# Advanced Genetic Mechanisms in ***Cannabis sativa***: Epistasis, Pleiotropy, and QTL Mapping for Cultivar Improvement

## 1. Introduction to Advanced Genetic Inheritance in Cannabis

### The Complex Genetic Landscape of ***Cannabis sativa***

*Cannabis sativa* L. is a species of profound genetic diversity, historically cultivated for a multitude of purposes, including the production of durable fibers, nutritious seed oil, and an extensive array of secondary metabolites, most notably cannabinoids and terpenes. The inheritance patterns of crucial agronomic traits—such as yield and plant architecture—and chemotypic characteristics—like cannabinoid content (e.g., \Delta^9-tetrahydrocannabinol and cannabidiol) and terpene profiles—are typically complex. These traits are often quantitative, meaning they exhibit continuous variation and are governed by the collective action of multiple genetic loci, their interactions with one another (gene-gene interactions or epistasis), and their interactions with the environment. The study of such quantitative traits necessitates sophisticated genetic approaches to dissect their underlying genetic architecture.

### Significance of Understanding Advanced Genetic Mechanisms

A comprehensive understanding of advanced genetic phenomena, such as epistasis and pleiotropy (wherein a single gene influences multiple distinct traits), is fundamental to unraveling the genetic basis of these complex characteristics in *Cannabis sativa*. Furthermore, the identification of Quantitative Trait Loci (QTLs)—specific genomic regions associated with variation in a quantitative trait—and the subsequent development of reliable molecular markers linked to these QTLs are of paramount importance for modern plant breeding. Such knowledge empowers breeders to make more informed selections, design more efficient breeding strategies, and ultimately accelerate genetic gain in cannabis cultivar development programs. This report aims to synthesize current research findings on these advanced genetic concepts as they pertain to *Cannabis sativa*, providing an expert-level overview of the field.

The path to understanding cannabis genetics has been shaped by its unique history. Unlike major agricultural crops that have benefited from decades of publicly funded research into quantitative genetics, cannabis research has faced significant historical legal restrictions. This has resulted in a comparative knowledge gap, particularly concerning its complex genetic architecture. Consequently, fundamental genetic studies in cannabis, including the exploration of epistasis, pleiotropy, and comprehensive QTL mapping, are not merely incremental advancements but are foundational pillars upon which robust and predictive breeding programs must be built. Discoveries in these areas may, therefore, have a more substantial initial impact on breeding strategies than similar findings in more extensively studied crops, as cannabis science is, in many respects, navigating a period of foundational discovery.

Adding to the complexity is the inherent biological nature of *Cannabis sativa*. The species is predominantly dioecious (having separate male and female plants) and exhibits an outcrossing mating system, which naturally contributes to high levels of heterozygosity within individuals and populations. This high heterozygosity, with estimates ranging from 12.5% to 40.5% , presents considerable challenges for genetic analysis. For instance, the study of recessive alleles, the detection of subtle epistatic interactions, and the development of stable, homozygous parental lines (essential for consistent F1 hybrid production or for creating certain types of mapping populations like Recombinant Inbred Lines) are more intricate and laborious processes than in predominantly self-pollinating species. Controlled crosses are obligatory, and achieving homozygosity requires multiple generations of managed pollination, which can also lead to inbreeding depression. From a genetic analysis perspective, high heterozygosity can mask the effects of alleles or confound their interpretation due to dominance and interactions occurring at heterozygous loci. Unraveling these complexities necessitates specific population structures (e.g., F2, RILs) and often larger population sizes for QTL mapping to achieve adequate statistical power. This, in turn, makes marker development more demanding, as markers must be informative across diverse and heterozygous genetic backgrounds.

## 2. Epistasis: Unmasking Gene Interactions in Cannabis

### Fundamental Principles of Epistasis

Epistasis describes a genetic interaction where the phenotypic expression of one gene is dependent on, or modified by, the alleles present at one or more other distinct gene loci. This phenomenon represents a deviation from simple Mendelian inheritance, where genes are assumed to contribute independently to the phenotype. Epistatic interactions can lead to modified phenotypic ratios in segregating populations, such as the F2 generation of a dihybrid cross, compared to the classical 9:3:3:1 ratio. Several types of epistatic interactions are recognized in genetics, each producing characteristic ratios:

* **Dominant Epistasis (12:3:1 ratio)**: Occurs when a dominant allele at one locus masks the phenotypic expression of alleles at a second locus. An example is fruit color in squash, where the dominant allele *W* (white) masks the effect of alleles at the *G/g* locus (yellow/green).
* **Recessive Epistasis (9:3:4 ratio)**: Characterized by a recessive genotype at one locus masking the expression of alleles at another locus. A classic example is coat color in Labrador retrievers.
* **Duplicate Recessive Epistasis (Complementary Gene Action; 9:7 ratio)**: This occurs when recessive alleles at either of two loci can mask the expression of dominant alleles at the other locus. Consequently, at least one dominant allele at *both* loci is required to produce a specific phenotype. Flower color in sweet peas, where functional enzymes from two different genes are needed for pigment production, exemplifies this interaction.
* **Duplicate Dominant Epistasis (Duplicate Gene Action; 15:1 ratio)**: In this scenario, a dominant allele at *either* of two loci is sufficient to produce the dominant phenotype. Kernel color in wheat, where a functional enzyme from either of two genes results in pigment, is a well-known example.
* **Dominant and Recessive Epistasis (Dominant Suppression; 13:3 ratio)**: This type of interaction occurs when a dominant allele at one locus and a recessive genotype at another locus produce the same phenotype, effectively suppressing the alternative phenotype unless a specific combination of alleles (dominant at the second locus and recessive at the suppressor locus) is present. Malvidin production in *Primula* illustrates this, where a dominant allele *D* suppresses malvidin production governed by allele *K*.

Epistatic interactions can be further categorized based on their effect on allelic contributions. **Magnitude epistasis** refers to interactions that alter the magnitude (size) of an allele's effect on a trait without changing its direction (e.g., beneficial vs. detrimental). In contrast, **sign epistasis** occurs when the interaction changes the direction of an allele's effect; an allele that is beneficial in one genetic background may become detrimental in another. Sign epistasis is particularly significant as it can lead to complex fitness landscapes and influence evolutionary trajectories.

The molecular underpinnings of epistasis can involve genes participating in the same biochemical pathway (e.g., enzyme-substrate relationships, sequential enzymatic steps, or feedback regulation) or genes in different pathways that converge to influence a common phenotype (e.g., a regulatory gene influencing the expression of a structural gene, or two signaling pathways crosstalking).

### Documented Epistatic Interactions in Cannabis

Research into the genetic architecture of complex traits in *Cannabis sativa* is increasingly revealing the importance of epistatic interactions, particularly for chemotypic traits.

**Cannabinoid and Terpene Biosynthesis:** Seminal work by McKay and colleagues, utilizing an F2 mapping population derived from a cross between the hemp cultivars 'Carmagnola' (fiber-type) and 'USO31' (dual-purpose), provided significant evidence for epistasis in the inheritance of cannabinoids and terpenes. Their analyses identified statistically significant epistatic interactions between QTLs located on linkage group 6 (specifically, a region designated LG6.35) and linkage group 9 (LG9.40) for 8 out of 17 measured biochemical traits. These interactions accounted for a considerable portion of the observed phenotypic variance, ranging from 3.89% to 15.11%. A notable example was the production of cannabidiolic acid (CBDa), where the specific combination of alleles at both the LG6.35 and LG9.40 loci significantly influenced the final CBDa concentration. Similar two-locus interactions involving these linkage groups were documented for cannabigerolic acid (CBGa), tetrahydrocannabinolic acid (THCa), and several terpenes, including \alpha-pinene, \alpha-terpinene, camphene, \gamma-terpinene, and 3-carene. The researchers posited that these interacting loci might harbor genes responsible for synthesizing precursor molecules common to both cannabinoid and terpene pathways, or they might contain genes involved in the interacting biosynthetic pathways themselves.

**THC Concentration:** Independent research by Campbell et al. (2020) also pointed towards non-additive genetic components in cannabinoid inheritance. Their study, using an information-theoretic approach to estimate composite genetic effects, suggested that autosomal additive-by-additive (AaAa) epistatic effects were likely contributors to the observed variation in THC concentration. This type of epistatic effect consistently featured in the top-ranking genetic models explaining THC variation in their datasets.

**General Considerations:** These findings underscore that epistasis is a prevalent genetic factor influencing both cannabinoid and terpene profiles in cannabis. The complex genetic architecture underlying cannabinoid inheritance appears to extend beyond simple additive gene action, with potential contributions from cytogenetic factors (genes in mitochondrial or plastid genomes) and maternal effects, which themselves can involve epistatic-like interactions with nuclear genes.

The observation that epistatic interactions in cannabis, such as those between LG6.35 and LG9.40, often involve QTLs that are themselves clusters of genes or exhibit pleiotropic effects (affecting multiple biochemical compounds) is particularly noteworthy. For example, the QTL on LG9.40 is associated with the *Olivetol Synthase* (OLS) gene, a critical upstream enzyme in the cannabinoid biosynthetic pathway. If a key enzyme like OLS, or a regulator of its activity located on LG9, interacts with another locus on LG6 that might control the supply of a shared precursor (like geranyl pyrophosphate) or regulate a parallel branch of terpene or cannabinoid synthesis, this interaction occurs at a significant pathway level, not merely between isolated structural genes. This suggests that epistasis in cannabis might frequently manifest as interactions between crucial regulatory hubs or key junctions in metabolic pathways. Consequently, attempts to manipulate a single key gene, such as OLS, could lead to cascading and interactive effects that are unpredictable without a thorough understanding of its epistatic partners. Breeding strategies aimed at achieving specific, complex chemotypes may therefore require selecting for particular *combinations* of alleles at these interacting hub loci, rather than focusing on individual genes in isolation.

### Implications of Epistasis for Cannabis Breeding and Genetic Prediction

The presence of significant epistatic interactions has profound implications for cannabis breeding and the prediction of genetic merit:

* **Complexity in Breeding:** Epistasis complicates breeding efforts because the phenotypic effect of an allele at one locus is contingent upon the allelic state of other interacting loci. This means that the breeding value of an allele can change depending on the genetic background into which it is introduced.
* **Efficacy of Marker-Assisted Selection (MAS):** Traditional MAS strategies that focus on accumulating favorable alleles at individual QTLs, assuming additive effects, may prove less effective if substantial epistatic interactions are not accounted for. The simple pyramiding of alleles from multiple loci might not result in the expected additive improvements if epistatic effects are significant and, particularly, if they are antagonistic (where the combined effect is less than the sum of individual effects).
* **Chemotype Optimization:** Breeding programs aiming to develop cultivars with highly specific cannabinoid and terpene ratios (i.e., complex chemotypes) must inherently consider these gene-gene interactions.
* **Genomic Selection (GS) Models:** GS models, which utilize genome-wide marker information to predict breeding values, may offer an advantage. Some GS models can implicitly capture non-additive genetic variance, including epistasis, potentially leading to more accurate predictions for complex traits in cannabis compared to MAS strategies based on a few major QTLs.
* **Heterosis and Combining Ability:** Understanding epistasis is also vital for predicting heterosis (hybrid vigor) and for assessing combining ability in hybrid cannabis breeding programs. Specific types of epistasis, notably additive × additive and dominance × dominance interactions, have been shown to directly contribute to the components of heterosis and combining ability. Ignoring these interactions can lead to erroneous inferences when identifying superior parental populations or predicting hybrid performance.

It is also crucial to recognize that detecting epistasis is statistically challenging, often requiring large population sizes and specialized experimental designs or analytical methods. Standard quantitative genetic analyses may have low statistical power to detect all but the strongest epistatic interactions, potentially leading to an underestimation of epistasis's true contribution to trait variation. Cannabis QTL studies, while advancing, have often utilized F2 populations of moderate sizes (e.g., 275 F2 individuals in the McKay study , 96 F2 individuals in Stack's work ). While these sample sizes are adequate for detecting major QTLs, they may not be sufficient to uncover the full spectrum of epistatic interactions, especially those with smaller effect sizes or involving alleles at low frequency in the mapping population. The epistatic interactions reported by McKay and colleagues, for instance, explained a considerable portion of phenotypic variance (up to ~15%), suggesting these are relatively strong interactions. It is plausible that numerous weaker, yet biologically relevant, interactions remain undetected. Therefore, the currently documented instances of epistasis in cannabis likely represent only the "tip of the iceberg." As cannabis genomics research matures, and larger, more diverse mapping populations are analyzed using increasingly sophisticated statistical tools (such as those designed to specifically model interaction effects), a more extensive and intricate network of epistatic interactions is expected to be unveiled. This will further refine our understanding of complex trait architecture in cannabis and underscore the limitations of breeding predictions based solely on additive gene action.

