CANDI: a web server for predicting molecular targets and pathways of cannabis-based therapeutics

Srinivasan Ekambaram[1,2] , Jian Wang[1,2] and Nikolay V. Dokholyan[1,2,3,4,5*]

Abstract

Background *Cannabis sativa* L. with a rich history of traditional medicinal use, has garnered significant attention in contemporary research for its potential therapeutic applications in various human diseases, including pain, inflammation, cancer, and osteoarthritis. However, the specific molecular targets and mechanisms underlying the synergistic effects of its diverse phytochemical constituents remain elusive. Understanding these mechanisms is crucial for developing targeted, effective cannabis-based therapies.

Methods To investigate the molecular targets and pathways involved in the synergistic effects of cannabis compounds, we utilized DRIFT, a deep learning model that leverages attention-based neural networks to predict compound-target interactions. We considered both whole plant extracts and specific plant-based formulations. Predicted targets were then mapped to the Reactome pathway database to identify the biological processes affected. To facilitate the prediction of molecular targets and associated pathways for any user-specified cannabis formulation, we developed CANDI (Cannabis-derived compound Analysis and Network Discovery Interface), a web-based server. This platform offers a user-friendly interface for researchers and drug developers to explore the therapeutic potential of cannabis compounds.

Results Our analysis using DRIFT and CANDI successfully identified numerous molecular targets of cannabis compounds, many of which are involved in pathways relevant to pain, inflammation, cancer, and other diseases. The CANDI server enables researchers to predict the molecular targets and affected pathways for any specific cannabis formulation, providing valuable insights for developing targeted therapies.

Conclusions By combining computational approaches with knowledge of traditional cannabis use, we have developed the CANDI server, a tool that allows us to harness the therapeutic potential of cannabis compounds for the effective treatment of various disorders. By bridging traditional pharmaceutical development with cannabis-based medicine, we propose a novel approach for botanical-based treatment modalities.

Keywords Cannabis, CANDI, Protein-targets and pathways

*Correspondence:
Nikolay V. Dokholyan
dokh@psu.edu
[1]Department of Pharmacology, Penn State College of Medicine, Hershey, PA 17033, USA
[2]Department of Neuroscience & Experimental Therapeutics, Penn State College of Medicine, Hershey, PA 17033, USA
[3]Department of Biochemistry & Molecular Biology, Penn State College of Medicine, Hershey, PA 17033, USA
[4]Department of Chemistry, Penn State University, University Park, PA 16802, USA
[5]Department of Biomedical Engineering, Penn State University, University Park, PA 16802, USA

Introduction

Cannabis sativa L. is among the most ancient cultivated plants, with evidence suggesting its utilization may date back nearly a million years (Ren et al. 2021). Its multifaceted advantages, particularly as a source of fiber, have resulted in its extensive function in both agricultural and industrial applications (Fordjour et al. 2023; H.-L. Li 1974). Currently, cannabis is consumed for medicinal, pharmaceutical, industry, food cosmetics and recreational purposes and is recognized for its various derived metabolites, including terpenoids, flavonoids, sterols, and phytocannabinoids(Simiyu et al. 2022). Phytocannabinoid compounds are being comprehensively reviewed and are stated to interrelate with a complex network of receptors and signaling pathways that play a crucial role in modulating various physiological processes, including pain perception, appetite, mood, and memory (Bonn-Miller et al. 2018; Pacher et al. 2006; Zou and Kumar 2018). The principal psychoactive constituent of cannabis, Δ9-tetrahydrocannabinol (THC), has been the focus of wide-ranging investigation and is the only approved cannabinoid-based prescription for the healing of chemotherapy-induced sickness in patients(Badowski 2017; Ng et al. 2024). However, the therapeutic potential of cannabis extends far beyond THC, with numerous other cannabinoids and terpenes exhibiting promising pharmacological activities (Alves et al. 2020).

