Understanding Saffron Biology using Omics- and Bioinformatics Tools- A Step Towards Genome Modified *Crocus sp.*

Amjad M. Husaini (amjadhusaini@skuastkashmir.ac.in)
Sher-E-Kashmir University of Agricultural Sciences and Technology Kashmir https://orcid.org/0000-0003-1293-3118
Syed Anam ul Haq
Sher-E-Kashmir University of Agricultural Sciences and Technology Kashmir
Alberto José López Jiménez
Universidad de Castilla-La Mancha

Research Article

Keywords: Crocus sativus L., genomics, transcriptomics, proteomics, metabolomics, bioinformatics, apocarotenoids, Fusarium, stigma

Posted Date: October 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-912096/v1

Abstract

Saffron is a unique plant in many respects and its cellular processes are regulated at multiple levels. The genetic makeup in the form of eight chromosome triplets (2n = 3x = 24) with a haploid generic content (genome size) of 3.45 Gbp is decoded into different types of RNA by transcription. The RNA then translates into peptides and functional proteins, sometimes involving post-translational modifications too. The interactions of genome, transcriptome, proteome and other regulatory molecules ultimately result in the complex set of primary and secondary metabolites of saffron metabolome. These complex interactions manifest in the form of a set of traits 'phenome' peculiar to saffron. The phenome responds to the environmental changes occurring in and around saffron and modify its response in respect of growth, development, disease response, stigma quality, apocarotenoid biosynthesis, etc. Understanding these complex relations between different yet interconnected biological activities is quite challenging in saffron where classical genetics has a very limited role owing to its sterility, and the whole genome sequencing has not been done. Omics-based technologies are immensely helpful in overcoming these limitations and develop a better understanding of saffron biology. In addition to creating a comprehensive picture of the molecular mechanisms involved in apocarotenoid synthesis, stigma biogenesis, corm activity, flower development, etc. omics-technologies will ultimately lead to the engineering of saffron plants with improved phenome. In this review, we discuss the role of omics-technologies and bioinformatics tools in studying saffron biology, and in its improvement.

Introduction

Saffron (*Crocus sativus* L.) is a sterile triploid plant. It is propagated vegetatively through daughter corms developing from mother corms. It belongs to the Iridaceae (Liliales, Monocots) whose genomes are relatively large [1]. The word "saffron" is derived from "zafran" the Arabic word that translates to "yellow". *Crocus sativus* L. is a herbaceous monocot plant propagated vegetatively using corms, and is prevalent throughout the tropical and subtropical regions of the northern hemisphere [2]. The major saffron growing regions of the world include Iran, Azerbaijan, Spain, Italy, India (Kashmir), Greece, and Turkey. The total world saffron production is estimated at 378.33 tons [3], of which about 90% is produced in Iran and the remainder in India (Kashmir), Greece, Afghanistan, Spain and Italy [3].

Saffron genome comprises of eight chromosome triplets (2n = 3x = 24). It has a genome size of 1C = 3.45 Gbp [4]. Saffron has countless medicinal properties like anticancer, antimutagenic, antioxidant, and even anti-covid [5, 6]. Saffron bioactive compounds have immense therapeutic properties useful for coronary artery diseases, neurodegenerative disorders, bronchitis, asthma, diabetes, fever, and colds. It has the potential to help tackle problems associated with severe acute respiratory syndrome coronavirus 2 (COVID-19) patients and post-covid-19 problems [5]. It can help manage stress and anxiety during isolation, quarantine and lockdowns. Its efficacy in managing depression is comparable to drugs like imipramine, fluoxetine, and citalopram. Owing to these properties and the glamour associated with it, saffron is one of the costliest spices in the world.

Saffron is propagated through corms [7], and does not produces fertilisable gametes [8] and is self-incompatible [9, 10]. This makes all modern saffron plants almost identical genetically. This is a bottleneck for the genetic improvement of this highly valued crop. The omics-based biology can be a benchmark for its genetic improvement [11]. The omics-based studies in saffron have broadly focused on the below-mentioned research areas.

1. Flower development and stigma apocarotenoid content

The most valued metabolites in *Crocus sativus* are synthesized in stigma tissue in a developmental stage-specific manner. Almost a decade ago we highlighted the importance of the saffron stigma transcriptome characterization for understanding the molecular basis of its flavour and colour biogenesis, the gynoecium developmental biology, and genomic organization [12, 13]. We expected functional genomics of *Crocus sativus* to play a vital role in finding candidate genes for producing stigma pigments and flavouring compounds. This would enable overexpression studies on saffron for enhancing the production of these pigments and flavouring compounds, and improve the quality of saffron.

Besides whole-genome sequencing, expressed sequence tags (ESTs) are a vital source for analyzing gene expression in specific organs, growth stages, developmental processes, and stress response in crops [13]. The first important database of ESTs for stigma biogenesis and apocarotenoid pathway contains 6768 ESTs [14]. The important contigs include those encoding non-heme-β-carotene-hydroxylase, putative glucosyltransferase, putative isoprenoid GTases, Myb-like protein, Myb305, and Cytochrome P450 [12]. Analysis of saffron stigma EST collections at different developmental stages has revealed that CsCCD2 (carotenoid cleavage dioxygenase) ESTs are predominant in the early stages [15].

Transcriptome analyses in saffron (including leaves, stamens, corm, tepals, and stigmas) have uncovered a large number of transcription factor-coding genes [16-18]. Approximately 105269 transcripts in leaf, corm, tepal, stamen and stigma [18], 64438 transcripts in flowers [17] in *C. sativus*, while 248099 transcripts in tepals of *Crocus ancyrensis* [16] have been reported so far. Transcripts encoding TFs involved in the secondary metabolite biosynthesis are the major ones up-regulated in stigma. Transcripts encoding MYB, MYB related, WRKY, C2C2-YABBY and bHLH transcription factors are differentially expressed [18]. Tissue-specific expression was shown by a total of 1075 transcripts, out of which 342 in stamen, 304 in leaf, 161 in tepal, 144 in stigma and 124 in corm.

Using deep transcriptomics analysis, a novel dioxygenase carotenoid cleavage dioxygenase (CCD2) which catalyzes the first step of crocin biosynthesis from carotenoid zeaxanthin has been identified [15]. Transcriptomic studies have led to the characterisation of glucosyltransferase [19] by dissecting carotenoid and flavonoid biosynthetic pathways of saffron [20]. The production of crocetin from phytoene and crocins from crocetin seems to be transcriptionally regulated [21]. A recent study has identified a new glycosytransferase, UGT91P3, as responsible of the last glycosylation step in the biosynthesis of crocins [22].

Genes encoding enzymes for volatile biosynthesis have been identified using in silico screening of the stigma cDNA database previously described [14, 19]. Comparison of the apocarotenoid content and expression profiles show that 1 deoxyxylulose 5 phosphate synthase (DXS) plays a vital role in apocarotenoid accumulation. DXS is expressed at all the developmental stages of *C. sativus* stigma, while 3 hydroxy 3 methylglutaryl CoA reductase (HMGR) is expressed at low levels only. Additionally, two putative terpene synthases (TS1 and TS2) showed differential expression, with TS2 having an important role in the biosynthesis of apocarotenoids. The expression of two carotenoid biosynthesis genes, CsPSY (phytoene synthase) and CsPDS (phytoene desaturase), also increased in the red stage. In another study, it was observed that with the transition from yellow to red stigmas, accumulation of zeaxanthin was accompanied by enhanced expression of phytoene synthase, phytoene desaturase and lycopene cyclase [20]. Massive accumulation of carotene hydroxylase and zeaxanthin cleavage dioxygenase transcripts also occurred.

A systematic comparative analysis of crocin data and transcriptomes of *C. sativus*, *C. ancyrensis* and *C. cartwrightianus,* has led to the identification of putative transcription factors affecting apocarotenoid accumulation during stigma development in saffron [23]. Expression levels of DXS-CLA1, ZDS, Z-ISO, PDS, CrtISO, BCH-2, LYC-B, CCD2, and UGT74AD2 and apocarotenoid levels had a positive correlation in the three species. In stigma, eleven TFs belonging to the bHLH, C2H2, ARF, HB, CBF/DREB1, NF-YC and ALFIN families show a correlation between expression and apocarotenoid levels in the 3 species. In another similar study, [24] compared the transcriptomes of cultivated *C. sativus* and wild *C. cartwrightianus*. The study found seven genes related to apocarotenoid biosynthesis, which showed differential expression between the samples. The seven genes are orthologues of carotenoid isomerase (CsTc091265), lycopene beta-cyclase (CsTc018497), zeaxanthin epoxidase (CsTc006236), UDP-glucosyltransferase (CsTc020060), phytoene synthase (CsTc009491), nine-cis-epoxy carotenoid dioxygenase (CsTc035409), and carotene beta-hydroxylase (CsTc000418). It is an important information for the saffron improvement program. The orthologue of gene UDP-glucosyltransferase (CsTc020060) is down-regulated in all individual saffron plants while it is up-regulated in all the *C. cartwrightianus* plants [24]. UDP-glucosyltransferase, being involved in the conversion of crocetin to crocin, could be a cause behind the difference in metabolite accumulation between Crocus species. Since triploidy and sterility help safeguard the favourable allele composition (regarding aroma and colour) from being segregated by recombination, modulation of gene expression using genome modification and advanced genetic engineering approaches can be a smart strategy to increase saffron apocarotenoid content in stigma, improve saffron quality and enhance its economic value.