### Table 1: Examples of Epistatic Gene Interactions in Plants and Cannabis

| Type of Epistasis | Organism | Genes/Loci Involved (Example) | Affected Trait(s) | Observed Phenotypic Ratio / % Variance Explained (Interaction) | Reference(s) |
| --- | --- | --- | --- | --- | --- |
| Duplicate Dominant | Wheat | Two unlinked genes (e.g., *A* and *B*) | Kernel Color | 15 (colored) : 1 (white) |  |
| Complementary Gene Action | Sweet Pea | Two unlinked genes (e.g., *C* and *P*) | Flower Color | 9 (colored) : 7 (white) |  |
| Dominant Epistasis | Squash | *W* (white) epistatic to *G/g* (yellow/green) | Fruit Color | 12 (white) : 3 (yellow) : 1 (green) |  |
| Dominant Suppression | *Primula* | *D* (suppressor) epistatic to *K* (malvidin production) | Malvidin Production | 13 (no malvidin) : 3 (malvidin) |  |
| Recessive Epistasis | Rodents | e.g., *e/e* (yellow) epistatic to *B/b* (black/brown) | Coat Color | 9 (agouti) : 3 (black) : 4 (yellow) | (General example) |
| Two-Locus QTL Interaction | *C. sativa* | LG6.35 x LG9.40 | CBDa concentration | 9.59% of variance explained by interaction |  |
| Two-Locus QTL Interaction | *C. sativa* | LG6.35 x LG9.40 | CBGa concentration | 8.97% of variance explained by interaction |  |
| Two-Locus QTL Interaction | *C. sativa* | LG6.35 x LG9.40 | THCa concentration | 8.36% of variance explained by interaction |  |
| Two-Locus QTL Interaction | *C. sativa* | LG6.35 (AP.3) x LG9.40 (AP.6) | \alpha-pinene (ppm) | 7.64% of variance explained by interaction |  |
| Two-Locus QTL Interaction | *C. sativa* | LG6.35 (3C.1) x LG9.40 (3C.2) | 3-Carene (ppm) | 15.11% of variance explained by interaction |  |
| Additive x Additive (AaAa) | *C. sativa* | Multiple autosomal loci | THC concentration | Significant contribution to variance (model dependent) |  |

This table provides a comparative overview of classical Mendelian epistatic interactions and quantitative epistatic effects identified through QTL mapping in cannabis, underscoring the importance of considering gene interactions in genetic analyses and breeding strategies.

## 3. Pleiotropy: One Gene, Multiple Cannabis Traits

### The Concept of Pleiotropy

Pleiotropy is a fundamental genetic principle describing the phenomenon where a single gene exerts influence over two or more distinct, and often seemingly unrelated, phenotypic traits. The molecular mechanisms underlying pleiotropy are diverse. A gene product, such as a protein, might participate in multiple biochemical pathways, possess several functional domains each interacting with different cellular components, or act as a regulatory molecule (e.g., a transcription factor or signaling protein) that affects the expression or activity of various downstream target genes or processes. Alternatively, an enzyme might catalyze reactions involving different substrates or produce multiple products, each contributing to a different phenotype.

Pleiotropy can manifest in ways that are either challenging or advantageous for plant breeding. **Antagonistic pleiotropy** occurs when a favorable effect of a gene on one trait is coupled with an unfavorable effect on another trait. This creates a genetic constraint, making it difficult to simultaneously improve both traits by selecting for that particular gene. Conversely, **adaptive pleiotropy** (or positive pleiotropy) describes situations where a single gene has beneficial effects on multiple desired traits, making it a highly valuable target for selection as it can lead to concerted improvement across those traits.

Examples of pleiotropy are widespread in the plant kingdom. Classic examples include Gregor Mendel's observation that a single gene in peas (*Pisum sativum*) controlled flower color, seed coat color, and the presence of axillary pigmentation. In wheat (*Triticum aestivum*), the *Q* gene influences multiple yield components, while the Green Revolution gene *Rht-B1b* affects not only plant height but also various other architectural and yield-related traits. Similarly, in rice (*Oryza sativa*), the *NAL1* gene impacts both yield and the rate of photosynthesis.

### Evidence of Pleiotropic QTLs and Genes in Cannabis

Accumulating evidence from genetic mapping studies in *Cannabis sativa* indicates that pleiotropy is a common feature of its genetic architecture, influencing a wide range of agronomic and chemotypic traits.

**Agronomic Traits:** Research led by McKay and colleagues on a 'Carmagnola' x 'USO31' hemp F2 population revealed that many QTLs associated with agronomic traits tended to co-localize into a few key genomic regions. This pattern suggests that a relatively small number of loci might be responsible for a significant proportion of the phenotypic distinctions between these cultivars. Specifically, QTLs on Linkage Group 3 (notably around 60 cM, designated LG3.60) and Linkage Group 5 (around 5-10 cM, designated LG5.05) were found to be associated with multiple agronomic characteristics, including leaf water content, plant height, stem diameter, stem biomass, total dry biomass, and days to maturity. A compelling candidate gene underlying a major QTL cluster for several agronomic traits on LG5.05 was identified as a homolog of the *Arabidopsis thaliana* gene *TINY*. *TINY* encodes an ethylene-responsive transcription factor known to influence overall plant growth. Allelic variation within the cannabis *TINY* homolog was significantly associated with differences across multiple growth-related traits, strongly supporting its pleiotropic role in determining plant architecture and vigor in cannabis.

**Biochemical Traits (Cannabinoids and Terpenes):** Similar patterns of QTL co-localization indicative of pleiotropy have been observed for biochemical traits. The same research group identified extensive clustering of QTLs for a diverse array of cannabinoids (including CBDa, CBGa, THCa, and CBC) and numerous terpenes (such as \alpha-pinene, \beta-caryophyllene, and camphene) on Linkage Groups 6 (LG6.35) and 9 (LG9.40). A particularly strong candidate gene for the prominent QTL cluster on LG9.40, affecting multiple cannabinoids and terpenes, is *Olivetol Synthase (OLS)*. OLS is a key enzyme in the cannabinoid biosynthetic pathway, catalyzing the formation of olivetolic acid, a direct precursor to cannabigerolic acid (CBGA). Functional validation experiments demonstrated that allelic variation in the cannabis *OLS* gene affected olivetol production levels, which could explain differences in overall cannabinoid accumulation between cultivars. The association of this *OLS* locus with a QTL cluster influencing not only multiple cannabinoids but also various terpenes suggests a pleiotropic effect. This could occur through the control of olivetolic acid availability (which combines with geranyl pyrophosphate, a precursor also utilized in terpene synthesis) or via closely linked regulatory networks that coordinate both pathways.

**Morphology, Flowering Time, and Fiber Quality:** In hemp, genes that regulate flowering time are well-documented to exert pleiotropic effects on other important traits such as plant height, seed yield, cell wall composition (which directly impacts fiber quality), and the extent of secondary growth. For instance, members of the *FLOWERING LOCUS T (FT)* gene family, which are central regulators of flowering, can influence the balance between vegetative growth and the transition to flowering, thereby affecting overall plant architecture and biomass accumulation. The developmental shift to flowering also triggers significant changes in carbon partitioning within the plant, leading to altered stem development and fiber characteristics, including increased lignification and the formation of secondary bast fibers.

Sex determination genes in dioecious hemp also exhibit pleiotropy, influencing plant morphology (male plants are often more slender and flower earlier), flowering time, and fiber quality (fibers from male plants can differ in properties from those of female plants). Furthermore, QTLs identified for fiber yield and quality in hemp often have candidate genes involved in lignin and pectin biosynthesis; selection targeting these genes could pleiotropically affect other cell wall-related traits and overall plant structure. An example of a specific gene with pleiotropic effects on reproductive morphology is *CsMIKC1*, which is associated with a major QTL for inflorescence number and also influences overall flower production and grain yield in cannabis. Transcription factors such as *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPLs)*, *EARLY FLOWERING (ELFs)*, *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1)*, and *CENTRORADIALIS-Like (CEN-Like)* genes exhibit distinct expression patterns during vegetative-to-reproductive phase transition and flowering, suggesting broad regulatory roles that could pleiotropically affect various developmental processes and, consequently, yield-related traits.

**Cannabis Use Disorder (CUD):** While these studies focus on human genetics rather than plant traits, research into Cannabis Use Disorder (CUD) in humans has identified pleiotropic genetic loci that are shared with other complex disorders and traits, such as Alzheimer's disease, neuroplasticity, and immune system responses. This illustrates the general biological principle that single genetic regions can influence multiple, seemingly disparate complex phenotypes, a concept that is equally relevant to understanding the intricate genetic architecture of traits in *Cannabis sativa*.

The prevalence of such pleiotropic QTL "hubs" in cannabis, like the region on LG9.40 influencing chemotype and the region on LG5.05 affecting agronomy, is a significant finding. This pattern suggests that a limited number of key regulatory genes or critical pathway enzymes may function as master controllers for entire suites of related traits. From a breeding perspective, this could imply that fine-tuning diverse cannabis phenotypes might be more effectively achieved by modulating the activity or expression of these few hub genes, rather than attempting to independently alter dozens of downstream structural genes. This simplifies the conceptual model for breeding: instead of targeting a multitude of individual genes, breeders might concentrate on identifying and selecting for specific alleles or regulatory variants of these "hub" genes. However, this modular control also presents a challenge: achieving very specific, independent modifications in traits that are co-regulated by the same pleiotropic hub could be difficult without employing advanced genetic tools, such as gene editing, to potentially decouple the unwanted pleiotropic effects from the desired ones.

### Challenges and Opportunities Presented by Pleiotropy in Breeding

The occurrence of pleiotropy presents both challenges and opportunities for cannabis breeding programs:

* **Correlated Responses to Selection:** Due to pleiotropy, selection for a specific trait can lead to unintended changes in other traits that are influenced by the same gene(s). If these pleiotropic effects are antagonistic (e.g., a gene that increases cannabinoid yield also reduces plant vigor or disease resistance), it creates a significant breeding challenge, as improving one trait may come at the expense of another.
* **Breaking Linkages versus Leveraging Pleiotropy:** If multiple desirable traits are influenced by different, unfavorably coupled alleles of the same pleiotropic gene, or by tightly linked genes where one has a positive effect and the other a negative effect, it can be very difficult to combine all desired characteristics in a single cultivar. Recombination to break tight linkages may be infrequent. Conversely, if a single gene exhibits favorable pleiotropic effects on several desired traits (adaptive pleiotropy), it becomes an extremely valuable target for selection, as it allows for simultaneous improvement of multiple characteristics.
* **Understanding Molecular Mechanisms:** Elucidating the precise molecular basis of pleiotropy (e.g., how a gene like *OLS* or *TINY* affects multiple distinct traits) is crucial. If different functional domains of a protein are responsible for its different pleiotropic effects, it might be possible, through advanced techniques like gene editing, to modulate specific effects independently.
* **Multi-trait Selection Indices:** Given the prevalence of pleiotropy, cannabis breeding programs must often employ multi-trait selection indices. These indices assign economic or breeding weights to various traits and aim to achieve a balanced improvement in overall genetic merit, rather than maximizing a single trait in isolation, especially when undesirable correlated responses are likely.

It is important to note that the observation of pleiotropy often stems from the co-localization of QTLs for different traits to the same genomic region. While true pleiotropy (a single gene causing multiple effects) is a strong possibility in such cases, the current resolution of QTL mapping in cannabis may not always definitively distinguish between true pleiotropy and the tight physical linkage of multiple distinct genes, where each gene affects a different trait but all are located very close together on the chromosome. Cannabis genomes are characterized by high heterozygosity and a significant proportion of repetitive DNA sequences , which can complicate fine-mapping efforts and broaden QTL confidence intervals. Linkage disequilibrium (LD), the non-random association of alleles at different loci, can also extend over considerable genomic distances in some populations, making it difficult to resolve closely linked genes. Therefore, while compelling candidate genes like *OLS* and *TINY* are proposed for the pleiotropic effects observed within their respective QTL regions, it remains formally possible that, for example, the QTL peak for plant height and the QTL peak for stem diameter on LG5.05 are attributable to two different, but very closely linked, genes rather than the *TINY* gene alone affecting both traits. Further fine-mapping of these QTL regions, combined with candidate gene validation studies (e.g., using association mapping in diverse germplasm, allele mining, gene expression analyses, transformation, or knockout/knockdown experiments), is essential to confirm true pleiotropy versus close linkage. This distinction is critical for breeding: if traits are governed by linked genes, there is a possibility (however small if linkage is very tight) that recombination could eventually separate them, offering more flexibility in breeding than if the traits are inextricably linked by a single pleiotropic gene.