On the other hand, cannabis, a formerly banned substance universally, has endured a substantial shift in perception, with various countries like the United States and Canada acknowledging its long-standing traditional medicinal use and legalizing its usage. This paradigm shift has been driven by scientific research and an emerging understanding of the potential therapeutic benefits of cannabis and its active compounds (Dalli et al., n.d). Modern computational and experimental studies on phytocannabinoids and other cannabis-derived compounds have elucidated their medicinal value in the treatment of diverse human disorders, including inflammatory bowel disease (IBD), cancer, Alzheimer's disease, Parkinson's disease, and multiple sclerosis(Abd-Nikfarjam et al. 2023; Carkaci-Salli et al. 2023; Cassano et al. 2020; Fadaka et al. 2022; Farrelly et al. 2021; Helcman and Šmejkal 2022; Hryhorowicz et al. 2021; Varshney et al. 2023). Consequently, the integration of cannabis-based therapeutics into conventional medical practice continues to expand, offering new treatment avenues and improved outcomes for patients with these debilitating conditions(Scherma et al. 2020). Hence, the historical significance and value of cannabis have further emphasized the importance of cannabis-based drug discovery, driving advancements in our understanding of its therapeutic potential and facilitating its integration into modern medical practice.

Naturally occurring chemical compounds from various sources are vital in diverse biological activities and are at the forefront of drug discovery studies. However, identifying the targets for these compounds remains a bottleneck in understanding their mechanisms of action(Li et al. 2021; Newman and Cragg 2016). Experimental techniques, such as affinity chromatography, protein microarrays, and genomic or proteomic studies, are typically employed for target identification, but they are highly time-consuming and relatively expensive(Cheng et al. 2011; Zhang et al. 2022). In contrast to the traditional drug development strategies, it is widely known that compounds often interact with multiple targets, presenting a potential limitation for the experimental approaches (Li et al. 2021). Computational methods offer an alternative by employing various algorithms to identify targets for compounds. For instance, models, such as network-based approaches, data mining, and machine learning, have been used to predict targets for compounds(Agamah et al. 2019; Ezzat et al. 2019; Nogueira and Koch 2019). Moreover, the recent development of deep learning networks has expanded the scope and improved the predictability of target identification from various biological databases that have grown enormously with abundant data on protein-ligand complexes. Deep learning models can effectively analyze large datasets and complex biological networks, making them increasingly valuable in modern drug target identification (Askr et al. 2023; Chen et al. 2024; Zeng et al., n.d.; Zhou et al. 2023). DRIFT is one such model that helps map the targets for the compounds using deep learning approaches by integrating neural network architecture to predict the target-compound binding affinity using the Yuel algorithm in the backend(Chirasani et al. 2022; Wang & Dokholyan, 2022). Hence, advancing computational methodologies for discerning compound-target correlations and extrapolating potential targets for pharmaceuticals and bioactive substances through amalgamating and integrating critical target data from myriad sources provides a valid approach to understanding the context of compound-target interactions. Furthermore, pathways play a pivotal role in elucidating the intricate nature of various diseases, as proteins function within complex networks of interactions(Liu and Chance 2013). Complex diseases often arise from the dysregulation of multiple targets within interconnected pathways or variations in different genes within the same pathways across diverse patient populations(Y.-A. Kim et al. 2011). Hence, elucidating the relationship between targets and disease-associated pathways is crucial for comprehending disease mechanisms and holds promise for developing efficacious treatments.

Despite significant advances, several critical knowledge gaps persist in our understanding of cannabis