Understanding saffron flower development is vital for improving its productivity and quality. The combination of class A genes (including APETALA1; CsAP1 and APETALA2; CsAP2), class B genes (including APETALA3; CsAP3 and PISTILLATA; CsPI) and class C genes (including AGAMOUS; CsAG), determines the identity of the organs developing in a whorl. An important gene in the stigma development of saffron is a C-class floral homeotic gene AGAMOUS (CsAG) gene [25]. Its expression began in the yellow stage of stigma, showed 16 folds increase as stigma turned from yellow to the orange stage and continued to increase up to the scarlet stage [25]. Similarly, the expression of transcript UGT85U1 increases from yellow stage to red stage and anthesis. However contrastingly, CsNCED, a regulatory gene encoding the enzyme involved in ABA biosynthesis, shows lower expression in all the developmental stages [26]. Relative transcript changes of *Cs*AP3 and NAC-like protein (*Cs*NAP) genes have also been studied during different stages of flower development [27, 28]. However, no direct correlation in the expression of these genes could be detected. CsAP3 expression was maximum during the late pre-anthesis of stigma development, while CsNAP expression increased abruptly at the scarlet stage of stigma. The study concluded that some factor(s) could regulate CsNAP expression, while CsAP3 gene could in turn regulate the factor(s). The promoter of CsAP3 gene consists of three CArG regions, which play a pivotal role in the expression of AP3 gene, of which CArG1 is the binding site for activator proteins, thus regulating floral growth. Given this [28] conducted a study to understand the interaction between nuclear factors with B class gene CsAP3 through its CArG1 promoter region. Nuclear proteins were isolated, and a

CArG1 sequence was synthesized artificially. Using Electrophoretic Mobility Shift Assay (EMSA), the binding interaction of CArG1 region with pure nuclear protein was studied, and the complex was used for protein identification using LCMS. CsNAP was identified as a conspicuous homeotic protein interacting with CArG1 region of AP3 promoter. Understanding the pathway and deciphering the complete mechanism of floral organ differentiation can pave the way for prolonged flowering of saffron by artificially manipulating the key genes. It will provide farmers ample time to collect the flowers and regulate flowering time/duration so that flower damage caused due to early frost in November can be avoided.

In a step forward to better understand the flowering mechanism, two sets of full-length transcriptomes of flowering and non-flowering saffron crocus have been generated using NGS and SMRT sequencing [29]. Recently, morphological, physiological and transcriptome analyses of apical bud samples of *C. sativus* were performed during the floral transition process, and a hypothetical model for the regulatory networks of the saffron flowering transition was proposed [30].

Proteomics is central to the understanding of saffron biology. However, not much work has been reported in saffron proteomics, unfortunately. Not many data sets are available in the PRIDE PRoteomics IDEntifications (PRIDE) Archive database [31], which is a member of the ProteomeXchange (PX) consortium [32]. The first dedicated protein database for saffron stigma (*Crocus sativus* L, taxonomy-id: 82528) samples at different developmental stages have been created only in the recent past [33]. The MS proteomics data can be accessed from the ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD009014 (https://www.ebi.ac.uk/pride/archive/projects/PXD009014. In another recent study, protein profiling of flowering and non-flowering saffron buds subjected to cold stress was done using isobaric tags for relative or absolute quantitation (iTRAQ). Out of 5,624 proteins identified in the study, 201 were differentially abundant protein species (DAPs) between these two groups. Upregulated DAPs play an important role in sucrose metabolism, lipid transport, glutathione metabolism, and gene silencing by RNA. Downregulated DAPs are involved in starch biosynthesis and oxidative stress response. Three new flower-related proteins, CsFLK, CseIF4a, and CsHUA1 were identified too [34].

A search in the GenBank protein database for saffron leads to just over 530 entries, with *C. sativus* (268), *C. cartwrightianus* (258), and *C. ancyrensis* (4) (http://www.ncbi.nlm.nih.gov/). Despite several tools available for predicting and visualising secondary and tertiary structures of proteins, there is no detailed analysis in saffron. A search on saffron crocus query in the UniProt Knowledgebase (UniProtKB) returns only 426 entries, out of which 420 are in Unreviewed (TrEMBL). Only six have been manually reviewed in Swiss-Prot, and include Crocetin glucosyltransferase 2, Crocetin glucosyltransferase 3, Profilin, Zeaxanthin 7,8(7',8')-cleavage dioxygenase, Carotenoid 9,10(9',10')-cleavage dioxygenase, and Pollen allergen Cro s 1. We could find only one 3-D x-ray diffraction-based crystal structure of saffron protein in the protein data bank (PDB) viz. Cysteine Protease (at 1.3 A Resolution) and is available at http://www.rcsb.org/pdb/explore/explore.dostructureId=3U8E.

As already highlighted, unlike rice, maize, wheat, tomato, etc., there are limited saffron-specific genomic resources available to explore its peculiar biology. There is a need to explore and utilize most modern technologies that can generate maximum useful information. Activity-based protein profiling (ABPP) is one such novel technique of chemical proteomics that has recently revolutionized proteomics. Besides its use in drug selectivity and diagnostics, it finds increased application in plant science [35-37]. ABPP uses small molecules as probes for labelling enzymes when these are in an active state. In saffron, the first report on ABPP demonstrated the multiplexing of probes and generated useful information about the active proteases involved at the different developmental stages of stigma [33]. The approach successfully identified and quantified sixty-seven differentially active glycosidases during the stigma development, implying that glycosidase activity is vital for stigma maturation. The results suggest potential candidate glycosidases involved in the conversion of picrocrocin into safranal.

GOLM and the MASSBANK databases are pretty popular for metabolomic profiling. Databases like KEGG, Reactome, MetaCyc and GO-ontology are important for biochemical pathways wherein these metabolites perform specific roles. Studying the metabolomics of the enzymes of flavonoid glucosylation and carotenoid biosynthesis [19, 38] is vital for understanding the dynamics of these pathways in saffron. Metabolic analysis of stigma at the yellow stage has shown low levels of crocetin, crocins, picrocrocin, and some unidentified compounds with maximum wavelengths around 250 nm. Picrocrocin and crocins have been detected early in the orange stage, increasing rapidly in the red stage. The glycosylated products of crocetin reach maximum levels in the red stage [39]. Picrocrocin level rises in the orange stage and achieves the maximum level at anthesis [40]

Besides apocarotenoids, saffron contains volatile compounds also. More than 160-volatile compounds have been detected using chromatography, spectroscopy and mass spectrometry techniques [41, 42]. In the yellow stigma (stage), the fatty acid derivatives predominate, while in the orange (stage), carotenoid derivatives too are present in addition to the fatty acid derivatives. In the red stigma

(stage), the volatiles derived from carotenoids accumulate to high levels, and β-cyclocitral, generated by the cleavage of β-carotene reaches maximum level. Just before anthesis at the scarlet stigma (stage), the volatile propanoic acid, 2-methyl-2,2-dimethyl-1-(2-hydroxy-1-methylethyl) propyl ester accumulate at high levels. However, their levels decrease at anthesis when monoterpenes and carotenoids reach their maximum levels [13]. Among the monoterpenes, linalool is emitted at high levels at anthesis and is responsible for floral odours [43, 44]. In the post-anthesis stage, the fatty acid-derived volatiles become the main volatile compounds.

2. Diversity of saffron and its characterization

Despite the advancement of sequencing technology and its affordability there is no whole-genome sequence available for any Crocus species, which is quite surprising! Some classical cytogenetic analyses involving chromosome counting and karyotyping have been done in saffron [45, 46]. Those studies have shown that saffron is a triploid with karyotype 2n=3x=24. It comprises of 8 triplets: two triplets are subacrocentric, three triplets are metacentric, two triplets are submetacentric and one triplet contains two kinds of chromosomes: chromosome 5(1), metacentric, and chromosomes 5(2,3), subacrocentric and smaller [13]. Some efforts have been made to improve our understanding of the genomic organisation of Crocus species. These studies are mostly based on RAPD [47], [48], Mir, Mansoor [49], IRAP markers [50, 51], Nuclear gene diversity [52-54], AFLP and SSR [55].

The barcode analysis of the 86 species of genus Crocus using rpoC1, matK and tmH-psbA regions has shown the importance of barcoding in the genetic diversity of Crocus [56]. Randomly amplified polymorphic DNA (RAPD) and inter simple sequence repeat (ISSR) marker profiles of 43 isolates of *C. sativus* collected from different geographical areas has been used to determine if this species is monomorphic or polymorphic. The results showed that the clones were identical at molecular level [57]. Surprisingly, ISSR markers showed no differences between *C. sativu*s and *C. cartwrightianus* [58]. In contrast, RAPD markers revealed considerable amount of genetic diversity among 10 elite saffron clones selected in Kashmir [48]. Long terminal repeats (LTRs), a retrotransposon (RTN)- based marker study in Iranian species of Crocus showed high diversity within and between species [50]. Using 12 microsatellite markers [59] succeeded in detecting good polymorphism within fifty Iranian individuals of *Crocus sativus*. A reasonable amount of polymorphism was detected in similar studies among Iranian *C. sativus* germplasms [60, 61].

There is ample evidence that epigenetics plays an important role in creating inheritable variation and contributes significantly to the traits in different plant species [62]. DNA methylation is the most widely studied epigenetic mark in plants as its genome-wide investigation is easier to accomplish [63]. In a study involving more than a hundred saffron accessions from WSCC (World Saffron and Crocus Collection, Spain), very low genetic variability was detected using 12 AFLP primer combinations. In contrast, very high epigenetic variability was detected with just 3 MS-AFLP primer combinations [64]. Five accessions from the WSCC germplasm having extremely low genetic variability were cultivated for three years in the same field. These accessions of different origins maintained different epigenotypes. It suggests that the epigenetic structure in saffron is highly stable [65]. The stability of saffron epigenotype over the years supports the idea that epigenetics may play a vital role in the constancy of saffron phenotype variability.

AFLP analysis using methylation-sensitive restriction enzyme-sequencing (MRE-seq) gives more insight into saffron's epigenome [66]. The study compared the epigenetic profile of 5 phenotypically different, but genetically similar accessions from the world saffron and crocus collection (WSCC) germplasm. Differential methylation of regions was detected in some genes encoding transcription factors, shaping the alternative phenotypes. Many SNPs and INDELs were identified, showing thereby that genetic polymorphism exists within the saffron species. Genetic variants were also detected in Gene Ontology (GO) terms, portraying a genetic basis for alternative phenotypes. A heatmap of the 50 highest polymorphic GOs shared between accessions highlighted the presence of two distinct clusters of Indian and Spanish accessions. Twelve GOs showed lower polymorphism in the Spanish accessions than Indian accessions [66].

