012) developed a high-density integrated linkage map (2,291 cM) of flue-cured tobacco that included 851 markers [238 diversity arrays technology (Dart) and 613 SSR] in 24 linkage groups. Gong et al. (2016) generated a high-density  $\sim 2,662$  cM length integrated genetic map of flue-cured tobacco containing 4,215 SNPs and 194 SSRs distributed on 24 linkage groups with an average distance of 0.60 cM between adjacent markers.

#### 10.7.5.4 Intra Type Genetic Maps

Ma et al. (2008) constructed an intra type genetic map of flue-cured and burley tobaccos, based on sequence related amplified polymorphisms (SRAPs) and ISSR markers, containing 26 linkage groups and 112 markers. Currently, the available

high-density maps in tobacco are constructed with SSR (Bindler et al. 2011) and SNP (Gong et al. 2016; Tong et al. 2020) markers as detailed above. The widely referred SSR map of Bindler et al. (2011) was constructed with 2,318 microsatellite markers covering a total length of 3,270 cM while the SNP map of Tong et al. (2020) covers ~3,487 cM with 45,081 SNPs. The combination of SNPs and genetic maps, if developed, helps in designing precise breeding strategies and genomic selection in tobacco. Diverse genetic maps existing at present constructed can be effectively utilized in mapping QTLs, positional cloning, comparative genomics analysis, marker-assisted breeding and genomic selection etc. It would be necessary to further build the genetic linkage maps of tobacco in different cultivating types for their effective utilization in breeding of those types.

### ***10.7.6 Enumeration of Mapping of Simply-Inherited Stress Related Traits***

The availability of *Nicotiana* genome sequences (Sierro et al. 2014; Edwards et al. 2017) information and high-density molecular maps in recent times is laying the foundation for trait discovery and fine mapping of trait of interest in tobacco (Yang et al. 2019). Tobacco plant cope with abiotic stresses through activation and regulation of specific stress-related genes. The genes involved in the whole-sequence of molecular responses to abiotic stresses include genes for signaling, transcriptional control, protection of membranes and proteins, and free-radical and toxic-compound scavenging (Wang et al. 2003; Xiang et al. 2016; Yang et al. 2016). Hence, resistance or tolerance to a specific stress is not controlled by a single gene. In general, abiotic stress are controlled by poly genes with low heritability and are influenced by environment. In view of these complications the studies showing mapping of abiotic stress tolerance in tobacco is scanty.

### ***10.7.7 Framework Maps and Markers for Mapping Resistance QTLs***

Framework maps constructed using SSR and SNP markers that were already identified and mapped to linkage groups in tobacco (Lu et al. 2012; Tong et al. 2020). High density SSR map of Bindler et al. (2011) and SNP map of Tong et al. (2020) can be the ideal ones for constructing framework maps while mapping various traits (Edwards et al. 2017). The SNP-based high density genetic map, *N. tabacum* 30 k Infinium HD consensus map 2015 can be one of the best resources for fine mapping any trait of interest ([https://solgenomics.net/cview/map.pl?map\\_version\\_id=178](https://solgenomics.net/cview/map.pl?map_version_id=178)).

In few cases for reliability and consistency RAPD markers are being converted into sequence characterized amplified region (SCAR) markers. SCAR markers can be

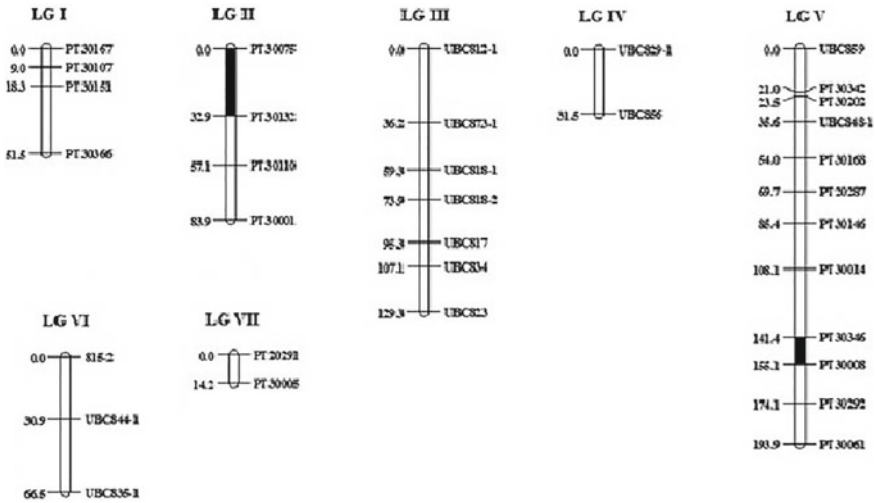
developed after sequencing the resultant RAPD bands and designing 18–25 base PCR primers that can specifically amplify the sequenced DNA segment. CAPS, conserved ortholog sequences (COS), random amplified microsatellite polymorphism (RAMP), ISSRs and target region amplification polymorphism (TRAP) are some of the other markers that can be used in map construction. CAPS markers developed are the primers designed on the known sequence of a gene of interest. COS primers used are universal primers based on sequence alignments of orthologs (genes that are conserved in sequence and copy number) from multiple solanaceous species. RAMP markers include SSR primers that amplify the genomic DNA in the presence or absence of RAPD primers. TRAPs are two PCR-based primers, one from target EST and the other is an arbitrary primer.

### ***10.7.8 QTL Mapping Software Used***

Mapmaker/Exp 3.0 is the widely used mapping software in mapping of various QTL traits in tobacco followed by various version of Join Map and Map Chart. Some of the other softwares used are Mapmaker/QTL, QTL IciMapping 4.1, QTL Network 2.1, R/QTL, AYMY-SS, Stat Graphics Plus 5.0 and QTL Cartographer V 2.5 (Sarala et al. 2021).

### ***10.7.9 Details on Trait Wise QTLs***

Biparental mapping populations have been used to identify QTLs controlling chloride accumulation rate in tobacco genome (Li-Hua et al. 2011; Hatami et al. 2013). Using family-based linkage mapping approach, Li-Hua et al. (2011) detected two QTLs for total chlorine concentration. Hatami et al. (2013) developed a genetic linkage map for oriental-type tobacco by using F<sub>2</sub> generation individuals from the cross between two divergent oriental-type tobacco genotypes viz., ‘Basma Series 31’ and ‘SPT 406’. Through single marker analysis of the above F<sub>2</sub> population, SSR marker PT30346 was found to be significantly associated with chloride accumulation. They have further identified two QTLs for chloride accumulation in the leaf, namely, Chl<sub>IM</sub> (on LG II) and Chl<sub>CIM</sub> (on LG V) with R<sup>2</sup> values (phenotypic variance) of 0.4 and 0.07, using interval mapping (IM) and composite interval mapping (CIM), respectively (Fig. 10.3). A LOD score of 2.3 was used for identifying these significant QTLs that were estimated to originate from maternal parent, Basma Series 31.



**Fig. 10.3** Linkage map of SSR and ISSR markers in oriental tobacco showing the location of QTLs affecting chlorine accumulation. Bars represent intervals associated with the QTLs (Hatami et al. 2013)

### 10.8 Marker-Assisted Breeding for Resistance Traits

Marker-assisted breeding (MAB) refers to the breeding procedure in which DNA marker detection and selection are integrated into a traditional breeding programs. The status of MAB and its prospects are discussed here. The advantage of using DNA markers is that they can be detected at any stage of plant growth in contrast to that of classical markers detection of which is usually limited to certain growth stages. Polymorphism for DNA markers is available throughout the genome, and their presence or absence is not affected by environments and usually do not directly affect the phenotype. If the markers are located in close proximity to the target gene or present within the gene, selection of the markers will ensure the success in selection of the gene. Thus, DNA markers are the major types of genetic markers for MAB (Jiang 2013). Most commonly used molecular markers in tobacco include RAPD, AFLP, SSR, SCAR, CAPS, dCAPS (derived CAPS), and KASP (Yang et al. 2019). Different DNA markers have own advantages and disadvantages for specific purposes. Comparatively, SSRs have most of the desirable features and availability of large number of SSRs make them markers of choice in tobacco. SNPs require more detailed knowledge of the specific, single nucleotide DNA changes responsible for genetic variation among individuals. Fairly large number of SNPs have become available in tobacco making them important choice of markers for MAB in tobacco.

### ***10.8.1 Germplasm Characterization and DUS***

Initiation of marker-assisted breeding program requires the identification of closely linked markers for target traits in germplasm. These tightly linked markers can be utilized in MAB while screening parents, F<sub>1</sub> and other segregating materials for selecting plants with target traits. Gene mapping, QTL analysis, association mapping, classical mutant analysis, linkage or recombination analysis, bulked segregant analysis, etc. provides information on marker trait associations in germplasm lines and mapping populations. It is also important to identify the linkage situation i.e. cis/trans (coupling or repulsion) linkage with the desired allele of the trait.

Large number of studies for screening germplasm and sources of resistance are made in tobacco to identify closely linked markers to various abiotic stresses for their introgression into cultivated varieties. QTL analysis and genetic mapping, through biparental or association mapping (AM) populations, have accelerated the dissection of genetic control of stress resistant traits in tobacco. Tightly linked markers are important tools for DUS characterization of varieties also. This information has the potential to make MAS a successful option for tobacco improvement.

### ***10.8.2 Marker-Assisted Gene Introgression***

Marker-assisted back crossing (MABC) is the simplest form of resistant trait introgression that is most widely and successfully used in transferring abiotic stress resistant genes/QTLs into elite cultivars. MABC aims to transfer one or a few genes/QTLs for resistance from one genetic source (donor parent) into a superior cultivar or elite breeding line (recurrent parent) to improve the stress resistance. Unlike traditional backcrossing, MABC is based on the alleles of markers associated with or linked to gene(s)/QTL(s) of interest instead of phenotypic performance of target trait. MABC program with foreground selection for the marker allele(s) of donor parent at the target locus (ex. resistance) ensures the transfer of target trait from donor parent. Background selection for the marker alleles of recurrent parent in all genomic regions of desirable traits (agronomic traits) except the target locus will takes care of the genome recovery of recurrent parent while effectively transferring of resistance into elite genotypes (Hospital 2003). MAS can be used when other characters are to be combined from two parents along with resistance trait. However, MAS will be highly effective for simply inherited character controlled by a few genes than for a complex character governed by large number of genes.

### 10.8.3 Gene Pyramiding

Pyramiding of several genes/QTLs can be achieved through multiple-parent crossing or complex crossing, multiple backcrossing, and recurrent selection. A number of factors are involved in deciding suitable breeding scheme for marker-assisted gene pyramiding such as the number of parents that contain the desired genes/QTLs, number of genes/QTLs required to be transferred, the heritability of traits of interest, and other factors (e.g. marker-gene association and genotyping costs). Pyramiding of three or four desired genes/QTLs exist in three or four different lines can be realized by three-way, four-way or double crossing. They can also be brought together by convergent backcrossing or stepwise backcrossing. Complex or multiple crossing and/or recurrent selection may often be preferred for pyramiding of more than four genes/QTLs.

Gene pyramiding can be achieved through three different strategies or breeding schemes namely stepwise, simultaneous/synchronized and convergent backcrossing or transfer. In the stepwise backcrossing, the target genes/QTLs are transferred from donor parents one after the other into the recurrent parent (RP). Gene pyramiding through stepwise backcrossing is easier and more precise to adopt as it transfer only one gene/QTL at a time requiring small population size and lower genotyping cost. But this method takes a longer time to complete. In the simultaneous (synchronized) backcrossing, the recurrent parent is initially crossed to each of the donor parents and the resultant single-cross  $F_1$ s are crossed with each other to produce two double-cross  $F_1$ s and then the two double-cross  $F_1$ s are crossed to produce a hybrid integrating all the target genes/QTLs in heterozygous state. This hybrid and its progeny with heterozygous markers for the target genes/QTLs is subsequently back crossed to the RP until the satisfactory recovery of the RP genome. Finally, the RP genome recovered line is selfed to achieve homozygosity. Simultaneous or synchronized backcrossing requires a large population and more number of genotypings as all target genes/QTLs are handled at the same time and takes shorter time to transfer multiple genes. Both stepwise and synchronized backcrossing strategies are employed in convergent backcrossing. First each of the target gene/QTLs from the donors are transferred separately into the RP through single crossing followed by backcrossing based on the linked markers to produce improved lines. The improved lines are crossed with each other and the consequential hybrids are intercrossed to pyramid all the genes/QTLs into the final improved line. Convergent backcrossing require less time (compared to stepwise transfer) and easily fix and pyramid genes (compared to simultaneous transfer).

If all the parents are improved cultivars with complementary genes or favorable alleles for the traits of interest, marker-assisted complex or convergent crossing (MACC) can be undertaken to pyramid multiple genes/QTLs. In this method, the hybrid of convergent crossing is selfed and MAS for target traits is performed for several consecutive generations until genetically stable lines with desired marker alleles and traits are achieved. For reducing the population size and avoiding the loss of most important genes/QTLs, the most important genes/QTLs can be detected and

selected first in early generations and less important markers later. Theoretically, it is possible to apply MABC and MACC for pyramiding target genes/QTLs in tobacco crop. Currently information is not available about the release of commercial varieties developed using these strategies.

#### ***10.8.4 Limitations and Prospects of MAS and MABC***

MAS and marker-assisted back cross (MABC) breeding may not be universally useful in spite of their advantages (Jiang 2013). Quick DNA extraction techniques and a high throughput marker detection system are essential to handle large number of samples and large-scale screening of multiple markers. Development of suitable bioinformatics and statistical software packages are required for efficient and quick labeling, storing, retrieving, processing and analyzing large data set requirements, and even for integrating data sets available from other programs. Hence, the startup expenses and labor costs involved in MAS and MABC are higher compared to conventional techniques making them not in the reach of all the researchers (Morris et al. 2003).

When the distance between the marker and the gene of interest is higher, the chance of recombination between gene and marker increases there-by make the selection of resistant plants based on marker ineffective due to false positives. Use of flanking markers on either side of the locus of interest will increase the probability that the desired gene is selected. Sometimes markers that were used to detect a locus may not be 'breeder-friendly'. Such markers viz., RFLP and RAPD may need to be converted into more reliable and easier to use markers. RFLP markers may be converted to STS (sequence tagged site) for detection via PCR protocols (Ribaut and Hoisington 1998) and RAPD markers into SCAR markers for reliable and repeatable amplifications (Lewis 2005; Milla et al. 2005). Inaccurate estimates of locations and effects of QTLs may result in slower progress than expected through MAS (Beavis 1998). Yet times, markers developed for MAS in one population may not be suitable for screening other populations, either due to lack of marker polymorphism or the absence of a marker-trait association.

With the increasing utilization of molecular markers in various fields viz., germplasm evaluation, genetic mapping, map-based gene discovery, characterization of traits etc., MAB is going to become a powerful and reliable tool in genetic manipulation of important traits in tobacco. Availability of high density linkage maps in tobacco provides a framework for identifying marker-trait associations and selecting markers for MAB. Markers linked to resistant traits can fruitfully be utilized in MAB in tobacco. Only the markers that are closely associated with the target traits or tightly linked to the gene of interest can provide sufficient guarantee for the success in practical breeding. Availability of new high-throughput marker genotyping platforms for the detection of SSR and SNP markers along with the sequencing information of cultivated and wild relatives of *Nicotiana* is going to have a great impact on discovering marker trait associations that can be used for MAS in the future. Array-based



methods such as DArT (Lu et al. 2012) and single feature polymorphism (SFP) detection (Rostoks et al. 2005) offer low-cost marker technologies for whole-genome scans in tobacco. Rapid growth in genomics research and huge data generated from functional genomics in tobacco in the recent years is leading to the identification of many candidate genes for numerous traits including abiotic stress resistance. SNPs within candidate genes could be extremely useful for ‘association mapping’ and circumvents the requirement for constructing linkage maps and performing QTL analysis for new genotypes that have not been previously mapped. The availability of large numbers of publicly available markers and the parallel development of user-friendly databases (Sol genome network, NCBI etc.) for the storage of marker and QTL data, increasing number of studies on genes and marker trait associations will undoubtedly encourage the more widespread use of MAS in tobacco.

Closely linked markers allows the selection of disease/pest resistance traits even without the incidence of pests and diseases. MAS based on markers tightly linked to the multiple genes/QTLs for traits of interest can be more effective in pyramiding desirable genes than conventional breeding. Selection for all kinds of traits at seedling stage in MAB helps to minimize the costs as undesirable genotypes are eliminated at early stages. Use of co-dominant markers (e.g. SSR and SNP) in MAB allow effective selection of recessive alleles in the heterozygous state without selfing or test crossing, thus saves time and accelerate breeding progress. As more and more newer techniques are available genotypic assays based on molecular markers may be faster, cheaper and more accurate than conventional phenotypic assays and thus MAB may result in higher effectiveness and higher efficiency in terms of time, resources and efforts saved in future.

MAB has brought great challenges, opportunities and prospects for breeding crops including tobacco. As a new member of the whole family of plant breeding, MAB cannot replace conventional breeding, but can be a supplementary addition to conventional breeding. Higher costs and technical demands of MAB will continue to be an obstacle for its large-scale use, especially in the developing countries (Collard and Mackill 2008). Integration of MAB into conventional breeding programs will be an optimistic strategy for tobacco improvement in the future.

## 10.9 Map-Based Cloning of Resistance Genes

### 10.9.1 Traits and Genes

Identification and subsequent mapping of interesting mutants became difficult in view of the high levels of redundancy between genes in the large and complex genome of tobacco with the absence of molecular markers and genomic resources till recent years. Having anchored 64% of the genome assembly to chromosomal locations in recent years, a possibility now exists for map-based cloning of abiotic stress resistant genes (Edwards et al. 2017). For the first time, successful map-based cloning in

tobacco was done by Edwards et al. (2017) for NtEGY1 and NtEGY2 homeologous candidate genes for YB1 and YB2 loci conferring white stem phenotype in recessive condition in burley tobacco.

### ***10.9.2 Strategies: Chromosome Landing and Walking***

Currently available high density genetic maps, genome sequences and bacterial artificial chromosome (BAC) clones are paving the way for map based cloning of resistance genes in tobacco. In general, chromosome landing and walking strategies are used in identification of clones carrying gene of interest for map based cloning. However, in the only reported case of map based cloning in tobacco, Edwards et al. (2017) used a specific technique to clone genotyped pairs of near-isogenic lines (NILs) carrying dominant or recessive alleles of the *YB1* and *YB2* loci (cultivars SC58, NC95, and Coker 1) with a custom 30 K Infinium iSelect HD Bead Chip SNP chip (Illumina Inc., San Diego, CA) that was used in developing a high density genetic map (*N. tabacum* 30 k Infinium HD consensus map 2015; [https://solgenomics.net/cview/map.pl?map\\_version\\_id=178](https://solgenomics.net/cview/map.pl?map_version_id=178)). Genomic regions comprising SNP polymorphisms distinguishing the nearly isogenic lines were identified and SNP markers closely linked to the loci were aligned to the genome assembly and predicted potential candidate genes. Coding regions of candidate genes were then amplified, using the primers specifically designed, from first-strand cDNA from tobacco cultivars K326 and TN90. The fragments, thus amplified were finally cloned into a vector.

### ***10.9.3 Libraries***

Physical mapping, comparative genome analysis, molecular cytogenetics etc. requires the availability of high-capacity libraries. Such resources are also powerful tools for large-scale gene discovery, elucidation of gene function and regulation, and map-based cloning of target trait loci or genes associated with important agronomic and resistant traits and use in crop improvement programs. BAC libraries are the large DNA insert libraries (inserts of DNA up to 200,000 base pairs) of choice for genomics research. Cloning of larger DNA segments (more than 1000 kb) are possible with yeast artificial chromosome (YAC) libraries and greatly facilitates chromosome walking and physical mapping around the target locus. While, transformation-competent artificial chromosome (TAC) libraries make it possible to clone and transfer genes efficiently into plants. In recent years, BAC libraries are constructed and utilized in tobacco for genome sequencing, mapping and comparative genome analysis. Reports currently not available regarding the construction of YAC and TAC libraries in tobacco.

Tobacco Genome Initiative (TGI) generated a BAC library (9.7-fold genome coverage) for assembling the partial genome of Hicks Broad leaf variety and used

425,088 BAC clone library for construction of physical map and ancestral annotation of tobacco cultivar, Hicks Broadleaf (Opperman et al. 2003; Rushton et al. 2008; Sierro et al. 2013b). Edwards et al. (2017) constructed two libraries having 150,528 BACs from K 326 variety using HindIII or EcoRI, with average insert sizes of 115 kb and 135 kb, respectively (representing  $\sim 8 \times$  coverage of the genome) for generating a whole-genome profile (WGP) map. Jingjing (2018) reported a tobacco genome sequence of the HongDa cultivar using the combination of BAC-to-BAC libraries and whole-genome shotgun technologies. Yuhe (2012) constructed a BAC library of wild tobacco, *N. tomentosiformis* (one of the parent of *N. tabacum*) with the average DNA inserted size of 110 kb.

#### **10.9.4 Test for Expression**

The function of a target gene can successfully be validated through the transformation of a cloned gene into mutant plant and looking for wild phenotype rescue. As on date, mutant complementation studies with cloned genes are not reported in tobacco in view of the absence of map based cloning of functional genes in general and abiotic stress resistant genes particular.

### **10.10 Genomics-Aided Breeding for Resistance Traits**

#### **10.10.1 Structural and Functional Genomic Resources Developed**

Structural genomics deals with the structure of the genome and the knowledge on genomic structure is essential for gene tagging, identification and cloning of novel genes for further genomic assisted breeding. The genome sequencing information of 12 *Nicotiana* species (*N. tabacum*, *N. rustica*, *N. attenuata*, *N. benthamiana*, *N. knightiana*, *N. obtusifolia*, *N. otophora*, *N. paniculata*, *N. sylvestris*, *N. tomentosiformis*, *N. undulata* and *N. glauca*) are available at NCBI website (<https://www.ncbi.nlm.nih.gov/>) and Sol Genome Network (SGN) (Asaf et al. 2016).

Advances in transcriptomic analysis and functional genomics in tobacco led to development of large data sets and tools. A data base of well characterized 2,513 transcription factors (TFs) was developed in tobacco using a dataset of 1,159,022 gene-space sequence reads (Rushton et al. 2008). Further, the transcriptional activity for thousands of tobacco genes in different tissues expression microarray from a set of over 40 k unigenes and gene expression in 19 different tobacco samples has been generated (Edwards et al. 2010). 772 transcription factors previously identified in tobacco were mapped to the array and 87% of them being expressed in at least one tissue in the generated tobacco expression atlas (TobEA). Putative transcriptional

networks were identified based on the co-expression of transcription factors. SGN contains transcriptome sequence collections of *N. sylvestris* (32,276), *N. tomentosiformis* (31,961) and *N. tabacum* (26,284) from transcriptome projects and unigenes data sets of *N. sylvestris* (6,300), *N. tabacum* (84,602) and *N. benthamiana* (16,024).

Large collection of data on nucleotides, genes and protein sequences on *Nicotiana* are available at NCBI site. More than 3 million nucleotide sequences of 20 *Nicotiana* spp. that includes genomic DNA/RNA, mRNA, cRNA, ncRNA, rRNA, tRNA and transcribed RNA are generated by various researchers as on 30.09.2021. Among them, around 895,700 sequences are comprehensive, integrated, non-redundant, well-annotated set of reference sequences including 456,507 ESTs and 1,420,639 genomic survey sequences (GSS). Further, a total of over 201,560 records of gene sequences belonging to 12 *Nicotiana* spp. viz., *N. tabacum*, *N. tomentosiformis*, *N. sylvestris*, *N. attenuata*, *N. undulate*, *N. otophora*, *N. suaveolens*, *N. glauca*, *N. stocktonii*, *N. repanda*, *N. amplexicaulis* and *N. debneyi* are available at NCBI website.