pharmacology and its therapeutic potential. There remains a need for further exploration into their mechanisms of action, efficacy, and safety profiles. Furthermore, the variability in cannabis strains, lack of standardized formulations, and potential adverse effects associated with long-term use pose significant challenges to the development of cannabis-based therapeutics. Addressing these gaps is imperative for advancing our understanding of cannabis pharmacology and translating it into safe and effective treatments for a wide range of disorders. In light of these considerations, we have focused the study on cannabis-based drug discovery, which aims to harness the synergistic effects of the plant's diverse phytochemical constituents, a phenomenon known as the "entourage effect." This strategy recognizes that the therapeutic efficacy of cannabis may not be solely attributable to a single compound but rather to the intricate interplay between phytocomplex present in the plant(Ferber et al. 2020; Koltai and Namdar 2020). Accruing data from numerous studies suggests that cannabis extracts or combinations of cannabis-derived compounds may elicit synergistic effects in alleviating pain, reducing inflammation, and mitigating the psychoactive effects(Anand et al. 2021; Bonn-Miller et al. 2018; Chacon et al. 2022; Namdar et al. 2020; Sepulveda et al. 2022). Hence, utilizing computational algorithms, we aim to shed light on the intricate interplay between cannabinoids, terpenes, and other compounds, with the ultimate goal of contributing to the development of novel and efficacious cannabis-based therapeutics. Therefore, leveraging computational algorithms, we seek to elucidate the complex synergistic interactions between cannabinoids, terpenes, and other bioactive constituents within the cannabis plant. This multi-faceted approach aims to identify potential therapeutic targets, optimize drug formulations, and ultimately contribute to the development of innovative and effective cannabis-based therapies for a wide range of medical conditions. Furthermore, we have developed a user-friendly web interface (CANDI, http://candi.dokhlab.org) to facilitate the prediction of targets and relevant pathways for cannabis compounds and formulations, thereby streamlining the drug discovery process and enhancing accessibility for researchers and clinicians alike. CANDI demonstrates versatility beyond traditional therapeutic applications, serving as a valuable tool for clinical development, drug repurposing initiatives, and experimental research. This comprehensive approach enables researchers to explore novel applications, optimize drug formulations, and identify potential repurposing opportunities for existing cannabis compounds. Hence, our study contributes to the advancement of drug discovery efforts aimed at harnessing the therapeutic potential of cannabis compounds for the effective treatment of various disorders.

Materials and methods

Data Curation and Compilation. The initial dataset comprising compounds sourced from the cannabis plant was curated from Pennsylvania state-approved keystone state testing – cannabis laboratory(Raup-Konsavage et al. 2020). These compounds were systematically classified into three main categories: cannabinoids, terpenes, and flavonoids. In total, 97 compounds, with 29 falling under cannabinoids, 45 under terpenes, and 23 under flavonoids, as detailed in Table 1. In addition, we have also updated with 2245 compounds obtained from the cannabis database (Wishart et al. 2024).

Target prediction. Targets for the cannabis compounds were initially determined using the DRIFT algorithm, with the SMILES notation serving as the input format. Subsequently, the obtained targets were refined to include only protein-related factors. The resultant sorted targets and their respective scores were then utilized for subsequent analyses. Further, the results were structured in a matrix (*C*, *T*), where *C* represents compounds and *T* represents targets, each with corresponding predicted scores. Subsequently, for each compound, the user-provided formulation was incorporated as weights (W_j), which were then multiplied with the corresponding scores in the matrix (*MScore*). The resulting values were summed over all targets ($j = 1$ to *n*), yielding a final score for each compound.

$$Target\ scores = \sum\nolimits_{j=1}^{n} MScore * W_j$$

This computation yielded the finalized results, presented in concatenated form, which were subsequently sorted according to normalization criteria. Ultimately, the targets associated with the user-provided formulation for the set of compounds were obtained, along with their normalized scores.

Pathway Mapping. We undertook a systematic curation process to map the pathways associated with the identified targets utilizing data from the REACTOME database(Milacic et al. 2024). Initially, the mapping of UniProt identifiers to pathways was facilitated through an in-house Python script. Subsequently, the UniProt identifiers and their corresponding normalized scores derived from the target analysis were employed as input for pathway prediction. Notably, these scores were utilized as weights during the prediction process. The mapping procedure involved querying the REACTOME database to retrieve pathways associated with the identified UniProt entries. The retrieved pathways were concatenated, forming a comprehensive list. To rank the pathways, we utilized pathway scores. To compute the pathway scores, the weights of the UniProt identifiers mapped to each

Table 1 A selected list of compounds extracted from *Cannabis sativa* through experimental studies. This dataset encompassed 29 cannabinoids, 45 terpenes, and 23 flavonoids