Phylogenetic analyses of nuclear loci and chloroplast genome, genome-wide DNA polymorphism indicate that Crocus sativus is genetically similar to *C. cartwrightianus* populations. Genome sequencing and Fluorescence in situ hybridisation (FISH) have demonstrated that genomes of two *Crocus cartwrightianus* individuals with slight chromosomal differences had gotten fused, and it could be the parental origin of saffron *Crocus sativus* L. [67, 68]. Another view is that the most likely ancestors of saffron are *C. cartwrightianus* and *C. pallasii* subsp. Pallasii (or close relatives) [69].

3. Saffron growth, development and disease

While there are ample omics-based studies on apocarotenoid biosynthesis pathway, the studies on the growth and development of saffron are limited. Proteomic analysis has led to the identification of differentially accumulated proteins (in somatic embryos) of *C.*

sativus. Thirty-six proteins have been identified, including those involved in protein synthesis, carbohydrate and energy metabolism, defence and stress response, nitrogen metabolism and secondary metabolism [70]. Metabolomic studies have provided insights into the corm composition of *C. sativus,* too [71]. At the sprouting stage (in corms), sugars like glucose, fructose, and maltose reveal a strong positive correlation with palmitate, turanose, oxalic acid, ethanolamine, linoleic acid, and tetronic acid; and a negative correlation with sitosterol mannoside and octadecanoic acid. At bud development, fatty-acid biosynthesis significantly relied on carbohydrate metabolism intermediates. Sucrose breakdown reached its maximum to begin the sprouting and bud growth in *C. sativus*.

Climate change and the associated biotic and abiotic stresses are the most daunting challenges to saffron cultivation [11]. Omics based biological studies of saffron crop shall pave the way for its sustainable production, especially given the climate change associated problems. MicroRNome of plants though ubiquitous and small in size, plays an important role in abiotic stress. MicroRNA sequencing, though ignored in *C. sativus*, can be vital in understanding the regulation of saffron genomic elements. These can also throw light on the regulatory networks underlying the apocarotenoid biosynthesis in *C. sativus*. A study on an EST library from mature *C. sativus* stigmas has helped detect two putative microRNAs, miR414 and miR837-5p, in saffron stigma [72]. Co-expression network analysis has revealed them to play vital roles in metabolic pathways. The predicted targets of the miR414 are: β-carbonic anhydrase 5, Transducin/WD40 repeat-like superfamily protein and three-transposable element genes AT2G13700.1, AT4G06613.1, AT3G29783.1. The predicted targets of miR837-5p are SEC14 cytosolic factor family protein/phosphoglyceride transfer family protein, Enhancer of polycomb-like protein, and F-box/RNI-like/FBDlike domains containing protein. In addition three more miRNAs viz., csa-miR1, csa-miR2 and csamiR3 have also been predicted by using in silico methods of EST analysis [73]. The predicted targets of these miRNAs are involved in regulating plant growth, senescence, stress responses, disease resistance, mRNA export, protein synthesis and post-translational modifications [73].

In an RNA-seq based transcriptome study, useful information was categorised in the form of small databases for -viruses, bacteria, fungi, and plants [74]. It used YeATS suite from the NCBI and Ensembl databases, and showed that the soybean mosaic virus is abundantly expressed in the corm, tepal, leaf, stigma, and stamen tissues [74]. Furthermore, it has been shown that there is a difference in fungal diversity between roots and corms of *C. sativus*. At the flowering stage, the dominant phylum in the rhizosphere is Zygomycota, while in the cormosphere Basidiomycota is dominant. In the cormosphere, Basidiomycota is prevalent at the flowering stage, while Zygomycota is dominant at the dormant stage. However, in the bulk soil, Ascomycota dominates during both stages [75].

Saffron corm rot caused by *Fusarium oxysporum* is a major disease, causing heavy losses in saffron-producing countries [11, 76]. ABPP, a chemical proteomics-based technique, has been upscaled by multiplexing diverse probes (targeting serine hydrolases, α-glycosidases, β-glycosidases and cysteine proteases) to give a broad snapshot of active proteases having a role in corm rot infection [33]. It has detected the suppressed activity of an α-glycosidase upon *F. oxysporum* infection, which is consistent with the view that *F. oxysporum* suppresses AGLU1 in the apoplast to overcome its antifungal activity [33], [77, 78]. While the activities of putative α-glycosidases (100-kD) and β-glucosidases (50-70 kD) increased upon infection, the activities of serine hydrolases (50, 60 kD) decreased. Additionally, many β-glucosidases (45-60 kD) appeared, while some (65-70 kD) disappeared. In the ABPP based chemical proteomics study, drastic changes were visualised in the activity profile of cysteine proteases, especially papain-like Cys proteases and vacuolar processing enzymes (Table 1).

4. Saffron adulteration & Spice quality

The molecular analysis involving a complete set of metabolites existing in a cell at a particular instant is the backbone of understanding metabolic pathways and is called metabolomics. It is highly significant for plants due to the crucial role of the secondary metabolites in plant survival. These metabolites are extracted from the tissues, separated and analysed in a high-throughput manner to generate metabolic fingerprints. Many tools available in the bioinformatics toolbox help identify and characterise these metabolites [79]. In saffron, metabolite fingerprinting (based on 1H NMR spectra) and chemometrics have helped in authenticating saffron as Italian or Iranian [80, 81]. These also help detect the presence of plant-based adulterants in saffron [82]. 1H NMR and chemometrics studies have shown that saffron can preserve its valuable characteristics up to four years [83].

High-performance thin-layer chromatography (HPTLC) helps study chemical diversity among saffron accessions. In a study in recent past, fifty-three saffron accessions from Khorasan Razavi were characterized for chemical diversity using HPTLC. Based on the heat maps generated at different wavelengths, crocin and picrocrocin content was found helpful in categorising saffron [84]. The third important bioactive molecule, safranal, is not among the major volatiles produced in the fresh tissue. It (safranal) is the primary aroma component comprising 60-70% essential oil content [85, 86]. Safranal gets produced by picrocrocin degradation during the dehydration of the stigma [87].

5. Medicinal value & drug development

Saffron bioactive compounds have immense therapeutic properties, including those beneficial against coronary artery diseases, neurodegenerative disorders, bronchitis, asthma, diabetes, fever, colds, and metabolic syndrome. A detailed analysis of its medicinal properties points to its immense untapped potential for easing the distress symptoms of severe acute respiratory syndrome coronavirus 2 (COVID-19) patients and managing the post-covid-19 syndrome [5]. Despite the importance of saffron in medicine and phytochemistry, modern approaches based on omics studies are relatively rare [4, 51, 55].

The metabolic and biochemical properties of saffron confirm its immense role in the pharmacognosy and pharma industry [5]. Studies on the binding potential of carotenoid pathway bioactive molecules for angiotensin-converting enzyme 2 (ACE2) receptor of SARS-CoV-2 show the possibility of using the saffron based remedy for novel coronavirus [88]. Flexible molecular docking followed by atomic level interaction study indicated that lutein and picrocrocin form various interactions with different amino acid residues of ACE2. In-depth analysis revealed that these interactions with the majority of the residues of ACE2 could be crucial for receptor-binding domain (RBD) binding and, therefore, can disrupt the interaction between RBD and ACE2. The study provides a clue for advanced studies involving *in vitro*, animal models and clinical studies. The efficacy of saffron in managing depression is comparable to drugs like imipramine, fluoxetine, and citalopram. The saffron metabolites can help manage stress and anxiety during the prolonged lockdown, isolation, and quarantine. Owing to all these beneficial properties and as an immunity booster, saffron extracts may be added in some drug formulations in future.

6. Bioinformatics for omics data analysis

Intricate regulatory networks of gene expression control the tissue and stage-specific accumulation of various metabolites. The systems biology approach integrates different omics technologies, including transcriptomics, proteomics, metabolomics, etc., so that biological systems are investigated in an integrated manner at different levels. The analysis of the complex datasets that get generated need to be integrated in the framework of known biological pathways, and corrections in any discrepancies that may have crept in because of the other simpler approaches are also made. Bioinformatics plays a crucial role during the data generation, analysis and interpretation of the different omics technologies for the mining of meaningful information. It is crucial for the interpretation of a massive amount of data generated through high throughput technologies, filtering out useful information for interpretation by the researchers for comprehensive views on systems functionality [89-91]. Moreover, it provides resources derived by exploiting -omics technologies [91, 92] or subsequent analyses, including sequence comparisons, gene family investigations, molecular modelling, etc. [91, 93-95].

Omics-based technologies and other molecular research tools have led to the generation of a huge amount of information, which has necessitated the advancement of bioinformatics. This acts like a ‘feedback promotion’ and causes advancement in omics technologies due to its better handling of the ‘big data’. Bioinformatics creates and advances algorithms, computational techniques, and databases to better solve problems in the analysis of huge biological data. It has a key role in the textual mining of biological literature and query biological data. Bioinformatics tools can easily compare genetic and genomic data to better understand the evolutionary relationships between organisms. At a more integrative level, it analyses the biological pathways and metabolic networks to give a better understanding into systems biology. It helps in conducting simulation and modelling studies on DNA, RNA and proteins to understand their molecular interactions better, thus strengthening structural biology. It has assisted evolutionary biologists to i) trace the evolution of organisms by calculating changes in their DNA; ii) build complex models of populations for predicting the outcome, and iii) share information about a large number of species [96].

Large-scale expression profiling studies in saffron have generated huge amounts of data and, the discipline of bioinformatics has been indispensable for ‘deriving information’ from these data. As predicted [12], characterisation of the saffron stigmas through omics-studies coupled with bioinformatics tools has generated vital novel information about the molecular basis of flavour, colour biogenesis, genomic organisation and the biology of the gynoecium of saffron (Table).