### Table 2: Documented Pleiotropic Genes and QTLs in Cannabis and Other Plants

| Gene/QTL Region (Example) | Organism | Multiple Traits Influenced | Putative Molecular Function (if gene) | Reference(s) |
| --- | --- | --- | --- | --- |
| *TINY* homolog (on LG5.05) | *C. sativa* | Plant height, stem diameter, stem biomass, dry biomass, leaf water content, days to maturity | Ethylene-responsive transcription factor |  |
| *Olivetol Synthase (OLS)* (on LG9.40) | *C. sativa* | Multiple cannabinoids (CBDa, CBGa, THCa, etc.) and multiple terpenes (\alpha-pinene, \beta-caryophyllene, etc.) | Polyketide synthase in cannabinoid pathway |  |
| QTL on LG3.60 | *C. sativa* | Plant height, stem diameter, stem biomass, dry biomass, leaf water content, seed yield, days to maturity | Unknown (region contains multiple genes) |  |
| QTL on LG6.35 | *C. sativa* | Multiple cannabinoids (CBDa, CBGa, THCa, etc.) and multiple terpenes (\alpha-pinene, camphene, etc.) | Unknown (region contains multiple genes) |  |
| *CsMIKC1* | *C. sativa* | Inflorescence number per branch, flower production, grain yield | MADS-box transcription factor |  |
| Flowering Time Genes (e.g., *FT* homologs) | *C. sativa* | Flowering time, plant height, seed yield, cell wall composition (fiber quality), secondary growth | Flowering pathway integrator / Transcription factor |  |
| Sex Determination Loci | *C. sativa* | Sex expression, plant morphology (stature), flowering time, fiber quality | Genes involved in sex chromosome differentiation |  |
| *A* gene | Pea | Flower color, seed coat color, axillary pigmentation | bHLH Transcription factor |  |
| *Q* gene | Wheat | Grain yield, grains per m$^{-2}$, 1000-grain weight, grains per spike/spikelet | AP2-like transcription factor |  |
| *Rht-B1b* / *Rht-D1b* | Wheat | Plant height (semi-dwarfism), yield components, coleoptile length, anther retention, disease susceptibility (e.g., FHB) | DELLA protein (GA signaling repressor) |  |
| *NAL1 (GPS)* | Rice | Grain number, plant architecture, panicle structure, leaf width, photosynthesis rate | Protein of unknown function, affects polar auxin transport |  |
| *FRI (FRIGIDA)* | *Arabidopsis* | Flowering time (late flowering), number of inflorescence nodes, number of branches | Protein involved in vernalization pathway |  |

This table systematically lists examples where single genetic factors are known to influence multiple downstream traits, highlighting the interconnectedness of phenotypes and the importance of considering such effects in selection programs.

## 4. Quantitative Trait Loci (QTL) Mapping for Complex Cannabis Traits

### Principles of QTL Mapping

Quantitative traits, such as yield, potency, and specific terpene profiles in *Cannabis sativa*, are characterized by continuous variation and are influenced by the combined effects of multiple genes (polygenes) and their interactions with the environment. Quantitative Trait Locus (QTL) mapping is a statistical methodology employed to identify specific genomic regions (loci) that contain genes contributing to this variation in a quantitative trait. The fundamental principle underlying QTL mapping is the detection of a statistical association between phenotypic variation for a trait and genotypic variation at polymorphic molecular marker loci distributed throughout the genome.

The process of QTL mapping typically involves several key steps:

1. **Development of a Mapping Population:** This is a critical initial step and involves creating a segregating population, usually by crossing two parental lines that exhibit significant phenotypic contrast for the trait(s) of interest and are genetically diverse enough to ensure sufficient marker polymorphism. Common types of mapping populations include F2 populations (progeny from selfing or intercrossing F1 individuals), backcross (BC) populations, Recombinant Inbred Lines (RILs), and Doubled Haploids (DHs). The size of the mapping population is a crucial factor for statistical power; populations of 200-300 individuals are often considered a minimum for detecting QTLs with moderate effects, with larger populations (500+) being preferable for higher resolution and detection of minor QTLs.
2. **Genotyping and Linkage Map Construction:** The individuals in the mapping population are genotyped using a suite of polymorphic molecular markers (e.g., Single Nucleotide Polymorphisms (SNPs), Simple Sequence Repeats (SSRs)). These markers are then used to construct a genetic linkage map, which illustrates the linear order of markers on each chromosome and the relative genetic distances between them, typically measured in centiMorgans (cM) based on recombination frequencies. A dense and accurate linkage map is essential for precisely locating QTLs.
3. **Phenotyping:** Each individual in the mapping population must be accurately phenotyped for the quantitative trait(s) under investigation. This often involves growing the population in replicated trials, potentially across multiple environments (locations and/or years), to obtain reliable phenotypic data and to assess the stability of QTL effects and the presence of QTL x Environment (GxE) interactions.
4. **Statistical Methods for QTL Detection:** Various statistical approaches are used to test for associations between marker genotypes and trait phenotypes:
   * **Single Marker Analysis (SMA):** This is the simplest method, involving tests like t-tests, Analysis of Variance (ANOVA), or simple linear regression to assess the association of each marker locus with the trait independently. While straightforward, SMA is less powerful if the QTL is not very tightly linked to the marker being tested and does not provide precise estimates of QTL location.
   * **Interval Mapping (IM):** This method evaluates genomic intervals flanked by pairs of linked markers for the presence of a QTL. It typically uses maximum likelihood or regression methods to estimate the probability (often expressed as a Logarithm of Odds, or LOD score) of a QTL residing within an interval. A common threshold for declaring a significant QTL is a LOD score greater than 3.0. IM is generally more powerful and provides better localization of QTLs than SMA.
   * **Composite Interval Mapping (CIM):** CIM enhances IM by incorporating additional markers (cofactors) located outside the test interval into the statistical model, typically through multiple regression. This helps to control for the effects of other QTLs segregating in the population, thereby increasing the precision of QTL detection and reducing bias in the estimation of QTL effects, especially when multiple QTLs are located on the same chromosome.
   * **Bayesian Methods and Genome-Wide Association Studies (GWAS):** While traditional QTL mapping is typically performed in structured biparental populations, GWAS is used to identify marker-trait associations in diverse, often unrelated individuals from natural populations or breeding germplasm collections. GWAS leverages historical recombination events and linkage disequilibrium (LD) across the genome.

### Key QTLs Identified in Cannabis

QTL mapping studies in *Cannabis sativa* have begun to identify genomic regions associated with important agronomic and chemotypic traits.

**Yield and Agronomic Characteristics:** Research by McKay and colleagues, using an F2 population from a 'Carmagnola' × 'USO31' hemp cross, identified 34 QTLs influencing eight distinct agronomic traits: leaf water content, plant height, thousand seed mass, stem diameter, stem biomass, seed yield, total dry biomass, and days to maturity. A significant finding was the co-localization of many of these QTLs, particularly to regions on linkage groups (LG) 3 and 5. The candidate gene *TINY*, an ethylene-responsive transcription factor homolog located on LG5, was associated with QTLs for multiple growth-related traits, suggesting a pleiotropic role in plant development. General approaches to mapping yield-related QTLs in other crops, such as rice, often reveal QTL clusters and identify trait-enhancing alleles from wild relatives or diverse germplasm, providing relevant strategies for hemp improvement programs.

**Cannabinoid Content and Potency (THC, CBD, CBG, etc.):** Early work by Weiblen et al. (2015) reported a single major QTL responsible for the THC:CBD chemotype ratio. Subsequent research by Grassa et al. (2021), using the same mapping population, further mapped loci associated with cannabinoid variation and proposed candidate genes. The comprehensive study by McKay's group identified 35 QTLs for 17 biochemical traits, including the major cannabinoids CBDa, THCa, CBGa, and cannabichromene (CBC). Key QTLs for these compounds were found to cluster on LG6 and LG9. *Olivetol synthase (OLS)*, an enzyme crucial for the biosynthesis of olivetolic acid (a cannabinoid precursor), was identified as a strong candidate gene underlying cannabinoid variation associated with the QTL cluster on LG9. Stack et al., in dissertation research, identified sixteen cannabinoid QTLs in an F2 population. This work highlighted that while variation in cannabinoid synthase loci (e.g., *THCAS*, *CBDAS*) directly affects the THC:CBD ratio, overall cannabinoid potency (total cannabinoid content) was also associated with QTLs on other chromosomes. For example, a potent QTL influencing total cannabinoid quantity was located on chromosome 3, unlinked to the cannabinoid synthase gene clusters on chromosome 9 (formerly often referred to as LG1 in some maps). Research from Southern Cross University also reported a major QTL cluster for principal cannabinoids on chromosome 7.

**Terpene Profiles and Diversity:** The study by McKay et al. also identified QTLs for a wide range of terpenes, including \alpha-pinene, \beta-caryophyllene, myrcene, and limonene. Many of these terpene QTLs co-localized with cannabinoid QTLs, particularly on LG6 and LG9, suggesting linked genetic control or pleiotropic effects. Stack et al. identified ten terpene QTLs, including a significant QTL influencing the ratio of limonene to \alpha-pinene, indicating genetic control over relative terpene proportions. The Southern Cross University group found a major QTL cluster for monoterpenes on chromosome 5, with candidate terpene synthase (*TPS*) genes identified within this region. Analogous QTL mapping methodologies in other aromatic plants, such as tea (*Camellia sinensis*), have successfully identified QTL clusters for volatile terpenes and candidate genes including *TPSs*, 1-deoxy-D-xylulose-5-phosphate synthase (*DXS*), and transcription factors, offering valuable insights and approaches applicable to cannabis terpene research.

The recurrent discovery of QTL clusters influencing functionally related traits—such as the co-localization of QTLs for multiple cannabinoids and terpenes on LG6 and LG9, and for various agronomic traits on LG3 and LG5 —is a compelling observation. This pattern suggests that the cannabis genome may be organized into "trait modules." Such modules could be controlled by sets of physically linked genes that contribute to a common physiological process or by single pleiotropic regulatory genes that orchestrate the expression of multiple downstream target genes involved in related traits. This modular genetic architecture could be an evolutionary adaptation or a consequence of human selection pressures that have historically targeted general chemotypes (e.g., high THC or high CBD) or overall morphotypes (e.g., fiber-type vs. drug-type). From a breeding standpoint, this modularity implies that it might be relatively straightforward to select for broad phenotypic profiles (e.g., "high overall cannabinoid content with a specific class of terpenes") by targeting these major QTL cluster regions. However, it also suggests that independently fine-tuning individual compounds or traits within such a module could be more challenging if the underlying genes are very tightly linked or if the module is governed by a single, highly pleiotropic gene.

### Methodological Considerations and Limitations in Cannabis QTL Studies

Despite progress, QTL mapping in cannabis faces several methodological challenges and limitations:

* **High Heterozygosity:** As a predominantly dioecious and outcrossing species, cannabis exhibits high levels of heterozygosity. This complicates accurate genotyping, often requiring greater sequencing depth for SNP calling and making the determination of linkage phase (which allele comes from which parent) more difficult.
* **Reference Genome Issues:** Historically, the lack of a single, universally adopted, high-quality annotated reference genome for *C. sativa* has been a significant impediment, confounding the comparison of QTL locations and candidate genes across different studies. Although several genome assemblies are now available (e.g., for 'Purple Kush', 'Finola', CBDRx), issues such as assembly gaps, potential misassemblies, variable annotation quality, and inconsistent chromosome nomenclature persist. The International Cannabis Research Consortium (ICRC) proposed the CBDRx Cs10 assembly as a reference to improve standardization.
* **Repetitive DNA Content:** The cannabis genome is characterized by a high proportion of repetitive DNA elements (estimated at ~70%). These repetitive sequences complicate de novo genome assembly, accurate read mapping for marker discovery (especially from short-read sequencing data), and can lead to errors in variant calling.
* **Population Structure:** In GWAS, underlying population structure (i.e., genetic relatedness among individuals that is not accounted for) can lead to spurious marker-trait associations. While some studies focusing on specific germplasm pools (e.g., Canadian drug-type accessions) have reported low levels of population structure , broader collections of cannabis germplasm can be highly structured due to geographic origin, historical breeding practices, and differentiation between hemp and drug types.
* **Phenotyping Accuracy and GxE Interactions:** The accurate and precise measurement of complex quantitative traits is paramount for reliable QTL detection. However, phenotyping can be laborious, expensive, and influenced by environmental factors. QTL effects can also vary across different environments (QTL x Environment or GxE interactions), meaning that a QTL detected in one environment may not be significant or may have a different effect size in another. Multi-environment trials are necessary to identify stable QTLs versus those with environment-specific effects.
* **The "Beavis Effect":** In initial QTL mapping studies, particularly those employing relatively small population sizes, the magnitude of the effects of detected QTLs tends to be overestimated. This statistical artifact necessitates validation of QTLs in larger populations or in different genetic backgrounds before they are widely used in breeding.
* **Limited Resolution of QTLs:** QTL mapping typically identifies relatively broad genomic regions (often spanning several centiMorgans and potentially containing hundreds of genes) that are associated with trait variation. Fine-mapping these QTLs to pinpoint the causal gene(s) and the specific causative polymorphism(s) requires significant additional resources, such as the development of larger mapping populations, near-isogenic lines (NILs), or the use of association mapping with very high marker density.

The evolution of genetic understanding in cannabis, particularly for traits like chemotype, mirrors the historical progression seen in other species for complex traits. Early models often proposed simple, Mendelian inheritance for traits like the THC:CBD ratio, frequently invoking a single major locus with co-dominant alleles (e.g., the 'B' locus). However, as more sophisticated QTL mapping studies have been conducted, it has become clear that such traits are typically polygenic, involving multiple QTLs, and are further complicated by epistatic interactions between these loci. For instance, the work by McKay's group explicitly contradicts the single-locus model for cannabinoid production, identifying at least four distinct loci contributing to variation in these chemotypes. This shift from simplistic Mendelian approximations to more nuanced quantitative genetic frameworks signifies the maturation of cannabis genetics research. It underscores the critical importance of applying robust quantitative genetic methodologies to accurately dissect trait architecture. Relying on oversimplified genetic models can be misleading for breeding efforts and can provide an incomplete picture of heritability and predictability. The "more complex scenario with several linked paralogs" for cannabinoid synthase genes, alluded to in earlier literature , is now being more clearly elucidated through detailed QTL mapping and genomic analyses.