Sequence read archive (SRA) data, available through multiple cloud providers and NCBI servers, is the largest publicly available repository of high throughput sequencing data. Nearly, 5,080 records of SRA data of 20 *Nicotiana* spp. are available at NCBI website (as on 30.09.2021). Raw sequencing data and alignment information in SRA are helpful in improving the reproducibility and facilitation of new discoveries through data analysis (<https://www.ncbi.nlm.nih.gov/sra>). Around 4,860 curated gene expression data sets as well as original series and platform records of 11 *Nicotiana* spp. are available at gene expression omnibus (GEO) repository of NCBI as on 30.09.2021 (<https://www.ncbi.nlm.nih.gov/gds>). Around 275,000 collection of protein sequences from several sources, including translations from annotated coding regions in GenBank, RefSeq and third party annotation (TPA) Sequence, as well as records from other data bases are available for 20 *Nicotiana* spp. at NCBI.

Genomic resource collection of SGN includes transcriptome sequences, mRNAs, and predicted proteins of five wild *Nicotiana* spp. namely *N. attenuata*, *N. benthamiana*, *N. tomentosiformis*, *N. sylvestris* and *N. otophora* and four versions of *N. tabacum* are available. SGN also hosts 39 transcript libraries of *N. tabacum* and two of *N. sylvestris*. The proteomic data generated globally is stored and can be accessed through the Universal Protein Resource (<https://www.uniprot.org/>). UniProt database contains 73,606 protein entries associated with *Nicotiana tabacum* proteome (UP000084051) as on 31.03.2021.

## 10.10.2 Details of Genome Sequencing

### 10.10.2.1 Nuclear Genome Sequencing

Tobacco has the largest genome size (4.5 Gbp) with large proportion of repetitive sequences compared to other solanaceous crops (tomato, potato, chilli and brinjal) in spite of similar basic chromosome number ( $n = 12$ ) (Zimmerman and Goldberg 1977;

Arumuganathan and Earle 1991; Kenton et al. 1993; Leitch et al. 2008; Sierro et al. 2014). Even it is 50% larger than that of human genome. In 2003, the first tobacco genome sequencing project was initiated through Tobacco Genome Initiative (TGI) with a purpose of sequencing open reading frames of *N. tabacum* cv. Hicks Broadleaf using methyl filtration method of complex reduction (Opperman et al. 2003; Rushton et al. 2008; Sierro et al. 2013b). The sequencing completed in 2007 and sequencing data are available at NCBI Gene Bank (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA29349>). However, the genome coverage (689 Mb) was limited because of employment of only enriched genes and portion of genes that are under methylated relative to TEs (Wang and Bennetzen 2015). Later with the advancement of next-generation sequencing (NGS) technologies, sequencing of entire genomes of three cultivated *N. tabacum* as well as 11 wild relatives (*N. knightiana*, *N. paniculata*, *N. rustica*, *N. glauca*, *N. obtusifolia*, *N. otophora*, *N. attenuata*, *N. sylvestris*, *N. tomentosiformis*, *N. undulata* and *N. benthamiana*) have been completed since 2014.

In 2013, the first sequences were released with assemblies of *N. sylvestris* and *N. tomentosiformis* having  $94.0 \times$  and  $146.0 \times$  genome coverage and a length of 2,222 and 1,688 Mb, respectively (Sierro et al. 2013a). Further, Philip Morris International released three genomes of *N. tabacum*, one at Scaffold level (cv. TN90) and two at contig (cv. K326 and Basma Xanthi-BX) level in 2014 (Seirro et al. 2014). The assembled sequences consist of 3,700 Mb with a GC content of 39%, covering 29-49X using TN 90 as a reference genome.

In 2017, the British American Tobacco released the scaffold level improved version of *N. tabacum* cv. K 326 with a genome coverage of 4,600 Mb of 86X and 33.5% GC content, an increase of 3.6 Gb (i.e. 81% of predicted genome size) compared to previous version (Seirro et al. 2014). Presently, 16 assemblies belonging to 12 *Nicotiana* species (*N. sylvestris*, *N. tomentosiformis*, *N. benthamiana*, *N. tabacum*, *N. otophora*, *N. attenuata*, *N. obtusifolia*, *N. glauca*, *N. knightiana*, *N. paniculata*, *N. rustica* and *N. undulata*) are available at NCBI genebank with range of 18-146X. Two assemblies of *N. attenuata* and four for *N. tabacum*. *N. attenuata* reference sequence (2,366 Mb) at chromosome level (12 haploid) while at contig level *N. tabacum* (K326, and Basma Xanthi) and *N. benthamiana* are available. The detail statistics including assembly level and their N50 and L50 values for each genome is also provided at Table 10.2.

Genomic resource collections consisting of five wild *Nicotiana* spp. and four of *N. tabacum* are available at Sol Genome Network (SGN). Under the TGI project, filtered genome sequences generated include contig level using *N. tabacum* cv. BX, *N. tabacum* cv. K326 and *N. tabacum* cv. TN90 as well as improved K326 assemblies of genome scaffolds, proteins and cDNA (Edwards et al. 2017). Further, scaffold level genome assemblies of five *Nicotiana* species (*N. attenuata*, *N. benthamiana*, *N. tomentosiformis*, *N. sylvestris* and *N. otophora*) also information on predicted proteins and mRNA are available at this database.

**Table 10.2** Genome sequencing details of *Nicotiana* species as (available at NCBI site)

Genome	Assembly level	Statistics	Genome coverage
<i>N. sylvestris</i>	Scaffold	Total length (Mb): 2,222 GC%: 39.80	94.0x
<i>N. tomentosiformis</i>	Scaffold	Total length (Mb): 1,688 GC%: 39.0999	146.0x
<i>N. benthamiana</i>	Contig	Total length (Mb): 62 GC%: 39.6	79x
<i>N. tabacum</i> cv. TN90 (RefSeq Genome)	Scaffold	Total length (Mb): 3,643 GC%: 39.2006	49.0x
<i>N. tabacum</i> cv. K326 (sequencing and assembly)	Contig	Total length (Mb): 3,733 GC%: 39	38.0x
<i>N. tabacum</i> cv.: BX (sequencing and assembly)	Contig	Total length (Mb): 3,736 GC%: 39	29.0x
<i>N. otophora</i>	Contig	Total length (Mb): 2,689 GC%: 39.1	66.0x
<i>N. attenuata</i> strain: UT (RefSeq, sequencing and assembly)	Chromosome (12 haploid)	Total length (Mb): 2,366 GC%: 41.333	30.0x
<i>N. obtusifolia</i>	Scaffold	Total length (Mb): 1,223 GC%: 38.4	50.0x
<i>N. attenuata</i> ecotype Arizona (sequencing and assembly)	Scaffold	Total length (Mb): 1,828 GC%: 29.1	18.0x
<i>N. tabacum</i> cv. K326 (sequencing and assembly)	Scaffold	Total length (Mb): 4,647 GC%: 33.5	86x
<i>N. glauca</i>	Scaffold	Total length (Mb): 3,223 GC%: 36.6	35.0x
<i>N. knightiana</i>	Scaffold	Total length (Mb): 2,299 GC%: 39.7	82.0x
<i>N. paniculata</i>	Scaffold	Total length (Mb): 2,191 GC%: 39.4	100.0x
<i>N. rustica</i>	Scaffold	Total length (Mb): 4,231 GC%: 40	90.0x

(continued)

**Table 10.2** (continued)

Genome	Assembly level	Statistics	Genome coverage
<i>N. undulata</i>	Scaffold	Total length (Mb): 1,914 GC%: 40	112.0x

### 10.10.2.2 Organelle Genome Sequencing

Chloroplast and mitochondrial genomes of tobacco are circular DNA molecules. The size of the plastid genome of *Nicotiana* species is 0.156 Mb. Shinozaki et al. (1986) for the first time sequenced the chloroplast genome of tobacco in 1986. Presently, sequencing of about 219 plastid genomes have been completed including 12 Popset data (DNA sequences derived from population, phylogenetic, mutation and ecosystem studies) related to five *Nicotiana* species. In addition, recently, Mehmood et al. (2020) assembled the plastid genomes and compared the five tobacco species namely *N. knightiana* (155,968 bp), *N. rustica* (155,849 bp), *N. paniculata* (155,689 bp), *N. obtusifolia* (156,022 bp) and *N. glauca* (155,917 bp). Reference plastid genomes of five *Nicotiana* species namely *N. tabacum* (155,943 bp), *N. attenuata* (155,886 bp), *N. tomentosiformis* (155,745 bp), *N. sylvestris* (155,941 bp), and *N. otophora* (156,073 bp) are available at NCBI (Table 10.2).

Starting from 2003, till date sequencing of eight mitochondrial complete genomes consist of five popsets of three *Nicotiana* species namely *N. tabacum* (430,597 bp), *N. attenuata* (394,341 bp) and *N. sylvestris* (430,597 bp) are completed and the details are available at NCBI site (Table 10.3). Further, reference mitochondrial genomes are made available for *N. tabacum*, *N. attenuata* and *N. sylvestris* at the NCBI.

**Table 10.3** *Nicotiana* organelle genome details available at NCBI (as on 31.03.2021)

Organelle reference genome	PopSet	Archived genomes	Sequence length (bp)	Genes	Proteins
<i>Plastid</i>					
<i>N. tabacum (RefSeq)</i>	10	124	155,943	144	98
<i>N. attenuata</i>	1	15	155,886	129	84
<i>N. tomentosiformis</i>	–	30	155,745	150	102
<i>N. sylvestris</i>	1	42	155,941	149	101
<i>N. otophora</i>	–	8	156,073	155	108
<i>Mitochondrial</i>					
<i>N. tabacum</i>	4	6	430,597	183	153
<i>N. attenuata</i>		1	394,341	68	40
<i>N. sylvestris</i>	1	1	430,597	64	37



### 10.10.3 Gene Annotation

Genome annotation aims at identifying functional elements along the sequence of a genome. Once a genome is sequenced, it must be annotated to understand its functions for its further successful use in genetic manipulation. In tobacco, both nuclear and organelle genomes are successfully annotated. Gene annotation records of *Nicotiana* sp data available at NCBI and SGN databases are presented in Tables 10.4 and 10.5.

*N. sylvestris* (2014) was the first species annotated followed by *N. tabacum* cv. TN 90 (2016), *N. attenuata* strain UT (2016) and *N. tomentosiformis* (2020). Presently, annotation reports of four *Nicotiana* species viz., *N. sylvestris*, *N. tabacum* cv. TN 90, *N. attenuata* strain UT and *N. tomentosiformis* are available at NCBI website (Table 10.4).

Predictions of genes annotated are made available for published genomes of *N. attenuata*, *N. benthamiana*, *N. tomentosiformis*, *N. sylvestris* and four versions of *N. tabacum* at SGN site also (Table 10.5). For *N. tabacum*, predicted proteins were ranging from 69,500 to 122,388 and mRNA from 145,503 to 189,413 in *N. tabacum*.

**Table 10.4** Gene annotation reports of *Nicotiana* species (as per NCBI)

Species	<i>N. tabacum</i>	<i>N. sylvestris</i>	<i>N. tomentosiformis</i>	<i>N. attenuata</i>
Genome assembly name	Ntab-TN90	Nsyl	Ntom_v01	NIATTr2
Annotation release number	100	100	102	100
Release date	04-05-2016	22-10-2014	21-04-2020	06-12-2016
RefSeq scaffolds	168,245	253,917	159,547	37,135
mRNAs	84,001	48,059	50,010	44,491
Genes and pseudogenes	73,946	40,317	45,485	39,977
– protein-coding	61,526	33,678	31,842	34,094
– non-coding	9,019	4,667	11,549	3,886
– pseudogenes	3,401	1,972	2094	1,997
– genes with variants	14,549	9,033	10,545	6,946

**Table 10.5** Gene annotation details of tobacco species (as per SGN)

<i>Nicotiana</i> species	Reference scaffolds	mRNA	Predicted proteins
<i>N. tabacum</i> cv. K326	382,373	145,503	85,994
<i>N. tabacum</i> cv. BX	420,216	146,748	86,009
<i>N. tabacum</i> cv. TN90	249,104	189,413	122,388
<i>N. tabacum</i> v4.5 (Edwards et al. 2017)	1,084,432	–	69,500
<i>N. tomentosiformis</i>	159,548	85,853	53,753
<i>N. sylvestris</i>	253,917	87,234	54,497
<i>N. benthamiana</i>	56,094	–	57,140
<i>N. attenuata</i>	37,182	33,449	33,449

However, a smaller number of proteins (33,449–54,497) and mRNA (33,449–87,234) were predicted for *Nicotiana* species compared to *N. tabacum*.

Uniprot database of *N. tabacum* (UP000084051) contains 73,606 protein entries (<https://www.uniprot.org/>) as on 31.03.2021. Edwards et al. (2017) identified predicted proteins exhibiting cross-over with the related Solanaceae species like tomato and potato in addition to other flowering plants based on gene ontology analysis.

Annotations are also available for published organelle genome sequences (Table 10.3). Predicted plastid genes in five *Nicotiana* spp. vary from 129 to 155 and proteins from 84 to 108. The predicted genes and proteins for mitochondrial genome are more for cultivated species, *N. tabacum* (183; 153, respectively) compared to wild species, *N. attenuata* (68; 40) and *N. sylvestris* (64; 37).

#### **10.10.4 Impact on Germplasm Characterization and Gene Discovery**

Availability of sequencing information of *Nicotiana* Species made it possible to compare sequences within *Nicotiana* Species and also with other solanaceous crops. Thus, the comparative genomics helped to understand the relationships between cultivated and wild species and their progenitors in terms of their sequence similarity and genome rearrangements (Wu et al. 2009; Siervo et al. 2014; Gong et al. 2016; Asaf et al. 2016; Edwards et al. 2017). Homologous genes were identified between the genomes of *N. tabacum* cv. TN90, K326 and BX and other solanaceous crops like tomato and potato (Siervo et al. 2014).

Gene annotation using published genome sequences has helped in identification of functional sequences and predicted mRNAs and proteins that can be expressed in tobacco. Siervo et al. (2014) recognized the genome assemblies and genomic regions responsible for virus resistance in draft genomes. The *N. glutinosa* *N* gene, source of TMV resistance in tobacco was found in the draft genome sequence of TN 90 cultivar and weak identity in susceptible genome of K326 and BX genome. Thus, genes responsible for abiotic stresses can also be identified using genome sequence information.

Using the available data on sequences of *Nicotiana* genomes and EST, large number of SSRs and SNPs were identified (Bindler et al. 2011; Tong et al. 2012, 2020; Cai et al. 2015; Xiao et al. 2015; Thimmegowda et al. 2018; Wang et al. 2018). The potential markers have been extensively used for characterization of germplasm and DUS testing for assessment of genetic diversity and genetic relatedness among cultivated varieties (Moon et al. 2008, 2009a, b; Davalieva et al. 2010; Fricano et al. 2012; Gholizadeh et al. 2012; Prabhakararao et al. 2012; Xiang et al. 2017; Binbin et al. 2020).

Core markers developed based on genotyping-by-sequencing were used for varietal identification and fingerprinting of cigar tobacco accessions (Wang et al. 2021).

The high-density maps developed based on SSR and SNP markers will be useful for germplasm characterization and identification of target traits including abiotic stresses.

Genome-wide SNP markers were identified and used for association analysis of leaf chemistry traits in natural populations of tobacco germplasms (Tong et al. 2020). The high density genetic maps developed in tobacco facilitate the tobacco genetic researchers to detect genome-wide DNA polymorphisms, fine map and clone their trait of interest. Genome-wide DNA polymorphisms could also be identified using the custom 30 K Infinium iSelect HD Bead Chip SNP chip (Edwards et al. 2017). Map based cloning of two homeologous candidate genes conferring white stem phenotype in recessive condition in burley tobacco proved that map based cloning of target traits is possible in tobacco (Edwards et al. 2017).

Discovery of novel genes/alleles for any given trait could be achievable with genotyping-by-sequencing and whole-genome re-sequencing methods. Genomics tools can be used for rapid identification and selection of novel beneficial genes and their incorporation into cultivated species.

Genome-wide association studies (GWAS) could be used to identify the novel genomic regions governing traits of interest by associating between DNA polymorphisms and trait variations in diverse germplasm collections that are phenotyped and genotyped.

### ***10.10.5 Application of Structural and Functional Genomics in Genomics-Assisted Breeding***

The latest advances in high throughput sequencing technologies and accurate phenotyping platforms are helpful in transforming molecular breeding to genomic assisted breeding (GAB). GAB has key role in developing future tobacco cultivars with rapid accumulation of beneficial alleles and removing deleterious alleles (Varshney et al. 2021). With the availability of *Nicotiana* draft genomes, transcriptome and metabolome profiles might be helpful in understanding the genomic locations, ancestral relationships, gene products and expression patterns responsible for abiotic stresses. Development of genetic maps (Sect. 10.7) using available molecular markers (Sects. 10.5.2 and 10.7.2) and identification of trait specific linked DNA markers with key abiotic stress tolerant traits can pave the way for the utilization of GAB in tobacco improvement.

Presently, in tobacco application of structural and functional genomics in GAB and application of DNA markers for MAS is in infant stage. Considerable progress made in genomics and generation of genotypic and phenotypic populations can be used to develop predictive models for estimation of breeding values for genomic selection. Breeding value can be used to predict the performance of parents in crossing and genetic advance in breeding lines based on genomic profile of target population in the target environment. Early generation selection of desirable lines without much

effect of environment can shorten the time required for breeding. The genotypic data generated from seed or seedling and the favorable alleles can be used to predict the performance of mature individuals without extensive phenotyping over years in different environments during GAB (Varshney et al. 2014).

## 10.11 Recent Concepts and Strategies Developed

Gene editing and nanotechnology are emerging as important tools for genomic designing in crop plants for various traits including abiotic stress tolerance.

### 10.11.1 Gene Editing

Gene editing involves precise modifications in the target genes to bring desirable changes in the phenotypes of organism. Gene editing technology alter gene expression of gene of interest through structural and epigenetic modifications of the DNA of target gene through various techniques like nuclease technologies, homing endonucleases, and certain chemical methods (Khan 2019). Other molecular techniques like meganuclease (MegaN), transcription activator-like effector nucleases (TALENs), and zinc-finger nucleases (ZFNs) have also emerged as important gene-editing technologies (Townsend et al. 2009). These initial technologies have limitation of lower specificity due to their off-targets side effects. The latest discovery of the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) nuclease system seems more efficient in improving the efficiency and feasibility to taking the genome-engineering techniques to the next higher level of molecular engineering (Ahmer et al. 2021). Either *Agrobacterium*-mediated or protoplast transformation are routinely used delivery systems in plants in most of the gene editing technologies. Editing of target genes was successfully demonstrated in tobacco using meganucleases (Puchta 1999; Honig et al. 2015), ZFNs (Townsend et al. 2009), TALENs (Zhang et al. 2013) and CRISPR/Cas9 system (Hirohata et al. 2019; Tian et al. 2020, 2021). These promising techniques can effectively be used in the improvement of tobacco for abiotic stress tolerance in future.

### 10.11.2 Nanotechnology

The genetic engineering technique basically modifies plant cell genomes, through efficient delivery of modifier biomolecules such as DNA, SiRNA and miRNA as genetic cargo to targeted plant (Cunningham et al. 2018; Ahmar et al. 2021). *Agrobacterium*-mediated transformation (AMT), biolistic delivery of DNA, electroporation, viral vectors and chemical delivery are most widely used delivery systems. Random

DNA integration, disruption of endogenous plant genes, variable gene expression arising out of inserted sites, tissue damage, cytotoxicity, requirement of protoplasts regeneration, less cargo sizes, high host specificity etc. are one or the other limitations associated with these systems (Niazian et al. 2017; Toda et al. 2019). These limitations can be effectively overcome by advanced nanotechnological tools.

Nanoparticles (NPs) are more useful for delivery of various genetic cargo (DNA, RNA, proteins and ribonucleoproteins) across plant systems in species independent passive manner (Cunningham et al. 2018). Thus, NP mediated delivery systems successfully used for increasing the efficiency, robustness and versatility of genetic engineering through use of nanocarriers by forming a binding complex with bio-modifier molecules (CRISPR/Cas) while delivering into plant cells (Demirer et al. 2021). This incomparable potential of the NP based delivery of biomolecules to plant cells has revolutionized the gene delivery systems (Ahmer et al. 2021) and emerged as cutting-edge technology that provides new insights and robustness of gene editing technology (Cunningham et al. 2018). Various barriers like species- and tissue-specific limitations of delivery of biomolecules to plant cells can be overcome by NPs because of smaller size and transverse the cell wall. Further NPs can also be engineered to mediate cargo delivery to various subcellular parts, where even AMT cannot target such as mitochondrial or chloroplast DNA.

NPs system of biomolecular delivery in both plant and animals are classified into five types based on the base material use namely bio-inspired, carbon-based, silicon-based, polymeric, and metallic/magnetic (Cunningham et al. 2018). Each NP type delivers different genetic cargos viz., carbon nanotubes (CNTs) can carry RNA and DNA; metallic NPs can only deliver DNA as genetic cargo; silicon-based NPs can carry DNA and proteins; polymeric NPs (e.g., PEG and polyethyleneimine) can transfer encapsulated RNA, DNA, and proteins into cells (Silva et al. 2010).

Cationic NPs type can bind to the plant cell wall (negatively charged) and successfully be used for gene transfer, whereas CNT NPs have been used to deliver plasmid DNA into various crops. NP mediated cargo delivery can be done either in physical or non-physical means. Physical methods include creating transient pores in the cell membrane with electric fields, soundwaves, or light, magnetofection, microinjection, and biolistic particle delivery. Non-physical methods embrace the use of cationic carriers, incubation, and infiltration. Various NPs can behave differently in specific plant cells, which require optimizing their application for different plant species and their dose and spatiotemporal tuning (Ahmar et al. 2021). The use of NPs for plant transformation enables an efficient method because NPs protect the genetic cargo from cellular enzymatic degradation (Ahmar et al. 2021).

Literature on NP mediated transformation system in tobacco mainly focusing on genetically modified tobacco mosaic virus-based metallic nanomaterial synthesis, NPs as pesticides, NP uptake, effects on plant growth, biomolecule delivery systems etc. (Burklew et al. 2012; Wang et al. 2016; Tirani et al. 2019). Torney et al. (2007) was first to demonstrate NP co delivery of DNA and chemicals in *N. tabacum* plants via biolistic delivery of 100–200 nm gold capped mesoporous silica NPs (MSNs).

Zinc NPs were used to deliver DNA plasmid into tobacco (Fu et al. 2012). Silicon carbide whiskers (SCW) and MSN, have been effectively used to transfer genes

into tobacco without using other physical methods (Golestanipour et al. 2018). In general, the SCW method has one disadvantage compared with other NP-mediated plant transformation in that an adequate protocol is required for plant regeneration from cell cultures. Silva et al. (2010) used polymer NPs to introduce siRNA into tobacco protoplasts, providing an alternative gene knockout mechanism in plant cells. Meanwhile, NP-mediated passive delivery of DNA plasmids has been reported with tobacco through CNTs (Burlaka et al. 2015; Kwak et al. 2019) and dsDNA through clay nano sheet NPs (Mitter et al. 2017). Demirer and colleagues (2019) have recently achieved passive delivery of DNA plasmids and protected siRNA using functionalized CNT NPs for transient silencing of constitutively expressed gene in transgenic *N. benthamiana* leaves with 95% efficiency.

With significant advantages, use of NPs in GE, possess few challenges like nanophytotoxicity effect on plant growth and environment (Ahmar et al. 2021). Cell structural stability and metabolic pathway disturbance and deposition and dispersal to other plant cells after application needs further research. Another limitation is efficient binding of biomolecules to NPs and the disintegration of the binding complex in plant cells because of different binding affinities with NPs based on their structure, charge, chemical composition, and surface area, making them ideal for a bioconjugation complex. So, their optimization for binding specific biomolecules requires further research to enhance their versatility as genetic cargo.