GenoType and GenoDive are two important programs to analyse the genotypic diversity in clonal/asexual organisms [97]. The significance of genetic differentiation between accessions of saffron in Iran through the calculation of clonal diversity indices and AMOVA has been done using these tools [64]. PIECE is a comprehensive plant gene comparison and evolution database containing all the annotated genes described from 25 plant species with available sequenced genomes. In saffron comparative analysis of gene structures was done with the comparative genomics database PIECE for Plant Intron and Exon Comparison [98]. MIcroSAtellite (MISA) microsatellite

finder is a tool for finding microsatellites in nucleotide sequences. Using MIcroSAtellite (MISA) Perl script in saffron [99] counted simple sequence repeats (SSRs), also known as microsatellites.

As *Crocus sativus* is a species without whole-genome sequencing, *de novo* transcriptome analysis provides an excellent and necessary platform to deepen the research on this plant at the molecular level [100]. Full length reconstruction of transcriptomes from short-reads generated by Illumina sequencing technologies is the most challenging step in RNA-seq studies. In the absence of a reference genome, most common assembly strategies rely on Bruijn graph, including packages such as Trinity, SOAPdenovo-Trans, Velvet, Rnnotator and Oases [101, 102] [103]. Many studies have used Trinity for de novo assembly of saffron transcriptomics data [30, 100, 104], whereas others rely on strategies combining some of the aforementioned packages [18]. Alternative methods to Illumina sequencing, such as PacBio long-read sequencing, imply specialized software such SMRT Analysis software suite [105], [29], [99].

One of the main downstream applications after de novo assembly is transcript expression estimation, which generally implies, in the absence of a reference genome, mapping reads against the assembled transcriptome. Algorithms that quantify expression from transcriptome mappings include RSEM, eXpress, Sailfish and kallisto, among others [106], [107], [108], [109]. These algorithms typically depend on short read alignment programs such as Bowtie, which enables ultrafast and memory-efficient alignment of large sets of sequencing reads to a reference sequence [110]. In a recent study, the differentially expressed genes in saffron were identified via pair-wise comparisons of gene expression patterns between stigma and the other four tissues (corm, leaf, tepal and stamen) by the 'DESeq' package [99]. The package is used for quantitative analysis of comparative RNA-seq data using shrinkage estimators for dispersion and fold change [111].

Functional annotation of the transcripts generated by the aforementioned methods is typically achieved using similarity-detection tools such as BLAST [112]. Blast2GO has become a popular tool, allowing massive annotation of complete transcriptome datasets against a variety of databases, as well as GO functional classification and KEGG pathway enrichment [113]. Other software, WEGO (Web Gene Ontology Annotation Plot), allows visualizing, comparing and plotting GO annotation results. These tools have been widely used in the functional classification of unigenes in RNA-seq studies on *Crocus sativus* [16], [23], [30], [38], [99]. Other annotation tools are based on the identification of specific domains in protein sequences. PlantTFcat is a high-performance web-based analysis tool that is designed to identify and categorise plant Transcription factor (TF)/Transcriptional regulator (TR)/Chromatin regulator (CR) genes from genome-scale protein and nucleic acid sequences by systematically analysing InterProScan domain patterns in protein sequences. Candidate transcription factors implicated in crocin biosynthesis in *Crocus sieberi* tepal and *C. sativus* stigma have been identified using PlantTFcat [16]. The Plant Transcription Factor Database (Pln TFDB), which is an integrative database that provides putatively complete sets of transcription factors and transcriptional regulators in plant species, has been used to identify genes encoding transcription factors in the network in saffron (*Crocus sativus*) [72].

RNA-sequencing is a valuable tool to gain knowledge on high-level functions in biological systems. KEGG is an integrated database resource for biological interpretation of genome sequences and other high-throughput data [114]. It is the reference knowledge base that integrates current knowledge on molecular interaction networks such as pathways and complexes (PATHWAY database), information about genes and proteins generated by genome projects (GENES/SSDB/KO databases) and information about biochemical compounds and reactions (COMPOUND/GLYCAN/REACTION databases) [115]. [30] performed KEGG pathway analysis of differentially expressed genes (DEGs) and mapped 8251 unigenes into 130 standard pathways using KEGG database in saffron (*Crocus sativus* L.). Moreover, 14,671 genes were also annotated using KEGG database in Saffron [30].

TransDecoder identifies candidate coding regions within transcript sequences, such as those generated by de novo RNA-Seq transcript assembly using Trinity, or constructed based on RNA-Seq alignments to the genome using Tophat and Cufflinks [116, 117]. It has been used in *Crocus sativus* L protein domain annotation [33], [99]. Open reading frame detection and domain annotation from de novo assembled transcripts of *Crocus sativus* L using TransDecorder along with two other algorithms (GeneMarkS-T, Prodigal) has led to the identification of 67 active glycosidases that are differentially active during stigma development, implying that glycosidase activity has a major role in the maturation of stigma [33]. Prodigal (PROkaryotic DYnamic programming Gene-finding Algorithm) is a fast, lightweight, open-source gene prediction program [118], while GeneMarkS-T is used for *ab initio* identification of protein-coding regions in RNA transcripts [119].

Different proteins that were either upregulated or down-regulated in saffron under cadmium toxicity have been putatively identified using the MASCOT software search engine [120]. MaxQuant is a proteomics software for analysing large mass-spectrometric data sets [121]. It has been used for peptide and protein identification in different developmental stages of saffron stigma [33]. Peptide relative quantification between different MS runs was based solely on the LFQs, as calculated by MaxQuant (MaxLFQ algorithm). Another

associated software platform (Perseus) supports researchers in the interpretation of protein quantification, interaction and post-translational modification, and is used for statistical analysis of MaxQuant output [121]. Saffron stigma spectra files submitted to an Andromeda search in MaxQuant were finally analysed and filtering of the results was done for post-translational modification, pattern recognition, time-series analysis in Perseus version 1.5.5.3. [33]. As discussed above, the identification and quantification of active glycosidases using ABPP could not have been possible without the support of bioinformatics tools [33]. Open Reading Frame Detection and Domain Annotation Softwares like Gene-MarkS-T [119], TransDecoder [116] and Prodigal [118] were used. Eight glycosidases (three GH3, three GH35, two GH116, and one GH1) were up-regulated more than 2-fold in stage 4 stigmas. Moreover, the differential 110-kD βGH (Glucoside hydrolase) detected with labeling is most likely the GH116 enzyme CsTc017194, because this enzyme has a predicted molecular mass of 106 kD and is 4.5-fold up-regulated in stage 4 stigmas. The study illustrates the power of ABPP with bioinformatic predictive algorithms for quantitative glycosidase activity profiling on non-model plant species, like saffron (Table 1).

There is immense scope for bioinformatics studies for elucidating biochemical functions of saffron proteins and bioactive compounds [5, 33, 88].

Table 1

Significant research findings and outcomes of omics-based research studies conducted in saffron and its allies.

S. No	Omics approach used	The gist of the main findings and outcomes	Reference
1.	Genomics	• Whole genome sequencing of Crocus sp has not been done. • There are contradictory results on the detection of polymorphisms using marker-based analysis. • Some studies conclude that saffron is a monomorphic species and whole genome sequencing is needed to discriminate between its isolates. • Some studies show that molecular markers are quite efficient in detecting polymorphism. Such studies conclude that saffron is not monomorphic and that there is diversity which could be useful for breeding purposes. • AFLP analysis using methylation-sensitive restriction enzyme-sequencing (MRE-seq) has shown that phenotypically different but genetically similar accessions vary in the methylation pattern of genomic regions encoding transcription factors and may result in alternative phenotypes. • Epigenetic structure in saffron is highly stable and may play a vital role in the constancy of saffron phenotype variability. • ISSR primers are reported to be capable of easily distinguishing genuine saffron from fake one.	[57]; [61]; [49]; [59]; [122]; [123]; [69]; [124]; [65, 66]
2.	Transcriptomics	• De novo transcriptome assemblies have been created from leaves, stamens, corm, tepals, and stigmas of *Crocus sativus*. • The most valued compounds of C. sativus are synthesised inside stigma in a developmental stage-specific manner. • During the transition from yellow stage to red stage stigmas there is an accumulation of zeaxanthin accompanied by sharp increase in the expression of phytoene synthase, phytoene desaturase, lycopene β cyclase, β carotene hydroxylase and zeaxanthin cleavage dioxygenase. • CsCCD2 (carotenoid cleavage dioxygenase) ESTs are prominent in the saffron stigma libraries obtained from early stages of stigma development. • UDP-glucosyltransferase is vital for conversion of crocetin to crocin, and therefore causes difference in metabolite accumulation between Crocus species. • 1 deoxyxylulose 5 phosphate synthase (DXS) plays a vital role in apocarotenoid accumulation in stigma. • There is no direct concordance in the expression of *Cs*AP3 and *Cs*NAP gene expression in saffron. • Identification, isolation, and biochemical characterisation of uridine diphosphate glycosyltransferase (UGT709G1), which catalyses the HTCC glucosyltransferase reaction to yield picrocrocin, can provide a vital lead for the industrial production of picrocrocin/safranal. • Differentially expressed full-length transcripts of flowering and non-flowering saffron crocus have been identified and characterised. • Stigma development in field- and indoor-cultivated saffron is similar with respect to apocarotenoid content and gene expression profiles of 12 genes involved in apocarotenoid biosynthesis. • Carotenoid cleavage dioxygenase (CCD2) catalyzes the first step of crocin biosynthesis from carotenoid zeaxanthin and gets expressed at an extremely high level in the stigma as compared to corm, leaf, tepal, and stamen. • A C-class floral homeotic gene AGAMOUS (CsAG) gene is vital for stigma development of saffron. Its expression begins at yellow stage of stigma and increases sharply to orange stage, and continues to increase upto scarlet stage. • CsAP3 expression is maximum at late preanthesis of stigma development, while CsNAP expression increases abruptly at the scarlet stage of stigma. • CsNAP protein binds to the CArG1 region of CsAP3 promoter, and might be regulating CsAP3 expression indirectly by modulating CArG1 promoter.	[20]; [18]; [12]; [13]; [15]; [24] ; [19]; [125]; [27]; [17]; [18]; [126]; [29]; [23]; [30]; [127]; [99]; [128]; [25]; [28]
3.	Metabolomics	• Two novel saponins namely Azafrine 1 and Azafrine 2 have been isolated, purified, and structurally elucidated from the external part of saffron corm, suggesting that they may be acting as phytoprotectans.	[129]; [83];