### Table 3: Summary of Key QTLs for Agronomic and Chemotype Traits in Cannabis

| Trait Category | Specific Trait | Linkage Group/Chromosome | QTL Name/Region (cM or bp) | LOD Score | PVE (%) | Candidate Gene(s) | Reference(s) |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Agronomic** | Plant Height | LG3 | ~60.3 cM (PLHT.1) | 19.41 | 22.35 | - |  |
|  | Plant Height | LG5 | ~9.3 cM (PLHT.2) | 14.01 | 15.38 | *TINY* (near SB.3) |  |
|  | Stem Biomass | LG5 | ~9.3 cM (SB.3) | 32.48 | 31.05 | *TINY* |  |
|  | Seed Yield | LG2 | ~98.5 cM (SY.1) | 12.59 | 11.35 | - |  |
|  | Seed Yield | LG5 | ~3.0 cM (SY.4) | 14.15 | 12.92 | *TINY* (near) |  |
|  | Days to Maturity | LG5 | ~3.0 cM (DTM.3) | 33.10 | 34.27 | *TINY* (near) |  |
| **Cannabinoid** | CBDa | LG6 | ~33.7 cM (CBDa.1) | 32.56 | 28.21 | - |  |
|  | CBDa | LG9 | ~41.1 cM (CBDa.2) | 44.85 | 44.79 | *Olivetol Synthase (OLS)* |  |
|  | THCa | LG6 | ~28.9 cM (THCa.1) | 20.58 | 24.54 | - |  |
|  | THCa | LG9 | ~41.1 cM (THCa.2) | 25.56 | 32.10 | *OLS* |  |
|  | CBGa | LG6 | ~33.4 cM (CBGa.1) | 20.32 | 26.46 | - |  |
|  | CBGa | LG9 | ~41.1 cM (CBGa.2) | 25.53 | 35.21 | *OLS* |  |
|  | Total Cannabinoid (Potency) | Chr 3 | Not specified | - | 17% | Not linked to cannabinoid synthase clusters |  |
|  | THC:CBD Ratio | Major QTL | Not specified | - | Large | Cannabinoid synthases (*THCAS*, *CBDAS*) |  |
|  | Major Cannabinoids | Chr 7 | QTL Cluster | - | Major | Candidate cannabinoid synthase |  |
| **Terpene** | \alpha-pinene | LG9 | ~43.0 cM (AP.6) | 32.09 | 31.04 | *OLS* (linked region) |  |
|  | \beta-caryophyllene | LG9 | ~41.3 cM (BC.3) | 35.06 | 43.88 | *OLS* (linked region) |  |
|  | Geraniol | LG9 | ~36.0 cM (GE.1) | 26.39 | 43.33 | *OLS* (linked region) |  |
|  | Limonene:\alpha-pinene ratio | Not specified | Significant QTL | - | - | - |  |
|  | Monoterpenes | Chr 5 | QTL Cluster | - | Major | Candidate terpene synthases |  |
|  | Volatile Terpenes (Tea) | LG05 (Tea genome) | QTL Cluster (e.g. qGER5.1) | >3.0 | ~10-18% | *TPS*, *DXS*, *GRAS* TF, *bHLH* TF |  |

This table serves as a centralized summary of current QTL knowledge for major cannabis traits, enabling researchers and breeders to identify genomic regions pertinent to their specific interests for further investigation or incorporation into breeding strategies. (Note: LG assignments can vary between studies depending on the reference map used; specific marker names are often numerous and not always reported concisely in summaries).

## 5. Genetic Markers for Targeted Cannabis Breeding

### Overview of Molecular Marker Types and Their Application in Cannabis

Molecular markers are identifiable DNA sequences that exhibit polymorphism (variation) among individuals within a population. These markers, by virtue of their linkage to genes controlling traits of interest, can be used to indirectly track the inheritance of those genes or genomic regions. Several types of molecular markers have been employed in cannabis research and breeding:

* **Simple Sequence Repeats (SSRs or Microsatellites):** These markers are based on variations in the number of tandemly repeated short DNA sequences (typically 2-6 base pairs long). SSRs are generally co-dominant (allowing heterozygotes to be distinguished from homozygotes), highly polymorphic, reproducible, and relatively inexpensive to assay without requiring high-throughput sequencing technology for analysis. They have proven useful for genetic diversity assessment, variety identification, parentage analysis, and the construction of genetic linkage maps in cannabis.
* **Single Nucleotide Polymorphisms (SNPs):** SNPs are variations at a single nucleotide position in the DNA sequence. They are the most abundant type of genetic variation in most genomes, including cannabis. SNPs are amenable to high-throughput genotyping platforms, such as arrays and Genotyping-by-Sequencing (GBS), making them ideal for constructing high-density linkage maps, performing genome-wide association studies (GWAS), and for genomic selection.
* **Other Marker Types:** Earlier research in cannabis also utilized markers such as Random Amplified Polymorphic DNA (RAPD) and Sequence Characterized Amplified Regions (SCARs). RAPDs are generated by PCR using arbitrary primers, while SCARs are derived from specific DNA fragments (often initially identified by RAPD or AFLP) that are converted into more robust, sequence-specific PCR assays. While RAPDs are less reproducible, SCARs have been valuable for targeting specific loci, such as those involved in chemotype or sex determination.

The primary applications of these molecular markers in *Cannabis sativa* include: assessing genetic diversity and population structure, managing germplasm resources, identifying and authenticating cultivars, determining plant sex at early developmental stages, mapping genes and QTLs associated with desirable traits (e.g., cannabinoid profiles, terpene content, yield, disease resistance, flowering time), and ultimately, facilitating Marker-Assisted Selection (MAS) in breeding programs.

### Known Markers Associated with Specific Desirable Traits

**Cannabinoid Chemotypes (THC:CBD ratio, specific cannabinoids):** The genetic control of the primary cannabinoid chemotype (i.e., THC-dominant, CBD-dominant, or intermediate) is largely determined by alleles at the cannabinoid synthase loci.

* **SNPs in Cannabinoid Synthase Genes:** Variations within the coding sequences of *THCA synthase (THCAS)* and *CBDA synthase (CBDAS)* are critical. Specific SNPs have been identified that can lead to non-functional or altered-function enzymes, thereby drastically changing the plant's chemotype. For example, a SNP at position 706 in the *THCAS* gene resulting in an amino acid change from glutamic acid to glutamine was associated with significantly reduced THCA production and accumulation of CBGA. Similarly, a C-to-T SNP at position 583 in the *CBDAS* coding sequence, found in several cultivars, introduces a premature stop codon, leading to an inactive CBDAS enzyme and consequently very low or no CBD production. Other mutations near the FAD binding site or catalytic site of these synthases are also implicated in altered enzyme activity.
* **SCAR Markers:** SCAR markers have been developed to differentiate major chemotypes. The B1080/B1192 multiplex PCR assay, based on *CBDAS* and *THCAS* gene sequences, proved effective in distinguishing chemotype I (THC-dominant, 1192 bp band), chemotype III (CBD-dominant, 1080 bp band), and chemotype II (intermediate, both bands). While an earlier marker, B190/B200, showed some utility, B1080/B1192 offered more reliable codominant scoring. Markers targeting the broader "B locus," historically considered the primary chemotype-determining region, have also been utilized.

**Terpene Profiles:** The development of specific, widely validated markers for individual terpenes or complex terpene profiles is less advanced than for major cannabinoids. However, progress is being made:

* QTL studies are identifying genomic regions associated with the content of various terpenes. Markers (typically SNPs) within these QTL regions are candidate tools for MAS.
* Genetic variation in *terpene synthase (TPS)* genes is known to be a primary driver of terpene diversity in cannabis. SNPs located within or near *TPS* genes are therefore prime targets for marker development.
* The identification of a significant QTL for the limonene:\alpha-pinene ratio suggests that markers could be developed to select for specific relative proportions of terpenes, which contribute significantly to aroma and potential entourage effects.

**Yield Components:** QTLs associated with various yield components (e.g., biomass, seed yield, plant height) have been identified in cannabis. Markers linked to these QTLs, often SNPs derived from WGS data, are potential tools for MAS. However, the polygenic nature and complex inheritance of yield traits make simple MAS challenging.

**Disease Resistance:** A significant focus has been on developing markers for resistance to powdery mildew (PM), a prevalent disease in cannabis cultivation.

* ***PM1* Locus:** A dominant PM resistance gene, designated *PM1*, has been mapped to a region on one of the largest cannabis chromosomes. This locus is rich in genes containing nucleotide-binding site (NBS) and leucine-rich repeat (LRR) domains, characteristic of plant R-genes. SNP markers linked to *PM1* have been identified.
* ***PM2* Locus:** A novel, single dominant PM resistance locus, *PM2*, has been identified on chromosome 9. SNP-based genetic markers (validated using PACE genotyping assays) capable of tracking *PM2* resistance in breeding populations have been developed. Candidate genes within the *PM2* region include those involved in hormonal regulation (salicylic acid pathway), ROS accumulation, cell death induction, and putative R-proteins (e.g., an LRR-Ser/Thr kinase).
* ***CsMLO1* Gene:** Mutations in the *Mildew Locus O (MLO)* gene *CsMLO1*, which typically acts as a susceptibility gene, can confer broad-spectrum resistance to PM. A 6.8-kb insertion in *CsMLO1* in the resistant cultivar 'FL 58' leads to a truncated, non-functional protein. Molecular marker assays (both fluorescence-based PACE and gel-based PCR) have been developed to detect this specific insertion and distinguish PM-resistant (*mlo*) from PM-susceptible genotypes.

**Flowering Time (including Day-Neutral/Autoflower trait):** Control over flowering time is crucial for adapting cultivars to different latitudes and production systems.

* GWAS and QTL mapping studies have identified several loci associated with variation in flowering time. Candidate genes in these regions include those involved in light perception (phytochromes, cryptochromes), transcription factors central to flowering pathways (e.g., homologs of *FT*, *SOC1*, *AP1*, *LFY*), and the microRNA *miR156*.
* **Day-Neutral (Autoflower) Trait:** This trait, where flowering is independent of photoperiod, is typically controlled by a single recessive gene located in a large region on Chromosome 1. Candidate genes implicated include *TARGET OF EARLY ACTIVATION TAGGED (TOE)/APETALA2 (AP2)* and *PSEUDO-RESPONSE REGULATOR 3 (PRR3)*, with disruptions in key domains observed in day-neutral plants. Specific SNP markers (e.g., contig504\_3790889, contig262\_2502527) and associated TaqMan-based qPCR assays have been developed for this trait. Mutations in a *CsPRR37* homolog have also been correlated with the autoflowering phenotype. Additionally, *FT-like* genes exhibit sex-specific expression and copy number variation, potentially contributing to flowering time differences.

**Sex Determination:** Distinguishing male (XY) from female (XX) plants at an early stage is critical for cannabinoid and floral production, as male plants are generally undesirable.

* Y-specific markers have been developed to identify male plants. Multiplex PCR assays combining markers for Y-specific coding regions with an autosomal control marker have demonstrated high accuracy in sexing cannabis seedlings.
* The SCAR119 marker has been widely used for male identification.
* PACE-PCR Allele Competitive Extension markers have also shown high success in sex identification.

**Plant Morphology:** Markers are used to assess genetic diversity related to broad morphological classifications, such as differentiating hemp-type from drug-type cannabis, or characterizing plants based on vernacular terms like "sativa" or "indica" (which often correlate with leaf morphology and plant stature). Markers linked to QTLs for specific morphological traits like plant height and leaf shape are also emerging from mapping studies.

### Development and Validation of Markers for Marker-Assisted Selection (MAS)

Marker-Assisted Selection (MAS) is a breeding strategy that utilizes DNA markers closely linked to genes or QTLs controlling desirable traits to select superior individuals within a breeding population. The goal is to improve the efficiency, speed, and precision of selection compared to relying solely on phenotypic evaluation. The general process involves identifying robust marker-trait associations through QTL mapping or association studies, and then using these markers to screen large numbers of plants, often at early developmental stages.

**Challenges in Cannabis MAS:**

* **Trait Complexity:** Many important traits in cannabis, such as overall yield or nuanced chemotype profiles (specific ratios of multiple cannabinoids and terpenes), are highly polygenic and influenced by epistatic interactions. Simple MAS targeting one or a few markers is often insufficient for such complex traits; it is most effective for simply inherited traits or those controlled by major QTLs with large effects.
* **Marker Validation:** A significant challenge is the need for robust validation of markers across diverse genetic backgrounds. A marker that is predictive of a trait in the specific population in which it was identified may not be informative or may have a different linkage phase with the causal gene in unrelated germplasm. The high genetic diversity within *C. sativa* exacerbates this issue.
* **Cost and Throughput:** While the costs of genotyping have decreased significantly, large-scale MAS programs still require investment in marker assays and laboratory infrastructure. High-throughput genotyping platforms, such as PACE genotyping for SNPs and Indels, are improving efficiency.
* **Standardization:** The historical lack of a standardized set of universally informative SNP markers for cannabis has been a limitation, hindering comparative studies and broader applicability of findings.