S.no	Cannabis Compounds
Cannabinoids	
1	Cannabichromene
2	Cannabichromenic acid
3	Cannabidiol
4	Cannabidiolic acid
5	Cannabidivarin
6	Cannabidivarinic acid
7	Cannabigerol
8	Cannabigerolic acid
9	Cannabicyclolic acid
10	Cannabinol
11	Cannabinolic acid
12	Delta-8-tetrahydrocannabinol
13	Delta-9-tetrahydrocannabinolic acid A
14	Delta-9-tetrhydrocannabivarin
15	Delta-9-tetrahydrocannabivarinic acid
16	Delta-9-tetrahydrocannabinol
17	Canniprene
18	Cannabiphorol
19	Tetrahydrocannabiphorol
20	11-Nor-9-carboxy-THC
21	Cannabicyclol
22	Cannabigerovarin
23	Cannabigerovarinic acid
24	Cannabichromevarin
25	Cannabichromevarinic acid
26	Cannabielsoin
27	Cannabifuran
28	Cannabicitran
29	Cannabigerolic acid monomethylether
Terpenes	
30	*trans*-ß-Farnesene
31	ß-Caryophyllene
32	Humulene
33	*trans-trans*-α-Farnesene
34	Epi-α-Bisabolol
35	ß-Myrcene
36	(-)-Limonene
37	Delta-3-carene
38	Fenchyl alcohol
39	(S)-α-Terpineol
40	Guaiol
41	α-pinene
42	Linalool
43	Caryophyllene oxide
44	Camphene
45	α-Terpinene
46	Eucalyptol
47	γ-Terpinene
48	Fenchone
49	(±)-*trans*-Nerolidol

Table 1 (continued)

S.no	Cannabis Compounds
50	(R)-Camphor
51	Valencene
52	(+) Cedrol
53	*cis*-Nerolidol
54	(+)-Fenchone
55	α-Phellandrene
56	Hexahydro Thymol
57	α-Cedrene
58	Geranyl Acetate
59	*cis*-ß-ocimene
60	Borneol
61	ß-Pinene
62	Terpinolene
63	Nerol
64	(E)-ß-Ocimene
65	Sabinene
66	Juniper camphor
67	(-)-Isoborneol
68	Borneol
69	ß-Eudesmol
70	γ-Eudesmol
71	*cis*-Sabinene hydrate
72	p-Cymene
73	α-Thujene
74	Betaine
Flavonoids	
75	Kaempferol
76	Luteolin
77	Vitexin
78	Rutin
79	Chrysin
80	Baicalin
81	Orientin
82	quercetin-o-glucoside
83	Isovexitin
84	Cynaroside
85	Apigenin-7-o-glucoside
86	Catechin
87	Quercetin
88	Epicatechin
89	Epigalocatechin
90	Cannflavin B
91	Sitosterol
92	Cannflavin A
93	Naringenin
94	Ferulic acid
95	Caffeic acid
96	Tyramine
97	Salicylic acid

pathway were aggregated and divided by the total number of UniProt identifiers provided as input.

$$Pathway\ score = \sum\ (MW) * \frac{NM}{NT}$$

where *MW*, *NM*, and *NT* represent mapped target weights, the number of mapped targets, and the number of total targets, correspondingly.

This systematic approach ensured the accurate prediction of pathways associated with the identified targets, enhancing our understanding of the biological processes influenced by the investigated compounds.

Compound-Target-Pathway Similarity Analysis. To assess the relationship between the compounds and targets, we have utilized the DRIFT predictions on the cannabis compounds to establish an indirect relationship between them. We leveraged the target information and scores to generate vector representations for each compound. These vectors served as the basis for computing cosine similarity scores, enabling the quantification of compound-target relationships.

$$\cos\theta = \frac{A.B}{\|A\|\,\|B\|}$$

where *A* and *B* represent the vectors corresponding to two compounds. The computed similarity scores were visualized as a heatmap using the Matplotlib library in Python. This visualization method provided users with an intuitive means to comprehend the degree of similarity between compounds and their associated targets. The approach allowed for a clear representation of the complex relationships within the dataset, enabling researchers to rapidly identify patterns and potential areas of interest. We extended our analysis by curating pathways associated with the compounds using the Reactome database, a comprehensive open-source database of human biological processes. This additional step allowed us to map the similarity between compounds and their related pathways, providing a more holistic view of the potential biological impacts of these substances.

Construction of CANDI Web Interface. We developed a user-friendly web interface using Flask, HTML5, CSS, and JavaScript. HTML5 was utilized to structure the content of the web pages, while CSS was employed for styling and layout customization. JavaScript was integrated to enhance user interactivity and functionality, ensuring a seamless browsing experience. Python Flask was utilized for handling data retrieval and processing tasks in the back-end development. For data storage, we used the local built-in storage system. The compatibility of CANDI was tested across popular web browsers such as Chrome and Firefox to ensure consistent performance

and rendering. Overall, CANDI provides users with an intuitive and versatile platform for accessing and analyzing cannabis compound data.