## 10.12 Brief on Genetic Engineering for Resistance Traits

Tobacco is extensively used as a model plant system in genetic engineering research (Jube and Borthakur 2007) for the study of basic biological functions, such as plant-pathogen interactions, environmental responses, growth regulation, senescence, etc. In view of this, number of studies have been undertaken incorporating genes from bacteria, animals and other plant species into tobacco and their functional roles validated. Genetic engineering of tobacco plants for resistance related traits are reviewed here.

### 10.12.1 Target Traits and Alien Genes

Tobacco was used as a model system for the functional validation of a number of abiotic stress responsive genes from different crop plants through their transgenic incorporation into tobacco. Large volume of information is available in literature on this area. Target traits for imparting resistance to various abiotic stresses that are validated in tobacco are briefly detailed here under.

### 10.12.1.1 Drought Tolerance

A number of drought resistant candidate gene were transferred to tobacco to study improved drought tolerance (Kolodyazhnaya et al. 2009). These include DREB and WRKY transcription factors, genes that alter the levels of trehalose and mannitol, and LEA genes (Tarczynski et al. 1992, 1993; Pilon-Smits et al. 1998; Kasuga et al. 2004; Wang et al. 2006; Wei et al. 2008; Rabara et al. 2015). Further, a number of reports on tobacco transgenics incorporated with novel genes/gene products/transcription factors viz., RING-finger protein (RFP1), NAC transcription factors 2a, atriplex hortensis choline monoxygenase (AhCMO),  $\Delta$ 1-pyrroline-5-carboxylate synthetase (P5CSF129A), neomycine phosphotransferase (nptII), pyrroline-5-carboxylate synthetase (P5CS), NADP-ME, Cox, Mannitol-1-phosphate dehydrogenase (mt1D), pleurotus sajor-caju trehalose, phosphorylase (PsTP), Trehalose-6-phosphate synthase (TPS1), *Triticum aestivum* ubiquitin 2, Boea hygrometrica late embryogenesis and abundant proteins (BhLEA1 BhLEA2) for validating their role in incorporating drought resistance are available in literature.

### 10.12.1.2 Salinity Tolerance

Salt stress tolerance responses induces extensive gene expression of various genes viz., ion transport, antioxidant systems, hormonal regulators and autotransporters. The overexpression of various genes from various sources showed tolerance against salinity stress in tobacco. Gene/gene products viz., Na<sup>+</sup>/H<sup>+</sup> antiporter (SOS1) and *PPase* TVP1 for altered accumulation of Na<sup>+</sup> and K<sup>+</sup> in shoot and root; CDH, BDH, COX, PEAMT genes for osmo- protection through glycine betaine proved enhance salt tolerance. During the evolution halophytes developed various morphological, anatomical, and physiological mechanisms to sustain under salinity condition. The salt responsive genes/antiporters (SeCMO, SbGSTU, SbNHX1, AINHx, AISAP) from halophytes (*Salicornia* spp. and *Aleuropus* spp.) have been well characterized and extensively used to develop salt stress tolerant transgenic tobacco.

### 10.12.1.3 Heat Tolerance

Thermo tolerance is a complex multigenic trait, which is influenced by genotype  $\times$  environment interactions. Heat shock proteins (Hsps) are molecular chaperones that maintains vast range of cellular functions ranging from protein aggregation and folding of proteins to membrane stability and transcription factors under extreme temperature conditions. Transgenics possessing heat tolerance have been developed using different genetic engineering techniques in tobacco. Various members of *HSPs* (*HsP* 101, *HsP* 70, *HsP* 16.9 and *HsP* 18.2) and several other genes/plant proteins such as FAD 7, Dank 1, ubiquitin, cytosolic Cu/Zn-SOD and Mn-POD have been characterized and developed transgenic plants showing enhanced tolerance to heat/cold stress in tobacco. Over expression of cytokinin oxidase/dehydrogenase



(CKX1) gene, modification of fatty acids in thylakoid membranes of tobacco chloroplasts osmoprotectants in transgenic tobacco found to impart temperature tolerance in tobacco.

#### **10.12.1.4 Cold Tolerance**

Considerable number of cold responsive genes and gene regulated networks have been reported in tobacco. Several studies have showed that transgenics developed with transformation of novel cold responsive genes from different sources (*Casell* spp., *Brachypodium* spp., *Populus* spp.) imparts cold tolerance through increase in the activities of leaf malondialdehyde, superoxide dismutase and antioxidant enzyme activity, and increased accumulation of inositol etc. in tobacco.

#### **10.12.1.5 Water Logging Tolerance**

Water logging condition triggers a series of morpho biochemical changes and gene expression leading to adaptation to the stress. Various regulatory networks and transcription factors were involved to adapt to water logging and low oxygen. The transgenes from *Brassica oleracea* and *Actinidia deliciosa* proved to accumulate more dry matter, leaf chlorophyll content and fresh and dry weight in tobacco under waterlogging conditions.

#### **10.12.1.6 Heavy Metal Tolerance**

To cope with heavy metal tolerance, plants employ various strategies which involve complex physiological and biochemical changes including changes in global gene expression to cope up with heavy metal tolerance. Recently nickel, cadmium and aluminum toxicity were addressed in tobacco through genetic engineering approaches using novel stress responsive gene like AtDHAR and SbMYB15 transcription factors.

#### **10.12.1.7 Engineering Multiple Stress Tolerance in Tobacco**

Stress tolerance mechanism in plants is controlled by complex transcriptional network and transcription factors (TFs) are the major player in this network. The cascade of molecular responses consists of perception of stress, transduction of signals to cellular machinery, gene expression, metabolic changes lead to stress tolerance. Plants exhibit both, stress-specific and correlated other shared responses that may protect them from several environmental stresses. Recent developments of stress biology showed evidence of crosstalk between abiotic and biotic stress responses in biological systems. Novel functional genes possessing multiple stress resistance have

been transformed into the tobacco plants successfully and their phenotypic effect were determined.

### ***10.12.2 Review on Achievements of Transgenics***

Tobacco has served as a model plant for producing large number of transgenics having abiotic stress tolerance and other economically important genes. However, no genetically transformed tobacco varieties (transgenic cultivars) are released for commercial cultivation in any of the countries, in view of the opposition faced by genetically modified (GM) tobacco in the global market (Bowman and Sisson 2000). Though GM Approval Database of International Service for the Acquisition of Agri-biotech Applications (ISAAA) reports two GM tobacco events viz. (1) oxynil herbicide tolerance and (2) nicotine reduction, antibiotic resistance (GM approval database 2021), none of them are cultivated on commercial scale in any of the countries. In contrast, millions of hectares of genetically engineered soybean, corn, cotton and canola are being grown throughout the world (ISAAA 2019). Thus, tobacco breeding efforts lag behind those of other crops in genetic engineering. In addition, the strong opposition from the European countries to genetically modified organisms (GMOs) is also acting as hindrance in transgenic tobacco breeding. Thus, genetic engineering of tobacco cultivars is on hold until the trade related obstacles are alleviated. However, this methodology holds great promise for improving tobacco cultivars in terms of disease and pest resistance, and possibly health-related constituents in the cured leaf.

### ***10.12.3 Organelle Transformation***

As organelles, plastids and mitochondria, containing genetic materials in small DNA genomes provide an opportunity for transformation in plants (Butow and Fox 1990). Plastid genomes of tobacco are typically 150 kb, and codes for about 140 genes. Plastids are the site of most important biosynthetic processes and pathways, such as photosynthesis, photorespiration, metabolism of amino acids, lipids, starch, carotenoids, other isoprenoids, phenol compounds, purines, pyrimidines, isoprenoids, starch, pigments, vitamins synthesis, and also are implicated in the metabolism of phytohormones such as cytokinins, abscisic acid, and gibberellins (Kuchuk et al. 2006; Rascon-Cruz et al. 2021). Compared with conventional nuclear genetic engineering, plastid genome transformation offers several benefits (Kuchuk et al. 2006; Li et al. 2021). High level of transgene expression is possible with chloroplasts as there are about 100 chloroplasts per cell, each containing about 100 copies of genome. Thus, there is possibility of 10,000 copies of transgenes per cell due to plastid transformation. Gene silencing, or the so-called position effects observed in nuclear transformation were not described for plastid genes. Hence, the level of expression is much

more predictable. Unlike integration into the nuclear genome, integration of heterologous DNA into a plastome occurs via a homologous recombination mechanism, thus allowing very precise genetic manipulations. Multigene engineering through stacking transgenes in synthetic operons in a single transformation event is possible through plastid transformation. Maternal inheritance of plastomes avoid the risk of uncontrolled transgene release into the environment (Kuchuk et al. 2006; Li et al. 2021).

Stable transformation of the plastome was first achieved in tobacco by Svab et al. (1990). Over the years, plastid transformation in tobacco has become more and more routine, with a transformation efficiency approaching that of nuclear transformation (Svab and Maliga 1993; Daniell et al. 2016; Li et al. 2021). Plastids of *N. tabacum* var. Petit Havana (Svab et al. 1990), *N. benthamiana* (Davarpanah et al. 2009) and *N. sylvestris* (Maliga and Svab 2010) were successfully transformed by different workers. The development of plastid transformation technology has paved the way to transgene expression, genome editing, and RNA editing analysis in plastids.

Though possibility of transformation exists for mitochondrial genome, reliable methods for the transformation of mitochondria using a biolistic device currently exist only for yeasts (Johnston et al. 1988) and green algae (Remacle et al. 2006). No successful transformation of mitochondria in plant systems has been reported to date (Li et al. 2021). A genetic transformation system for plant mitochondria would allow functional analyses of the mitochondrial genome and its products, and would open the way for engineering of the genome to modify mitochondrial metabolism, or to introduce cytoplasmic male sterility (CMS) into new crops and varieties (European Commission 1989; Wang et al. 2020).

#### **10.12.4 Biosynthesis**

Major changes occur in physiology, metabolites, mRNA levels, and promoter activities during the tobacco response to abiotic stresses (Rabara et al. 2015). Components of a core metabolic response at the gene, metabolite, plant hormone, transcription factor, and promoter levels and are regulated by family-specific changes in transcription factor activity. Numerous biochemical pathways leading to tens of thousands of primary and secondary metabolites are involved in abiotic stress resistance in plants (Nascimento and Fett-Neto 2010). Primary metabolic pathways producing sugars (trehalose, sucrose and fructan), amino acids (tryptophan and proline), and ammonium compounds (polyamines and glycine betaine) serve as osmoprotectants under stressed condition. Enhanced biosynthesis of these osmoprotectants will improve the abiotic stress tolerance (Rathinasapathi 2000; Rontein et al. 2002).

### 10.12.5 *Metabolic Engineering Pathways and Gene Discovery*

A metabolic pathway is defined as any sequence of feasible and observable biochemical-reaction steps connecting a specified set of input and output metabolites. The pathway flux is the rate at which input metabolites are processed to form output metabolites. Metabolic engineering involves beneficial alteration of metabolic pathways using recombinant DNA technology to better understand and utilize the cellular pathways for the production of useful metabolites (Bailey 1991). This method involves overexpression or downregulation of certain proteins in a metabolic pathway, such that the cell produces a new product.

Complete understanding of metabolic pathway and genes involved in the pathway and host cell for genetic modifications are essential for the successful engineering (Fuentes et al. 2018). Multiple transgenes expressing more than one gene involved in the pathway are frequently required in metabolic engineering which is a challenge with traditional transgenic approaches. New technological options such as combinatorial transformation (large-scale co-transformation of the nuclear genome) and transformation of the chloroplast genome with synthetic operon constructs (Bock 2013) offers straight forward multigene engineering by pathway expression from operons, high transgene expression levels, and increased transgene containment due to maternal inheritance of plastids. It also provides direct access to the large and diverse metabolite pools available in chloroplasts and non-green plastid types.

The transcription factors are often present as gene families and regulate target genes in tissue- and species-related patterns (Bovy et al. 2002). Transcription factors (TFs) tend to control multiple pathway steps and hence, considered as powerful tools for the manipulation of complex metabolic pathways for engineering the levels of metabolites (Broun 2004; Grotewold 2008). Analysis of changes in transcriptomes and metabolomes should provide clues related to regulation by transcription factors in heterologous systems. As a group, flavonoids are involved in many aspects of plant growth and development, such as pathogen resistance, pigment production, UV light protection, pollen growth, and seed coat development (Harborne 1986). Hence, manipulation of phenylpropanoid pathway responsible for flavonoid production can be a strategy for biotic and abiotic stress resistance.

Metabolic engineering was successfully demonstrated for enhancing the target molecules in tobacco altering multiple genes through plastid transformation (Lücker et al. 2004; Lu et al. 2013, 2017). Grafting the transplastomic tobacco onto the non-transformable species *N. glauca* facilitated the horizontal transfer of the transgenic plastid genomes across the graft junction (Lu et al. 2017). Thus, grafting may be helpful in the transplastomic engineering of plant species that are otherwise not amenable. Metabolic engineering of artemisinic acid biosynthetic pathway provided a proof of concept for combining plastid and nuclear transformation to optimize product yields from complex biochemical pathways in chloroplasts (Fuentes et al. 2016). Transplastomic tobacco that expressed the core artemisinic acid biosynthetic pathway from two synthetic operons accumulated only low levels of the metabolite.

However, super transformation of the transplastomics lines using the COSTREL (combinatorial super transformation of transplastomic recipient lines) approach, increased the artemisinic acid content up to 77-fold. Reduction in photorespiration could be obtained through the introduction of three distinct alternative glycolate metabolic pathways into tobacco chloroplasts (South et al. 2019). Coupling the alternative photorespiratory pathway with reduced expression of a glycolate and glycerate transporter to limit glycolate flux out of the chloroplast raised biomass productivity by >40% under field conditions (South et al. 2019). In this study, a total of 17 constructs were designed for nuclear transformation; these multienzyme pathways could potentially be introduced into the chloroplast by plastid transformation in the form of operons.

### 10.12.6 Gene Stacking

Gene stacking or gene pyramiding or multigene transfer refers to incorporation of two or more genes of interest into a single plant. The combined traits resulting from this process are called stacked traits. A biotech crop variety that bears stacked traits is called a biotech stack or simply stack (ISAAA 2020). Biotech stacks are engineered to overcome the myriad of problems such as insect pests, diseases, weeds, and environmental stresses so that farmers can increase their productivity. Insect resistance based on multiple genes confers stable resistance than single gene which may breakdown due to co-evaluation of pests. The main methods for genetically engineering plants with gene stacking involve (i) the simultaneous introduction, by the co-transformation process, and (ii) the sequential introduction of genes using the re-transformation processes or the sexual crossing between separate transgenic events.

Though, stacked products are promising and technically feasible in tobacco (Bakhsh et al. 2018; Li et al. 2018; Boccardo et al. 2019), till date, none of the stacks are approved for commercial cultivation in tobacco mainly because of their transgenic tag. Gene pyramiding events in tobacco are mainly used as proof of concepts or for gene function and interaction studies. Regulatory principles and procedures for approval and release of biotech stacks differ globally (ISAAA 2020). No separate or additional regulatory approval is necessary in countries like the USA and Canada for commercializing hybrid stacks developed through crossing a number of already approved biotech lines. This policy is based on the argument that interactions between individual trait components in a stack that have been shown to pose no environmental or health hazard would not result in new or altered hazards. However, in Japan and European Union the stacks are considered new events, even if individual events have market approval, and must pass through regulatory approval process including safety assessment (ISAAA 2020). Risk assessment of stacks is focused on the identification of additional risks that may arise from the combined genes.

### **10.12.7 Gene Silencing**

Gene silencing is the prevention or reduction in the expression of a certain gene through regulation of its gene expression in a cell. Gene silencing strategies are particularly useful in functional characterization of abiotic stress resistant genes. RNA interference (RNAi)-based silencing of stress responsive genes and studying the knockdown plants for their response to stress can be options for assessing functional significance of these genes and their utilization (Senthil-Kumar et al. 2010). Functional characterization of stress responsive genes helps to understand the role of specific genes in stress tolerance so as to manipulate them in designing stress resistant cultivars. Primarily, the mechanism of RNAi mediated gene silencing is based on the exogenous production of short interfering RNAs/microRNAs (siRNAs/miRNAs) by an organism to control the expression of genes. Expression or introduction of double-stranded (ds) RNA in eukaryotic cells can trigger sequence-specific gene silencing of transgenes and endogenes.

### **10.12.8 Prospects of Cisgenics**

In cisgenesis, the extra DNA originates from a donor plant with which the acceptor plant can cross-breed (Schouten et al. 2006). Combining both traditional breeding techniques and modern biotechnology, this approach dramatically speed up the breeding process. Introduction of desired genes through cisgenesis overcome the linkage drag and prevents hazards such as induced translocation or mutation breeding (Telem et al. 2013; Hou et al. 2014). Using cisgenesis both abiotic and biotic stress resistance genes can be pyramided to provide broader and long-lasting forms of resistance. Cisgenesis reduces the time required for transferring a single gene or more so with multiple genes compared to conventional breeding that require several backcrossed generations to remove undesired genes (Telem et al. 2013).

Introduction of exogenous transfer process related genes in cisgenesis can be avoided with the use of new transformation protocols without bacterial selection markers and use of species-specific plastic DNAs (P-DNAs) instead of bacterial T-DNAs for insertion of isolated genes (de Vetten et al. 2003; Rommens et al. 2004; Schaart et al. 2004). Techniques such as promoter trapping and RNA fingerprinting for the isolation of native regulatory elements can be exploited for the precise expression of target traits (Meissner et al. 2000; Trindade et al. 2003). Majority of the methods for production of cisgenic crops have been patented, and therefore scientists need appropriate approvals to use these patents or design new methods to eliminate the undesired DNA sequences from host genomes (Holme et al. 2013).

The prospects for cisgenesis are enormous in tobacco crop in view of the availability of large number of wild species and germplasm resources, genome sequence information of cultivated tobacco and few wild relatives, and comparative genomic techniques, the development of efficient gene isolation techniques like map-based

cloning and allele mining for identification and cloning of abiotic stress resistant traits from tobacco and their wild relatives. Cisgenesis was successfully demonstrated in tobacco through various gene editing techniques such as ZFNs (Townsend et al. 2009), TALENs (Zhang et al. 2013) and CRISPR-Cas (Upadhyay et al. 2013; Ali et al. 2015). Single or multiple biotic stress resistant cisgenes can be successfully identified, cloned and transferred into tobacco in future with the ever-improving gene technologies.

Cisgenesis introduce only genes of interest from the plant itself or from a crossable species which otherwise could also be transferred by traditional breeding techniques. Hence, cisgenesis is more similar to traditional plant breeding than transgenesis. Release of cisgenic plants into the environment is as safe as that of traditionally bred plants and there is no environmental risk evoked. Therefore, cisgenic plants can be considered as non-transgenic, in spite of using the methods of genetic engineering. There is a need to distinguish cisgenesis from transgenesis as any restrictions on cisgenesis could hamper further research and application of improved crop varieties, especially at a time when more number of genes from crops and their crossable wild relatives are being isolated.

Surveys indicated that cisgenic plants are more acceptable to common people than transgenic plants (Viswanath and Strauss 2010; Gaskell et al. 2011; Mielby 2011). However, GMO regulations in majority of the countries do not distinguish transgenic plants from cisgenic plants. Product-based regulation system rather than a process-based one followed in Canada making it legally possible to control cisgenic plants less strictly compared to transgenic plants. In Australia, cisgenic plants are treated differently under GMO regulations (Russell and Sparrow 2008). While, European Food Safety Authority (EFSA 2012) validated that cisgenic plants are similar to the traditionally bred plants in terms of environment, food and feed security.

In spite of the availability of abiotic stress resistant tobacco transgenics, worldwide GMO regulations making it difficult to utilize them for commercial cultivation. In such situation, differential treatment to cisgenics are treated differently, that will boost the cisgenesis research in tobacco for improving tobacco yields and resistant factors.

### **10.13 Brief Account on Role of Bioinformatics as a Tool**

Advances in sequencing technology applications have resulted in the accumulation of large volumes of biological data in terms of nucleic acid sequences. To store and analyze these data, number of general and crop specific databases were created. The databases may contain the information covering one or more than one type of omics in an integrated way. The information pertaining to tobacco are being stored and accessed through quite several databases, globally. However, some of the important key databases that are covering tobacco data information are discussed here under.



### 10.13.1 Gene and Genome Databases

Advances in sequencing technologies, gene mapping and tagging projects, and phylogenetic studies have resulted in accumulation of large volumes of genomic data in tobacco. The genomic databases serve as hubs for storing, sharing and comparison of accumulated data across research studies, data types, individuals and organisms.

Among the various databases, the key genome databases harboring *Nicotiana* genome and gene information are NCBI Genome, Sol Genome Networks (SGN), Kyoto Encyclopedia of Genes and Genomes (KEGG genome), EnsemblPlants, *Nicotiana attenuata* data hub (NaDH), The International Nucleotide Sequence Database Collaboration (INSDC), Gramene etc. (Table 10.6). NCBI and SGN together are the important databases that covers all the information on genomes and genes of various *Nicotiana* species. At present, genome sequences of 12 *Nicotiana* spp. viz., *N. tabacum*, *N. tomentosiformis*, *N. sylvestris*, *N. attenuata*, *N. undulate*, *N. otophora*, *N. suaveolens*, *N. glauca*, *N. stocktonii*, *N. repanda*, *N. amplexicaulis* and *N. debneyi* at scaffold or contigs level, chloroplast genomes of five species and mitochondrial genomes of three species are available with one or the other databases. Further, more than two lakh records of gene sequences belonging to 12 *Nicotiana* species are accumulated at various data bases. These databases are sharing the stored information with other databases and providing extensive tools for the analysis of sequences and annotation. INSDC is a long-standing foundational initiative that operates between DNA Databank of Japan (DDBJ), European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) and NCBI and covers the spectrum of data raw reads, through alignments and assemblies to functional annotation. Other databases mentioned above provide access to the *Nicotiana* resources, mostly, through collaboration with other databases along with additional analysis tools existing with their databases. *Nicotiana attenuata* data hub (NaDH) covers the exclusive information on *N. attenuata* and its similarities with other *Nicotiana* species and 11 published dicot species. A website of Boyce Thompson Institute's for *N. benthamiana* resources provide access to *N. benthamiana* genomic resources available at SGN including gene and protein data, markers, genes to phenotypes database etc. (<https://btscience.org/our-research/research-facilities/research-resources/nicotiana-benthamiana>). It is also providing tools for alignment, annotation, designing siRNAs for VIGS, CRISPR designing etc.

In addition to above databases, the Gene Ontology resource database provides access to scientific information about the molecular functions of genes (or, more properly, the protein and noncoding RNA molecules produced by genes) from many different organisms, from humans to bacteria, their cellular locations and processes those gene products may carry out (Table 10.6). Currently, 25,761 genes and gene products are found to be associated with the term *Nicotiana* in The Gene Ontology resource database.