**Successes and Potential:** Despite these challenges, MAS holds considerable potential in cannabis breeding, particularly for:

* **Chemotype Selection:** Markers for major cannabinoid synthase alleles (e.g., functional vs. non-functional *THCAS* or *CBDAS*) are relatively effective for selecting plants with desired THC:CBD ratios (e.g., ensuring THC levels remain below legal thresholds in hemp, or maximizing THC or CBD in drug-type cultivars).
* **Sex Determination:** Markers that reliably identify male plants at the seedling stage allow for their early removal, which is crucial for maximizing yield and quality in unpollinated female flower production systems.
* **Major Disease Resistance Genes:** Markers for simply inherited resistance genes, such as those for powdery mildew (*PM1*, *PM2*, *mlo* alleles), enable efficient introgression of resistance into susceptible elite cultivars.

The trajectory of marker development in cannabis appears to prioritize traits that have significant regulatory or economic implications and relatively simpler genetic control. For instance, markers for "negative" traits—such as those to ensure low THC content in hemp , to identify and eliminate male plants in sinsemilla production , or to select for major disease resistance genes —have seen more rapid and successful development. This is likely driven by the strong economic incentives and regulatory pressures associated with these characteristics, as well as their often monogenic or oligogenic inheritance. In contrast, markers for highly complex "positive" traits, such as the precise combination of multiple terpenes contributing to a specific aroma profile, or subtle components of overall yield, are emerging more slowly from QTL studies. This suggests a phased adoption of MAS in cannabis breeding: initial successes will continue to focus on traits with simpler genetic underpinnings, while the improvement of highly polygenic traits will likely necessitate more advanced approaches like genomic selection, even as individual QTL markers for components of these complex traits are progressively identified and validated.

Furthermore, the vast genetic diversity and complex population structure within *Cannabis sativa* as a species pose a significant hurdle to the universal applicability of many currently developed markers, especially SNPs discovered through GBS or WGS projects in specific populations or QTL studies. As demonstrated by the day-neutral trait markers that were not universally applicable across unrelated cultivars , a marker's linkage phase with a QTL, or even its presence and polymorphism, can vary substantially in different genetic backgrounds. This necessitates either the development of "universal" markers, perhaps targeting highly conserved functional regions that harbor causative polymorphisms across diverse germplasm, or the creation and validation of population-specific marker sets. It also implies that marker information from one research study or breeding program may not be directly transferable to another's unique germplasm without careful re-validation, underscoring the need for cautious interpretation and application of published marker-trait associations.

### Table 4: Genetic Markers Associated with Desirable Traits in Cannabis

| Trait | Marker Type(s) | Specific Marker Name/Locus (Example) | Associated Gene(s) (if known) | Key Findings/Utility for MAS | Reference(s) |
| --- | --- | --- | --- | --- | --- |
| **Chemotype (THC/CBD Ratio)** | SNP, SCAR | SNPs in *THCAS/CBDAS*, B1080/B1192 SCAR | *THCAS*, *CBDAS* | Distinguishes THC-dom, CBD-dom, intermediate types; selection for low THC in hemp or high THC/CBD in drug-types. |  |
| **Powdery Mildew Resistance** | SNP, PACE assay, Gel PCR | *PM1* linked SNPs, *PM2* linked SNPs, *CsMLO1* insertion/SNP markers | *PM1* (NBS-LRR type), *PM2* (LRR-kinase type), *CsMLO1* | Selection for resistance to powdery mildew; *mlo* provides broad-spectrum resistance. Enables pyramiding of R-genes. |  |
| **Flowering Time (Day-Neutral/Autoflower)** | SNP, qPCR assay | contig504\_3790889, contig262\_2502527, SNPs in *CsPRR37* | *TOE/AP2*, *PRR3*, *CsPRR37* | Early selection for day-neutral flowering, accelerating breeding cycles and adapting cultivars to diverse environments. |  |
| **Sex Determination** | SCAR, Multiplex PCR (Y-specific SNPs) | SCAR119, MADC1/MADC2 (Y-specific) | Y-chromosome specific sequences | Early identification and removal of male plants in populations grown for floral products. |  |
| **Terpene Profile (Ratio)** | SNP (from QTL) | QTL for Limonene:\alpha-pinene ratio | Putative *TPS* genes in QTL region | Potential for selecting specific terpene ratios influencing aroma/flavor. |  |
| **Agronomic (e.g., Plant Height)** | SNP (from QTL) | Markers linked to QTL on LG5 | *TINY* homolog | Selection for plant architecture traits, though complex and polygenic. |  |

This table provides a practical summary of some of the molecular tools currently available or under active development for cannabis breeding, directly addressing key traits of interest.

## 6. Gene Interactions in Cannabinoid and Terpene Biosynthetic Pathways

### Overview of Cannabinoid Biosynthesis Pathway

The biosynthesis of cannabinoids in *Cannabis sativa* is a complex enzymatic process primarily occurring in the glandular trichomes of female flowers. The pathway initiates with precursors from two major metabolic routes: geranyl pyrophosphate (GPP), derived from the plastidial methylerythritol 4-phosphate (MEP) pathway, and olivetolic acid (OA), synthesized via a polyketide pathway involving the condensation of hexanoyl-CoA with malonyl-CoA units.

Key enzymes in the formation of OA include olivetol synthase (OLS) and olivetolic acid cyclase (OAC). Subsequently, an aromatic prenyltransferase, often referred to as cannabigerolic acid synthase (CBGAS) or geranylpyrophosphate:olivetolate geranyltransferase (GPT), catalyzes the crucial alkylation step, condensing OA with GPP to form cannabigerolic acid (CBGA). CBGA serves as the central precursor, or "mother cannabinoid," from which other major cannabinoids are derived.

From CBGA, the pathway branches out, with specific oxidocyclase enzymes converting CBGA into the acidic forms of the principal cannabinoids:

* **Tetrahydrocannabinolic acid synthase (THCAS)** converts CBGA to \Delta^9-tetrahydrocannabinolic acid (THCA).
* **Cannabidiolic acid synthase (CBDAS)** converts CBGA to cannabidiolic acid (CBDA).
* **Cannabichromenic acid synthase (CBCAS)** converts CBGA to cannabichromenic acid (CBCA). These three synthases compete for the common substrate CBGA, and their relative activities and expression levels largely determine the primary cannabinoid profile (chemotype) of a given cannabis plant. The acidic cannabinoids (THCA, CBDA, CBCA, etc.) are the predominant forms found in fresh plant material. They undergo non-enzymatic decarboxylation (loss of a carboxyl group) upon exposure to heat or prolonged storage, converting them into their pharmacologically active neutral forms (THC, CBD, CBC, respectively).

### Overview of Terpene Biosynthesis Pathway

Terpenes (or terpenoids) are a vast and diverse class of volatile organic compounds responsible for the characteristic aromas and flavors of cannabis, and they are also believed to contribute to its therapeutic effects. Their biosynthesis also originates from fundamental isoprenoid units:

* The **Methylerythritol 4-Phosphate (MEP) pathway**, located in plastids, is the primary source of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) for the synthesis of monoterpenes (C$*{10}$ compounds) and diterpenes (C$*{20}$ compounds).
* The **Mevalonic Acid (MVA) pathway**, occurring in the cytosol, primarily contributes IPP and DMAPP for the synthesis of sesquiterpenes (C$*{15}$ compounds) and triterpenes (C$*{30}$ compounds). While these pathways are largely distinct, some crosstalk in terms of precursor exchange can occur between them.

IPP and DMAPP are condensed to form geranyl diphosphate (GPP, C$*{10}), the direct precursor for monoterpenes like limonene, myrcene, and pinene. Further addition of an IPP unit to GPP yields farnesyl diphosphate (FPP, C*{15}$), the precursor for sesquiterpenes such as \beta-caryophyllene and humulene. The final step in producing the diverse array of individual terpenes involves a large family of enzymes called **terpene synthases (TPSs)**. These enzymes catalyze complex cyclization and modification reactions of GPP and FPP. *Cannabis sativa* possesses a suite of *TPS* genes, and genetic variation within these genes is a major contributor to the wide chemical diversity of terpene profiles observed across different cannabis strains.

### Evidence of Co-regulation and Interactions

The cannabinoid and terpene biosynthetic pathways, while leading to structurally distinct classes of compounds, are intricately linked and exhibit evidence of co-regulation and interaction at multiple levels:

* **Shared Precursors:** A critical point of intersection is the precursor **geranyl pyrophosphate (GPP)**. GPP, primarily derived from the MEP pathway, is essential for the synthesis of CBGA (the cannabinoid precursor) and is also the direct precursor for all monoterpenes. This shared reliance on GPP creates a metabolic branch point where the flux of this intermediate can be channeled towards either cannabinoid or monoterpene production, implying a potential for competition and metabolic trade-offs.
* **Co-expression Networks:** Transcriptomic studies of cannabis glandular trichomes—the primary sites for both cannabinoid and terpene synthesis—have revealed significant co-expression networks involving genes from both pathways. For instance, a co-expression module containing *THCA synthase* was found to be enriched not only for other cannabinoid pathway genes (like *CBGA synthase*) but also for genes involved in the MEP pathway (supplying GPP), sesquiterpene production, and even upstream glycolysis (providing initial carbon skeletons). This coordinated expression pattern strongly suggests shared or interconnected transcriptional regulatory mechanisms.
* **QTL Co-localization:** As discussed previously (Section 4), QTLs for various cannabinoids and terpenes frequently co-localize to the same genomic regions, notably on linkage groups 6 and 9 in the 'Carmagnola' x 'USO31' population. The *OLS* locus on LG9, a key determinant of cannabinoid levels, also appears to influence terpene profiles. Such co-localization can indicate either tight physical linkage of distinct causal genes or, more compellingly, pleiotropic effects of single genes or regulatory elements within these regions that affect both pathways.
* **Oxylipin Pathway Interaction:** Emerging research suggests a link between the oxylipin pathway and cannabinoid biosynthesis. The oxylipin pathway, involving enzymes like lipoxygenase (LOX) and hydroperoxide lyase (HPL), is hypothesized to provide hexanoyl-CoA, an early precursor for olivetolic acid synthesis. The co-expression of *LOX* and *HPL* genes with cannabinoid biosynthetic genes within trichomes lends support to this metabolic connection.
* **Epistatic Interactions:** The epistatic interactions observed between QTLs on LG6 and LG9, which affect the accumulation of multiple cannabinoids and terpenes , further point to a complex interplay in the regulation or biochemical steps of these pathways.

The competition for the common precursor GPP between the monoterpene and cannabinoid biosynthetic pathways represents a significant metabolic reality. This implies a potential trade-off: if cellular resources or enzymatic activities are heavily skewed towards producing a large quantity of monoterpenes, the availability of GPP for CBGA synthase could be diminished, potentially leading to lower overall cannabinoid production, and vice-versa. This dynamic assumes that the upstream MEP pathway's capacity to produce GPP is not infinitely adjustable or perfectly synchronized with the demands of both downstream pathways. For cannabis breeders aiming to develop cultivars with specific chemotypes (e.g., high cannabinoids *and* a rich monoterpene profile), this presents both a challenge and an opportunity. It suggests that simply upregulating the terminal synthase for one class of compounds might negatively impact the other. A more effective strategy might involve understanding and manipulating the regulation of GPP production itself (e.g., by enhancing MEP pathway output) or influencing the partitioning of GPP between the competing cannabinoid and terpene synthases. Genetic modifications or breeding selections that boost overall MEP pathway flux might be necessary to simultaneously elevate both cannabinoid and monoterpene levels.

### Epigenetic and Transcriptional Regulation

While the direct enzymatic steps are becoming clearer, the overarching regulation of these pathways involves complex transcriptional and potentially epigenetic mechanisms:

* **Epigenetic Regulation:** Epigenetic modifications, such as DNA methylation, histone modifications (acetylation, methylation), and the activity of non-coding RNAs (e.g., microRNAs), can influence gene expression without altering the underlying DNA sequence. These mechanisms are known to play roles in regulating cannabinoid receptor function in animals and are increasingly recognized as important in plant secondary metabolism, often mediating responses to environmental stimuli. Exposure of cannabis plants or cells to certain conditions or compounds has been shown to alter DNA methylation patterns , suggesting that epigenetic factors could modulate the expression of cannabinoid and terpene biosynthetic genes.
* **Transcriptional Regulation:** The coordinated expression of genes within these pathways strongly implies the involvement of transcription factors (TFs). While specific TFs directly controlling clusters of cannabinoid or terpene biosynthetic genes in cannabis are still being elucidated, TFs are generally the master regulators of such pathways in plants. The *TINY* gene, an ethylene-responsive TF, was identified as a candidate for agronomic QTLs with pleiotropic effects. Other TFs, such as those from the MYB and HD-ZIP IV families, are known to be involved in trichome development —the very sites of synthesis—and thus indirectly control the capacity for cannabinoid and terpene production. Future research will likely uncover specific TFs that bind to promoter regions of synthase genes (e.g., *THCAS*, *CBDAS*, *TPSs*) and orchestrate their expression levels.
* **Developmental Regulation:** Cannabinoid and terpene profiles are not static but change dynamically throughout the plant's life cycle, particularly during the flowering stage. This indicates robust developmental control over the activity of these biosynthetic pathways, likely mediated by changes in gene expression programs triggered by developmental cues and hormonal signals.