		• ^{1}H NMR-based metabolomics is useful to determine quality deterioration of saffron upon storage and for quality control. • Liquid chromatography coupled to electrospray ionisation time-offlight mass spectrometry is an important tool for assessing saffron authenticity. • Tepals may have nutrition value owing to the presence of phytosterols and fatty acids, and can be processed as a source of flavonoids. • Metabolite profiling of stigma, tepal and stamen of *Crocus sativus* flower by ultra-performance liquid chromatography-quadrupole time-of-flight mass spectrometry (UPLC-QTof-MS/MS) has shown that coniferin and crocin-2 are special components in stigmas, while flavonoids are high in tepals. • High resolution mass spectrometry metabolomic studies in saffron from several countries has revealed that the phytochemical content varies among the samples of different countries. • At the yellow stage of stigma there are very low levels of crocetin, crocins, picrocrocin. • Picrocrocin and crocins are detected early in the orange stigma stage and increase rapidly in the red stigma stage. • The glycosylated products of crocetin reach maximum levels in the red stigma stage. • Saffron bioactive compounds are useful against coronary artery diseases, neurodegenerative disorders, bronchitis, asthma, diabetes, fever, colds, and metabolic syndrome. • Saffron can alleviate the symptoms of severe acute respiratory syndrome coronavirus 2 (COVID-19) patients and manage post-covid-19 syndrome. • The efficacy of saffron in managing depression is comparable to drugs like imipramine, fluoxetine, and citalopram. • Saffron can be used as an adjuvant in drug formulations as it acts as an immunity booster and anti-depressant.	[130]; [131]; [132]; [71]; [84]; [133]; [134]; [39]; [5]
4.	Proteomics	• Thirty-six differentially accumulated proteins have been detected during somatic embryogenesis in *Crocus sativus* and involvement of ascorbate-glutathione cycle has been suspected in somatic embryo establishment. • Saffron protein database of stigma at different developmental stages is available through ProteomeXchange Consortium via the PRIDE partner repository with the data set identifier PXD009014 https://www.ebi.ac.uk/pride/archive/projects/PXD009014. • Two hundred and one differentially abundant protein species (DAPs) under cold stress affecting the floral initiation of saffron have been revealed using iTRAQ-based proteomics followed by real-time qPCR. • Saffron dormant corms exposed to low temperature stress do not bloom perhaps due to changes in the 'reactive oxygen species–antioxidant system–starch/sugar interconversion homeostasis flowering pathway'.	[70]; [33]; [34]
5	ABPP	• Drastic changes in the activity profile of cysteine proteases especially papain-like Cys proteases and vacuolar processing enzymes occur in the corms infected with *Fusarium oxysporum*. • The activity of α-glycosidase AGLU1 gets suppressed upon *Fusarium oxysporum* infection in saffron corms irrespective of the F.o strain. • Activities of putative α-glycosidases (100-kD) and β-glucosidases (50-70 kD) increase upon *F. oxysporum* infection, while the activities of serine hydrolases (50, 60 kD) decrease. • Many β-glucosidases (45-60 kD) appear, while some (65-70 kD) disappear during *F. oxysporum* infection. • Glycosidase activity has a major role in maturation and development of stigma. • Sixty-seven active glycosidases that are differentially active during stigma development have been identified and quantified.	[33]
6.	miRNomics	• Five miRNAs csa-miR1, csa-miR2, csamiR3, miR414 and miR837-5p have been reported in *Crocus sativus* using in silico methods of EST analysis. These miRNAs may play roles in plant growth, disease resistance, senescence, stress responses, etc.	[73]; [72]

Table 2

Bioinformatic tools and databases useful for omics data analysis

S. No	Bioinformatic Tools	Web address	Role	Reference
1.	SAM and BCF tools	https://www.htslib.org/ https://github.com/samtools/samtools https://www.htslib.org/ https://github.com/samtools/bcftools	Tools for processing and analysing sequencing data	[135]
2.	MEGA	http://www.megasoftware.net/	Comparative analysis and inferring evolutionary relationships of homologous sequences.	[136-139]
3.	Trinity	https://github.com/trinityrnaseq/trinityrnaseq/releases/tag/v2.8.6	Tool for de novo transcriptome assembly of RNA-seq data	[116, 140]
4.	SMART 9	https://smart.embl.de/	Database for Identification and analysis of protein domains within protein sequences	[141]
5.	MPI bioinformatics toolkit	http://toolkit.tuebingen.mpg.de/	Web service for comprehensive and collaborative protein bioinformatic analysis	[142, 143]
6.	BiGGEsTS	http://kdbio.inesc-id.pt/software/biggests	Tool for revealing local coexpression of genes in specific intervals of time	[144]
7.	PlantGDB	http://www.plantgdb.org/	Database for comparative genomics/ genomic database encompassing sequence data for plants	[145]
8.	KEGG	http://www.kegg.jp/ http://www.genome.jp/kegg/	Database resource for biological interpretation of genome sequences and other high-throughput data	[114]

9.	TrichOME	http://www.planttrichome.org/	Comparative Omics database for plant trichomes	[146]
10.	PlantTFcat	https://www.zhaolab.org/PlantTFcat/	Tool for Identification and categorisation of plant transcription factors and transcriptional regulators	[147]
11.	Pln TFDB	http://plntfdb.bio.uni-potsdam.de/v3.0/	Database for functional and evolutionary study of plant transcription factors	[148, 149]
12.	Ensembl Plants	http://plants.ensembl.org	Database for visualising, mining and analysing plant genomic data	[150]
13.	Wego	http://wego.genomics.org.cn/	Web tool for plotting GO annotations	[151]
14.	edgeR	http://bioconductor.org/packages/edgeR/	Package for differential expression analysis of digital gene expression data	[152, 153]
15.	Bowtie	http://bowtie.cbcb.umd.edu/ https://sourceforge.net/projects/bowtie-bio/	Ultrafast, memory-efficient alignment program for aligning short DNA sequence reads to large genomes.	[154]
16.	KaPPA-View	http://kpv.kazusa.or.jp/kpv4/	Web-based database for analysing omics data in plants	[155, 156]
17.	Transcriptogramer	http://bioconductor.org/packages/transcriptogramer	R package for transcriptional analysis based on protein–protein interaction	[157]
18.	Cufflinks	http://cole-trapnell-lab.github.io/cufflinks	Open-source software for RNA-Seq data analysis	[158, 159]
19.	Paintomics	http://www.paintomics.org/	Web based tool for joint visualization of transcriptomics and metabolomics data	[160]
20.	PIECE	https://probes.pw.usda.gov/piece/index.php	Database for	[161]

			plant gene structure comparison and evolution	
21.	MISA-Web	http://misaweb.ipk-gatersleben.de/	Tool/web server for microsatellite prediction and counting	[162]
22.	Prodigal	https://github.com/hyattpd/Prodigal	Protein-coding gene prediction software tool	[118]
23.	GeneMarkS-T	http://topaz.gatech.edu/GeneMark/license_download.cgi	Tool for identification of protein-coding regions in RNA transcripts.	[119]
24.	MaxQuant	https://maxquant.net/maxquant/	Quantitative proteomics software package for analysing large mass-spectrometric data sets.	[121]
25.	Perseus	https://maxquant.net/perseus/	Software platform for interpreting protein quantification, interaction and post-translational modification data.	[163]
26.	GenAlex	https://biology-assets.anu.edu.au/GenAlEx/Welcome.html	Platform for population genetic analysis.	[164]
27.	DnaSP	http://www.ub.edu/dnasp	Software package for DNA sequence polymorphism analysis of large data sets.	[165]
28.	TransDecoder	https://github.com/TransDecoder/TransDecoder	Tool for Identification of potential coding regions within reconstructed transcripts.	[116, 117]
29.	RepeatMasker package	https://www.repeatmasker.org/	Program to screen DNA sequences for interspersed repeats and low complexity DNA sequences	[166]
30.	GenoType and GenoDive	http://www.patrickmeirmans.com/software	Programs for the analysis of genetic diversity of asexual organisms.	[97]

31.	psRNATarget	https://www.zhaolab.org/psRNATarget/	A small RNA target analysis server	[167]
32.	DESeq 2 package	http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html	Package for differential analysis of gene expression in plants.	[111]

7. Conclusion

Omics based technologies have revolutionized biology, and saffron is no exception. Such studies have helped better understand the molecular mechanisms of flower development in saffron and could lead to the creation of such saffron flowers that would have carpels in place of stamens, therefore doubling the yield. This could be an ambitious target but is certainly achievable. Except for saffron whole-genome sequencing, which is still awaited, a lot of useful information about saffron biology has been generated using omics-based techniques. These novel technologies helped discover new genes, study their expression, function, evolutionary relationships, etc. and made a plethora of information available to the scientific community. It has taken us closer to achieving the goal of developing engineered saffron. It will not be too far when these techniques enable editing genes encoding apocarotenoid biosynthesis through novel genome editing tools like CRISPR-Cas, making saffron breeding programs successful.

Omics tools can be useful in locating sources of resistance and agronomically interesting traits for transfer to saffron by appropriate biotechnological tools. Such tools can also help appreciate the extent of the diversity of various geographic or genetic groups of cultivated saffron to infer relationships between groups and accessions. The information derived can be utilised for constructing biological pathways involved in the biosynthesis of principal components of saffron. Saffron metabolomics studies have revealed many peculiar properties of this interesting spice. However, the major challenge remains in identifying the incongruities in the biochemical pathways and the metabolic networks and correlating them with the phenotype.