Most of the gene and genomic databases provide tools for searching, alignment and comparison of sequences with other *Nicotiana* species. Apart from analysis of genome sequence data, various genome databases are facilitating the analysis of gene

**Table 10.6** Important genomic resource databases providing information on *Nicotiana*

S. No.	Name	Url
1	National Center for Biotechnology Information (NCBI)	<a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a>
2	Sol Genomics Networks (SGN)	<a href="http://solgenomics.net/">http://solgenomics.net/</a>
3	<i>Nicotiana attenuata</i> Data Hub (NaDH)	<a href="http://nadh.ice.mpg.de">http://nadh.ice.mpg.de</a>
4	Kyoto Encyclopedia of Genes and Genomes (KEGG genome)	<a href="https://www.genome.jp/kegg/">https://www.genome.jp/kegg/</a>
5	EnsemblPlants	<a href="https://plants.ensembl.org">https://plants.ensembl.org</a>
6	Gramene	<a href="https://www.gramene.org/">https://www.gramene.org/</a>
7	VISTA	<a href="https://genome.lbl.gov/">https://genome.lbl.gov/</a>
8	The Gene Ontology (GO) Resource	<a href="http://geneontology.org/">http://geneontology.org/</a>
9	Gene Expression Omnibus (GEO)	<a href="https://www.ncbi.nlm.nih.gov/geo/">https://www.ncbi.nlm.nih.gov/geo/</a>
10	ArrayExpress (AE)	<a href="https://www.ebi.ac.uk/arrayexpress/">https://www.ebi.ac.uk/arrayexpress/</a>
11	Genomic Expression Archive (GEA)	<a href="https://www.ddbj.nig.ac.jp/gea/index-e.html">https://www.ddbj.nig.ac.jp/gea/index-e.html</a>
12	Universal Protein Resource (UniProt)	<a href="https://www.uniprot.org/">https://www.uniprot.org/</a>
13	Pfam database	<a href="http://pfam.xfam.org/">http://pfam.xfam.org/</a>
14	SolCyc Biochemical Pathways	<a href="https://solcyc.solgenomics.net/">https://solcyc.solgenomics.net/</a>
15	REACTOME	<a href="https://reactome.org/">https://reactome.org/</a>
16	The Golm Metabolome Database (GMD)	<a href="http://gmd.mpimp-golm.mpg.de/">http://gmd.mpimp-golm.mpg.de/</a>
17	MoNA (Massbank of North America)	<a href="https://mona.fiehnlab.ucdavis.edu/">https://mona.fiehnlab.ucdavis.edu/</a>
18	Database for Annotation, Visualization, and Integrated Discovery (DAVID)	<a href="https://david.ncifcrf.gov/">https://david.ncifcrf.gov/</a>
19	The International Nucleotide Sequence Database Collaboration (INSDC)	<a href="https://www.insdc.org/">https://www.insdc.org/</a>
20	The BioStudies Database	<a href="https://www.ebi.ac.uk/biostudies/">https://www.ebi.ac.uk/biostudies/</a>

variation and expression, analysis and prediction of gene and protein structure and function, prediction and detection of gene regulation networks, etc.

### 10.13.2 Comparative Genome Databases

The increasing availability of genomic sequence from multiple organisms has provided large dataset for orthologous-sequence comparisons. The rationale for using cross-species sequence comparisons is to identify biologically active regions of a genome based on the observation that sequences that perform important functions are often conserved between evolutionarily distant species, distinguishing them

from nonfunctional surrounding sequences. This is most readily apparent for protein-encoding sequences but also holds true for the sequences involved in the regulation of gene expression.

Comparison of whole-genome sequences at the level of nucleotide or protein provides a detailed narration of syntenic relation at genetic level. Comparative genome studies will identify the types of genes, gene families, and their location also provide information on the history of evolutionary rearrangements of the gene including duplications that might be responsible for the identified genetic variation. By carefully comparing genome characteristics that define various organisms, researchers can pinpoint regions of similarity and difference. This information can be used to identify putative genes and regulatory elements for various traits that may lead to their cloning and further utilization.

A variety of tools for comparison of complete genome sequences of within or between the different species are available in different databases. All the gene and genome databases of tobacco namely NCBI, SGN, NaDH etc. are offering tools for comparative genome analysis. VISTA is a comprehensive suite of programs and databases for comparative analysis of genomic sequences. There are two ways of using VISTA—one can submit their own sequences and alignments for analysis (VISTA servers) or examine pre-computed whole-genome alignments of different species.

Gramene, a knowledge base was founded on comparative functional analyses of genomic and pathway data for model plants and major crops including tobacco. The current release, #64 (September 2021), hosts 114 reference genomes, and round 3.0 million genes from 90 plant genomes with 3,256,006 input proteins in 1,23,064 families with orthologous and paralogous classifications. Comparative genomics collection totals 340 pairwise DNA alignments and 80 synteny maps. Plant Reactome portrays pathway networks using a combination of manual biocuration and orthology-based projections to 106 species. The Reactome platform facilitates comparison between reference and projected pathways, gene expression analyses and overlays of gene–gene interactions. Gramene integrates ontology-based protein structure–function annotation; information on genetic, epigenetic, expression, and phenotypic diversity; and gene functional annotations extracted from plant-focused journals.

Various online/web applications can provides comparative analyses at both the genomic and genic levels tools, such as BRIG (Alikhan et al. 2011), Mauve (Darling et al. 2004), Artemis Comparison Tool (ACT) (Carver et al. 2005), geneCo (Jung et al. 2019) etc. can be used for comparative genomics apart from various tools provided various databases. At *Nicotiana attenuata* data hub, genes of 11 published dicot species were compared and found to cluster into 23,340 homologous groups (HG) based on their sequence similarity with at least two homolog sequences. The phylogenetic trees were also constructed for all these HG.

Comparative analyses of *Nicotiana* plastid genomes among themselves and with currently available Solanaceae genome sequences indicated the existence of similar GC and gene content, codon usage, simple sequence and oligonucleotide repeats, RNA editing sites, and substitutions (Asaf et al. 2016). Such analysis also revealed

that *N. otophora* is a sister species to *N. tomentosiformis* within the *Nicotiana* genus, and *Atropa belladonna* and *Datura stramonium* are their closest relatives (Asaf et al. 2016).

Comparison of whole nuclear and plastid genomes made it possible to identify and confirm of wild progenitor species and their relative genome contributions in the evaluation of cultivated tobacco genomes (Murad et al. 2002; Lim et al. 2004, 2005; Leitch et al. 2008; Sierro et al. 2014, 2018; Edwards et al. 2017). Whole-genomic sequence comparison indicated that the genome of *N. sylvestris* and *N. tomentosiformis* contributes 53 and 47%, respectively, for *N. tabacum* specifying a larger biased genome reduction in T genome (Siirro et al. 2014). In case of *N. rustica*, 41% of genome originated from the paternal donor (*N. undulata*), while 59% originated from the maternal donor (*N. paniculata*/*N. knightiana*) (Siirro et al. 2018). Chloroplast genome comparisons revealed that *N. otophora* is a sister species to *N. tomentosiformis* within *Nicotiana* genus and *Atropa belladonna* and *Datura stramonium* are the closest relatives (Asaf et al. 2016). Maternal parent of the tetraploid *N. rustica* was found to be the common ancestor of *N. paniculata* and *N. knightiana*, and the later species is more closely related to *N. rustica*. Gene clustering analysis revealed the commonality of 14,623 ortholog groups among the *Nicotiana* species and 207 specific to *N. rustica* (Siirro et al. 2018). It was speculated from the results the higher nicotine content of *N. rustica* leaves is the result of the progenitor genomes combination and of a more active transport of nicotine to the shoot.

### 10.13.3 Gene Expression Databases

Large volume of data is being generated on gene expression patterns in tobacco ranging from seed to senescence under varied conditions in response to abiotic stresses. With initiatives of the Tobacco Genome Initiative (TGI) resulted in enrichment of the sequence information of transcriptionally active regions of the tobacco genome in the form of ESTs, short, single pass sequence reads derived from complementary DNA (cDNA) libraries and methyl filtered genome space sequence reads (GSRs). Kamalay and Goldberg (1980) measured the extent of structural gene expression in an entire tobacco plant. Matsuoka et al. (2004) observed the changes in gene expression during the growth of tobacco BY-2 cell lines and isolated 9,200 EST fragments corresponding to about 7,000 genes. Rushton et al. (2008) identified 2,513 TFs covering the 64 well-characterized plant TF families and these were used to create a database of tobacco transcription factors (TOBFAC). Edwards et al. (2010) designed tobacco expression microarray using Affymetrix platform from a set of 40 k unigenes and measured the gene expression in 19 different tobacco samples to produce the tobacco expression atlas (TobEA). TobEA provides a snapshot of the transcriptional activity of tobacco genes in different tissues throughout the lifecycle of the plant. Expression profiling of tobacco leaf trichomes resulted in the identification of putative genes involved in resistance to biotic and abiotic stresses (Harada et al. 2010; Cui et al. 2011).

The expression databases covers transcript/RNA information of different genes under varied native or test conditions along with the relevant software tools for analysis and retrieval of the data. ESTs is relatively expensive and time consuming. However, Microarrays provide a faster less costly alternative for measuring gene expression simultaneously that can be more easily and reproducibly applied across varied range of conditions to identify genes specific expression patterns or responses. At present the gene expression data have been stored as microarray and RNA-seq datasets in the public databases such as Gene Expression Omnibus (GEO), ArrayExpress (AE) and Genomic Expression Archive (GEA) (Table 10.6). These databases act as useful resources for the functional interpretation of genes and their expression. GEO contained 4,860 curated gene expression data sets as well as original series and platform records of 11 *Nicotiana* spp. (<https://www.ncbi.nlm.nih.gov/gds>). Genomic Expression Archive has 205 gene expression records of 11 *Nicotiana* spp. SGN is maintaining 39 transcript libraries of *N. tabacum* and two of *N. glauca*. Further, there are exclusive expression databases for *Nicotiana attenuata* (NaDH) and *N. benthamiana* (<https://btscience.org/our-research/research-facilities/research-resources/nicotiana-benthamiana>) along with the Sol genome networks for expression analysis among solanaceous members.

#### 10.13.4 Protein or Metabolome Databases

Proteome is the study of proteins thought to be expressed by an organism in its life cycle. The metabolome deals with the metabolites of small size (<1500 Da) in a specific cell of an organ or organism. The metabolome of the plant act as link between genotype and phenotype. It also indicates the stage/organ specific response of the plants through gene expression in response to external environment. It not only influences the gene expression but also affects the protein functions of the plant which make metabolomics a central component in elucidating cellular systems and decoding gene functions.

Proteomics and metabolomics approaches play significant role in functional genomics and have been essentially required for understanding plant development and abiotic stress tolerance. Proteome and metabolome profiling are potential tool for phenotyping plants under varied environmental changes and biotic stresses. Such studies contribute significantly to the study of abiotic stress biology by distinguishing different compounds such as auxiliary products of stress metabolism from biosynthetically complex abiotic pathways, stress induced signal molecules, molecules that are part of plant acclimation process etc.

The resultant metabolic compounds could be further studied by direct measurement or correlating with the changes in transcriptome and proteome expression during stress condition and can be confirmed by mutant analysis. Thus, metabolome study aid in unravelling the different pathways related to plant development and response to stresses. With the advent of high throughput-based systems, proteome and metabolome profiling was extensively carried out in the model plant like tobacco to

examine stress signaling pathways, cellular and developmental processes (Xiang et al. 2016). Principal databases hosting tobacco proteome information are UniProt, Pfam, KEGG, SGN, NCBI, etc. and metabolome databases are SolCyc, REACTOME, The Golm Metabolome Database (GMD), MoNA (Massbank of North America), etc. The salient features of these data bases are briefed below.

UniProt, an association between the European Bioinformatics Institute (EBI), the Swiss Institute of Bioinformatics (SIB), and the Protein Information Resource (PIR), is comprised of three databases, each optimized for different uses. Universal Protein Resource (UniProt), a comprehensive resource for protein sequence and annotation data base generated globally from proteomic data. The UniProt Knowledge base (UniProtKB) is the central access point for extensively curated protein information, function, classification and cross-references. The UniProt Reference Clusters (UniRef) combine closely related sequences into a single record to speed up sequence similarity searches. The UniProt Archive (UniParc) is a comprehensive repository of all protein sequences, consisting only of unique identifiers and sequences. UniProt provides several sets of proteins thought to be expressed by organisms whose genomes have been completely sequenced, termed “proteomes”. There are 73,606 protein entries associated with proteome *N. tabacum* and a total of 154,728 entries for all *Nicotiana* species as on 30.09.2021. The Pfam database is a large collection of protein families, each represented by multiple sequence alignments. This database provides tools for protein alignments and annotation, domain organization of a protein sequence etc. There are about 4,970 unique results for the search term *Nicotiana* in this data base as on 30.09.2021 indicating the protein entries in the database. KEGG database in addition to providing protein information, projects the biological processes from various organisms onto pathways consolidated into large network schemes. At present, the KEGG database have the information of annotated proteins of *N. tabacum* (61,780 No.), *N. tomentosiformis* (30,989) *N. sylvestris* (33,816) and *N. attenuata* (34,218). In the Sol genomics network database also providing the data of proteins annotated based on the draft genome sequences of *N. tabacum*, *N. benthamiana* and *N. attenuata*. Around 275,000 collection of protein sequences from several sources are available for 20 *Nicotiana* spp. at NCBI along with annotated reports of four *Nicotiana* spp. (Table 10.4). SolCyc is a collection of Pathway Genome Databases (PGDBs) for Solanaceae species generated using Pathway Tools. It is a database hub at SGN for the manual curation of metabolic networks and includes annotated metabolic, regulatory and signaling processes in Solanaceous plants based on Omics data obtained from multiple resources. It has species-specific databases for *N. tabacum* (K326Cyc), *N. attenuata* (NattCyc), *N. sylvestris* (NiSylCyc), *N. tomentosiformis* (NiTomCyc), *N. benthamiana* (BenthaCyc); and multi-species databases for Combined *Nicotiana* genus (Nicotiana Cyc) and Combined Solanaceae database (Solana Cyc). Apart from the proteomic data, NaDH is providing metabolome database with analysis tools for *N. attenuata* with facilities for the search of metabolites and fragments based on annotation and measured values.

REACTOME offers bioinformatics tools for the visualization, interpretation and analysis of pathway knowledge to support basic and clinical research, genome analysis, modeling, systems biology and education. GMD, an open access metabolome database provides public access to custom mass spectral libraries, metabolite profiling experiments as well as additional information and tools, e.g. with regard to methods, spectral information or compounds. MoNA (Massbank of North America) is a centralized and collaborative database of metabolite mass spectra, metadata and associated compounds. MoNA currently contains over 200,000 mass spectral records from experimental and in-silico libraries as well as from user contributions.

The proteomic studies in tobacco revealed different stress responses (Amme et al. 2005). Analysis of the proteome of glandular trichomes revealed the enrichment of proteins belonging to components of stress defense responses. Metabolome study under water stress in tobacco identified a useful marker for drought stress for members of Solanaceae (Rabara et al. 2017).

### 10.13.5 Integration of Different Data

Analysis of omics data provides biological understanding at a specific molecular layer such as genome, proteome, transcriptome and metabolome. Understanding of agronomic traits requires the complete knowledge of complex crosstalk between different molecular layers, such as genome, proteome, transcriptome and metabolome. An integrative analysis of multiple layers of molecular data or system biology helps to discover and elucidate molecular mechanisms of phenotypic traits and resistance responses to various abiotic stresses (Singh et al. 2016). With the advent of high-throughput techniques and availability of multi-omics data generated from a large set of samples, lot of promising tools and methods have been developed for data integration and interpretation. Most of the biological databases collect and integrate data from different sources.

Databases namely INSDC, NCBI, SGN, NaDH, KEGG genome, EnsemblPlants, DAVID (Database for Annotation, Visualization, and Integrated Discovery), The BioStudies Database etc. are some of the integrated databases and resources that are collecting and integrating the omics data from different plants and tobacco (Table 10.6). As a collaborative foundational initiative, INSDC covers the spectrum of data raw reads, through alignments and assemblies to functional annotation, enriched with contextual information relating to samples and experimental configurations. NCBI in addition to providing wide-ranging tobacco data information in its different databases, offers tools for integration of structural and functional genomic data and their annotations. The curated proteome data of *Nicotiana* species in Uniport, and metabolome data from different resources are being integrated in new datahubs like SGN, KEGG, NaDH etc. to provide holistic information from gene to pathway for the researchers. KEGG is an integrated resource database consisting of 16 databases including genes and proteins, metabolites and other chemical substances, biochemical reactions, enzyme, disease-related network variations etc.



EnsemblPlants is an integrative resource that includes genome-scale information for sequenced plant species (currently 33 in no.). Data provided includes genome sequence, gene models, functional annotation, and polymorphic loci. DAVID is a web-accessible database that integrates functional genomic annotations with intuitive graphical summaries. Lists of gene or protein identifiers are rapidly annotated and summarized according to shared categorical data for gene ontology, protein domain, and biochemical pathway membership. Numerous public sources of protein and gene annotation information have been integrated into DAVID database for over 1.5 million genes from more than 65,000 species. European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) is building BioStudies Database, a resource for accepting and archiving data generated in “multi-omics” experiments. Building of archives, databases and analysis tools in an integrated approach have proven successful for better understanding and comparing omic resources of tobacco.

## **10.14 Brief Account on Social, Political and Regulatory Issues**

### ***10.14.1 Concerns and Compliances***

Tobacco is high value commercial crop generating higher farm income and revenue to national government and farmers. The crop is providing livelihood security to the people involved in tobacco production, processing and marketing in one side and on the other side it is posing serious health risks to consumers. In another way use of huge quantity of forest wood for FCV curing causing deforestation, smoking related environment pollution and also spitting habits of chewing tobacco are causes for concern. Climate change patterns, emerging biotic and abiotic stresses, pesticide residues, consumer preferences and tobacco regulatory policies are becoming increasingly complex and posing challenges for tobacco cultivation (ICAR 2015). WHO-Framework Convention on Tobacco Control (FCTC), with membership of 182 countries, envisages non-price, price and tax measures to reduce the supply and demand for tobacco in the world. On May 31st every year, the world observes World No Tobacco Day (WNTD) promoted by the World Health Organization (WHO) with a primary focus on encouraging users to refrain from tobacco consumption and its related products at least for a period of 24 h.

### ***10.14.2 Patent and IPR Issues***

Researchers are developing various management strategies for minimizing crop yield losses at field level due to abiotic stresses. This includes development of abiotic stress



tolerant varieties through conventional breeding and biotechnological interventions. The advances in biotechnology and bioinformatics generated various genome-based tools, techniques, genes and gene constructs in the field of abiotic stress tolerance (Dangl and Jones 2001). Intellectual property rights (IPR) for plants help protect inventions made in research and development of new tobacco varieties (CORESTA 2005). In turn, this encourages investment and helps continue the development of new varieties that increases economic returns throughout the tobacco supply chain. Over the past 25 years, an increasing number of governments and international organizations have enacted laws, regulations, or policies that acknowledge the need for IPR.

Patents provides the ownership right to make, use, sell, offer for sale, or import for those purposes a patented product. The ability to patent plant varieties is recognized in some countries, but in many countries such as India it is disallowed. Instead India provided farmers rights considering their historical contribution of preserving and protecting valuable genetic resources. There have been few patents granted for tobacco varieties in the world (CORESTA 2005). United States permitted the use of utility patents to protect plant varieties. European countries that are members of the European Patent Office (EPO) ban the patenting of plant varieties, but recent determinations support patent claims directed to plants of more than one variety. Thus, a utility patent and the ruling by the EPO may suggest a type of broad protection for a novel plant trait that is not recognized under PVP.

Recent advances in molecular biology, plant genomics and crop science have brought about a paradigm shift of thought regarding how tobacco plants can be utilized for commercial and medicinal uses. There is a scope for patenting of novel methods of making abiotic stress resistant tobacco genotypes, methods of introgressing nucleic acid molecules associated with stresses, genes conferring resistance to various stresses (Hefferon 2010). Patenting activity in resistance genes in tobacco is initiated in 1992 and there is considerable progress in patenting from the year 2000 and it was more prominent from 2010 onwards (Prabhakararao et al. 2016). Majority of these patent documents (around 60%) are in the jurisdiction of United States of America (USA) and China.

The intellectual information generated in the frontier areas is available in both non-patent and patent literature. Nearly 80% of all the technical information available in the world is hidden in the patent documents and other IP assets (Prabhakararao et al. 2016). Patent mapping helps in retrieving and exploring the information protected in the intellectual property documents. Collections of patent documents are available in a number of patent information databases (<https://guides.library.queensu.ca/patents/databases>). Most patent offices provide free access to patent documents via public databases. Patent information can be used to decide the patentability of an invention, avoid re-invention and infringement, provide the current state of the art in a given field of technology, find the latest trends in R and D being pursued by the peers and competitors etc.

### ***10.14.3 Disclosure of Sources of GRs, Access and Benefit Sharing***

Genetic resources (GRs) are a key for the number of biotechnological innovations (Steward 2018). History reveals that less than 1% of species have provided the basic resources for the development of all civilization. It is not possible to predict which genes, species or ecosystems will become valuable in the future. Over the last decades, regulations have been developed that aim to improve the sustainable use of GRs and benefit sharing i.e., Access-Benefit Sharing (ABS). Convention on Biological Diversity (CBD) adopted in 1992 serves as initiation point in many countries for biodiversity conservation and use. The more recent Nagoya protocol, a 2010 supplementary agreement to the CBD, aims to improve the fair and equitable sharing of benefits arising out of the utilization of genetic resources. The existing ABS systems vary widely and GR-rich countries tending to organize their systems more strictly focusing on acquiring an equitable share of the benefits related to products resulting from the use of GR. Over the years, several governments introduced disclosure requirements (DRs) in the patent system as an extra component to enhance ABS compliance.

#### ***10.14.4 Framers' Rights***

Farmers around the world have been the custodians, innovators and protectors of agricultural biodiversity (Craig 2004; FAO 2017) since the dawn of cultivation of crop plants. Farmers are involved in collecting the best seeds and cultivating various types of tobacco species/types throughout the world. Through the careful selection of their best seeds, propagation of material, and exchange with each other farmers lead to develop of new innovations/varieties as well as to diversify crop varieties. Country like India, households traditionally raising different tobacco landraces in their kitchen gardens since generations from the seeds collected from their own crops, thus maintaining and protecting biodiversity. Considering the past, present and future contributions of farmers in conserving, improving, and making available plant genetic resources, farmer's rights are to be protected. Farmers' access to use and exchange the seed and propagating material are to be protected from seed regulations (variety release and seed marketing regulations), legislation concerning to intellectual property rights (patents and plant breeders' rights), and regulations dealing with bio-prospecting of genetic resources.

The concept of farmers' rights was emerged in international negotiations within FAO in 1986 to counter the increased demands for plant breeders' rights (PBR) being voiced in international negotiations. In 1987, solutions were being proposed, serving as the foundation for all further negotiations on Farmers' Rights. In 1989, farmers' Rights gained formal recognition by the FAO Conference. In 1991, the Conference

decided to set up a fund for the realization of these rights, but this has never materialized. CBD with its resolution on the promotion of sustainable agriculture urged FAO to commence negotiations for a legally binding international regime on the management of plant genetic resources and to resolve the question of Farmers' Rights. Agenda 21 approved at the UN Conference on Environment and Development held in Rio de Janeiro in 1991 had voiced similar demands. In November 1996, Global Plan of Action for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture (Global Plan of Action) acknowledges the need to realize Farmers' Rights and was endorsed by the FAO Council, by the Conference of the Parties to the CBD, and by the World Food Summit at FAO. The Second Global Plan of Action prepared under the aegis of the Commission on Genetic Resources for Food and Agriculture was adopted by the FAO Council in 2011. This action plan contains a set of recommendations and activities intended as a framework, guide and catalyst for action at community, national, regional and international levels.

The International Treaty on Plant Genetic Resources for Food and Agriculture (International Treaty) adopted in 2001 addressed the issue of Farmers' Rights in its Preamble and Article 9. The International Treaty recommends Contracting Parties to protect and promote Farmers' Rights in accordance with their national laws. In Article 9, the Treaty recognize the enormous contribution made and will continue to make by the farmers of all regions of the world for the conservation and development of plant genetic resources as the basis of food and agricultural production throughout the world. Measures are suggested that covering the protection of traditional knowledge, benefit-sharing and participation in decision-making. The rights of farmers to save, use, exchange and sell farm-saved seeds and propagating material are addressed in the International Treaty, but without any legally binding provisions on how to implement Farmers' Rights at national level.