The emerging link between the oxylipin pathway and cannabinoid biosynthesis, where oxylipin pathway enzymes like LOX and HPL may provide an initial substrate (hexanoyl-CoA) for olivetolic acid synthesis , suggests a deeper integration of primary metabolism (fatty acid metabolism, from which oxylipins derive) and specialized secondary metabolism than has been traditionally emphasized in discussions of cannabis chemotype. Oxylipins, including jasmonates, are well-established signaling molecules in plants, mediating responses to wounding, pathogen attack, and other stresses. If cannabinoid synthesis is metabolically connected to the oxylipin pathway, it implies that environmental stresses or agronomic practices that trigger oxylipin production might also modulate cannabinoid levels, not just through generalized stress physiology but potentially through specific alterations in precursor supply. This opens new avenues for investigating how cultivation conditions affecting plant stress and fatty acid metabolism could be optimized to influence cannabinoid profiles. Furthermore, natural genetic variation in oxylipin pathway genes, which has largely been unexplored in cannabis, could represent an additional source of variation contributing to differences in cannabinoid content among cultivars.

### Table 5: Key Genes and Enzymes in Cannabinoid and Terpene Biosynthetic Pathways and their Interactions

| Pathway | Key Gene/Enzyme | Substrate(s) | Product(s) | Potential Points of Interaction/Co-regulation | Reference(s) |
| --- | --- | --- | --- | --- | --- |
| **Shared Precursor** | MEP Pathway enzymes (e.g., *DXS*, *DXR*) | Pyruvate, Glyceraldehyde-3-phosphate | IPP, DMAPP | Supplies precursors for GPP; Co-expression with cannabinoid/terpene pathways. |  |
|  | Geranyl Pyrophosphate Synthase (GPPS) | IPP + DMAPP | Geranyl Pyrophosphate (GPP) | **Crucial branch point**: GPP is used by both cannabinoid and monoterpene synthases. Competition for GPP. |  |
| **Cannabinoid** | Olivetol Synthase (OLS) / Olivetolic Acid Cyclase (OAC) | Hexanoyl-CoA, Malonyl-CoA | Olivetolic Acid (OA) | Key precursor for CBGA. *OLS* locus (LG9) shows pleiotropic effects on cannabinoids & terpenes and epistatic interactions. Co-expressed with oxylipin pathway genes. |  |
|  | Cannabigerolic Acid Synthase (CBGAS / GPT) | Olivetolic Acid + GPP | Cannabigerolic Acid (CBGA) | Central precursor for major cannabinoids. Competes with monoterpene synthases for GPP. |  |
|  | THCA Synthase (THCAS) | CBGA | THCA | Competes with CBDAS & CBCAS for CBGA. SNPs affect activity. Co-expressed with terpene pathway genes. |  |
|  | CBDA Synthase (CBDAS) | CBGA | CBDA | Competes with THCAS & CBCAS for CBGA. SNPs affect activity. |  |
|  | CBCA Synthase (CBCAS) | CBGA | CBCA | Competes with THCAS & CBDAS for CBGA. |  |
| **Terpene** | Various Terpene Synthases (TPSs) e.g., Limonene Synthase, Myrcene Synthase | GPP (for monoterpenes), FPP (for sesquiterpenes) | Diverse monoterpenes & sesquiterpenes | Genetic variation in *TPS* genes drives terpene diversity. Some TPSs co-expressed with cannabinoid pathway genes. Competes with CBGAS for GPP. |  |
|  | Farnesyl Pyrophosphate Synthase (FPPS) | GPP + IPP | Farnesyl Pyrophosphate (FPP) | Precursor for sesquiterpenes. |  |
| **Interacting Pathway** | Lipoxygenase (LOX), Hydroperoxide Lyase (HPL) | Fatty acids (e.g., Linolenic acid) | Hexanal / Hexanoic acid (potential precursor for OA) | Hypothesized to supply early precursors for cannabinoid pathway. Co-expression with cannabinoid genes in trichomes. |  |

This table visually outlines the core biosynthetic steps and highlights the shared components and enzymes, illustrating the biochemical basis for the observed genetic interactions, co-regulation, and pleiotropic effects that collectively determine the final chemotype of *Cannabis sativa*.

## 7. Integrative Strategies for Cannabis Cultivar Improvement

The growing understanding of advanced genetic mechanisms like epistasis and pleiotropy, coupled with progress in QTL mapping and marker development, provides a foundation for more sophisticated and efficient strategies in cannabis cultivar improvement. Moving beyond simple phenotypic selection, integrative approaches that leverage genomic information are becoming increasingly crucial.

### Leveraging Epistatic and Pleiotropic Understanding in MAS

Marker-Assisted Selection (MAS) aims to accelerate genetic gain by selecting for desirable alleles indirectly using linked molecular markers. However, the complexities introduced by epistasis and pleiotropy necessitate refinements to traditional MAS strategies:

* **Beyond Single QTLs:** For traits influenced by epistasis, selection based on individual QTLs assuming additive effects may be suboptimal. The value of an allele at one locus can be masked or modified by alleles at interacting loci. Therefore, MAS strategies must evolve to consider these interactions.
* **Selection for Favorable Allelic Combinations:** Effective MAS for epistatic traits should focus on identifying and selecting for specific *combinations* of alleles at the interacting loci that together produce the desired phenotype. This might involve developing marker haplotypes that span these interacting genomic regions or using models that explicitly account for interaction effects when calculating marker scores.
* **Managing Pleiotropic Effects:** When a targeted QTL exhibits pleiotropy, breeders must carefully evaluate all significant trait effects. If antagonistic pleiotropy is present (e.g., a QTL allele increases cannabinoid content but reduces disease resistance), breeders face a trade-off. Strategies might involve accepting the compromise, attempting to break the undesirable linkage if the effects are due to closely linked genes rather than true pleiotropy (though this can be difficult), or seeking alternative alleles with more favorable pleiotropic profiles. Conversely, if a QTL has favorable pleiotropic effects on multiple desired traits (e.g., enhancing both yield and a specific quality attribute), it becomes a highly valuable target for selection, simplifying the breeding process.
* **Pyramiding Genes/QTLs:** MAS is a powerful tool for pyramiding (combining) multiple desirable genes or QTLs into a single elite background. This is particularly relevant for stacking genes for different disease resistances, or combining QTLs for chemotype control with those for agronomic performance. However, a nuanced understanding of epistasis is critical during gene pyramiding, as the combined phenotypic effect of the pyramided genes may not be purely additive; synergistic or antagonistic interactions can occur.

### Marker-Assisted Recurrent Selection (MARS) and F2 Enrichment

For complex traits governed by multiple QTLs, including those with epistatic effects, recurrent selection strategies enhanced by molecular markers can be effective:

* **Marker-Assisted Recurrent Selection (MARS):** MARS aims to gradually increase the frequency of favorable alleles for multiple QTLs within a breeding population over several cycles of intermating and selection. In each cycle, individuals are genotyped, and selection decisions are based on marker scores that can incorporate information from multiple QTLs, potentially weighted by their estimated effects (including interaction terms if known). This approach is particularly suited for improving quantitative traits with moderate to low heritability.
* **F2 Enrichment:** This strategy involves using markers to cull undesirable genotypes in early segregating generations (e.g., F2). By eliminating individuals lacking key favorable alleles or possessing undesirable linkage blocks, breeders can enrich the population for superior genotypes before committing resources to extensive phenotypic evaluation in later generations.

### Genomic Selection (GS)

For highly complex traits influenced by numerous small-effect QTLs and pervasive epistatic interactions, where individual QTL effects are difficult to resolve or model effectively, Genomic Selection (GS) is emerging as a powerful alternative or complement to traditional MAS.

* GS utilizes dense, genome-wide marker data (typically thousands to millions of SNPs) to predict the genetic merit (Genomic Estimated Breeding Value, or GEBV) of individuals. Unlike MAS, which focuses on specific, validated QTLs, GS models simultaneously estimate the effects of all markers across the genome, assuming that most markers are in linkage disequilibrium with at least one QTL.
* Importantly, GS models can, in principle, capture not only additive genetic variance but also components of non-additive variance, including dominance and epistasis, depending on the model used. This can lead to more accurate predictions of overall genetic performance for complex traits.

### Addressing Challenges in Cannabis Breeding Programs

The application of these advanced strategies in cannabis breeding must contend with several inherent challenges:

* **High Genetic Diversity & Population Structure:** The vast genetic diversity within *C. sativa* and the often-strong population structure (e.g., between hemp and drug types, or among different landraces and modern cultivars) necessitate careful selection of parental material for developing mapping populations and training populations for GS. Markers and QTL effects identified in one genetic background may not be transferable to another without validation.
* **Long Generation Times & Dioecy:** Cannabis can have a relatively long generation interval compared to some annual crops, and its dioecious nature complicates controlled crossing and the development of inbred lines. MAS can help accelerate selection, particularly by enabling selection at the seedling stage for traits expressed later in development (e.g., flower chemistry) or for sex determination. The use of techniques to produce feminized seeds (by inducing male flowers on female plants) can aid in managing dioecy for specific breeding objectives.
* **Cost of Genotyping and Phenotyping:** While the costs of genotyping have decreased substantially, large-scale MAS or GS programs still represent a significant investment, particularly for high-density SNP genotyping. Moreover, accurate and high-throughput phenotyping for complex traits (e.g., detailed chemotype analysis, multi-environment yield trials) remains a major bottleneck and expense.
* **Genotype x Environment (GxE) Interactions:** The effects of many QTLs, especially those for complex traits like yield or secondary metabolite production, can be significantly influenced by environmental conditions. Breeding programs must conduct multi-environment trials to identify QTLs that are stable across diverse environments or to select for QTLs that confer specific adaptation to target environments.
* **Integration with Conventional Breeding:** MAS and GS are not replacements for conventional breeding methods but are powerful tools to enhance their efficiency and precision. Their most effective application comes from thoughtful integration into well-designed, long-term breeding schemes.

A crucial aspect often emphasized in marker-assisted backcrossing (MABC) in other crops is "background selection," which involves using genome-wide markers to select individuals that not only carry the desired target allele(s) (foreground selection) but also possess the highest proportion of the elite recurrent parent's genome. Given the high heterozygosity of cannabis and the common goal of introgressing specific traits (e.g., a novel cannabinoid profile or disease resistance) into established elite cultivars, the systematic application of background selection will become increasingly important. This helps to rapidly recover the desirable genetic background of the recurrent parent, minimize "linkage drag" (the unwanted co-introgression of deleterious alleles linked to the target gene from the donor parent), and reduce the number of backcross generations needed. While current MAS efforts in cannabis often highlight foreground selection for traits like chemotype , sex , or disease resistance , a more comprehensive approach incorporating background selection will be vital for efficiently developing improved cultivars that retain the overall quality and performance of the elite parental lines.

The successful and widespread application of MAS and GS for complex traits in cannabis will also heavily depend on the continued development and accessibility of robust, publicly available, and well-curated genomic resources. These include high-quality reference genome assemblies, comprehensive databases of validated molecular markers, QTL atlases summarizing marker-trait associations across diverse germplasm, and user-friendly bioinformatics tools. The "translation gap"—the lag between research discoveries in genomics and their practical application in breeding programs—has been noted for many other crop species. This gap could be particularly pronounced in cannabis, given its historically fragmented research landscape, unless concerted efforts are made in collaborative resource building, data standardization, and open data sharing initiatives. Without such foundational infrastructure, individual breeding programs may struggle to implement advanced molecular breeding techniques effectively or will face the substantial cost of developing these resources independently.

## 8. Conclusion and Future Perspectives

### Summary of Current Understanding

The genetic architecture of *Cannabis sativa* is proving to be remarkably complex, with phenomena such as epistasis and pleiotropy playing significant roles in shaping key agronomic and chemotypic traits. Epistatic interactions, where the effect of one gene is conditional upon alleles at other loci, have been documented for cannabinoid and terpene biosynthesis, indicating that these pathways are not regulated in isolation but are subject to intricate gene-gene interplay. Pleiotropy, where single genes or QTLs influence multiple distinct traits, is also evident, with specific genomic regions and candidate genes like *TINY* and *Olivetol Synthase* implicated in controlling suites of related agronomic or biochemical characteristics, respectively.

Considerable progress has been made in QTL mapping for traits of economic importance, including yield components, cannabinoid profiles (THC, CBD, CBG), and terpene diversity. These studies are beginning to pinpoint genomic regions and, in some cases, candidate genes responsible for variation in these complex traits. Concurrently, the development of molecular markers, particularly SNPs and SSRs, is advancing. Markers are now available, or are in advanced stages of development, for critical traits such as cannabinoid chemotype, sex determination, powdery mildew resistance, and flowering time (including the day-neutral/autoflower characteristic). These markers are paving the way for the implementation of marker-assisted selection (MAS) in cannabis breeding. Furthermore, research is highlighting the interconnectedness of biosynthetic pathways, such as those for cannabinoids and terpenes, which share precursors and exhibit co-regulation at the transcriptional level, often involving epistatic and pleiotropic genetic control.