CANDI Web Interface. CANDI offers a suite of interactive modules, each tailored to address distinct stages of cannabis-based drug discovery (Fig. 1A).

Compound Search: This module serves as a comprehensive repository of information on individual cannabis-derived compounds (Fig. 1B). Users may search for specific compounds using various formats, including generic names, SMILES strings, and PubChem IDs. Upon searching, users can access detailed data, including the function to download the results in table format.

Predicted Molecular Targets: A curated list of proteins or receptors likely to interact with the compound is provided based on experimental evidence and computational predictions with corresponding predicted scores. The interface also includes a bar plot to represent the targets and their scores visually.

Similarity Search Results: A list of structurally similar compounds and similarity scores calculated using the FP2 fingerprint and SMILES strings are provided to explore potential analogs with enhanced or altered pharmacological profiles.

Assay Data: When available, results from relevant biological assays are presented, offering insights into the compound's potency with a value alongside the assay method.

Formulation: Recognizing the importance of the entourage effect, this module allows users to input a specific formulation of multiple cannabis compounds. CANDI then leverages its underlying algorithms to efficiently predict the target and map its relevant pathways (Fig. 1C). Potential molecular targets that the specific combination of compounds in the formulation may uniquely or preferentially modulate are identified. The associated biological pathways likely to be impacted by the formulation are mapped highlighting potential therapeutic applications.