#### ***10.14.5 Traditional Knowledge***

Local communities uses traditional knowledge (TK) generated using long standing traditions and practices in coping with extreme weather and adapting to climate change from the centuries for their survival (Swiderska et al. 2011). The diversity of traditional varieties maintained by farmers around the world are important for adaptation to climate changes and emerging issues of abiotic stresses. Local communities use wild foods to supplement their diets and thus conserve wild species which are valuable sources of abiotic stress resistant genes. The traditional varieties or landraces are genetically more diverse than modern varieties and are good sources of resistance to abiotic stresses (<https://www.cbd.int/traditional/what.shtml>). Because of their long experience in cultivating crops under varied changing climates, the traditional farmers are well placed to identify resilient crop species and resistant varieties for abiotic stresses with the available accumulated TK. Traditional Knowledge about resilient properties, such as abiotic and biotic stress resistance traits and wild crop relatives

can be a valuable information for developing stress tolerant tobacco varieties (Jarvis et al. 2008).

### **10.14.6 *Treaties and Conventions***

International agreements that have special significance in the context of agricultural sector in general and biotechnology particular are CBD, International Treaty for Plant Genetic Resources for Food and Agriculture (ITPGRFA) and The International Union for the Protection of New Varieties of Plants (UPOV). The salient features and provisions of them are briefly discussed hereunder.

#### **10.14.6.1 Convention on Biological Diversity (CBD)**

The Convention on Biological Diversity (CBD), widely known as the Biodiversity Convention, is an international legally binding treaty and was ratified in 1992 at the Rio earth summit (<https://www.cbd.int/convention/>). It is often seen as the key document concerning the sustainable development and currently has 196 Parties (168 Signatures). The Convention has three main goals viz., conservation of biological diversity (or biodiversity), sustainable use of its components and fair and equitable sharing of benefits arising from genetic resources. The convention recognized for the first time in international law that the conservation of biological diversity is “a common concern of humankind”. It is an integral part of the development process and covers all ecosystems, species, and genetic resources. The objective of CBD is to develop national strategies for the conservation and sustainable use of biological diversity. The Contracting Parties shall, in accordance with national legislation and policies, encourage and develop methods of cooperation for the development and use of technologies, including indigenous and traditional technologies, in pursuance of the objectives of this Convention. It links traditional conservation efforts to the economic goal of using biological resources sustainably. The Contracting Parties shall promote international technical and scientific cooperation in the area of conservation and sustainable use of biological diversity. For this purpose, the Contracting Parties shall also promote cooperation in the training of personnel and exchange of experts. Article 19 of the Convention deals with the issues in respect of handling of living modified organism resulting from biotechnology.

CBD has three important protocols viz., The Nagoya Protocol on Access and Benefit-sharing, The Cartagena Protocol on Biosafety and The Nagoya—Kuala Lumpur Supplementary Protocol on Liability and Redress to the Cartagena Protocol on Biosafety under CBD. The essential features of these protocols are briefed below.

### 10.14.6.2 Cartagena Protocol on Biosafety

The Cartagena Protocol on Biosafety is an international treaty governing the movements of living modified organisms (LMOs) resulting from modern biotechnology from one country to another (<http://bch.cbd.int/protocol>). Adopted as a supplementary agreement to CBD, it came into force on 11 September 2003. The objective of the Protocol is to ensure an adequate level of protection in the field of the safe transfer, handling and use of ‘living modified organisms resulting from modern biotechnology’ that may have adverse effects on the conservation and sustainable use of biological diversity, taking also into account risks to human health, and specifically focusing on trans-boundary movements. The Protocol provides for Parties to enter into bilateral, regional and multilateral agreements and arrangements regarding international trans-boundary movements of living modified organisms.

The Protocol establishes a Bio-Safety Clearing-House to: (a) Facilitate the exchange of scientific, technical, environmental and legal information on, and experience with, living modified organisms; and, (b) Assist Parties to implement the Protocol, taking into account the special needs of developing country Parties, in particular the least developed and small island developing States among them, and countries with economies in transition as well as countries that are centres of origin and centres of genetic diversity.

### 10.14.6.3 Nagoya Protocol

The Nagoya Protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their Utilization (ABS, is a supplementary international agreement to CBD, adopted on 29 October 2010 in Nagoya, Japan and entered into force on 12 October 2014 (<https://www.cbd.int/abs/>). Currently, it has 131 Parties (132 ratifications) (92 signatories). The protocol provides for a transparent legal framework for the effective implementation of fair and equitable sharing of benefits arising out of the utilization of genetic resources. It applies mainly to genetic resources and traditional knowledge (TK) associated with genetic resources that are covered under CBD, and to the benefits arising from their utilization.

The Nagoya—Kuala Lumpur Supplementary Protocol on Liability and Redress to the Cartagena Protocol on Biosafety

Adopted as a supplementary agreement to the Cartagena Protocol, it aims to contribute to the conservation and sustainable use of biodiversity by providing international rules and procedures in the field of liability and redress relating to living modified organisms (<http://bch.cbd.int/protocol/supplementary/>). The Protocol was entered into force on 5 March 2018 and currently, it has 49 Parties. The Protocol requires that response measures are taken in the event of damage resulting from living modified organisms which find their origin in a transboundary movement, or where

there is sufficient likelihood that damage will result if timely response measures are not taken. A definition of ‘damage’, referring to an adverse effect on the conservation and sustainable use of biological diversity that is measurable or otherwise observable and significant, taking also into account risks to human health is provided in the Protocol. The Protocol necessitates a causal link between the damage and the living modified organism be established. While imposing the requirement for response measures, the Protocol obliges the Parties to continue to apply existing legislation on civil liability or to develop specific legislation concerning liability for material or personal damage associated with the conservation and sustainable use of biological diversity.

#### **10.14.6.4 The International Treaty on Plant Genetic Resources for Food and Agriculture**

The International Treaty on Plant Genetic Resources for Food and Agriculture Adopted in 2001 (FAO 2009) with the objectives of conservation and sustainable use of all plant genetic resources for food and agriculture and the fair and equitable sharing of the benefits arising out of their use, in harmony with CBD, for sustainable agriculture and food security. This legally binding Treaty covers all plant genetic resources and vital in ensuring the continued availability of the plant genetic resources that countries will need to feed their people. The Treaty recognizes the enormous contribution made by the local and indigenous communities and farmers of all regions of the world and takes measures for protecting Farmers’ Rights. The Contracting Parties agree to establish an efficient, effective, and transparent multilateral system to facilitate access to plant genetic resources for food and agriculture, and to share, in a fair and equitable way, the benefits arising from their utilization. The treaty takes care of (a) protection of traditional knowledge; (b) the right to equitably participate in sharing benefits arising from the utilization of plant genetic resources; and (c) the right to participate in making decisions, at the national level, on matters related to the conservation and sustainable use of plant genetic resources for food and agriculture.

#### **10.14.6.5 The UPOV Convention**

The UPOV (International Union for Protection of Plant Varieties) Convention was adopted for the protection of plant variety in Paris in 1961 and entered into force in 1968. It was subsequently revised in 1972, 1978 and 1991. The 1978 Act entered into force in 1981, and the 1991 Act in 1998 ([www.upov.int](http://www.upov.int)). An inter-governmental organization was established with headquarters in Geneva, Switzerland. UPOV’s mission is to provide and promote an effective system of plant variety protection, with the aim of encouraging the development of new varieties of plants, for the benefit of society. Plant Variety Protection (PVP) under UPOV is enabled in the 77 members countries. The current act of the convention adopted in 1991 recognizes breeder’s

rights to a variety if the variety is: (1) new; (2) distinct; (3) uniform; and (4) stable. The breeder's rights as per 1991 Act require authorization of the breeder to perform the following: (1) production or reproduction (multiplication) (2) conditioning for the purpose of propagation, (3) offering for sale, (4) selling or marketing, (5) exporting, (6) importing, (7) stocking for any purpose mentioned in (1) to (6) above. Breeder's rights to a variety remain in effect for a period of 20 years from the date on which the rights were granted. The Act of 1991, for the first time included the protection against "essentially derived" varieties, which are derived from the protected variety, that is not clearly distinguishable from the protected variety, or which requires repeated use of the protected variety for production purposes. An essentially derived variety could be developed from a protected variety through: (1) the selection of a natural or induced mutant, or of a somaclonal variant, (2) the selection of a variant individual from that of the initial variety, (3) backcrossing, and (4) transformation by genetic engineering. Exceptions to breeder's rights were granted for: (1) acts done privately and for non-commercial purposes, (2) acts done for experimental purposes, and (3) acts done for the purpose of breeding other varieties, except for the generation of essentially derived varieties.

#### **10.14.7 Participatory Breeding**

For thousands of years prior to 1800s, during and after the domestication of *Nicotiana* species, one of the principal method of tobacco improvement was through conservation of diversity and selection of naturally occurring high yielding and stress resistant variants by cultivators. The systematic varietal improvement started by scientists in 1990s in established research organizations globally has led to release of number of high yielding stress resistant tobacco varieties through conventional plant breeding techniques.

In view of the limitations to formal breeding and the threats to farmers' seed systems, participatory plant breeding (PPB) emerged as a means to overcome some of the limitations of formal system and to bring farmers back into the breeding process as active participants (Greenberg 2018). The role played by farmers in agricultural biodiversity conservation and use is taken as an advantage while making them as important partners in breeding plant varieties. In the development of improved varieties, PPB ensures the improvement of the adapted local genetic materials using the diversity available either with them or public gene banks to suit the farmer needs prevailing under their conditions. This also empower the farmer in terms of technical and organizational skills in maintaining and developing plant materials under their control, their on-farm management, and local creativity/innovation. PPB involves the active participation of farmers in few or all the steps of sequenced breeding program viz., priority setting, production and sharing of genetic materials and knowledge, acquisition of genetic material and selection, crossing, selection at early/advanced stages, and evaluation. PPB a complementary breeding process and may not be a

substitute for station-based research or scientist-managed on-farm trials (Hardon et al. 2005; Aguilar-Espinoza 2007; Ceccarelli et al. 2009).

## 10.15 Future Perspectives

### 10.15.1 Potential for Expansion of Productivity

The improvements in crop yield through conventional approaches is still achievable in tobacco (Sarala et al. 2016). The genotype improvements in tobacco for increasing and stabilizing yields can further be accelerated through the combination of traditional breeding techniques with genome designing strategies. Advancements in genomic research would assist in designing of appropriate genome assisted breeding strategies for attaining maximum potential yields with good quality and required stress resistance in a short span of time.

### 10.15.2 Potential for Expansion into Nontraditional Areas

Tobacco is an important commercial crop that plays a significant role in the economies of many countries (FAO 2019). Another way, it is one of the most important model systems in plant biotechnology till date and going to continue further. *Nicotiana* species are investigated for aspects concerning the elucidation of principles of disease resistance, synthesis of secondary metabolites and basic questions of plant physiology. In view of its higher level of biomass accumulation, tobacco is a promising crop to produce commercially important substances (e.g., antigens, antibodies, drugs and vaccines) through molecular farming and cultivation of tobacco for its valuable native phyto-chemicals viz., nicotine, solanesol, proteins and organic acids (Sarala 2019). Hence, tailoring tobacco for molecular farming is going to be an important objective for tobacco improvement programs.

## References

- Abdurakhmonov IY, Abdukarimov A (2008) Application of association mapping to understanding the genetic diversity of plant germplasm resources. *Intl J Plant Genom.* <https://doi.org/10.1155/2008/574927>
- Abrams MD, Kubiske ME Steiner KC (1990) Drought adaptation and responses in five genotypes of *Fraxinus pennsylvanica* Marsh.: photosynthesis, water relations and leaf morphology. *Tree Physiol* 6:305–315
- Aguilar-Espinoza E (2007) Evaluating a participatory plant breeding approach in a local seed system of Northern Nicaragua: challenges for upscaling technologies and institutions. Master's thesis, Norwegian University of Life Sciences



- Ahmar S, Mahmood T, Fiaz S et al (2021) Advantage of nanotechnology-based genome editing system and its application in crop improvement. *Front Plant Sci* 12:663849. <https://doi.org/10.3389/fpls.2021.663849>
- Ajab H, Yaqub A, Malik SA et al (2014) Characterization of toxic metals in tobacco smoke and cigarette ash from selected imported and local brands in Pakistan. *Sci World J*. <https://doi.org/10.1155/2014/413614D>
- Akehurst BC (1981) *Tropical agricultural series tobacco*, 2nd edn. New York
- Ali MA, Abbas A, Niaz S et al (2009) Morpho-physiological criteria for drought tolerance in sorghum (*Sorghum bicolor*) at seedling and post-anthesis stages. *Intl J Agri Biol* 11:674–680
- Ali Z, Abul-Faraj A, Piatek M et al (2015) Activity and specificity of TRV-mediated gene editing in plants. *Plant Signal Behav* 10(10):e1044191. <https://doi.org/10.1080/15592324.2015.1044191>
- Alikhan NF, Petty NK, Zakour NLB et al (2011) BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402
- Amme S, Rutten T, Melzer M et al (2005) A proteome approach defines protective functions of tobacco leaf trichomes. *Proteomics* 5(10):2508–2518. <https://doi.org/10.1002/pmic.200401274>
- Anderson E, Dorothea DW (1931) The genetic analysis of an unusual relationship between self-sterility and self-fertility in *Nicotiana*. *Ann MO Bot Gard* 18(1):97–116. <https://doi.org/10.2307/2394046>
- Anderson J, Wagner M, Rushworth C et al (2014) The evolution of quantitative traits in complex environments. *Heredity* 112:4–12. <https://doi.org/10.1038/hdy.2013.33>
- Anuradha M, Sivaraju K, Krishnamurthy V (2013) Effect of waterlogging on physiological characteristics, yield and quality of flue-cured tobacco. *Indian J Plant Physiol* 18:67–70
- Arslan B, Okumus A (2006) Genetic and geographic polymorphism of cultivated tobaccos (*Nicotiana tabacum*) in Turkey. *Genetika* 42(6):818–823. <https://doi.org/10.1134/S10227954060123>
- Arumuganathan K, Earle ED (1991) Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* 9:208–218
- Asaf S, Khan AL, Khan AR et al (2016) Complete chloroplast genome of *Nicotiana otophora* and its comparison with related species. *Front Plant Sci* 7:843. <https://doi.org/10.3389/fpls.2016.00843>
- Baghyalakshmi K, Sarala K, Prabhakara Rao K et al (2018) Genetic diversity in tobacco genotypes evaluated under All India Network Project. *Multilogic Sci* 8 (issue special):30–34
- Bailey JE (1991) Towards a science of metabolic engineering. *Science* 21:252(5013):1668–1675. <https://doi.org/10.1126/science.2047876>. PMID: 2047876
- Bakhsh A, Dinç T, Hussain T (2018) Development of transgenic tobacco lines with pyramided insect resistant genes. *Turk J Biol* 42(2):174–186. <https://doi.org/10.3906/biy-1710-71>
- Baloglu MC, Oz MT, Oktem HA et al (2012) Expression analysis of TaNAC69-1 and TtNAMB-2, Wheat NAC family transcription factor genes under abiotic stress conditions in durum wheat (*Triticum turgidum*). *Plant Mol Biol Rep* 30:1246–1252
- Barchi L, Bonnet J, Boudet C et al (2007) A high-resolution, intraspecific linkage map of pepper (*Capsicum annum* L.) and selection of reduced recombinant inbred line subsets for fast mapping. *Genome* 50(1):51–60
- Barnes JD, Kevin E, Paul ND et al (1996) The influence of UV-B radiation on the physicochemical nature of tobacco (*Nicotiana tabacum* L.) leaf surfaces. *J Exp Bot* 47(1):99–109
- Barrett BA, Kidwell KK (1998) AFLP-based genetic diversity assessment among wheat cultivars from the Pacific Northwest. *Crop Sci* 38:1261–1271
- Basimia A, Hatami MH, Darvishzadeh R et al (2014) Mixed linear model association mapping for low chloride accumulation rate in oriental-type tobacco. *J Plant Interact* 9:666–672
- Beavis WD (1998) QTL analysis: power, precision, and accuracy. In: Paterson A (ed) *Molecular dissection of complex traits*. CRC Press, Boca Raton, Florida
- Belknap WR, Garbarino JE (1996) The role of ubiquitin in plant senescence and stress response. *Trends Plant Sci* 1:331–335

- Berbec A, Doroszewska T (2020) The use of *Nicotiana* species in tobacco improvement. In: Ivanov NV, Sierro N, Peitsch MC (eds) The tobacco plant genome, compendium of plant genomes. Springer Nature, Switzerland AG. [https://doi.org/10.1007/978-3-030-29493-9\\_8](https://doi.org/10.1007/978-3-030-29493-9_8)
- Binbin H, Geng R, Cheng L et al (2020) Genetic diversity and fingerprinting of 33 standard flue-cured tobacco varieties for use in distinctness, uniformity and stability testing. *BMC Plant Biol* 20:378
- Bindler G, Plieske J, Bakaher N et al (2011) A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large-scale microsatellite marker development. *Theor Appl Genet* 123:219–230
- Bindler G, van der Hoeven R, Gunduz I et al (2007) A microsatellite marker-based linkage map of tobacco. *Theor Appl Genet* 114:341–349
- Bliss LC, Kramer PJ, Wolf FA (1957) Drought Resistance in Tobacco. *Tob Sci* 1–29:120–123
- Boccardo NA, Segretin ME, Hernandez I et al (2019) Expression of pathogenesis-related proteins in transplastomic tobacco plants confers resistance to filamentous pathogens under field trials. *Sci Rep* 9:2791. <https://doi.org/10.1038/s41598-019-39568-6>
- Bock R (2013) Strategies for metabolic pathway engineering with multiple transgenes. *Plant Mol Biol* 83(1–2):21–31. <https://doi.org/10.1007/s11103-013-0045-0>
- Boursiac Y, L eran S, Corratg -Faillie C et al (2013) ABA transport and transporters. *Trends Plant Sci* 18(6):325–333
- Bovy A, de Vos R, Kemper M et al (2002) High-flavonol tomatoes resulting from the heterologous expression of the maize transcription factor genes LC and C1. *Plant Cell* 14:2509–2526
- Bowman DT, Sisson V (2000) A historical overview of flue-cured tobacco breeding in the U.S.A. *Tob Sci* 44:59–64. <https://doi.org/10.3381/0082-4623-44.1.59>
- Bradbury PJ, Zhang Z, Kroon DE et al (2007) TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23(19):2633–2635
- Bretting PK, Widrechner MP (1995) Genetic markers and plant genetic resource management. *Plant Breed Rev* 13:11–86
- Brieger FG (1935) Genetic analysis of the cross between self-fertile *N. Langsdorfii* and the self-sterile *N. Sanderae*. *J Genet* 30(1):79–100
- Broun P (2004) Transcription factors as tools for metabolic engineering in plants. *Curr Opin Plant Biol* 2:202–209
- Brown JKM (2002) Yield penalties of disease resistance in crops. *Curr Opin Plant Biol* 5:339–344
- Bruner WB (1931) Root development of cotton, peanuts and tobacco in central Oklahoma. *Proc Oklahoma Acad Sci* 1931:20–37
- Burbridge TN (1960) The Australian species of *Nicotiana* L. (Solanaceae). *Aust J Bot* 8:342–380
- Burklew CE, Ashlock J, Winfrey WB et al (2012) Effects of Aluminum oxide Nanoparticles on the growth, development and microRNA expression of Tobacco (*Nicotiana tabacum*). *PLoS One* 7(5):e34783. <https://doi.org/10.1371/journal.pone.0034783>
- Burlaka OM, Pirko YV, Yemets AI et al (2015) Plant genetic transformation using carbon nanotubes for DNA delivery. *Cytol Genet* 49:3–12. <https://doi.org/10.3103/S009545271506002X>
- Buschmann C, Lichtenthaler HK (1998) Principles and characteristics of multi-colour fluorescence imaging of plants. *J Plant Physiol* 152:297–314. [https://doi.org/10.1016/s0176-1617\(98\)80144-2](https://doi.org/10.1016/s0176-1617(98)80144-2)
- Butow RA, Fox TD (1990) Organelle transformation: shoot first, ask questions later. *Trends Biochem Sci* 15(12):465–468. [https://doi.org/10.1016/0968-0004\(90\)90299-Q](https://doi.org/10.1016/0968-0004(90)90299-Q)
- Cai C, Yang Y, Cheng L et al (2015) Development and assessment of EST-SSR marker for the genetic diversity among tobaccos (*Nicotiana tabacum* L.). *Russ J Genet* 51:591–600
- Cai CC, Chai LG, Wang Y et al (2009) Construction of genetic linkage map of burley tobacco (*Nicotiana tabacum* L.) and genetic dissection of partial traits. *Acta Agron Sin* 35:1646–1654
- Cameron KD, Teece MA, Smart LB (2006) Increased accumulation of cuticular wax and expression of lipid transfer protein in response to periodic drying events in leaves of tree tobacco. *Plant Physiol* 140:176–183. <https://doi.org/10.1104/pp.105.069724>
- Campbell RB (1973) Flue cured tobacco yield and oxygen content of soil in lysimeters flooded for various periods. *Agron J* 65:783–786

- Carver TJ, Rutherford KM, Berriman M et al (2005) ACT: the Artemis comparison tool. *Bioinformatics* 21(16):3422–3423. <https://doi.org/10.1093/bioinformatics/bti553>
- Ceccarelli S, Guimaraes E, Weltzien E (eds) (2009) Plant breeding and farmer participation. FAO, Rome
- Champoux MC, Wang G, Sarkarung S et al (1995) Locating genes associated with root morphology and drought avoidance in rice via linkage to molecular markers. *Theor Appl Genet* 90:969–981
- Chaplin JF, Mann TJ (1978) Evaluation of tobacco mosaic resistance factor transferred from burley to flue-cured tobacco. *J Hered* 69:175–178
- Chaplin JF, Matzinger DF, Mann TJ (1966) Influence of homozygous and heterozygous mosaic-resistance factor on quantitative characters of flue-cured tobacco. *Tob Sci* 10:81–84
- Chari MS (1995) Role of research in the improvement of productivity and quality of Indian flue-cured Virginia tobacco. Central Tobacco Research Institute, Rajahmundry, India, p 26
- Chase MW, Knapp S, Cox AV et al (2003) Molecular systematics, GISH and the origin of hybrid taxa in *Nicotiana* (Solanaceae). *Ann Bot* 92:107–127
- Cheng L, Chen X, Jiang C et al (2019) High-density SNP genetic linkage map construction and quantitative trait locus mapping for resistance to cucumber mosaic virus in tobacco (*Nicotiana tabacum* L.). *Crop J* 7(4):539–547. <https://doi.org/10.1016/j.cj.2018.11.010>
- Chéour F, Arul J, Makhlof J et al (1992) Delay of membrane lipid degradation by calcium treatment during cabbage leaf senescence. *Plant Physiol* 4:1656–1660
- Chuan Yin D, Zhizhong H, Xuliang Y (2009) AFLP analysis of genetic diversity of tobacco germplasms. Paper presented at 63rd tobacco science research conference, Florida, 27–30 Sept 2009
- Clark JM, Townley-Smith TF, McCaig TN et al (1984) Growth analysis of spring wheat cultivars of varying drought resistance. *Crop Sci* 24:537–541
- Clarkson JJ, Knapp S, Garcia VF et al (2004) Phylogenetic relationships in *Nicotiana* (Solanaceae) inferred from multiple plastid DNA regions. *Mol Phylogenet Evol* 33:75–90
- Clarkson JJ, Lim KY, Kovarik A et al (2005) Long-term genome diploidization in allopolyploid *Nicotiana* section *Repandae* (Solanaceae). *New Phytol* 168:241–252
- Clausen RE, Cameron DR (1944) Inheritance in *Nicotiana tabacum* XVIII. Monosomic analysis. *Genetics* 29(5):447–477
- Clausen RE, Goodspeed TH (1926) Inheritance in *Nicotiana tabacum* VII. The monosomic character, fluted. *Univ Calif Pub Bot* 11:61–82
- Cocucci M, Negrini N (1988) Changes in the levels of calmodulin and of a calmodulin inhibitor in the early phases of radish (*Raphanus sativus* L.) seed germination—effects of ABA and fusicoccin. *Plant Physiol* 88:910–914
- Collard BC, Mackill DJ (2008) Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Phil Trans Roy Soc Lond B Biol Sci* 363:557–572
- CORESTA (2005) Tobacco seed integrity task force final report—May 2005, 1–14 pp. [https://www.coresta.org/sites/default/files/technical\\_documents/main/Seed-Integrity-TF-May05.pdf](https://www.coresta.org/sites/default/files/technical_documents/main/Seed-Integrity-TF-May05.pdf). Accessed 13 Jul 2021
- Craig B (2004) Farmers' rights: intellectual property regimes and the struggle over seeds. *Politics Society* 32(4):511–543. <https://doi.org/10.1177/0032329204269979>
- Cramer GR, Nowak RS (1992) Supplemental manganese improves the relative growth, net assimilation and photosynthetic rates of salt-stressed barley. *Physiol Planta* 84:600–605
- Cui H, Zhang ST, Yang HJ et al (2011) Gene expression profile analysis of tobacco leaf trichomes. *BMC Plant Biol* 11:76. <https://doi.org/10.1186/1471-2229-11-76>
- Cunningham FJ, Goh NS, Demirer GS, Matos JL (2018) Nanoparticle-mediated delivery towards advancing plant genetic engineering. *Trends Biotechnol* 36:882–897. <https://doi.org/10.1016/j.tibtech.2018.03.009>
- D'hoop BB, Paulo MJ, Kowitzanich K et al (2010) Population structure and linkage disequilibrium unravelled in tetraploid potato. *Theor Appl Genet* 121:1151–1170
- Dangl JL, Jones JDG (2001) Plant pathogens and integrated defense responses to infection. *Nature* 411:826–833