### Outlook on Future Research Directions

Despite these advances, much remains to be explored to fully harness the genetic potential of *Cannabis sativa*. Key future research directions include:

* **Fine-Mapping and Causal Gene Validation:** A critical next step is to move beyond the identification of broad QTL regions to the fine-mapping and functional validation of the specific causal genes and polymorphisms underlying trait variation. This will necessitate the use of techniques such as high-resolution linkage mapping in larger populations, association mapping in diverse germplasm, detailed transcriptomic and proteomic analyses of contrasting genotypes, metabolomic profiling to link genetic changes to biochemical outcomes, and targeted gene editing technologies (e.g., CRISPR/Cas9) to confirm gene function.
* **Expanded Genome-Wide Association Studies (GWAS):** Conducting GWAS in larger and more diverse cannabis germplasm collections, encompassing landraces, feral populations, and modern cultivars from various geographical origins, will be essential for discovering a wider array of alleles and loci contributing to complex traits. This will also provide a better understanding of population-specific genetic variations and adaptation.
* **Elucidating Regulatory Networks:** A major frontier is the elucidation of the complex transcriptional and epigenetic regulatory networks that govern quantitative traits, particularly the coordinated expression of genes in the cannabinoid and terpene biosynthetic pathways. Identifying key transcription factors, microRNAs, and epigenetic modifications (e.g., DNA methylation, histone marks) that modulate these pathways will offer new targets for manipulating chemotypes.
* **Multi-omics Integration:** Integrating data from multiple "omics" platforms—genomics (DNA sequence), transcriptomics (gene expression), proteomics (protein abundance and modifications), and metabolomics (metabolite profiles)—will be crucial for building comprehensive, systems-level models of how genetic variation translates into phenotypic outcomes.
* **Improving Genomic Resources:** Continued investment in developing high-quality, chromosome-level genome assemblies for a wider range of diverse cannabis genotypes is needed. The creation of a cannabis pangenome, representing the entire gene repertoire of the species, along with improved genome annotation and curated databases for variants and QTLs, will significantly enhance the power and accuracy of genomic research.
* **Developing Advanced Breeding Tools and Strategies:** Refining MAS strategies to effectively incorporate information on epistasis and pleiotropy, implementing genomic selection (GS) for complex traits, and exploring novel breeding techniques such as speed breeding (to shorten generation times) and the development of true F1 hybrid breeding systems informed by an understanding of heterotic groups, will be key to accelerating genetic gain.

The future of cannabis breeding, particularly for optimizing complex chemotypes, will likely necessitate a "systems genetics" approach. This paradigm moves beyond focusing on individual genes or QTLs in isolation and instead aims to understand and manipulate entire networks of interacting genes, proteins, and metabolites. The prevalence of epistasis, pleiotropy, co-localized QTL clusters, and co-regulated biosynthetic pathways strongly indicates that cannabis traits are not determined by independent genetic units but are emergent properties of a complex, interconnected biological system. Predictive models for breeding will therefore need to incorporate these network properties. Breeding strategies might evolve to target key regulatory nodes within these networks or aim to reconfigure network connections to achieve desired chemotypic outcomes, rather than simply attempting to stack individual "favorable" alleles in an additive fashion.

### Technological Advancements

The rapid pace of innovation in genomics and biotechnology will continue to drive cannabis research:

* **Sequencing Technologies:** Continuously declining sequencing costs and the increasing accuracy and read-length of next-generation and third-generation sequencing platforms will make whole-genome sequencing and resequencing more accessible, facilitating comprehensive variant discovery and pangenome construction.
* **High-Throughput Genotyping and Phenotyping:** Advances in automated, high-throughput genotyping (e.g., SNP arrays, GBS, amplicon sequencing) and phenotyping (e.g., image-based analysis, non-destructive chemical sensing) will enable the analysis of larger populations with greater precision.
* **Bioinformatics and Computational Biology:** Sophisticated bioinformatics pipelines, machine learning algorithms, and artificial intelligence will be indispensable for analyzing the vast and complex datasets generated by multi-omics studies and for developing accurate predictive models for GS and trait dissection.

### Bridging the Gap to Application

Translating fundamental research discoveries into practical breeding tools and improved cannabis cultivars remains a key objective and a significant challenge. This will require:

* **Collaboration:** Stronger collaborations between academic researchers, public institutions, and private sector breeders are essential to ensure that research priorities align with industry needs and that new technologies are effectively disseminated and implemented.
* **Data Sharing and Standardization:** Promoting open data sharing, establishing standardized protocols for phenotyping and genotyping, and developing curated, publicly accessible databases for cannabis genomic information will be crucial for accelerating progress across the entire community.
* **Addressing Regulatory Frameworks:** Navigating and adapting to evolving regulatory landscapes for cannabis cultivation and research will continue to be an important factor influencing the pace and direction of scientific advancement.

As cannabis genomics continues to mature, an increasing ethical and practical imperative will be the systematic characterization, conservation, and sustainable utilization of the vast genetic diversity residing within *Cannabis sativa*. Feral populations and traditional landraces, often found in diverse geographic regions, may harbor unique alleles for valuable traits such as abiotic stress tolerance, resistance to local pests and diseases, or novel cannabinoid and terpene profiles that are not present in highly selected commercial cultivars. Breeding efforts, particularly in the early stages of crop domestication or intensification, can inadvertently lead to a narrowing of genetic diversity by focusing on a limited set of elite lines. The genomic tools and QTL mapping approaches discussed in this report are not only useful for breeding but can also be powerfully applied to characterize genetic diversity within germplasm collections and to prioritize unique genetic resources for conservation. Proactive and well-funded initiatives for germplasm collection, comprehensive molecular and phenotypic characterization, and long-term ex situ and in situ conservation are essential. These genetic resources will be invaluable for addressing future breeding challenges, such as adapting cannabis to changing climatic conditions, combating new pest and disease pressures, and developing cultivars with entirely new chemotypic or agronomic profiles.

In conclusion, the sophisticated understanding of advanced genetic inheritance mechanisms—including epistasis, pleiotropy, and the genomic location of QTLs—is rapidly transforming *Cannabis sativa* from a species with a historically limited scientific knowledge base into one where modern molecular breeding techniques can be applied with increasing precision and efficacy. This ongoing revolution in cannabis genomics and genetics holds immense promise for the development of tailored cultivars designed to meet a wide spectrum of industrial, medicinal, and agricultural requirements, ultimately unlocking the full potential of this versatile plant.