Compound-target-pathway Similarity: This module facilitates target-based drug discovery by enabling users to identify cannabis compounds that can act as new agonists for specific targets, based on their relationships to particular targets and pathways (Fig. 1D). Users can input a set of cannabis compounds, and CANDI employs a cosine similarity algorithm to assess the similarity between the input compounds, known targets and pathways. This analysis identifies cannabis compounds predicted to interact with similar and potentially distinct targets and pathways. This method enhances our understanding of individual compounds and illuminates the complex network of interactions within biological systems. By providing a holistic view of the relationships between compounds and their targets, this feature aids in discerning combinations of compounds that may
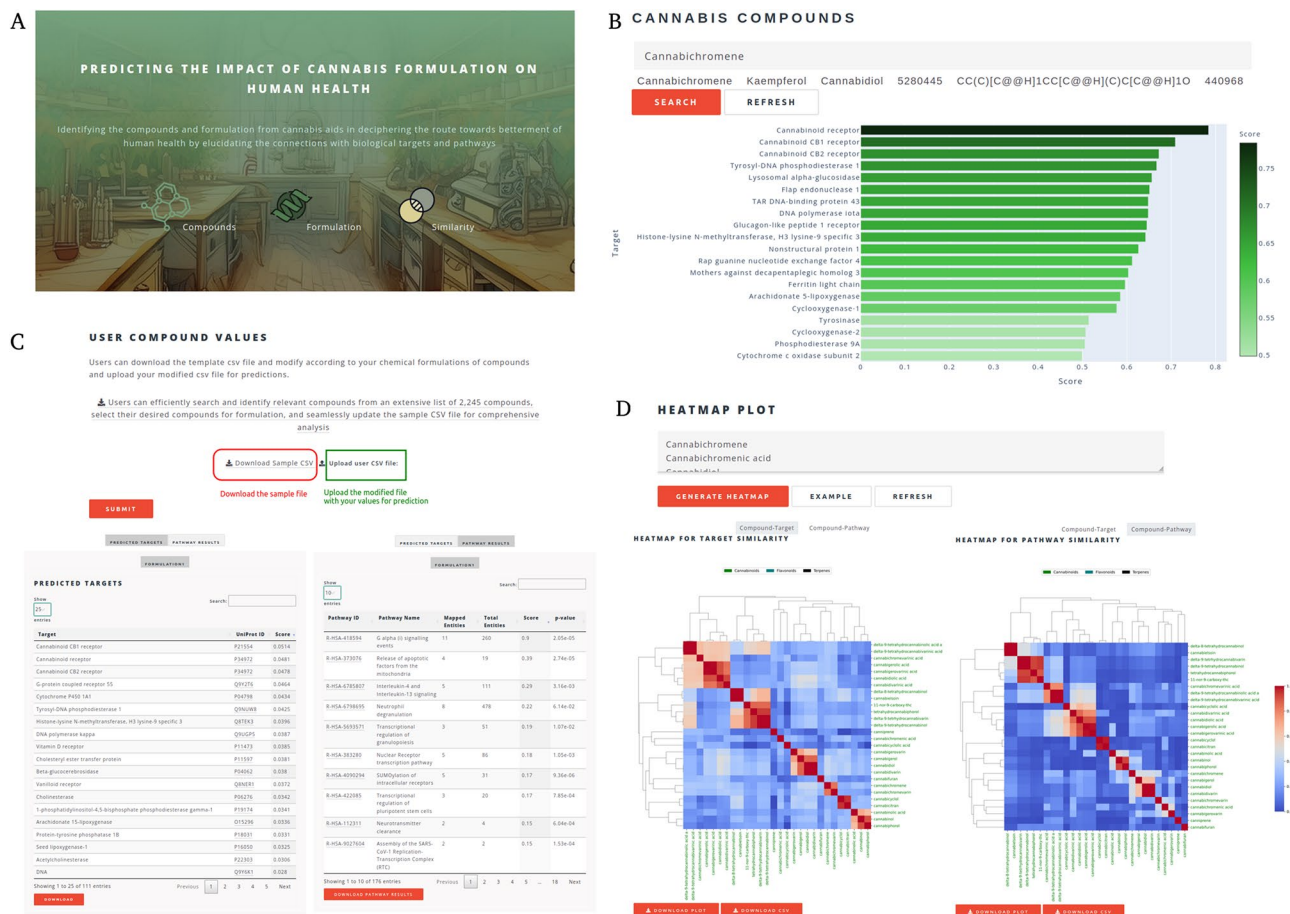
Fig. 1 The overall functioning of the CANDI webserver. (**A**) Front page of the CANDI user interface for the compounds, formulation and similarity functions. (**B**) The compound information data could be obtained for all 2245 cannabis compounds using common names, SMILES, and PubChem IDs. By browsing the compound information, users can intuitively obtain targets for the compound with a predicted score, pharmacophore-similar compounds with similarity values, assay value and type, and graphical representation of the targets vs. score as a bar plot. (**C**) The Formulation page lets users download the file to add user values and upload the file for the target and pathway prediction. (i) Predicted targets are ranked according to the score and linked to their corresponding uniport entries. (ii) Pathways were mapped for the predicted targets from the Reactome database, and the pathway score was shown as an interactive table. (**D**) The compound-target-pathway similarly page allows the user to provide input for the cannabis compounds to identify the relationship between the compound-target and compound-pathway. The output is a heatmap with the download option for the image file and data in CSV format

synergistically modulate multiple targets within a given pathway.

Results and discussion

Cannabis sativa exhibits promising therapeutic potential, substantiated by accumulating scientific evidence. However, the development of standardized cannabis-based therapeutics is hampered by challenges inherent to the plant's phytochemical complexity. We employ a deep learning computational approach to predict molecular targets and associated pathways for cannabis formulations, elucidating the synergistic effect. This research is facilitated by CANDI, a user-friendly web server designed to analyze compound-target interactions and therapeutic mechanisms comprehensively.