Declarations

Acknowledgements:

AMH is grateful to National Mission on Himalayan Studies, Ministry of Environment, Forest and Climate Change, Government of India for funding saffron research in the form of a research grant to him. He is also thankful to the Scientists and staff of SKUAST-K for their help and support.

Funding:

Not applicable

Availability of data and material:

Not applicable

Code availability

Not applicable

Compliance with Ethical Standards:

Not applicable

Conflict of interest:

The authors have no conflict of interest and nothing to disclose

Ethical approval:

This article does not contain any studies with human participants or animals performed by any of the authors

Informed consent:

Not applicable

References

1. Fernández J-A, *Biology, biotechnology and biomedicine of saffron.* Recent research developments in plant science. Vol. 2 (2004) p. 127–159
2. Kafi M et al (2018) An expensive spice saffron (Crocus sativus L.): a case study from Kashmir, Iran, and Turkey. In: Global perspectives on underutilized crops. Springer, pp 109–149
3. Shahnoushi N et al (2020) *Economic analysis of saffron production*, in *Saffron*. Elsevier, pp 337–356
4. Pandita D (2021) *Saffron (Crocus sativus L.): Phytochemistry, therapeutic significance and omics-based biology*, in *Medicinal and Aromatic Plants*. Elsevier, pp 325–396
5. Husaini AM, Jan KN, Wani GA, *Saffron: A potential drug-supplement for severe acute respiratory syndrome coronavirus (COVID) management.* Heliyon (2021) p. e07068
6. Premkumar K, Ramesh A (2010) Anticancer, antimutagenic and antioxidant potential of saffron: An overview of current awareness and future perspectives. Functional plant science technology 4:91–97
7. Husaini AM et al (2010) Saffron (Crocus sativus Kashmirianus) cultivation in Kashmir: practices and problems. Functional Plant Science Biotechnology 4(2):108–115
8. Ghaffari S (1986) Cytogenetic studies of cultivated Crocus sativus (Iridaceae). Plant Syst Evol 153(3):199–204
9. Caiola MG, Somma DD, Lauretti P, *Comparative study of pollen and pistil in Crocus sativus L.(Iridaceae) and allied species*. Annali Di Botanica, 2000. 58
10. Caiola MG, *Embryo origin and development in Crocus sativus L.(Iridaceae).* Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology, 2005. 139(3): p. 335–343
11. Husaini AM (2014) Challenges of climate change: Omics-based biology of saffron plants and organic agricultural biotechnology for sustainable saffron production. GM crops food 5(2):97–105
12. Husaini AM et al., *Bioinformatics for saffron (Crocus sativus L.) improvement*. Communications in Biometry & Crop Science, 2009. 4(1)
13. Husaini AM, Ashraf N (2010) Understanding Saffron biology using bioinformatics tools. Saffron Functional Plant Science Biotechnology 4:31–37
14. D'Agostino N et al (2007) An EST database from saffron stigmas. BMC plant biology 7(1):1–8
15. Frusciante S et al., *Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis.* Proceedings of the National Academy of Sciences, 2014. **111**(33): p. 12246–12251
16. Ahrazem O et al (2018) Transcriptome analysis in tissue sectors with contrasting crocins accumulation provides novel insights into apocarotenoid biosynthesis and regulation during chromoplast biogenesis. Scientific reports 8(1):1–17
17. Baba SA et al (2015) Comprehensive transcriptome analysis of Crocus sativus for discovery and expression of genes involved in apocarotenoid biosynthesis. BMC Genomics 16(1):1–14
18. Jain M et al (2016) De novo transcriptome assembly and comprehensive expression profiling in Crocus sativus to gain insights into apocarotenoid biosynthesis. Scientific reports 6(1):1–13
19. Moraga ÁR et al (2009) Metabolite and target transcript analyses during Crocus sativus stigma development. Phytochemistry 70(8):1009–1016
20. Castillo R, Fernández J-A, Gómez-Gómez L (2005) Implications of carotenoid biosynthetic genes in apocarotenoid formation during the stigma development of Crocus sativus and its closer relatives. Plant physiology 139(2):674–689

21. Ahrazem O et al (2015) Saffron: its phytochemistry, developmental processes, and biotechnological prospects. J Agric Food Chem 63(40):8751–8764
22. López AJ et al (2021) A New Glycosyltransferase Enzyme from Family 91, UGT91P3, Is Responsible for the Final Glucosylation Step of Crocins in Saffron (Crocus sativus L.). Int J Mol Sci 22(16):8815
23. Ahrazem O et al (2019) Multi-species transcriptome analyses for the regulation of crocins biosynthesis in Crocus. BMC Genomics 20(1):1–15
24. Nemati Z, *The origin of saffron*. 2018
25. Wafai AH et al (2017) RELATIVE EXPRESSION ANALYSIS OF CSAG GENE DURING DIFFERENT STAGES OF STIGMA DEVELOPMENT IN CROCUS SATIVUS. L (SAFFRON). Journal of Emerging Technologies Innovative Research 4:272–274
26. Gómez-Gómez L et al (2017) Unraveling massive crocins transport and accumulation through proteome and microscopy tools during the development of saffron stigma. Int J Mol Sci 18(1):76
27. Wafai AH et al (2015) Comparative expression analysis of senescence gene CsNAP and B-class floral development gene CsAP3 during different stages of flower development in saffron (Crocus sativus L.). Physiology Molecular Biology of Plants 21(3):459–463
28. Wafai AH, Husaini AM, Qadri RA (2019) Temporal expression of floral proteins interacting with CArG1 region of CsAP3 gene in Crocus sativus L. Gene Reports 16:100446
29. Qian X et al (2019) Single-molecule real-time transcript sequencing identified flowering regulatory genes in Crocus sativus. BMC Genomics 20(1):1–18
30. Hu J et al (2020) Transcriptome profiling of the flowering transition in saffron (Crocus sativus L.). Scientific reports 10(1):1–14
31. Perez-Riverol Y et al (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic acids research 47(D1):D442–D450
32. Griss J et al (2016) Recognizing millions of consistently unidentified spectra across hundreds of shotgun proteomics datasets. Nature methods 13(8):651–656
33. Husaini AM et al (2018) Multiplex fluorescent, activity-based protein profiling identifies active α-glycosidases and other hydrolases in plants. Plant physiology 177(1):24–37
34. Chen J et al (2021) Screening of Key Proteins Affecting Floral Initiation of Saffron Under Cold Stress Using iTRAQ-Based Proteomics. Frontiers in plant science 12:708
35. Serim S, Haedke U, Verhelst SH (2012) Activity-based probes for the study of proteases: recent advances and developments. ChemMedChem 7(7):1146–1159
36. Willems LI, Overkleeft HS, van Kasteren SI (2014) Current developments in activity-based protein profiling. Bioconjugate chemistry 25(7):1181–1191
37. Morimoto K, van der Hoorn RA (2016) The increasing impact of activity-based protein profiling in plant science. Plant Cell Physiol 57(3):446–461
38. Ahrazem O et al (2010) The expression of a chromoplast-specific lycopene beta cyclase gene is involved in the high production of saffron's apocarotenoid precursors. J Exp Bot 61(1):105–119
39. Rubio A, Fernández J-A, Gómez L. *Biosynthesis of carotenoids in saffron*. in *I International Symposium on Saffron Biology and Biotechnology 650*. 2003
40. Carmona M et al (2007) A new approach to saffron aroma. Critical reviews in food science nutrition 47(2):145–159
41. Assimiadis MK, Tarantilis PA, Polissiou MG (1998) UV-Vis, FT-Raman, and 1H NMR spectroscopies of cis-trans carotenoids from saffron (Crocus sativus L.). Appl Spectrosc 52(4):519–522
42. Van Calsteren M-R et al (1997) Spectroscopic characterization of crocetin derivatives from Crocus sativus and Gardenia jasminoides. J Agric Food Chem 45(4):1055–1061
43. Dobson HE, *Floral volatiles in insect biology.* Insect-plant interactions (2017) p. 47–82
44. Knudsen JT, Tollsten L, Bergström LG (1993) Floral scents—a checklist of volatile compounds isolated by head-space techniques. Phytochemistry 33(2):253–280
45. Brighton CA (1977) Cytology of Crocus sativus and its allies (Iridaceae). Plant Syst Evol 128(3):137–157
46. Agayev Y, Zarifi E. *Peculiar evolution of saffron (Crocus sativus L.): prosperity and decline*. in *III International Symposium on Saffron: Forthcoming Challenges in Cultivation, Research and Economics 850*. 2009