- Daniell H, Lin CS, Yu M et al (2016) Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol* 17:134. <https://doi.org/10.1186/s13059-016-1004-2>
- Darling AC, Mau B, Blattner FR et al (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 14(7):1394–1403. <https://doi.org/10.1101/gr.2289704>
- Darlington CD, Janaki Ammal EK (1945) *Chromosome Atlas of cultivated plants*. George Allen and Unwin Ltd., London
- Darvishzadeh R, Alvi SR, Sarrafi A (2011) Genetic variability for chlorine concentration in oriental tobacco genotypes. *Arch Agron Soil Sci* 57(2):167–177
- Davalieva K, Maleva I, Filiposki K et al (2010) Genetic variability of Macedonian tobacco varieties determined by microsatellite marker analysis. *Diversity* 2(4):439–449. <https://doi.org/10.3390/d2040439>
- Davarpanah SJ, Jung SH, Kim YJ et al (2009) Stable plastid transformation in *Nicotiana benthamiana*. *J Plant Biol* 52(3):244–250. <https://doi.org/10.1007/s12374-009-9023-0>
- de Vetten N, Wolters AM, Raemakers K et al (2003) A transformation method for obtaining marker free plants of a cross-pollinating and vegetatively propagated crop. *Nat Biotechnol* 21:439–442
- Del Piano L, Abet M, Sorrentino C et al (2000) Genetic variability in *Nicotiana tabacum* and *Nicotiana* species as revealed by RAPD markers: 1. Development of the RAPD procedure. *Beitr Tab Int/contrib Tob Res* 19:1–15
- Demirer GS, Zhang H, Matos JL et al (2019) High aspect ratio nanomaterials enable delivery of functional genetic material without DNA integration in mature plants. *Nat Nanotechnol* 14:456–464. <https://doi.org/10.1038/s41565-019-0382-5>
- Demirer GS, Silva TN, Jackson CT et al (2021) Nanotechnology to advance CRISPR-Cas genetic engineering of plants. *Nat Nanotechnol* 12:1–8
- Denduangboripant J, Piteekan T, Nantharat M (2010) Genetic polymorphism between tobacco cultivar-groups revealed by amplified fragment length polymorphism analysis. *J Agri Sci* 2(2):41–48
- Dias de Oliveira E, Bramley H, Siddique KH et al (2013) Can elevated CO<sub>2</sub> combined with high temperature ameliorate the effect of terminal drought in wheat? *Funct Plant Biol* 40:160–171. <https://doi.org/10.1071/FP12206>
- Djanaguiraman M, Prasad PVV (2010) Ethylene production under high temperature stress cause premature leaf senescence in soybean. *Funct Plant Biol* 37:1071–1084
- Doganlar S, Frary A, Daunay MC et al (2002) A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the Solanaceae. *Genetics* 161:1697–1711
- East EM (1932) Studies of self-sterility. IX. The behavior of crosses between self-sterile and self-fertile plants. *Genetics* 17:175–202
- Edwards KD, Bombarely A, Story GW et al (2010) Allen F, Mueller LA, Coates SA et al. *TobEA: an atlas of tobacco gene expression from seed to senescence*. *BMC Genomics* 11:142
- Edwards KD, Fernandez-Pozo N, Drake-Stowe K et al (2017) A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics* 18:448. <https://doi.org/10.1186/s12864-017-3791-6>
- EFSA (2012) Scientific opinion addressing the safety assessment of plants developed through cisgenesis and intragenesis. *EFSA J* 10(2):2561
- El-Morsy SHI, Dorra MDM, Abd El-Hady EAA (2009) Comparative studies on diploid and tetraploid levels of *Nicotiana glauca*. *Acad J Plant Sci* 2(3):182–188
- Esau K (1960) *Anatomy of seed plants*. *Soil Sci* 90:149
- European Commission (1989) Genetic transformation of plant mitochondria: development of a general strategy. [https://cordis.europa.eu/project/id/BAP\\*0102](https://cordis.europa.eu/project/id/BAP*0102). Accessed 21 Sept 2021
- Evanno G, Reganut E, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14:2611–2620
- Fan W, Sun X, Yang A (2015) Exploring high-potassium favorable allele mutation of tobacco based on genome-wide association analysis. *Acta Tab Sin* 25:11–19

- FAO (2009) International treaty on plant genetic resources for food and agriculture. Food and Agriculture Organization of the United Nations, Rome
- FAO (2017) The international treaty on plant genetic resources for food and agriculture: farmers' rights. Food and Agriculture Organization of the United Nations, Rome. <http://www.fao.org/3/I7820EN/i7820en.pdf>
- FAO (2019) FAOSTAT statistical database. Food and Agriculture Organization of the United Nations. Rome. <http://www.fao.org/faostat/en/#data/QCL>
- Fita A, Rodríguez-Burruezo A, Boscaiu M et al (2015) Breeding and domesticating crops adapted to drought and salinity: a new paradigm for increasing food production. *Front Plant Sci* 6:978. <https://doi.org/10.3389/fpls.2015.00978>
- Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. *Annu Rev Plant Biol* 54:357–374. <https://doi.org/10.1146/annurev.arplant.54.031902.134907>
- Flowers T (2004) Improving crop salt tolerance. *J Exp Bot* 55:307–319
- Flowers TJ, Colmer TD (2008) Salinity tolerance in halophytes. *New Phytol* 179:945–963. <https://doi.org/10.1111/j.1469-8137.2008.02531.x>
- Fontes RLS, Cox FR (1998) Zinc toxicity in soybean grown at high iron concentration in nutrient solution. *J Plant Nutr* 21:1723–1730
- Franco-Navarro JD, Brumós J, Rosales MA et al (2016) Chloride regulates leaf cell size and water relations in tobacco plants. *J Exp Bot* 67:873–891
- Franco-Navarro JD, Rosales MA, Cubero-Font P et al (2019) Chloride as macronutrient increases water use efficiency by anatomically-driven reduced stomatal conductance and increased mesophyll diffusion to CO<sub>2</sub>. *Plant J* 99:1–17
- Fricano A, Bakaheer N, Corvo MD et al (2012) Molecular diversity, population structure, and linkage disequilibrium in a worldwide collection of tobacco (*Nicotiana tabacum* L.) germplasm. *BMC Genetics* 13:18. <https://doi.org/10.1186/1471-2156-13-18>
- Friebe B, Jiang J, Raupp WJ et al (1996) Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status. *Euphytica* 91:59–87
- Fu YQ, Li LH, Wang PW et al (2012) Delivering DNA into plant cell by gene carriers of ZnS nanoparticles. *Chem Res Chin Univ* 28:672–676
- Fuentes P, Armarego-Marriott T, Bock R (2018) Plastid transformation and its application in metabolic engineering. *Curr Opin Biotechnol* 49:10–15. <https://doi.org/10.1016/j.copbio.2017.07.004>
- Ganal MW, Altmann T, Roder MS (2009) SNP identification in crop plants. *Curr Opin Plant Biol* 12:211–217
- Ganesh CT, Saiprasad GVS, Mohan Raju B et al (2014) Genetic structure in FCV tobacco population as assessed by multi-locus genotyping using SSR markers. *Adv Crop Sci Technol* 2:3. <https://doi.org/10.4172/2329-8863.1000127>
- Gangwar S, Singh VP (2011) Indole acetic acid differently changes growth and nitrogen metabolism in *Pisum sativum* L seedlings under chromium (VI) phytotoxicity: implication of oxidative stress. *Sci Hort* 129:321–328. <https://doi.org/10.1016/j.scienta.2011.03.026>
- Gaskell G, Allansdotir A, Allum N et al (2011) The 2010 Eurobarometer on the life sciences. *Nat Biotechnol* 29:113–114
- Gholizadeh S, Darvishzadeh R, Mandoulakani BA et al (2012) Molecular characterization and similarity relationships among flue-cured tobacco (*Nicotiana tabacum* L.) genotypes using simple sequence repeat markers. *Not Bot Hort Agrobot* 40(1):247–253
- Gier LJ (1940) Root systems of bright belt tobacco. *Amer J Bot* 27(9):780–787
- GM approval database (2021) International Service for the Acquisition of Agri-biotech Applications (ISAAA), New York. GM Crops List—GM Approval Database/ISAAA.org. Accessed 21 Sept 2021
- Goggin FL, Lorence A, Topp CN (2015) Applying high-throughput phenotyping to plant–insect interactions: picturing more resistant crops. *Curr Opin Insect Sci* 9:69–76. <https://doi.org/10.1016/j.cois.2015.03.002>

- Golestanipour A, Nikkha M, Aalami A et al (2018) Gene delivery to tobacco root cells with single-walled carbon nanotubes and cell-penetrating fusogenic peptides. *Mol Biotechnol* 60:863–878. <https://doi.org/10.1007/s12033-018-0120-5>
- Gong D, Huang L, Xu X et al (2016) Construction of a high-density SNP genetic map in flue-cured tobacco based on SLAF-seq. *Mol Breed* 36:100
- Goodspeed TH (1954) The genus *Nicotiana*: origins, relationships and evolution of its species in the light of their distribution, morphology and cytogenetics. *Chronica Botanica*, Waltham, Mass, USA
- Goodspeed TH, Thompson MC (1959) Cytotaxonomy of *Nicotiana* II. *Bot Rev* 25:385–415
- Gopalachari NC (1984) Tobacco. Indian Council of Agricultural Research, New Delhi
- Grattan S (2002) Irrigation water salinity and crop production. UCANR Publications, Oakland, CA. <https://doi.org/10.1007/s00425-003-1105-5>
- Greenberg S (2018) A review of participatory plant breeding and lessons for African seed and food sovereignty movements. Published by The African Centre for Biodiversity, Johannesburg, South Africa. <https://www.researchgate.net/publication/323771637>
- Grotewold E (2008) Transcription factors for predictive plant metabolic engineering: are we there yet? *Curr Opin Biotechnol* 19(2):138–144. <https://doi.org/10.1016/j.copbio.2008.02.002>
- Guardiola JM, Perez O, Diaz L (1987) Effect of chlorine and potassium on combustibility from fine plantations. *Tabacco* 10:29–43
- Guha A, Rasineni GK, Reddy AR (2010) Drought tolerance in mulberry (*Morus* spp.) a physiological approach with insight in to growth dynamics and leaf yield production. *Exp Agri* 46:471–488
- Guo JR, Li YD, Han GL et al (2018) NaCl markedly improved the reproductive capacity of the euhalophyte *Suaeda salsa*. *Funct Plant Biol* 45:350–361. <https://doi.org/10.1071/fp17181>
- Guo JR, Suo SS, Wang BS (2015) Sodium chloride improves seed vigour of the euhalophyte *Suaeda salsa*. *Seed Sci Res* 25:335–344. <https://doi.org/10.1017/s0960258515000239>
- Guo YH, Jia WJ, Song J et al (2012) *Thellungiella halophila* is more adaptive to salinity than *Arabidopsis thaliana* at stages of seed germination and seedling establishment. *Acta Physiol Plant* 34:1287–1294
- Gyula P, Schäfer E, Nagy F (2003) Light perception and signalling in higher plants. *Curr Opin Plant Biol* 6:446–452
- Harada E, Kim J, Meyer et al (2010) Expression profiling of tobacco leaf trichomes identifies genes for biotic and abiotic stresses. *Plant Cell Physiol* 51(10):1627–1637
- Harborne JB (1986) The flavonoids. *Advances in research since 1986*. Chapman Hall, London
- Hardon J, Norbu U, Nguyen S et al (2005) Evaluation report of the Biodiversity Use and Conservation in Asia Programme (BUCAP). NORAD, Oslo
- Hatami MH, Karimzadeh G, Darvishzadeh R et al (2013) Identification of QTLs associated with low chloride accumulation in oriental tobacco. *Genetika* 45:55–864
- Havlov M, Dobrev PI, Motyka V et al (2008) The role of cytokinins in responses to water deficit in tobacco plants over-expressing *trans*-zeatin O-glucosyltransferase gene under 35S or *SAG12* promoters. *Plant Cell Environ* 31:341–353
- He M, He C-Q, Ding N-Z (2018) Abiotic stresses: general defenses of land plants and chances for engineering multistress tolerance. *Front Plant Sci* 9:1771
- Hefferon KL (2010) Plant biotechnology patents: applications in agriculture and medicine. *Recent Pat Biotechnol* 4. <https://doi.org/10.2174/187220810791110723>
- Heggstad HE (1966) Ozone as a tobacco toxicant. *J Air Poll Contr Assoc* 16:691–694
- Heggstad HE, Middleton JT (1959) Ozone in high concentration as cause of tobacco leaf injury. *Science* 129:208–210
- Higase S (1959) Physiological studies on the flower formation of tobacco: effects of water-logging after transplanting stage on the developmental responses of tobacco. *Jpn J Crop Sci* 27(2):299–300
- Hirohata A, Sato I, Kaino K et al (2019) CRISPR/Cas9-mediated homologous recombination in tobacco. *Plant Cell Rep* 38:463–473
- Holme IB, Wendt T, Holm PB (2013) Intragenesis and cisgenesis as alternatives to transgenic crop development. *Plant Biotechnol J* 11(4):395–407. <https://doi.org/10.1111/pbi.12055>



- Honig A, Marton I, Rosenthal M et al (2015) Transient expression of virally delivered meganuclease in planta generates inherited genomic deletions. *Mol Plant* 8(8):1292–1294. <https://doi.org/10.1016/j.molp.2015.04.001>
- Hospital F (2003) Marker-assisted breeding. In: Newbury HJ (ed) *Plant molecular breeding*. Blackwell Publishing and CRC Press, Oxford and Boca Raton, pp 30–59. [https://doi.org/10.1007/978-3-642-21201-7\\_10](https://doi.org/10.1007/978-3-642-21201-7_10)
- Hou H, Atlihan N, Lu Z-X (2014) New biotechnology enhances the application of cisgenesis in plant breeding. *Front Plant Sci* 5:389. <https://doi.org/10.3389/fpls.2014.00389>
- Huang W, Yang YJ, Hu H (2016) Response of the water–water cycle to the change in photorespiration in tobacco. *J Photochem Photobiol B* 157:97–104
- Huang WK, Guo JY, Wan FH et al (2008) AFLP analyses on genetic diversity and structure of *Eupatorium adenophorum* populations in China. *Chin J Agri Biotechnol* 5:33–41
- Hurd EA (1971) Can we breed for drought resistance? In: Larson LL, Eastin JD (eds) *Drought injury and resistance in crops*. Crop Sci Soc America USA, pp 77–78
- ICAR (2015) *Vision 2050*: Central Tobacco Research Institute. Indian Council of Agricultural Research, New Delhi
- ICAR-CTRI (2016) *Annual report 2015–16*. ICAR-Central Tobacco Research Institute, Rajahmundry, AP, India
- ICAR-CTRI (2021) List of released/identified varieties of various tobacco types for cultivation in India. <https://ctri.icar.gov.in/files/varieties.pdf>. Accessed 25 Feb 2021
- IPCC (2019) Summary for policymakers. In: Shukla PR, Skea J, Calvo Buendia E et al (eds) *Climate change and land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*. World Meteorological Organization, Geneva, Switzerland (in press)
- ISAAA (2019) Global status of commercialized biotech/GM crops. ISAAA Brief 55–2019. <https://www.isaaa.org/resources/publications/briefs/55/default.asp9>. Accessed 30 Sept 2021
- ISAAA (2020) Pocket K No. 42 stacked traits in biotech crops. <https://www.isaaa.org/resources/publications/pocketk/42/default.asp>. Accessed 30 Sept 2021
- Jacobs JME, Van Eck HJ, Arens P et al (2004) A genetic map of potato (*Solanum tuberosum*) integrating molecular markers, including transposons, and classical markers. *Theor Appl Genet* 91:289–300
- Jafar M, Aggarwal R, Ramesh B et al (2012) SNP markers and their impact on plant breeding. *Int J Plant Genom* 2012. <https://doi.org/10.1155/2012/728398>
- Janardhan KV, Nataraju SP, Gurumurthy BR et al (1994) Screening flue-cured tobacco genotypes to drought using line source technique. *Tob Res* 20(1):16–20
- Jarup L, Berglund M, Elinder CG et al (1998) Health effects of cadmium exposure—a review of the literature and a risk estimate. *Scand J Work Environ Health* 24:1–51
- Jarvis A, Lane A, Hijmans RJ (2008) The effect of climate change on crop wild relatives. *Agri Ecosyst Environ* 126:13–23
- Jenkins GI (2014) The UV-B photoreceptor UVR8: from structure to physiology. *Plant Cell* 26:21–37
- Jiang G-L (2013) Molecular markers and marker-assisted breeding in plants. In: Andersen SB (ed) *Plant breeding from laboratories to fields*. IntechOpen. <https://doi.org/10.5772/52583>
- Jingjing JIN (2018) Tobacco genomic resource. <https://www.coresta.org/abstracts/tobacco-genomic-resource-31854.html>. Accessed 30 Sept 2021
- Jinks JL, Jayasekara NEM, Boughey H (1977) Joint selection for both extremes of mean performance and of sensitivity to a macro—environmental variable. II Single seed descent. *Heredity* 39:345–355
- Johnson GR, Frey KJ (1967) Heritability of quantitative attributes of Oats at varying levels of environmental stress. *Crop Sci* 7:43–46
- Johnston S, Anziano P, Shark K et al (1988) Mitochondrial transformation in yeast by bombardment with microprojectiles. *Science* 240:1538–1541

- Jones KJ, Shew HD (1995) Early season root production and zoospore infection of cultivars of flue-cured tobacco that differ in level of partial resistance to *Phytophthora parasitica* var. *nicotianae*. *Plant Soil* 172(1):55–61
- Juan CM, Mi J, Shreya A, Kössler S et al (2020) Expression of a carotenogenic gene allows faster biomass production by redesigning plant architecture and improving photosynthetic efficiency in tobacco. *Plant J Cell Mol Biol*. <https://doi.org/10.1111/tpj.14909>
- Juan R, del CN (1986) Irrigation water management and chemical and physical characteristics of covered dark tobacco. *Riego y Drenaje* 9:71–83
- Jube S, Borthakur D (2007) Expression of bacterial genes in transgenic tobacco: methods, applications and future prospects. *Elect J Biotechnol* 10(3):452–467. <https://doi.org/10.2225/vol10-issue3-fulltext-4>
- Julio E, Denoyes-Rothan B, Verrier JL et al (2006) Detection of QTLs linked to leaf and smoke properties in *Nicotiana tabacum* based on a study of 114 recombinant inbred lines. *Mol Breed* 18:69–91
- Jung J, Kim JI, Yi G (2019) geneCo: a visualized comparative genomic method to analyze multiple genome structures. *Bioinformatics* 35(24):5303–5305. <https://doi.org/10.1093/bioinformatics/btz596>
- Kaličanin B, Velimirović D (2012) Potentiometric stripping analysis of zinc cadmium and lead in tobacco leaves (*Nicotiana Tabacum* L) and soil samples. *Intl J Electrochem Sci* 7:313–323
- Kamalay JC, Goldberg RB (1980) Regulation of structural gene expression in tobacco. *Cell* 19:935–946. [https://doi.org/10.1016/0092-8674\(80\)90085-9](https://doi.org/10.1016/0092-8674(80)90085-9)
- Kami C, Lorrain S, Hornitschek P et al (2010) Light-regulated plant growth and development. *Curr Topics Dev Biol* 91:29–66
- Karaivazoglou NA, Papakosta DK, Divanidis S (2005) Effect of chloride in irrigation water and form of nitrogen fertilizer on Virginia (flue cured) tobacco. *Field Crops Res* 92:61–74
- Karaivazoglou NA, Tsotsolis NC, Tsadilas CD (2007) Influence of liming and form of Sci 100 nitrogen fertilizer on nutrient uptake growth yield and quality of Virginia (flue-cured) tobacco. *Field Crops* 100:52–60
- Kasuga M, Miura S, Shinozaki K et al (2004) Combination of the Arabidopsis DREB1A gene and stress-inducible rd29A promoter improved drought- and low-temperature stress tolerance in tobacco by gene transfer. *Plant Cell Physiol* 45:346–350
- Kenton A, Parokonny AS, Gleba YY et al (1993) Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. *Mol Gen Genet* 240:159–169
- Khan SH (2019) Genome-editing technologies: concept, pros and cons of various genome-editing techniques and bioethical concerns for clinical application. *Mol Ther Nucleic Acids* 16:326–334. <https://doi.org/10.1016/j.omtn.2019.02.027>
- Kim JW, Shon JY, Lee CK et al (2011) Relationship between grain filling and leaf senescence of temperate rice under high temperature. *Field Crop Res* 122:207–213
- King MJ (1990) Tobacco. In: Stewart BA, Nielsen DR. *Irrigation of agricultural crops*. Agronomy Series, vol 30. American Society of Agronomy Inc, Madison (WI), pp 811–833
- Knapp S, Chase MW, Clarkson JJ (2004) Nomenclatural changes and a new sectional classification in *Nicotiana* (Solanaceae). *Taxon* 52:73–82
- Kole C (2017) Aims and scope of the volume, genomic designing of climate smart cereal crops. Proposal submitted to Springer Nature.
- Kolodyazhnaya YS, Kutsokov NK, Levenko BA et al (2009) Transgenic plants tolerant to abiotic stresses. *Cytol Genet* 43:132–149
- Komarnitsky SI, Komarnitsky IK, Cox A et al (1998) Molecular phylogeny of the nuclear 5.8S ribosomal RNA genes in 37 species of *Nicotiana* genus. *Genetika* 34:883–889
- Kostoff D (1943) An androgenic *Nicotiana* haploid. *Zeits. F Zelloysch U mikr Anat* 9:640–642
- Kramer PJ (1951) Causes of injury to plants resulting from flooding of the soil. *Plant Physiol* 26:722–736
- Kramer PJ, Jackson WT (1954) Causes of injury to flooded tobacco plants. *Plant Physiol* 2(3):241–245