#### Works cited

1. Recent advances in cannabis biotechnology - Sci-Hub, https://sci-hub.se/downloads/2020-10-24/dd/10.1016@j.indcrop.2020.113026.pdf 2. (PDF) The genetics of Cannabis -genomic variations of key ..., https://www.researchgate.net/publication/346355269\_The\_genetics\_of\_Cannabis\_-genomic\_variations\_of\_key\_synthases\_and\_their\_effect\_on\_cannabinoids\_content 3. From fibers to flowering to metabolites: unlocking hemp (Cannabis sativa) potential with the guidance of novel discoveries and tools - Oxford Academic, https://academic.oup.com/jxb/article/76/1/109/7776000 4. Highly Predictive Genetic Markers Distinguish Drug-Type from Fiber-Type Cannabis sativa L, https://pmc.ncbi.nlm.nih.gov/articles/PMC6918397/ 5. Pleiotropic Loci for Cannabis Use Disorder Severity in Multi ..., https://pmc.ncbi.nlm.nih.gov/articles/PMC10247496/ 6. Cannabinoid Inheritance Relies on Complex Genetic Architecture ..., https://pmc.ncbi.nlm.nih.gov/articles/PMC7173683/ 7. Quantitative trait loci controlling agronomic and biochemical traits in ..., https://pmc.ncbi.nlm.nih.gov/articles/PMC9335937/ 8. justagriculture.in, https://justagriculture.in/files/newsletter/2021/june/11.%20QTL%20Mapping%20in%20Crop%20Plants%20Principles%20and%20Applications.pdf 9. QTL Analysis and Mapping - Illumina, https://www.illumina.com/techniques/popular-applications/qtl-analysis.html 10. Pleiotropy, epistasis and the genetic architecture of quantitative traits - PubMed Central, https://pmc.ncbi.nlm.nih.gov/articles/PMC11330371/ 11. Quantitative Trait Loci Controlling Agronomic and Biochemical Traits in Cannabis sativa, https://www.researchgate.net/publication/352774618\_Quantitative\_Trait\_Loci\_Controlling\_Agronomic\_and\_Biochemical\_Traits\_in\_Cannabis\_sativa 12. Creating New Cannabis Varieties: An Overview of Breeding ..., https://floraflex.com/default/blog/post/creating-new-cannabis-varieties-an-overview-of-breeding-methods1 13. CCA | The Science of Cannabis Genetics and Breeding, https://cca2go.com/seeds-of-change-the-science-of-cannabis-genetics-and-breeding/ 14. (PDF) Genetics and Marker-assisted Selection of the Chemotype in ..., https://www.researchgate.net/publication/225148108\_Genetics\_and\_Marker-assisted\_Selection\_of\_the\_Chemotype\_in\_Cannabis\_sativa\_L 15. Developing and Testing Molecular Markers in Cannabis sativa ..., https://pmc.ncbi.nlm.nih.gov/articles/PMC8540786/ 16. Bridging the Gap Between Cannabis Breeding and Genetics: Creating Advanced Cannabis Strains - FloraFlex Media, https://floraflex.com/default/blog/post/bridging-the-gap-between-cannabis-breeding-and-genetics-creating-advanced-cannabis-strains 17. Plant Breeding & Marker-Assisted Selection - 3CR Bioscience, https://3crbio.com/blog/allele-specific-pcr/plant-breeding-and-marker-assisted-selection-spotlight-on-hemp-cannabis-sativa/ 18. IJPB | Special Issue : Challenges in Cannabis sativa: Breeding and ..., https://www.mdpi.com/journal/ijpb/special\_issues/6G8H6HB21Q 19. Cannabis sativa research trends, challenges, and new-age ..., https://pmc.ncbi.nlm.nih.gov/articles/PMC8605354/ 20. Genetic insights into agronomic and morphological traits of drug ..., https://pmc.ncbi.nlm.nih.gov/articles/PMC11033274/ 21. Genetic Variability of Morphological, Flowering, and ... - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2020.00102/full 22. Establishing robust tools for Cannabis propagation and variety improvement, https://researchportal.scu.edu.au/esploro/outputs/doctoral/Establishing-robust-tools-for-Cannabis-propagation/991013224113502368 23. Gene Interactions - Mendelian Genetics, https://www.ndsu.edu/pubweb/~mcclean/plsc431/mendel/mendel6.htm 24. Epistasis and functional genomics - Wikipedia, https://en.wikipedia.org/wiki/Epistasis\_and\_functional\_genomics 25. Epistasis | PPT - SlideShare, https://www.slideshare.net/slideshow/epistasis-16756180/16756180 26. Epistasis and cryptic QTL identified using modified bulk segregant analysis of copper resistance in budding yeast | Genetics | Oxford Academic, https://academic.oup.com/genetics/advance-article/doi/10.1093/genetics/iyaf026/8037412 27. Epistasis and cryptic QTL identified using modified bulk segregant analysis of copper resistance in budding yeast | bioRxiv, https://www.biorxiv.org/content/10.1101/2024.10.28.620582v2.full-text 28. Mapping complex traits in hemp - Genes to Genomes, https://genestogenomes.org/mapping-complex-traits-in-hemp/ 29. Cannabinoid Inheritance Relies on Complex Genetic Architecture | Request PDF - ResearchGate, https://www.researchgate.net/publication/333330625\_Cannabinoid\_Inheritance\_Relies\_on\_Complex\_Genetic\_Architecture 30. The impact of epistasis in the heterosis and combining ability analyses - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2023.1168419/full 31. Beyond the single gene: How epistasis and gene-by-environment effects influence crop domestication | PNAS, https://www.pnas.org/doi/10.1073/pnas.1308940110 32. Epistasis Is a Major Determinant of the Additive Genetic Variance in Mimulus guttatus - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC4422649/ 33. The Role of epistasis in controlling seed yield and other agronomic traits in an Andean × Mesoamerican cross of common bean, https://psfaculty.plantsciences.ucdavis.edu/gepts/Johnson%20Gepts%202002.pdf 34. Marker-assisted selection: new tools and strategies - OAR@ICRISAT, https://oar.icrisat.org/5764/1/TPS\_3\_236-239\_1998.pdf 35. Marker-Assisted Selection and Genomic Selection - Plant Breeding E-Learning in Africa, https://pbea.agron.iastate.edu/files/2021/11/Marker-Assisted-Selection-and-Genomic-Selection.pdf 36. Chapter 7: Marker Assisted Selection and Genomic Selection – Molecular Plant Breeding, https://iastate.pressbooks.pub/molecularplantbreeding/chapter/marker-assisted-selection-and-genomic-selection/ 37. Molecular genetics, genomics, and quantitative genetics of ..., https://conservancy.umn.edu/items/6d1634b2-c581-48f4-8ae2-484572b1de08 38. Epistasis and pleiotropy‐induced variation for plant breeding - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC11536456/ 39. Antagonistic pleiotropic effects reduce the potential adaptive value of the FRIGIDA locus, https://www.pnas.org/doi/10.1073/pnas.0708209104 40. Antagonistic pleiotropy can maintain fitness variation in annual plants - ResearchGate, https://www.researchgate.net/publication/320403483\_Antagonistic\_pleiotropy\_can\_maintain\_fitness\_variation\_in\_annual\_plants 41. Pleiotropy and the evolution of floral integration - University of Colorado Boulder, https://www.colorado.edu/smithlab/sites/default/files/attached-files/smith2015newphy.pdf 42. Dissection of Pleiotropic QTL Regions Controlling Wheat Spike Characteristics Under Different Nitrogen Treatments Using Traditional and Conditional QTL Mapping - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2019.00187/full 43. Quantitative trait loci controlling agronomic and biochemical traits in Cannabis sativa | Genetics | Oxford Academic, https://academic.oup.com/genetics/article/219/2/iyab099/6310019 44. The Complex Interactions Between Flowering Behavior and Fiber ..., https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2019.00614/full 45. The Complex Interactions Between Flowering Behavior and Fiber Quality in Hemp, https://www.researchgate.net/publication/333154928\_The\_Complex\_Interactions\_Between\_Flowering\_Behavior\_and\_Fiber\_Quality\_in\_Hemp 46. From fibers to flowering to metabolites: unlocking hemp (Cannabis ..., https://pmc.ncbi.nlm.nih.gov/articles/PMC11659183/ 47. CsMIKC1 regulates inflorescence development and grain ..., https://academic.oup.com/hr/article/11/8/uhae161/7691881 48. Transcriptomic analysis reveals differential gene expression patterns during cannabis leaf morphogenesis and phase transition | Request PDF - ResearchGate, https://www.researchgate.net/publication/384145917\_Transcriptomic\_analysis\_reveals\_differential\_gene\_expression\_patterns\_during\_cannabis\_leaf\_morphogenesis\_and\_phase\_transition 49. Characterizing the expression patterns of short-day flowering genes, the manipulation of their expression using red-far red wavelengths, and the resulting phenotype/ chemotype in Cannabis sativa | bioRxiv, https://www.biorxiv.org/content/10.1101/2024.03.28.587211v1.full-text 50. Comparative genomics of flowering behavior in Cannabis sativa - PMC - PubMed Central, https://pmc.ncbi.nlm.nih.gov/articles/PMC10421669/ 51. Pleiotropy and genetically inferred causality linking multisite chronic pain to substance use disorders. - FirstWord Pharma, https://firstwordpharma.com/story/5828529 52. Pleiotropic Loci for Cannabis Use Disorder Severity in Multi-Ancestry High-Risk Populations - medRxiv, https://www.medrxiv.org/content/10.1101/2022.11.25.22282743v1.full.pdf 53. Mapping of QTLs for Yield Traits Using F2:3:4 Populations Derived From Two Alien Introgression Lines Reveals qTGW8.1 as a Consistent QTL for Grain Weight From Oryza nivara - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2022.790221/full 54. Genome-Wide Dissection of Novel QTLs and Genes Associated with Weed Competitiveness in Early-Backcross Selective Introgression-Breeding Populations of Rice (Oryza sativa L.) - MDPI, https://www.mdpi.com/2079-7737/14/4/413 55. Full article: Genomic and Chemical Diversity in Cannabis - Taylor & Francis Online, https://www.tandfonline.com/doi/full/10.1080/07352689.2016.1265363 56. Composite genetic linkage and quantitative trait locus (QTL) map... - ResearchGate, https://www.researchgate.net/figure/Composite-genetic-linkage-and-quantitative-trait-locus-QTL-map-derived-from-a-marijuana\_fig4\_328647528 57. Identification of QTL controlling volatile terpene contents ... - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2023.1130582/full 58. pmc.ncbi.nlm.nih.gov, https://pmc.ncbi.nlm.nih.gov/articles/PMC8605354/#:~:text=Despite%20several%20cannabis%20genome%20assemblies,et%20al.%2C%202021). 59. Comparative genomics of flowering behavior in Cannabis sativa - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2023.1227898/full 60. Comparison of Recombination Rate, Reference Bias, and Unique Pangenomic Haplotypes in Cannabis sativa Using Seven De Novo Genome Assemblies - MDPI, https://www.mdpi.com/1422-0067/26/3/1165 61. Genetic Diversity and Population Structure of Cannabis Based on the Genome-Wide Development of Simple Sequence Repeat Markers - Frontiers, https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2020.00958/full 62. Functional QTL mapping and genomic prediction of canopy height in wheat measured using a robotic field phenotyping platform | Journal of Experimental Botany | Oxford Academic, https://academic.oup.com/jxb/article/71/6/1885/5757976 63. Marker-Assisted Selection as a Component of Conventional Plant Breeding - Regulations.gov, https://downloads.regulations.gov/AMS-NOP-15-0002-0875/attachment\_14.pdf 64. QTL-Seq and Transcriptome Analysis Disclose Major QTL and Candidate Genes Controlling Leaf Size in Sesame (Sesamum indicum L.) - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2021.580846/full 65. Pleiotropy and epistasis within and between signaling pathways defines the genetic architecture of fungal virulence - PLOS, https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1009313 66. (PDF) Developing and Testing Molecular Markers in Cannabis sativa (Hemp) for Their Use in Variety and Dioecy Assessments - ResearchGate, https://www.researchgate.net/publication/355204956\_Developing\_and\_Testing\_Molecular\_Markers\_in\_Cannabis\_sativa\_Hemp\_for\_Their\_Use\_in\_Variety\_and\_Dioecy\_Assessments 67. Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes - PMC - PubMed Central, https://pmc.ncbi.nlm.nih.gov/articles/PMC8397809/ 68. Potentials and Challenges of Genomics for Breeding ... - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2020.573299/full 69. Making Sense of Cannabis Strains Through Chemometrics in Review, https://www.cannabissciencetech.com/view/making-sense-cannabis-strains-through-chemometrics-review 70. Broad-Scale Genetic Diversity of Cannabis for Forensic Applications | PLOS One, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0170522 71. Analysis of Morphological Traits, Cannabinoid Profiles, THCAS Gene Sequences, and Photosynthesis in Wide and Narrow Leaflet High-Cannabidiol Breeding Populations of Medical Cannabis - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2022.786161/full 72. Volatile Compounds in Cannabis: From Synthesis to Scent - JumpLights®, https://jumplights.com/volatile-compounds-in-cannabis-from-synthesis-to-scent 73. (PDF) MINI-REVIEW CANNABIS TERPENES IN RELATION TO HUMAN HEALTH, https://www.researchgate.net/publication/366325090\_MINI-REVIEW\_CANNABIS\_TERPENES\_IN\_RELATION\_TO\_HUMAN\_HEALTH 74. Terpene synthases from Cannabis sativa | PLOS One, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0173911 75. Molecular Mechanisms Underlying Potential Pathogen Resistance in Cannabis sativa - PMC - PubMed Central, https://pmc.ncbi.nlm.nih.gov/articles/PMC10420965/ 76. Genetic Mapping, Identification, and Characterization of a Candidate ..., https://apsjournals.apsnet.org/doi/10.1094/MPMI-04-23-0043-R 77. Discovery and Genetic Mapping of PM1, a Powdery Mildew Resistance Gene in Cannabis sativa L. - Frontiers, https://www.frontiersin.org/journals/agronomy/articles/10.3389/fagro.2021.720215/full 78. NLR- and mlo-Based Resistance Mechanisms against Powdery Mildew in Cannabis sativa, https://www.mdpi.com/2223-7747/13/1/105 79. Mapping and characterization of a novel powdery mildew ... - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2025.1543229/full 80. Genetic Architecture of Flowering Time and Sex ... - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2020.569958/full 81. (PDF) Genetic Mapping of SNP Markers and Candidate Genes ..., https://www.researchgate.net/publication/370108054\_Genetic\_Mapping\_of\_SNP\_Markers\_and\_Candidate\_Genes\_Associated\_with\_Day-Neutral\_Flowering\_in\_Cannabis\_sativa\_L 82. FT-like genes in Cannabis and hops: sex specific expression and copy-number variation may explain flowering time variation | bioRxiv, https://www.biorxiv.org/content/10.1101/2024.10.04.616617v1 83. Promising high fidelity genetic markers for sexing Cannabis sativa seedlings | G3 Genes, https://academic.oup.com/g3journal/advance-article/doi/10.1093/g3journal/jkaf077/8109470?searchresult=1 84. High fidelity genetic markers for sexing Cannabis sativa seedlings - bioRxiv, https://www.biorxiv.org/content/10.1101/2024.09.10.612257v1.full.pdf 85. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century, https://pmc.ncbi.nlm.nih.gov/articles/PMC2610170/ 86. The Impact of Marker-Assisted Selection on Soybean Yield and, https://bioscipublisher.com/index.php/bm/article/html/3947/ 87. breeding in the twenty-first century Marker-assisted selection: an approach for precision plant - Purdue Department of Statistics, https://www.stat.purdue.edu/~doerge/BIOINFORM.D/SPRING13/CollardMackill2008.pdf 88. Characterization and expression analysis of MATEs in Cannabis sativa L. reveals genes involving in cannabinoid synthesis - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2022.1021088/full 89. Profiling Cannabinoid Contents and Expression Levels of Corresponding Biosynthetic Genes in Commercial Cannabis (Cannabis sativa L.) Cultivars - MDPI, https://www.mdpi.com/2223-7747/11/22/3088 90. Gene Networks Underlying Cannabinoid and Terpenoid ..., https://pmc.ncbi.nlm.nih.gov/articles/PMC6670104/ 91. Simulation of single-allele introgression in a Cannabis sativa-like model, https://dr.lib.iastate.edu/bitstreams/41045580-19a9-41d2-b51e-a565bb8a2fc1/download 92. Manipulation of Cannabinoid Biosynthesis via Transient RNAi Expression - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2021.773474/full 93. Structural and Chemical Biology of Terpenoid Cyclases - ACS Publications, https://pubs.acs.org/doi/10.1021/acs.chemrev.7b00287 94. Oxylipin biosynthetic gene families of Cannabis sativa | PLOS One, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0272893 95. Molecular Insights into Epigenetics and Cannabinoid Receptors - MDPI, https://www.mdpi.com/2218-273X/12/11/1560 96. Molecular Insights into Epigenetics and Cannabinoid Receptors - PMC - PubMed Central, https://pmc.ncbi.nlm.nih.gov/articles/PMC9687363/ 97. Neurogenetic and Epigenetic Aspects of Cannabinoids - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC9498086/ 98. The Cannabis-Induced Epigenetic Regulation of Genes Associated with Major Depressive Disorder - PMC - PubMed Central, https://pmc.ncbi.nlm.nih.gov/articles/PMC9407536/ 99. Cannabis use and the sperm epigenome: a budding concern? | Environmental Epigenetics, https://academic.oup.com/eep/article/6/1/dvaa002/5810192 100. Lifetime cannabis use and childhood trauma increase risk of psychosis in carriers of <i>CNR1</i> genetic variants: findings from the STREAM study - SciELO, https://www.scielo.br/j/rbp/a/rGWMGNvZQMzzct55YrmDcRv/ 101. Cannabis Pharmacogenomics: A Path to Personalized Medicine - MDPI, https://www.mdpi.com/1467-3045/45/4/228 102. ELUCIDATING THE GENETIC CONTROL OF QUALITATIVE TRAITS IN HEMP A Dissertation Presented to the Faculty of the Graduate School of - Cornell eCommons, https://ecommons.cornell.edu/server/api/core/bitstreams/dfb13426-ea4f-4ca4-beec-f15f6b23ad0a/content 103. Challenges for effective marker-assisted selection in plants - ResearchGate, https://www.researchgate.net/publication/23164584\_Challenges\_for\_effective\_marker-assisted\_selection\_in\_plants 104. Leveraging trait and QTL covariates to improve genomic prediction of resistance to Fusarium head blight in Central European winter wheat - Frontiers, https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2024.1454473/epub 105. Marker-Assisted Selection in Plant Breeding: From Publications to Practice - ResearchGate, https://www.researchgate.net/publication/228868218\_Marker-Assisted\_Selection\_in\_Plant\_Breeding\_From\_Publications\_to\_Practice 106. Marker-Assisted Selection in Plant Breeding: From Publications to Practice, http://image.sciencenet.cn/olddata/kexue.com.cn/upload/blog/file/2010/12/2010129145320339859.pdf 107. Molecular Plant Breeding as the Foundation for 21st Century Crop Improvement - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC2442525/ 108. Genomics-Assisted Breeding in Maize: Techniques and Outcomes | Zhou - CropSci Publisher, https://cropscipublisher.com/index.php/mgg/article/html/3903/ 109. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century | Philosophical Transactions of the Royal Society B - Journals, https://royalsocietypublishing.org/doi/10.1098/rstb.2007.2170 110. Implementation of molecular markers for quantitative traits in breeding programs - challenges and opportunities - USDA ARS, https://www.ars.usda.gov/ARSUserFiles/60701500/Publications/Holland/Holland4thInt'lCropSciCong04.pdf 111. Advances and Perspectives in Tissue Culture and Genetic Engineering of Cannabis - MDPI, https://www.mdpi.com/1422-0067/22/11/5671 112. Challenges and potentials of new breeding techniques in Cannabis sativa - PubMed Central, https://pmc.ncbi.nlm.nih.gov/articles/PMC10285108/ 113. Genetic Diversity, Population Structure, and Cannabinoid Variation in Feral Cannabis sativa Germplasm from the United States | bioRxiv, https://www.biorxiv.org/content/10.1101/2025.03.10.642411v1.full-text