47. Caiola MG, Caputo P, Zanier R (2004) RAPD analysis in Crocus sativus L. accessions and related Crocus species. Biol Plant 48(3):375–380
48. Imran S et al. *Studies in relation to molecular variability in saffron*. in *III International Symposium on Saffron: Forthcoming Challenges in Cultivation, Research and Economics 850*. 2009
49. Mir MA et al., *Deciphering genetic diversity analysis of saffron (Crocus sativus L.) using RAPD and ISSR markers*. Saudi Journal of Biological Sciences, 2020
50. Alavi-Kia S et al (2008) Analysis of genetic diversity and phylogenetic relationships in Crocus genus of Iran using inter-retrotransposon amplified polymorphism. Biotechnology Biotechnological Equipment 22(3):795–800
51. Alsayied NF et al (2015) Diversity and relationships of Crocus sativus and its relatives analysed by inter-retroelement amplified polymorphism (IRAP). Ann Botany 116(3):359–368
52. Caiola MG, Canini A (2010) Looking for saffron's (Crocus sativus L.) parents. Functional Plant Science Biotechnology 4(2):1–14
53. Tsaftaris A et al (2011) The study of the E-class SEPALLATA3-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (Crocus sativus L.) and its putative progenitors. Journal of plant physiology 168(14):1675–1684
54. Harpke D et al (2013) Phylogeny of Crocus (Iridaceae) based on one chloroplast and two nuclear loci: ancient hybridization and chromosome number evolution. Mol Phylogenet Evol 66(3):617–627
55. Larsen B et al (2015) Large intraspecific genetic variation within the Saffron-Crocus group (Crocus L., series Crocus; Iridaceae). Plant Syst Evol 301(1):425–437
56. Seberg O, Petersen G (2009) How many loci does it take to DNA barcode a crocus? PloS one 4(2):e4598
57. Rubio-Moraga A et al (2009) Saffron is a monomorphic species as revealed by RAPD, ISSR and microsatellite analyses. BMC Res Notes 2(1):1–5
58. Moraga AR et al (2010) Intersimple sequence repeat markers for molecular characterization of Crocus cartwrightianus cv. albus. Ind Crops Prod 32(2):147–151
59. Nemati Z et al (2012) Isolation and characterization of a first set of polymorphic microsatellite markers in saffron, Crocus sativus (Iridaceae). Am J Bot 99(9):e340–e343
60. Beiki AH, Keifi F, Mozafari J (2010) Genetic differentiation of Crocus species by random amplified polymorphic DNA. Genet Eng Biotechnol J 18:1–10
61. Keify F, Beiki AH (2012) Exploitation of random amplified polymorphic DNA (RAPD) and sequence-related amplified polymorphism (SRAP) markers for genetic diversity of saffron collection. Journal of Medicinal Plants Research 6(14):2761–2768
62. Springer NM, Schmitz RJ (2017) Exploiting induced and natural epigenetic variation for crop improvement. Nat Rev Genet 18(9):563–575
63. Seymour DK, Becker C (2017) The causes and consequences of DNA methylome variation in plants. Curr Opin Plant Biol 36:56–63
64. Busconi M et al (2015) AFLP and MS-AFLP analysis of the variation within saffron crocus (Crocus sativus L.) germplasm. PloS one 10(4):e0123434
65. Busconi M et al (2018) Epigenetic stability in Saffron (Crocus sativus L.) accessions during four consecutive years of cultivation and vegetative propagation under open field conditions. Plant Sci 277:1–10
66. Busconi M et al (2021) Epigenetic Variability Among Saffron Crocus (Crocus sativus L.) Accessions Characterized by Different Phenotypes. Frontiers in plant science 12:349
67. Grilli Caiola M *Saffron reproductive biology*. in *I International Symposium on Saffron Biology and Biotechnology 650*. 2003
68. Schmidt T et al (2019) Adding color to a century-old enigma: multi-color chromosome identification unravels the autotriploid nature of saffron (Crocus sativus) as a hybrid of wild Crocus cartwrightianus cytotypes. New Phytol 222(4):1965–1980
69. Torricelli R et al (2019) Morphological and molecular characterization of Italian, Iranian and Spanish saffron (Crocus Sativus L.) accessions. Appl Ecol Environ Res 17:1875–1887
70. Sharifi G et al (2012) Identification of differentially accumulated proteins associated with embryogenic and non-embryogenic calli in saffron (Crocus sativus L.). Proteome science 10(1):1–15
71. Bagri J et al (2017) Metabolic shift in sugars and amino acids regulates sprouting in Saffron corm. Scientific reports 7(1):1–10
72. Zinati Z, Shamloo-Dashtpagerdi R, Behpouri A (2016) *In* silico Identification of miRNAs and their target genes and analysis of gene co-expression network in saffron (Crocus sativus L.) stigma. Molecular Biology Research Communications 5(4):233

73. Guleria P, Goswami D, Yadav KS (2012) Computational Identification of miRNAs and their targets from Crocus sativus L. Archives of Biological Sciences 64(1):65–70
74. Chakraborty S, *Transcriptome from saffron (Crocus sativus) plants in Jammu and Kashmir reveals abundant soybean mosaic virus transcripts and several putative pathogen bacterial and fungal genera.* bioRxiv (2016) p. 079186
75. Ambardar S et al (2016) Comparative metagenomics reveal phylum level temporal and spatial changes in mycobiome of belowground parts of Crocus sativus. PloS one 11(9):e0163300
76. Cappelli C (1994) Occurrence of Fusarium oxysporum f. sp. gladioli on saffron in Italy. Phytopathologia Mediterranea 33(1):93–94
77. Xiao J-z et al (1994) Extracellular glycoprotein (s) associated with cellular differentiation in Magnaporthe grisea. Molecular plant-microbe interactions 7(5):639–644
78. Monroe JD et al (1999) Structure, properties, and tissue localization of apoplastic α-glucosidase in crucifers. Plant physiology 119(2):385–398
79. Edwards D, Batley J (2004) Plant bioinformatics: from genome to phenome. Trends in biotechnology 22(5):232–237
80. Cagliani LR et al (2015) NMR investigations for a quality assessment of Italian PDO saffron (Crocus sativus L.). Food control 50:342–348
81. Yilmaz A et al (2010) 1 H NMR metabolic fingerprinting of saffron extracts. Metabolomics 6(4):511–517
82. Petrakis EA et al (2015) Evaluation of saffron (Crocus sativus L.) adulteration with plant adulterants by 1H NMR metabolite fingerprinting. Food Chem 173:890–896
83. Ordoudi SA et al (2015) 1H NMR-based metabolomics of saffron reveals markers for its quality deterioration. Food research international 70:1–6
84. Vahedi M et al (2018) Quantitative HPLC-based metabolomics of some Iranian saffron (Crocus sativus L.) accessions. Ind Crops Prod 118:26–29
85. Alonso GL et al (1996) Determination of safranal from saffron (Crocus sativus L.) by thermal desorption – gas chromatography. J Agric Food Chem 44(1):185–188
86. Tarantilis PA, Polissiou MG (1997) Isolation and identification of the aroma components from saffron (Crocus sativus). J Agric Food Chem 45(2):459–462
87. Raina BL et al (1996) Changes in pigments and volatiles of Saffron (Crocus sativusL) during processing and storage. J Sci Food Agric 71(1):27–32
88. Ganai SA, Husaini AM (2021) Investigating binding potential of carotenoid pathway bioactive molecules for ACE2 receptor of SARS-CoV-2: Possibility of a saffron based remedy for novel coronavirus! Journal of Horticulture and Postharvest Research
89. Kumar A et al (2015) Systems biology for smart crops and agricultural innovation: filling the gaps between genotype and phenotype for complex traits linked with robust agricultural productivity and sustainability. Omics: a journal of integrative biology 19(10):581–601
90. Chiusano ML et al (2008) ISOL@: an Italian SOLAnaceae genomics resource. BMC Bioinform 9(2):1–11
91. Ambrosino L et al (2020) Bioinformatics Resources for Plant Abiotic Stress Responses: State of the Art and Opportunities in the Fast Evolving-Omics Era. Plants 9(5):591
92. Choi H-K (2019) Translational genomics and multi-omics integrated approaches as a useful strategy for crop breeding. Genes genomics 41(2):133–146
93. Licciardello C et al (2014) Characterization of the glutathione S-transferase gene family through ESTs and expression analyses within common and pigmented cultivars of Citrus sinensis (L.) Osbeck. BMC plant biology 14(1):1–15
94. López de Maturana, E et al (2019) Challenges in the integration of omics and non-omics data. Genes 10(3):238
95. Monticolo F, Colantuono C, Chiusano ML (2017) Shaping the evolutionary tree of green plants: evidence from the GST family. Scientific reports 7(1):1–9
96. Ogbe RJ, Ochalefu DO, Olaniru OB (2016) Bioinformatics advances in genomics-A review. Int J Curr Res Rev 8(10):05–11
97. Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Molecular ecology notes 4(4):792–794
98. Ahrazem O et al (2020) Differential interaction of Or proteins with the PSY enzymes in saffron. Scientific reports 10(1):1–11
99. Yue J et al (2020) Full-length transcriptome sequencing provides insights into the evolution of apocarotenoid biosynthesis in Crocus sativus. Computational structural biotechnology journal 18:774–783

100. Tan H et al (2019) Transcriptome analysis reveals novel enzymes for apo-carotenoid biosynthesis in saffron and allows construction of a pathway for crocetin synthesis in yeast. J Exp Bot 70(18):4819–4834
101. Zerbino D, Zerbino D, Birney E, *E. Birney.* (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Research 18:821
102. Martin J et al (2010) Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. BMC Genomics 11(1):1–8
103. Schulz MH et al (2012) Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. Bioinformatics 28(8):1086–1092
104. Mahmodi P et al (2014) Analysis of saffron stigma (Crocus sativus L.) transcriptome using SOAPdenovo and Trinity assembly software. Crop Biotechnology 4(6):35–46
105. Rhoads A, Au KF (2015) PacBio sequencing and its applications. Genom Proteom Bioinform 13(5):278–289
106. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinform 12(1):1–16
107. Roberts A, Pachter L (2013) Streaming fragment assignment for real-time analysis of sequencing experiments. Nature methods 10(1):71–73
108. Anders S, Pyl PT, Huber W (2015) HTSeq—a Python framework to work with high-throughput sequencing data. bioinformatics 31(2):166–169
109. Bray N et al., *Near-optimal RNA-Seq quantification.* arXiv preprint arXiv:1505.02710, 2015
110. Langmead B, *Aligning short sequencing reads with Bowtie.* Current Protocols in Bioinformatics (2010) **32**(1): p. 11.7. 1-11.7. 14
111. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):1–21
112. Altschul SF et al (1990) Basic local alignment search tool. Journal of molecular biology 215(3):403–410
113. Conesa A et al (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21(18):3674–3676
114. Kanehisa M et al (2016) KEGG as a reference resource for gene and protein annotation. Nucleic acids research 44(D1):D457–D462
115. Kanehisa M et al (2004) The KEGG resource for deciphering the genome. Nucleic acids research 32(suppl_1):D277–D280
116. Haas BJ et al (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature protocols 8(8):1494–1512
117. Haas BJ, Papanicolaou A, *TransDecoder 5.5.0.* 2019
118. Hyatt D et al (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform 11(1):1–11
119. Tang S, Lomsadze A, Borodovsky M (2015) Identification of protein coding regions in RNA transcripts. Nucleic acids research 43(12):e78–e78
120. Rao J, Lv W, Yang J, *Proteomic analysis of saffron (Crocus sativus L.) grown under conditions of cadmium toxicity*. Bioscience Journal, 2017. 33(3)
121. Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nature protocols 11(12):2301–2319
122. Zheng H-j et al (2013) Construction of DNA finger printing for dry saffron. African Journal of Pharmacy Pharmacology 7(43):2807–2812
123. Gedik A et al (2017) Genetic diversity of Crocus sativus and its close relative species analyzed by iPBS-retrotransposons. Turkish Journal of Field Crops 22(2):243–252
124. Zarini HN et al (2019) A comparative assessment of DNA fingerprinting assays of ISSR and RAPD markers for molecular diversity of Saffron and other Crocus spp. in Iran. The Nucleus 62(1):39–50
125. IqbaLMzr J et al (2013) Relative expression of apocarotenoid biosynthetic genes in developing stigmas of Crocus sativus L. Journal of Crop Science Biotechnology 16(3):183–188
126. Diretto G et al (2019) UGT709G1: a novel uridine diphosphate glycosyltransferase involved in the biosynthesis of picrocrocin, the precursor of safranal in saffron (Crocus sativus). New Phytol 224(2):725–740
127. Zhou G et al., Flower cultivation regimes affect apocarotenoid accumulation and gene expression during the development of saffron stigma. Horticulture, Environment Biotechnology, 2020. 61(3): 473–484