- Kuchuk N, Sytnyk K, Vasylenko M et al (2006) Genetic transformation of plastids of different Solanaceae species using tobacco cells as organelle hosts. *Theor Appl Genet* 113:519–527. <https://doi.org/10.1007/s00122-006-0318-y>
- Kwak SY, Lew TTS, Sweeney CJ et al (2019) Chloroplast-selective gene delivery and expression in planta using chitosan-complexed single-walled carbon nanotube carriers. *Nat Nanotechnol* 14:447–455. <https://doi.org/10.1038/s41565-019-0375-4>
- Lambers H, Chapin FS III, Pons TL (2008) *Plant physiological ecology*, 2nd edn. Springer, New York
- Lander ES, Green P, Abrahamson J et al (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1(2):174–181. [https://doi.org/10.1016/0888-7543\(87\)90010-3](https://doi.org/10.1016/0888-7543(87)90010-3)
- Lecours N, Almeida GEG, Abdallah JM et al (2012) Environmental health impacts of tobacco farming: a review of the literature. *Tob Control* 21:191–196
- Legg PD, Litton CC, Collins GB (1981) Effects of the *Nicotiana debneyi* black root rot resistance factor on agronomic and chemical traits in burley tobacco. *Theor Appl Genet* 60:365–368
- Legg PD, Smeeton BW (1999) Breeding and genetics. In: David DL, Nielsen MT (eds) *Tobacco production, chemistry and technology*. Blackwell Science, Malden, MA, USA
- Leitch IJ, Hanson L, Lim KY et al (2008) The ups and downs of genome size evolution in polyploid species of *Nicotiana* (Solanaceae). *Ann Bot* 101:805–814
- Lester RN, Hawkes JG (2001) Solanaceae. In: Hanelt P (ed) *Mansfeld's encyclopedia of agricultural and horticultural crops (except ornamentals)*, vol 4. Springer Berlin
- Levitt J (1972) *Responses on plants to environmental stresses*. Academic Press, New York, NY
- Lewis RS (2005) Transfer of resistance to potato virus Y (PVY) from *Nicotiana africana* to *Nicotiana tabacum*: possible influence of tissue culture on the rate of introgression. *Theor Appl Genet* 110:678–687
- Lewis RS (2011) *Nicotiana*. In: Kole C (ed) *Wild crop relatives: genomic and breeding resources, plantation and ornamental crops*. Springer, Berlin, Heidelberg
- Li S, Chang L, Zhang J (2021) Advancing organelle genome transformation and editing for crop improvement. *Plant Commun* 2(2):100141. <https://doi.org/10.1016/j.xplc.2021.100141>
- Li XH, Xing XX, Xu SX et al (2018) Genome-wide identification and functional prediction of tobacco lncRNAs responsive to root-knot nematode stress. *PLoS One* 13(11):15
- Lidon FJC (2012) Micronutrients' accumulation in rice after supplemental UV-B irradiation. *J Plant Interact* 7:19–28
- Liharska TB, Koornneef M, van Wordragen M et al (1996) Tomato chromosome 6: effect of alien chromosomal segments on recombinant frequencies. *Genome* 39:485–491
- Li-Hua L, Mei-Xia C, Dong-Xin Z et al (2011) QTL analysis of six important traits in tobacco (*Nicotiana tabacum* L.). *Acta Agron Sin* 37:1577–1584
- Lim KY, Kovarik A, Matyasek R et al (2007) Sequence of events leading to near complete genome turnover in allopolyploid *Nicotiana* within five million years. *New Phytol* 175:756–763
- Lim KY, Matyasek R, Kovarik A et al (2004) Genome evolution in allotetraploid *Nicotiana*. *Biol Linn Soc* 82:599–606
- Lim KY, Matyasek R, Lichtenstein CP, Leitch AR et al (2000) Molecular cytogenetic analyses and phylogenetic studies in the *Nicotiana* section *Tomentosae*. *Chromosoma* 109:245–258
- Lim KY, Matyasek R, Kovarik A et al (2005) Molecular cytogenetics and tandem repeat sequence evolution in the allopolyploid *Nicotiana rustica* compared with diploid progenitors *N. paniculata* and *N. undulate*. *Cytogenet Genome Res* 109:298–309. <https://doi.org/10.1159/000082413>
- Lin TY, Kao YY, Lin S et al (2001) A genetic linkage map of *Nicotiana plumbaginifolia/Nicotiana longiflora* based on RFLP and RAPD markers. *Theor Appl Genet* 103:905–911
- Liu LH, Shao ZY, Zhang M et al (2015) Regulation of carotenoid metabolism in tomato. *Mol Plant* 8:28–39
- Liu XZ, He CS, Yang YM et al (2009) Genetic diversity among flue-cured tobacco cultivars on the basis of AFLP markers. *Czech J Genet Plant Breed* 45(4):155–159

- Liu XZ, Zhang HY (2008) Advance of molecular marker application in the tobacco research. *Afr J Biotechnol* 7(25):4827–4831
- Livingstone KD, Lackney VK, Blauth JR et al (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the Solanaceae. *Genetics* 152:1183–1202
- Long SP, Marshall-Colon A, Zhu XG (2015) Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* 161(1):56–66. <https://doi.org/10.1016/j.cell.2015.03.019>
- Lu JP (1997) The application of PAGE in the cultivars identification of flue-cured tobacco. *Seed* 5:30–32
- Lu XP, Gui YJ, Xiao BG et al (2012) Development of DArT markers for a linkage map of flue-cured tobacco. *Chin Sci Bull* 58:641–648
- Lu Y, Stegemann S, Agrawal S et al (2017) Horizontal transfer of a synthetic metabolic pathway between plant species. *Curr Biol* 19:3034–3041
- Lu YH, Rijzaani H, Karcher D et al (2013) Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. *Proc Natl Acad Sci USA* E623–E632
- Lücker J, Schwab W, Hautum B et al (2004) Increased and altered fragrance of tobacco plants after metabolic engineering using three monoterpene synthases from lemon. *Plant Physiol* 134:510–519. <https://doi.org/10.1104/pp.103.030189>
- Ludlow MM, Muchow RC (1990) A critical evaluation of traits for improving crop yields in water-limited environments. In: Brady NC (ed) *Advances in agronomy*, vol 43. Academic Press, pp 107–153
- Lugon-Moulin N, Martin F, Krauss MR et al (2008) Arsenic concentration in tobacco leaves: a study on three commercially important tobacco (*Nicotiana tabacum* L) types. *Water Air Soil Pollut* 192:315–319
- Lugon-Moulin N, Ryan L, Donini P et al (2006) Cadmium content of phosphate fertilizers used for tobacco production. *Agron Sustain Dev* 26:151–155
- Ma HB, Qi JM, Li YK et al (2008) Construction of a molecular genetic linkage map of tobacco based on SRAP and ISSR markers. *Acta Agron Sin* 34:1958–1963
- Macková H, Hronková M, Dobrá J et al (2013) Enhanced drought and heat stress tolerance of tobacco plants with ectopically enhanced cytokinin oxidase/dehydrogenase gene expression. *J Exp Bot* 64(10):2805–2815. <https://doi.org/10.1093/jxb/ert131>
- Madhav MS, Raju KS, Gaikwad K et al (2015) Development of new set of microsatellite markers in cultivated tobacco and their transferability in other *Nicotiana* spp. *Mol Plant Breed* 6(16):1–13
- Maliga P, Svab Z (2010) Engineering the plastid genome of *Nicotiana sylvestris*, a diploid model species for plastid genetics. In: Birchler J (ed) *Plant chromosome engineering: methods and protocols*. Humana Press, Totowa. [https://doi.org/10.1007/978-1-61737-957-4\\_2](https://doi.org/10.1007/978-1-61737-957-4_2)
- Manly KF, Cudmore RH Jr, Meer JM (2001) Map Manager QTX, cross-platform software for genetic mapping. *Mamm Genome* 12(12):930–932. <https://doi.org/10.1007/s00335-001-1016-3>
- Matsuoka K, Demura T, Galis I et al (2004) A comprehensive gene expression analysis toward the understanding of growth and differentiation of tobacco BY-2 cells. *Plant Cell Physiol* 45(9):1280–1289. <https://doi.org/10.1093/pcp/pch155>
- Matyasek R, Lim KY, Kovarik A et al (2003) Ribosomal DNA evolution and gene conversion in *Nicotiana rustica*. *Heredity* (Edinb) 91(3):268–275. <https://doi.org/10.1038/sj.hdy.6800333>
- McEvoy ET (1957) The growth and mineral content of flue-cured tobacco as influenced by reaction of nutrient solutions with ionic forms of nitrogen. *Can J Soil Sci* 37:79–83
- McKee CG (1968) Maryland 64, a high yielding, later maturing variety of Maryland tobacco. *Md Agrl Exp Sta Misc Publ* 654:4
- Meharg AA, Macnair MR (1992) Suppression of the high affinity phosphate uptake system: a mechanism of arsenate tolerance in *Holcus lanatus* L. *J Exp Bot* 43:519–524
- Mehmood F, Abdullah, Ubaid Z et al (2020) Plastid genomics of *Nicotiana* (Solanaceae): insights into molecular evolution, positive selection and the origin of the maternal genome of Aztec tobacco (*Nicotiana rustica*) *Peer J* 8:e9552. <https://doi.org/10.7717/peerj.9552>

- Meissner R, Chague V, Zhu Q et al (2000) Technical advance: a high throughput system for transposon tagging and promoter trapping in tomato. *Plant J* 22:265–274. <https://doi.org/10.1046/j.1365-313x.2000.00735.x>
- Menser HA, Street OE (1962) Effects of air pollution, nitrogen levels, supplemental irrigation, and plant spacing on weather fleck and leaf losses of Maryland tobacco. *Tob Sci* 6:165–169
- Mielby H (2011) Public attitudes to Cisgenic crops. PhD thesis, University of Copenhagen
- Milla SR, Levin JS, Lewis RS et al (2005) RAPD and SCAR markers linked to an introgressed gene conditioning resistance to *Peronospora tabacina* D.B. Adam. in tobacco. *Crop Sci* 45:2346–2354
- Miransari M (2011) Hyperaccumulators arbuscular mycorrhizal fungi and stress of heavy metal. *Biotechnol Adv* 29:645–653. <https://doi.org/10.1016/j.biotechadv.2011.04.006>
- Moon HS, Nicholson JS, Heineman A et al (2009a) Changes in genetic diversity of U.S. flue-cured tobacco germplasm over seven decades of cultivar development. *Crop Sci* 49(2):498–508. <https://doi.org/10.2135/cropsci2008.05.0253>
- Moon HS, Nicholson JS, Lewis RS (2008) Use of transferable *Nicotiana tabacum* L. microsatellite markers for investigating genetic diversity in the genus *Nicotiana*. *Genome* 51:547–559
- Moon HS, Nifong JM, Nicholson JS et al (2009b) Microsatellite-based analysis of tobacco (*Nicotiana tabacum* L.) genetic resources. *Crop Sci* 49:1–11. <https://doi.org/10.2135/cropsci2009.01.0024>
- Morgan JM (1984) Osmoregulation and water stress in higher plants. *Annu Rev Plant Physiol* 35:299–319
- Morris M, Dreher K, Ribaut JM (2003) Money matters (II): costs of maize inbred line conversion schemes at CIMMYT using conventional and marker-assisted selection. *Mol Breed* 11:235–247. <https://doi.org/10.1023/A:1022872604743>
- Munns R (2005) Genes and salt tolerance: bringing them together. *New Phytol* 167(3):645–663
- Munns R, Tester M (2008) Mechanisms of salinity tolerance. *Annu Rev Plant Biol* 59:651–681
- Murad L, Lim KY, Christopodulov V et al (2002) The origin of tobacco's T genome is traced to a particular lineage within *Nicotiana tomentosiformis* (Solanaceae). *Amer J Bot* 89:921–928
- Muttaqin AS, Suarma U, Nurjani E et al (2019) The impact of climate variability on tobacco productivity over Temanggung Regency, Indonesia. *E3S Web Confer* 76:04003. <https://doi.org/10.1051/e3sconf/20197604003>. ICST 2018
- Myles S, Peiffer J, Brown PJ et al (2009) Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* 21:2194–2202. <https://doi.org/10.1105/tpc.109.068437>
- Narayan RK (1987) Nuclear DNA changes, genome differentiation and evolution in *Nicotiana* (Solanaceae). *Plant Syst Evol* 157(04):161–180
- Narayanan AI, Murthy TGK, Sarala K et al (2003) Division of crop improvement. Information bulletin. Central Tobacco Research Institute, Rajahmundry
- Nascimento NC, Fett-Neto AG (2010) Plant secondary metabolism and challenges in modifying its operation: an overview. In: Fett-Neto A (ed) *Plant secondary metabolism engineering. Methods in molecular biology (methods and protocols)*, vol 643. Humana Press, Totowa
- Navabpour S, Morris K, Allen R et al (2003) Expression of senescence-enhanced genes in response to oxidative stress. *J Exp Bot* 54:2285–2292
- Ndunguru BJ, Ntare BR, Williams JH, Greenberg DC (1995) Assessment of groundnut cultivars for end-of-season drought tolerance in a Sahelian environment. *J Agri Sci* 125:79–85
- Niazian M, Sadatnoori SA, Galuszka P et al (2017) Tissue culture-based Agrobacterium-mediated and in Planta transformation methods. *Czech J Genet Plant Breed* 53:133–143. <https://doi.org/10.17221/177/2016-CJGPB>
- Nisar N, Li L, Lu S et al (2015) Carotenoid metabolism in plants. *Mol Plant* 8:68–82
- Nishi T, Tajima T, Noguchi S et al (2003) Identification of DNA markers of tobacco linked to bacterial wilt resistance. *Theor Appl Genet* 106(4):765–770. <https://doi.org/10.1007/s00122-002-1096-9>
- Nishiyama R, Watanabe Y, Fujita Y et al (2011) Tran. LSP Analysis of cytokinin mutants and regulation of cytokinin metabolic genes reveals important regulatory roles of cytokinins in drought,

- salt and ABA responses, and ABA biosynthesis. *Plant Cell* 23:2169–2183. <https://doi.org/10.1105/tpc.111.087395>
- Nobel PS (1980) Leaf anatomy and water use efficiency. In: Turner NC, Kramer PJ (eds) *Adaptation of plants to water and high temperature stress*. Wiley, New York, pp 43–55
- Nordberg G, Nogawa K, Nordberg M et al (2007) *Handbook on toxicology of metals*. Academic Press, New York
- Norom IC, Osibanjo O, Oji-nnorom CG (2005) Cadmium determination in cigarettes available in Nigeria. *Afr J Biotechnol* 4:1128–1132
- Nurhidayati T, Wardhani SP, Purnobasuki H et al (2017) Response morphology and anatomy of tobacco (*Nicotiana tabacum* L.) plant on waterlogging. *Conf Proc* 1908:040009. <https://doi.org/10.1063/1.5012723>
- Oeung S, Voleak N, Huykhim U et al (2017) Phytochemical analysis of different extracts of leaves of *Nicotiana tabacum* L of Cambodia. *Asian J Pharm* 1(3):18–26
- Okumus A, Gulumser A (2001) Chiasma frequency chromosome pairing at tobacco genotypes in the Middle Black Sea region. *Online J Biol Sci* 1(7):612–613
- Opperman CH, Lommel SA, Sosinski BR et al (2003) The tobacco genome initiative Paper presented at CORESTA Meeting, Agronomy/Phytopathology, Bucharest, Romania, 12–17 Oct 2003
- Ort DR, Merchant SS, Alric J et al (2015) Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc Natl Acad Sci USA*. <https://doi.org/10.1073/pnas.1424031112>
- Pallardy SG (2010) *Physiology of woody plants*. Academic Press, London
- Passioura JB (1983) Roots and drought resistance. *Agri Water Manage* 7:265–280
- Paterson AH, DeVerna JW, Lanini B et al (1990) Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. *Genetics* 124(3):735–742. <https://doi.org/10.1093/genetics/124.3.735>
- Pereira JS, Chaves MM (1993) Plant water deficits in Mediterranean ecosystems. In: Smith JAC, Griffiths H (eds) *Plant responses to water deficits—from cell to community*. BIOS Scientific Publishers, Oxford, pp 237–251
- Perez-Alfocea F, Balibrea ME et al (1996) Agronomical and physiological characterization of salinity tolerance in a commercial tomato hybrid. *Plant Soil* 180:251–257
- Pilon-Smits EAH, Terry N, Sears T et al (1998) Trehalose-producing transgenic tobacco plants show improved growth performance under drought stress. *J Plant Physiol* 152:525–532
- Popov VN, Antipina OV, Burakhanova EA (2013) Involvement of cell-wall invertase in low-temperature hardening of tobacco plants. *Russ J Plant Physiol* 2:221–226
- Povilaitis B, White FH (1966) Delcrest 66, a new variety of bright tobacco. *Can J Plant Sci* 46:458
- Prabhakararao K, Sarala K, Murthy TGK (2016) Patent mapping of disease resistance genes in tobacco. *J Agri Crop Sci* 3:01–07
- Prabhakararao K, Sarala K, Murthy TGK et al (2012) Application of Simple Sequence Repeat (SSR) markers in diversity analysis of tobacco types. *Tob Res* 38(1):29–33
- Pritchard JK, Stephens M, Rosenberg NA et al (2000) Association mapping in structured populations. *Am J Hum Genet* 67(1):170–181
- Puchta H (1999) Use of I-Sce I to induce DNA double-strand breaks in *Nicotiana*. *Methods Mol Biol* 113:447–451
- Purrington CB (2000) Costs of resistance. *Curr Opin Plant Biol* 3:305–308
- Qi JM, Wang T, Chen SH et al (2006) Genetic diversity and genetic relatives analysis of tobacco germplasm based on inter-simple sequence repeat (ISSR). *Acta Agron Sin* 32:373–378
- Rabara RC, Tripathi P, Reese RN et al (2015) Tobacco drought stress responses reveal new targets for Solanaceae crop improvement. *BMC Genomics* 16:484
- Rabara RC, Tripathi P, Rushton PJ (2017) Comparative metabolome profile between tobacco and soybean grown under water-stressed conditions. *BioMed Res Intl*. <https://doi.org/10.1155/2017/3065251>
- Rabbani MA, Maruyama K, Aba H et al (2003) Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiol* 133:1755–1767

- Rafalski JA (2010) Association genetics in crop improvement. *Curr Opin Plant Biol* 13:174–180. <https://doi.org/10.1016/j.pbi.2009.12.004>
- Rascon-Cruz Q, Gonzalez-Barriga CD, Blanca Flor Iglesias-Figueroa BF et al (2021) Plastid transformation: advances and challenges for its implementation in agricultural crops. *Electron J Biotechnol* 51:95–109. <https://doi.org/10.1016/j.ejbt.2021.03.005>
- Rastas P (2017) Lep-MAP3: robust linkage mapping even for low-coverage whole genome sequencing data. *Bioinformatics* 33(23):3726–3732. <https://doi.org/10.1093/bioinformatics/btx494>
- Rathinasabapathi B (2000) Metabolic engineering for stress tolerance: installing osmoprotectant synthesis pathways. *Ann Bot* 86:709–716
- Reddy AM, Kumar SG, Jyonthsnakumari G et al (2005) Lead induced changes in antioxidant metabolism of horsegram (*Macrotyloma uniflorum* (Lam) Verdc) and bengalgram (*Cicer arietinum* L). *Chemosphere* 60:97–104
- Regassa G, Chandravanshi B (2016) Levels of heavy metals in the raw and processed Ethiopian tobacco leaves. *Springer Plus* 5:232
- Remacle C, Cardol P, Coosemans N et al (2006) High-efficiency biolistic transformation of *Chlamydomonas* mitochondria can be used to insert mutations in complex I genes. *Proc Natl Acad Sci USA* 103:4771–4776
- Ren M, Zhang C, Jiang C (2014) Association analysis of tobacco aroma constituents based on high density SSR linkage group. *Acta Tab Sin* 20:88–93
- Ren N, Timko MP (2001) AFLP analysis of genetic polymorphism and evolutionary relationships among cultivated and wild *Nicotiana* species. *Genome* 44:559–571
- Renny-Byfield S, Michael C, Ales et al (2011) Next generation sequencing reveals genome downsizing in allotetraploid *Nicotiana tabacum*, predominantly through the elimination of paternally derived repetitive DNAs. *Mol Biol Evol* 28:2843–2854
- Rhodes D, Samaras Y (1994) Genetic control of osmoregulation in plants. In: Strange K (ed) *Cellular and molecular physiology*. CRC Press, Boca Raton, pp 347–367
- Ribaut J-M, Hoisington D (1998) Marker-assisted selection: new tools and strategies. *Trends Plant Sci* 3:236–239. [https://doi.org/10.1016/S1360-1385\(98\)01240-0](https://doi.org/10.1016/S1360-1385(98)01240-0)
- Rivero RM, Shulaev V, Blumwald E (2009) Cytokinin-dependent photorespiration and the protection of photosynthesis during water deficit. *Plant Physiol* 150:1530–1540. <https://doi.org/10.1104/pp.109.139378>
- Robertson MJT, Fukai S, Peoples MB (2004) The effect of timing and severity of water deficit on growth, development, yield accumulation and nitrogen fixation of mung bean. *Field Crops Res* 86:67–80
- Rolland F, Baena-Gonzalez E, Sheen J (2006) Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annu Rev Plant Biol* 57:675–709
- Rommens CM, Humara JM, Ye J et al (2004) Crop improvement through modification of the plant's own genome. *Plant Physiol* 135:421–431. <https://doi.org/10.1104/pp.104.040949>
- Rontein D, Basset G, Hanson AD (2002) Metabolic engineering of osmoprotectant accumulation in plants. *Metab Eng* 4:49–56
- Rossi L, Bindler G, Pijnenburg H et al (2001) Potential of molecular marker analysis for variety identification in processed tobacco. *Plant Var Seeds* 14:89–101
- Rostoks N, Borevitz J, Hedley P et al (2005) Single-feature polymorphism discovery in the barley transcriptome. *Genome Biol* 6:R54. <https://doi.org/10.1186/gb-2005-6-6-r54>
- Rushton PJ, Bokowiec MT, Han S et al (2008) Tobacco transcription factors: novel insights into transcriptional regulation in the Solanaceae. *Plant Physiol* 147(1):280–295
- Russell AW, Sparrow R (2008) The case for regulating intragenic GMOs. *J Agri Environ Ethic* 21:153–181. <https://doi.org/10.1007/s10806-007-9074-5>
- Sarala K (2019) Next gen breeding strategies for enhancing FCV Tobacco productivity and quality. Paper presented at national symposium on approaches and strategies for augmenting tobacco farmers' income, Indian Society of Tobacco Sciences, Rajahmundry, 19–20 July 2019

- Sarala K, Baghyalakshmi K, Prabhakara Rao K et al (2018) Genetic diversity among mutant germplasm accessions of *Nicotiana tabacum* as determined by morphological parameters. *Tob Res* 44(2):47–53
- Sarala K, Murthy TGK, Krishnamurthy V (2012) Tobacco (*Nicotiana tabacum* L.). In: Bharadwaj DN (ed) Breeding of field crops. Agrobios (India), Jodhpur, India
- Sarala K, Nanda C, Baghyalakshmi K et al (2021) Designing tobacco genomes for resistance to biotic stresses. In: Kole C (ed) Genomic designing for biotic stress resistant crops. Springer International Publishing, Switzerland (in press)
- Sarala K, Prabhakara Rao K, Murthy TGK et al (2016) Yield gains in Indian flue-cured virginia tobacco: a case study. *Int J Agric Sci* 8(52):2538–2541
- Sarala K, Rao RVS (2008) Genetic diversity in Indian FCV and burley tobacco cultivars. *J Genet* 87:159–163
- Sarala K, Reddy DD, Rao KP (2020) FCJ 11: a high yielding FCV tobacco cultivar for Northern light soils of Andhra Pradesh. Central Tobacco Research Institute, Rajahmundry. Doc No: D1/Pub/FL2
- Sarala K, Swamy AVSR, Nageswara Rao K et al (1998) Identifying drought tolerant FCV tobacco genotypes. Paper presented at the national symposium on Indian Tobacco-Problems and Prospects, ISTS, 20–23 January 1998
- Sauter A, Davies WJ, Hartung W (2001) The long-distance abscisic acid signal in the droughted plant: the fate of the hormone on its way from root to shoot. *J Exp Bot* 52:1991–1997
- Schaart JG, Krens FA, Pelgrom KTB et al (2004) Effective production of marker-free transgenic strawberry plants using inducible site-specific recombination and a bifunctional selectable marker gene. *Plant Biotechnol J* 2:233–240. <https://doi.org/10.1111/j.1467-7652.2004.00067.x>
- Schouten HJ, Krens FA, Jacobsen E (2006) Cisgenic plants are similar to traditionally bred plants: international regulations for genetically modified organisms should be altered to exempt cisgenesis. *EMBO Rep* 7:750–753
- Soccianti V, Crinelli R, Tirillini B et al (2006) Uptake and toxicity of Cr (Cr3+) in celery seedlings. *Chemosphere* 64:1695–1703
- Seki M, Kamei A, Yamaguchi-Shinozaki K et al (2003) Molecular responses to drought, salinity and frost: common and different paths for plant protection. *Curr Opin Biotechnol* 14:194–199
- Senthil-Kumar M, Hema R, Suryachandra TR et al (2010) Functional characterization of three water deficit stress-induced genes in tobacco and Arabidopsis: an approach based on gene down regulation. *Plant Physiol Biochem* 48(1):35–44. <https://doi.org/10.1016/j.plaphy.2009.09.005>
- Sharma P, Dubey RS (2005) Lead toxicity in plants. *Braz J Plant Physiol* 17:35–52
- Sharma V, Goel P, Kumar S et al (2019) An apple transcription factor, MdDREB76, confers salt and drought tolerance in transgenic tobacco by activating the expression of stress-responsive genes. *Plant Cell Rep* 38:221–241
- Shen J, Jiang CQ, Yan YF et al (2017) Effect of increased UV-B radiation on carotenoid accumulation and total antioxidant capacity in tobacco (*Nicotiana tabacum* L.) leaves. *Genet Mol Res* 16(1). <https://doi.org/10.4238/gmr16018438>. PMID: 28290611
- Shilpim M, Narendra T (2005) Cold, salinity and drought stresses: an overview. *Arch Biochem Biophys* 444:139–158
- Shinozaki K, Ohme M, Tanaka M et al (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J* 5(9):2043–2049
- Shinozaki K, Yamaguchi-Shinozaki K (1996) Molecular responses to drought and cold stress. *Curr Opin Biotechnol* 7:161–167. [https://doi.org/10.1016/S0958-1669\(96\)80007-3](https://doi.org/10.1016/S0958-1669(96)80007-3)
- Sierro N, Battey JN, Ouadi S et al (2013a) Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. *Genome Biol* 4(6):R60. <https://doi.org/10.1186/gb-2013-14-6-r60>
- Sierro N, Battey JND, Bovet L et al (2018) The impact of genome evolution on the allotetraploid *Nicotiana rustica*—an intriguing story of enhanced alkaloid production. *BMC Genomics* 19(1):855. <https://doi.org/10.1186/s12864-018-5241-5>
- Sierro N, Battey JND, Ouadi S et al (2014) The tobacco genome sequence and its comparison with those of tomato and potato. *Nat Commun* 5. <https://doi.org/10.1038/ncomms4833>



- Sierro N, Oeveren JV, Eijk MJTV et al (2013b) Whole genome profiling physical map and ancestral annotation of tobacco Hicks Broadleaf. *Plant J* 75(5):880–889. <https://doi.org/10.1111/tpj.12247>
- Silva AT, Nguyen A, Ye C et al (2010) Conjugated polymer nanoparticles for effective siRNA delivery to tobacco BY-2 protoplasts. *BMC Plant Biol* 10:291. <https://doi.org/10.1186/1471-2229-10-291>
- Singh S, Parihar P, Singh R et al (2016) Heavy metal tolerance in plants: role of transcriptomics proteomics metabolomics and ionomics. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2015.01143>
- Sivaraju K, Madhav MS, Sharma RK et al (2008) Molecular diversity in Indian tobacco types as revealed by randomly amplified DNA polymorphism. *J Plant Biochem Biotechnol* 17:51–56
- Smith KM (1937) A textbook of plant virus diseases. P Blakiston's Son & Co Inc., Philadelphia
- South PF, Cavanagh AP, Liu HW et al (2019) Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science* 363(6422):eaat9077. <https://doi.org/10.1126/science.aat9077>
- Srivastava S, Chaudhry V, Mishra A et al (2012) Gene expression profiling through microarray analysis in *Arabidopsis thaliana* colonized by *Pseudomonas putida* MTCC5279 a plant growth promoting rhizobacterium. *Plant Signal Behav* 7:235–245. <https://doi.org/10.4161/psb.18957>
- Stam P, Zeven C (1981) The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. *Euphytica* 30:227–238
- Steward R (2018) Economic impact of disclosure requirements in patent applications for 'genetic resources'-based innovation. International Federation of Pharmaceutical Manufacturers and Associations (IFPMA), Geneva Switzerland and Crop Life International, Brussels, Belgium
- Stojanovic D, Nikic D, Lazarevic K (2004) The level of nickel in smoker's blood and urine. *Cent Eur J Public Health* 12:187–189
- Suen DF, Wang CK, Lin RF et al (1997) Assignment of DNA markers to *Nicotiana sylvestris* chromosomes using monosomic alien addition lines. *Theor Appl Genet* 94:331–337
- Suzuki N, Koussevitzky S, Mittler R et al (2012) ROS and redox signaling in the responses of plants to abiotic stress. *Plant Cell Environ* 35:259–270
- Svab Z, Hajdukiewicz P, Maliga P (1990) Stable transformation of plastids in higher plants. *Proc Natl Acad Sci USA* 87(21):8526–8530. <https://doi.org/10.1073/pnas.87.21.8526>
- Svab Z, Maliga P (1993) High-frequency plastid transformation in tobacco by selection for a chimeric aadA gene. *Proc Natl Acad Sci USA* 90(3):913–917. <https://doi.org/10.1073/pnas.90.3.913>
- Swiderska K, Reid H, Song Y et al (2011) The role of traditional knowledge and crop varieties in adaptation to climate change and food security in SW China, Bolivian Andes and coastal Kenya. <https://pubs.iied.org/sites/default/files/pdfs/migrate/G03338.pdf>. Accessed 30 Sept 2021
- Tan SL, Yang YJ, Huang W (2020) Moderate heat stress accelerates photoinhibition of photosystem I under fluctuating light in tobacco young leaves. *Photosynth Res* 144(3):373–382
- Tang C, Asseng S, Diatloff E, Rengel Z (2003) Modelling yield losses of aluminium-resistant and aluminium-sensitive wheat due to subsurface soil acidity: effects of rainfall, liming and nitrogen application. *Plant Soil* 254:349–360
- Tanksley SD, Ganai MW, Prince JP et al (1992) High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132(4):1141–1160. <https://doi.org/10.1093/genetics/132.4.1141-1160>
- Tarczynski MC, Jensen RG, Bohnert HJ (1992) Expression of a bacterial mtlD gene in transgenic tobacco leads to production and accumulation of mannitol. *Proc Natl Acad Sci USA* 89:2600–2604
- Tarczynski MC, Jensen RG, Bohnert HJ (1993) Stress protection of transgenic tobacco by production of the osmolyte mannitol. *Science* 259:508–510
- Tavakol E, Pakniyat H (2007) Evaluation of some drought resistance criteria at seedling stage in wheat (*Triticum aestivum* L.) Cultivars. *Pak J Biol Sci* 10:1113–1117. <https://doi.org/10.3923/pjbs.2007.1113.1117>
- Telem RS, Wani SH, Singh NB et al (2013) Cisgenics—a sustainable approach for crop improvement. *Curr Genom* 14(7):468–476. <https://doi.org/10.2174/13892029113146660013>

- Terashima I (1992) Anatomy of non-uniform photosynthesis. *Photosynth Res* 31:195–212. <https://doi.org/10.1007/BF00035537>
- Tezuka T, Tsutomu K, Toshiaki M et al (2010) Seven of eight species in *Nicotiana* section *Suaveolentes* have common factors leading to hybrid lethality in crosses with *Nicotiana tabacum*. *Ann Bot* 106(2):267–276. <https://doi.org/10.1093/aob/mcq114>
- Tezuka T, Marubashi W (2012) Genes in S and T Subgenomes are responsible for hybrid lethality in interspecific hybrids between *Nicotiana tabacum* and *Nicotiana occidentalis*. *PLoS One* 7(2). <https://doi.org/10.1371/journal.pone.0036204>
- Thimmegowda GC, Ramadoss SK, Kaikala V et al (2018) Whole genome resequencing of tobacco (*Nicotiana tabacum* L.) genotypes and high-throughput SNP discovery. *Mol Breed* 38:121
- Tian Y, Chen K, Li X et al (2020) Design of high-oleic tobacco (*Nicotiana tabacum* L.) seed oil by CRISPR-Cas9-mediated knockout of NtFAD2-2. *BMC Plant Biol* 20:233
- Tian Y, Liu X, Fan C et al (2021) Enhancement of tobacco (*Nicotiana tabacum* L.) seed lipid content for biodiesel production by CRISPR-Cas9-mediated knockout of NtAn1. *Front Plant Sci* 11:599474. <https://doi.org/10.3389/fpls.2020.599474>
- Tirani MM, Haghjou MM, Ismaili A (2019) Hydroponic grown tobacco plants respond to zinc oxide nanoparticles and bulk exposures by morphological, physiological and anatomical adjustments. *Funct Plant Biol* 46:360–375
- Toda E, Koiso N, Takebayashi A et al (2019) An efficient DNA- and selectable-marker-free genome-editing system using zygotes in rice. *Nat Plants* 5:363–368. <https://doi.org/10.1038/s41477-019-0386-z>
- Tong Z, Jiang S, He W et al (2021) Construction of high-density genetic map and QTL mapping in *Nicotiana tabacum* backcrossing BC4F3 population using whole-genome sequencing. *Czech J Genet Plant Breed* 57:102–112. <https://doi.org/10.17221/8/2021-CJGPB>
- Tong Z, Zhou J, Xiu Z (2020) Construction of a high-density genetic map with whole genome sequencing in *Nicotiana tabacum* L. *Genomics* 112:2028–2033. <https://doi.org/10.1016/j.ygeno.2019.11.015>
- Tong Z, Xiao B, Jiao F et al (2016) Large-scale development of SSR markers in tobacco and construction of a linkage map in flue-cured tobacco. *Breed Sci* 66(3):381–390. <https://doi.org/10.1270/jsbbs.15129>
- Tong ZJ, Yang ZM, Chen XJ et al (2012) Large-scale development of microsatellite markers in *Nicotiana tabacum* and construction of a genetic map of flue-cured tobacco. *Plant Breed* 131:674–680
- Torney F, Trewyn BG, Lin VSY et al (2007) Mesoporous silica nanoparticles deliver DNA and chemicals into plants. *Nat Nanotechnol* 2:295–300
- Townsend JA, Wright DA, Winfrey RJ et al (2009) High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature* 459:442–445
- Trindade LM, Horvath B, Bachem C et al (2003) Isolation and functional characterization of a stolon specific promoter from potato (*Solanum tuberosum* L.). *Gene* 303:77–87. [https://doi.org/10.1016/S0378-1119\(02\)01147-2](https://doi.org/10.1016/S0378-1119(02)01147-2)
- Tso TC (1990) Production, physiology, and biochemistry of tobacco plant. *Ideals, Inc. Beltsville, Maryland*, pp 322–323
- Tso TC, Chaplin JF, Adams JD et al (1983) Simple correlation and multiple regression among leaf and smoke characteristics of burley tobaccos. *Beitr Tabakforsch Intl* 11(03):141–150
- Tuberosa R (2012) Phenotyping for drought tolerance of crops in the genomics era. *Front Physiol* 3:347. <https://doi.org/10.3389/fphys.2012.00347>
- Tuberosa R, Sanguineti MC, Landi P et al (2002) Identification of QTLs for root characteristics in maize grown in hydroponics and analysis of their overlap with QTLs for grain yield in the field at two water regimes. *Plant Mol Biol* 48:697–712
- Turner NC (1986) Adaptation to water deficits: a changing perspective. *Aust J Plant Physiol* 13:175–190
- United Nations (2019) Goal 13: take urgent action to combat climate change and its impacts. <https://www.un.org/sustainabledevelopment/climate-change/>. Accessed 31 Sept 2021



- Upadhyay SK, Kumar J, Alok A et al (2013) RNA guided genome editing for target gene mutations in wheat. *G3 (Bethesda)* 3:2233–2238
- Vajpayee P, Tripathi RD, Rai UN, Ali MB, Singh SN (2000) Chromium accumulation reduces chlorophyll biosynthesis nitrate reductase activity and protein content in *Nymphaea alba* L. *Chemosphere* 41:1075–1082
- Van Ooijen JW (2006) Joinmap<sup>®</sup> 4, Software for the calculation of genetic linkage maps in experimental population. Plant Research International, Wageningen
- Van Ooijen JW, Voorrips RE (2001) JoinMap<sup>®</sup> version 3.0: software for the calculation of genetic linkage maps. Plant Research International, Wageningen
- Vardi V, Venkatrayulu C (2019) Assessment of heavy metals in water of selected places of Nellore coast of Andhra Pradesh through atomic absorption spectroscopy. *Intl J Life Sci.* <https://doi.org/10.21276/SSR-IJLS2019553>
- Varshney RK, Bohra A, Yu J et al (2021) Designing future crops: genomics-assisted breeding comes of age. *Trends Plant Sci* 26(6):631–649. <https://doi.org/10.1016/j.tplants.2021.03.010>
- Varshney RK, Terauchi R, McCouch SR (2014) Harvesting the promising fruits of genomics: applying genome sequencing technologies to crop breeding. *PLoS Biol* 12(6):e1001883. <https://doi.org/10.1371/journal.pbio.1001883>
- Verma S, Yadav S, Singh I (2010) Trace metal concentration in different Indian tobacco products and related health implications. *Food Chem Toxicol* 48:2291–2297
- Viswanath V, Strauss SH (2010). Modifying plant growth the cisgenic way. *Inform Syst Biotechnol. News Rep* 2010:1–4
- Wagner GJ (1991) Secreting glandular trichomes: more than just hairs. *Plant Physiol* 96:675–679
- Wang H, Zhao P, Hölscher D et al (2012) Nighttime sap flow of *Acacia mangium* and its implications for nighttime transpiration and stem water storage. *J Plant Ecol* 5(3):294–304
- Wang RW, Cai X, Hu S et al (2020) Comparative analysis of the mitochondrial genomes of *Nicotiana tabacum*: Hints toward the key factors closely related to the cytoplasmic male sterility mechanism. *Front Genet* 11:257. <https://doi.org/10.3389/fgene.2020.00257>
- Wang W, Vinocur B, Altman A (2003) Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta* 28:1–14
- Wang X, Bennetzen JL (2015) Current status and prospects for the study of *Nicotiana* genomics, genetics, and nicotine biosynthesis genes. *Mol Genet Genom* 290(1):11–21. <https://doi.org/10.1007/s00438-015-0989-7>
- Wang X, Yang S, Chen Y et al (2018) Comparative genome-wide characterization leading to simple sequence repeat marker development for *Nicotiana*. *BMC Genomics* 19(1):500. <https://doi.org/10.1186/s12864-018-4878-4>
- Wang Y, Jiang J, Zhao X et al (2006) A novel LEA gene from *Tamarix androssowii* confers drought tolerance in transgenic tobacco. *Plant Sci* 171(6). <https://doi.org/10.1016/j.plantsci.2006.06.011>
- Wang Y, Lv H, Xiang X et al (2021) Construction of a SNP fingerprinting database and population genetic analysis of cigar tobacco germplasm resources in China. *Front Plant Sci.* <https://doi.org/10.3389/fpls.2021.618133>
- Wang Y, Sun C, Xu C (2016) Preliminary experiments on nano-silver against tobacco mosaic virus and its mechanism. *Tob Sci Technol* 49:22–30
- Wania SH, Kumar V, Shriram V et al (2016) Phytohormones and their metabolic engineering for abiotic stress tolerance in crop plants. *Crop J* 4:162–176
- Waugh R, Jannik JL, Muehlbauer GJ et al (2009) The emergence of whole genome association scans in barley. *Curr Opin Plant Biol* 12:218–222. <https://doi.org/10.1016/j.pbi.2008.12.007>
- Wei W, Zhang Y, Han L, Guan Z, Chai T (2008) A novel WRKY transcriptional factor from *Thlaspi caerulescens* negatively regulates the osmotic stress tolerance of transgenic tobacco. *Plant Cell Rep* 27(4):795–803
- Weil C, Schaart J, Niks R et al (2010) Traditional plant breeding methods. Wageningen UR Plant Breeding report 338, Wageningen

- Wenping LI, Sabaghnia N, Ranjbar R et al (2009) Correlation and path coefficient analysis and euclidean distance clustering for several characters in tobacco germplasm resource. *Chin Tob Sci* 30:59–63
- Wilson D (1975) Stomatal diffusion resistances and leaf growth during droughting of *Lolium perenne* plants selected for contrasting epidermal ridging. *Ann Appl Biol* 79(1):83–94
- Worland AJ, Gale MD, Law CN (1987) Wheat genetics. In: Lupton FJH (ed) *Wheat breeding: its scientific principals*. Chapman and Hill, London
- Wu FN, Eannetta NT, Xu Y et al (2010) COSII genetic maps of two diploid *Nicotiana* species provide a detailed picture of synteny with tomato and insights into chromosome evolution in tetraploid *N. tabacum*. *Theor Appl Genet* 120(4):809–827. <https://doi.org/10.1007/s00122-009-1206-z>
- Wu FN, Eannetta NT, Xu YM et al (2009) A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus *Capsicum*. *Theor Appl Genet* 118:1279–1293
- Xi L, Xiong S, Ren Y et al (2011) Virtual visualization system for growth of tobacco root. Paper presented at 4th conference on computer and computing technologies in agriculture (CCTA), Oct 2010, Nanchang, China
- Xiang D, Yao Z, Yanhua LY et al (2017) Analysis on solanesol content and genetic diversity of Chinese flue-cured tobacco (*Nicotiana tabacum* L.). *Crop Sci* 57(2):847–855. <https://doi.org/10.2135/cropsci2016.07.0568>
- Xiang XH, Wu XR, Chao JT et al (2016) Genome-wide identification and expression analysis of the WRKY gene family in common tobacco (*Nicotiana tabacum* L.). *Yi Chuan* 38(9):840–856
- Xiao B, Yuntao TY, Ni LN et al (2015) SNP-based genetic linkage map of tobacco (*Nicotiana tabacum* L.) using next-generation RAD sequencing. *J Biol Res (Thessalon)* 22:11. <https://doi.org/10.1186/s40709-015-0034-3>
- Xiao BG, Xu ZL, Chen XJ et al (2006) Genetic linkage map constructed by using a DH population for the flue-cured tobacco. *Acta Tab Sin* 12:35–40
- Xu MH, Zheng MH, Liu GT (1998) Genetic diversity in cultivars of *Nicotiana tabacum* revealed by RAPD markers. *J Agril Biotechnol* 6:281–285
- Yamori W, Evans JR, Von CS (2010) Effects of growth and measurement light intensities on temperature dependence of CO<sub>2</sub> assimilation rate in tobacco leaves. *Plant Cell Environ* 3:332
- Yang BC, Xiao B, Chen XJ et al (2007) Assessing the genetic diversity of tobacco germplasm using inter simple sequence repeat and inter-retrotransposon amplification polymorphism markers. *Ann Appl Biol* 150:393–401
- Yang BC, Xiao BG, Chen XJ et al (2005a) Genetic diversity of flue-cured tobacco varieties based on ISSR markers. *Hereditas (beijing)* 27(5):753–758
- Yang L, Li J, Ji J et al (2016) High temperature induces expression of tobacco transcription factor NtMYC2a to regulate nicotine and JA biosynthesis. *Front Physiol*. <https://doi.org/10.3389/fphys201600465>
- Yang LY, Shuang LY, Jun YL et al (2018) Effects of different growth temperatures on growth development and plastid pigments metabolism of tobacco (*Nicotiana tabacum* L) plants. *Bot Stud* 59 (5). <https://doi.org/10.1186/s40529-018-0221-2>
- Yang S, Li D, Fisher AJ (2019) CORESTA guide no. 16 molecular markers in tobacco. CORESTA, France
- Yang X, Feng Y, He Z, Stoffell PJ (2005b) Molecular mechanisms of heavy metal hyperaccumulation and phytoremediation. *J Trace Elem Med Biol* 18:339–353. <https://doi.org/10.1016/j.jtemb200502007>
- Yoshida S (2003) Molecular regulation of leaf senescence. *Curr Opin Plant Biol* 1:79–84
- Young ND, Tanksley SD (1989) Restriction fragment length polymorphism maps and the concept of graphical genotypes. *Theor Appl Genet* 77(1):95–101
- Yu YW, Xia YS, Li RH (2014) Association analysis of tobacco specific nitrosamines content in tobacco germplasm. *Acta Tab Sin* 20:48–55

- Yuan F, Lyv MJ, Leng BY et al (2015) Comparative transcriptome analysis of developmental stages of the *Limonium bicolor* leaf generates insights into salt gland differentiation. *Plant Cell Environ* 38:1637–1657
- Yuhe S (2012) BAC library construction of villi tobacco (*Nicotiana tomentosiformis*). *Chin Tob Sci. Corpus ID: 87250212*
- Zeba N, Isbat M (2011) Multivariate analysis for yield and yield contributing traits in F0 and F1 generations in tobacco (*Nicotiana tabacum*). *J Exp Biosci* 2(1):101–106
- Zhang DM (1994) Study on surface structure scanning of some tobacco seeds with electrical-glass observing. *Chin Tob Sin* 2:2–15
- Zhang HY, Liu XZ, He CS et al (2005) Random amplified DNA polymorphism of *Nicotiana tabacum* L. cultivars. *Biol Plant* 49:605–607
- Zhang HY, Liu XZ, He CS et al (2008) Genetic diversity among flue-cured tobacco cultivars based on RAPD and AFLP markers. *Braz Arch Biol Technol* 51(6):1097–1101. <https://doi.org/10.1590/S1516-89132008000600003>
- Zhang JS, Wang RG, Yang CY et al (2012) Genetic diversity of agronomic traits and association analysis with SRAP markers in flue-cured tobacco (*Nicotiana tabacum*) varieties from China and abroad. *Acta Agron Sin* 38:1029–1041
- Zhang Y, Zhang F, Li X et al (2013) Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol* 161(1):20–27. <https://doi.org/10.1104/pp.112.205179>
- Zhou ZS, Huang SQ, Guo K et al (2007) Metabolic adaptations to mercury-induced oxidative stress in roots of *Medicago sativa* L. *J Inorg Biochem* 101:1–9
- Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. *Plant Genome* 1:1–20. <https://doi.org/10.3835/plantgenome2008.06.0007ed>
- Zhu JK (2003) Regulation of ion homeostasis under salt stress. *Curr Opin Plant Biol* 6(5):441–445
- Zhu XG, Long SP, Ort DR (2010) Improving photosynthetic efficiency for greater yield. *Ann Rev Plant Biol* 61(1):235–261
- Zimmerman JL, Goldberg RB (1977) DNA sequence organization in the genome of *Nicotiana tabacum*. *Chromosoma* 59:227–252