128. Gao G et al (2021) Transcriptomic analysis of saffron at different flowering stages using RNA sequencing uncovers cytochrome P450 genes involved in crocin biosynthesis. Mol Biol Rep 48(4):3451–3461
129. Rubio-Moraga Á et al (2011) Triterpenoid saponins from corms of Crocus sativus: localization, extraction and characterization. Ind Crops Prod 34(3):1401–1409
130. Guijarro-Díez M et al (2015) Metabolomic fingerprinting of saffron by LC/MS: novel authenticity markers. Analytical bioanalytical chemistry 407(23):7197–7213
131. Consonni R et al (2016) On the traceability of commercial saffron samples using 1H-NMR and FT-IR metabolomics. Molecules 21(3):286
132. Feizy J, Reyhani N (2016) Gas chromatographic determination of phytosterols and fatty acids profile in saffron petals. Can Chem Trans 4(3):389–397
133. Xu S et al., Discrimination of Different Parts of Saffron by Metabolomic-Based Ultra-Performance Liquid Chromatography Coupled with High-Definition Mass Spectrometry. Chemistry & Biodiversity, 2019. 16(10): p e1900363
134. Gikas E, Koulakiotis NS, Tsarbopoulos A (2021) Phytochemical Differentiation of Saffron (Crocus sativus L.) by High Resolution Mass Spectrometry Metabolomic Studies. Molecules 26(8):2180
135. Danecek P et al (2021) Twelve years of SAMtools and BCFtools. GigaScience 10(2):giab008
136. Kumar S et al (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular biology evolution 35(6):1547–1549
137. Kumar S, Stecher G, Tamura K, *MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets.* Molecular biology and evolution (2016) **33**(7): p. 1870–1874
138. Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. Molecular biology evolution 38(7):3022–3027
139. Tamura K et al (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology evolution 28(10):2731–2739
140. Grabherr MG et al (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29(7):644
141. Letunic I, Khedkar S, Bork P (2021) SMART: recent updates, new developments and status in 2020. Nucleic acids research 49(D1):D458–D460
142. Zimmermann L et al (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. Journal of molecular biology 430(15):2237–2243
143. Gabler F et al (2020) Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. Current Protocols in Bioinformatics 72(1):e108
144. Gonçalves JP, Madeira SC, Oliveira AL (2009) BiGGEsTS: integrated environment for biclustering analysis of time series gene expression data. BMC Res Notes 2(1):1–11
145. Duvick J et al (2007) PlantGDB: a resource for comparative plant genomics. Nucleic acids research 36(suppl_1):D959–D965
146. Dai X et al (2010) TrichOME: a comparative omics database for plant trichomes. Plant physiology 152(1):44–54
147. Dai X et al (2013) PlantTFcat: an online plant transcription factor and transcriptional regulator categorization and analysis tool. BMC Bioinform 14(1):1–6
148. Pérez-Rodríguez P et al (2010) PlnTFDB: updated content and new features of the plant transcription factor database. Nucleic acids research 38(suppl_1):D822–D827
149. Riaño-Pachón DM et al (2007) PlnTFDB: an integrative plant transcription factor database. BMC Bioinform 8(1):1–10
150. Bolser DM et al (2017) *Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomic data*, in *Plant Genomics Databases*. Springer, pp 1–31
151. Ye J et al (2006) WEGO: a web tool for plotting GO annotations. Nucleic acids research 34(suppl_2):W293–W297
152. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1):139–140
153. McCarthy DJ, Chen Y, Smyth GK (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic acids research 40(10):4288–4297
154. Langmead B et al (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10(3):1–10

155. Tokimatsu T et al (2006) KappA-View: a tool for integrating transcriptomic and metabolomic data on plant metabolic pathway maps. In: Plant Metabolomics. Springer, pp 155–163
156. Tokimatsu T et al (2005) KaPPA-View. A web-based analysis tool for integration of transcript and metabolite data on plant metabolic pathway maps. Plant physiology 138(3):1289–1300
157. Morais DA, Almeida RM, Dalmolin RJ (2019) Transcriptogramer: an R/Bioconductor package for transcriptional analysis based on protein–protein interaction. Bioinformatics 35(16):2875–2876
158. Trapnell C et al (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature biotechnology 28(5):511–515
159. Trapnell C et al (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols 7(3):562
160. García-Alcalde F et al (2011) Paintomics: a web based tool for the joint visualization of transcriptomics and metabolomics data. Bioinformatics 27(1):137–139
161. Wang Y et al (2013) PIECE: a database for plant gene structure comparison and evolution. Nucleic acids research 41(D1):D1159–D1166
162. Beier S et al (2017) MISA-web: a web server for microsatellite prediction. Bioinformatics 33(16):2583–2585
163. Tyanova S et al (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods 13(9):731–740
164. Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular ecology notes 6(1):288–295
165. Rozas J et al (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. Molecular biology evolution 34(12):3299–3302
166. Smit A, Hubley R, Green P (2013) *RepeatMasker. 2013.* Institute for Systems Biology. repeatmasker. org], Seattle. http://www
167. Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. Nucleic acids research 39(suppl_2):W155–W159

Figures

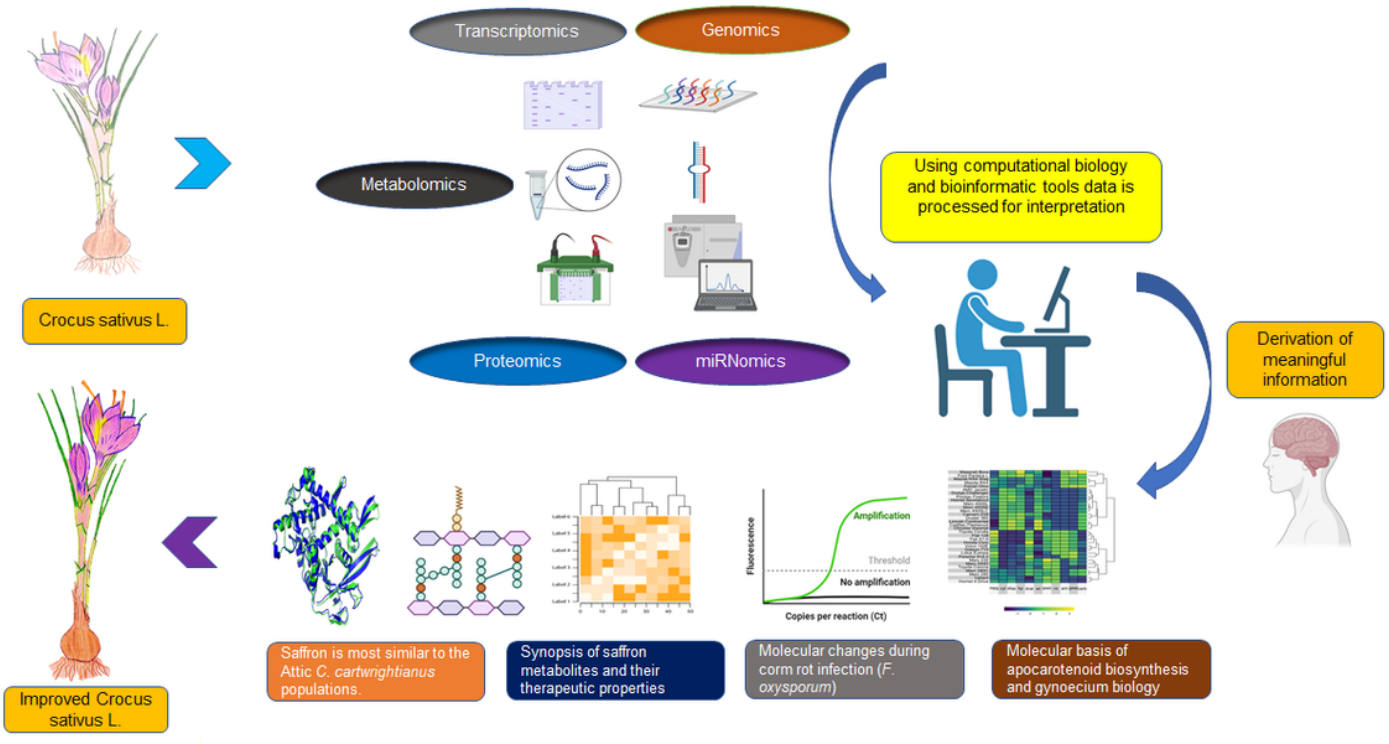

Figure 1

Major characteristics studied using omics-based approaches for understanding saffron biology