Architecture of CANDI. The CANDI web server is an integrated computational platform designed to facilitate the identification of molecular targets and associated pathways for user-specified formulations of cannabis-derived compounds (Fig. 2). The workflow is instigated by user input, wherein the specific combination and concentrations of cannabinoids, terpenes, and other relevant molecules of interest are defined. Leveraging the DRIFT algorithm(Chirasani et al. 2022), a deep learning model trained on structural and chemical properties, the platform predicts potential targets for the compounds. It assigns scores based on the likelihood of interaction. These scores are normalized and re-ranked, considering the user-specified formulation composition and concentrations. The ranked targets are mapped to their corresponding UniProt identifiers(The UniProt Consortium
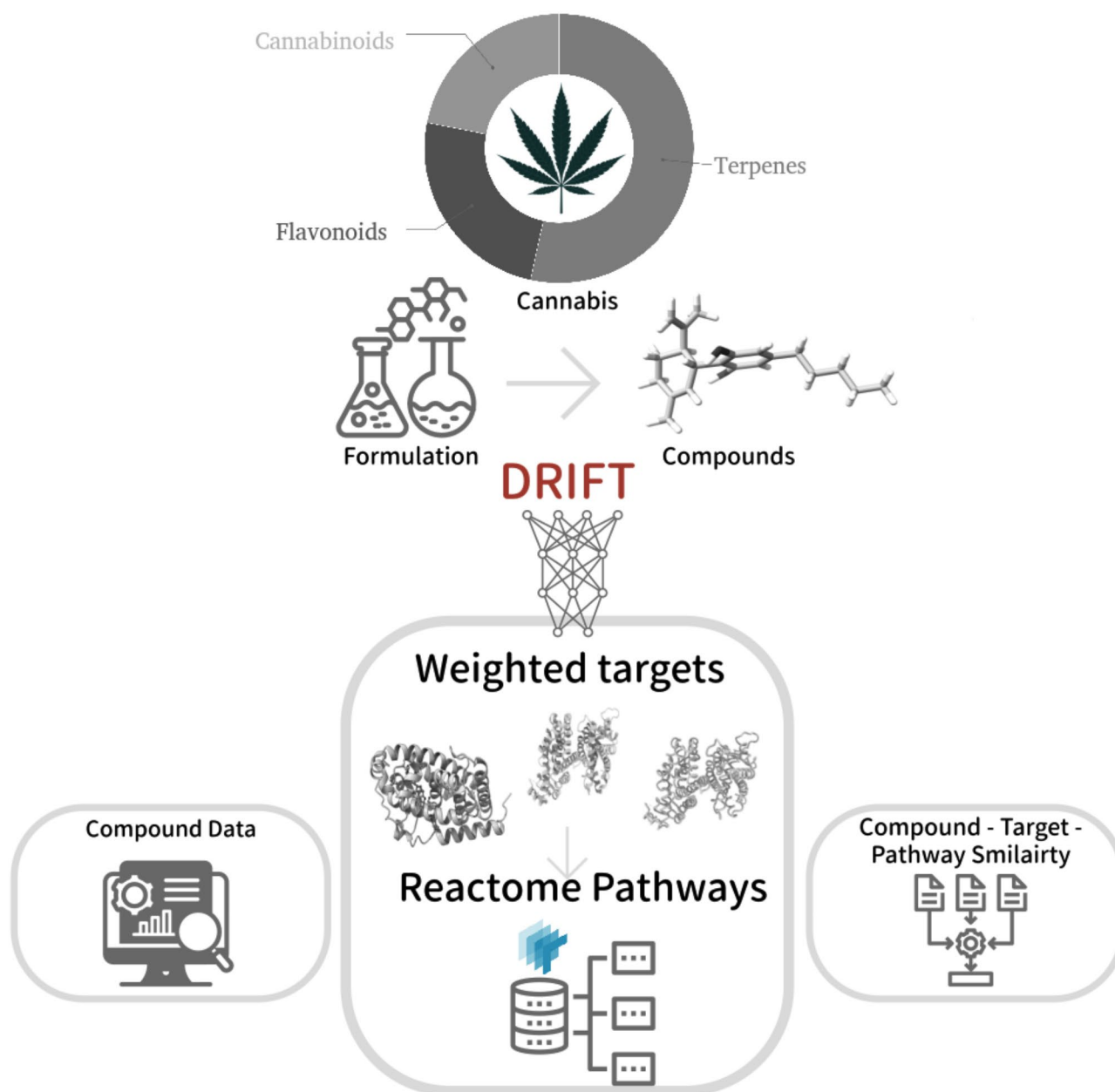
Fig. 2 The workflow adapted in this study to identify the targets and map associated pathways for the formulation of cannabis compounds

2023), enabling the identification of relevant pathways within the Reactome database(Milacic et al. 2024), a comprehensive resource of biological pathways and processes. The final output provided by CANDI is a ranked list of pathways, weighted by the number and scores of associated targets, offering insights into the potential mechanisms underlying the therapeutic effects of the specified cannabis formulation. This integrated computational approach enables researchers to systematically explore the intricate interplay between cannabis compounds and their molecular targets, accelerating the development

of targeted therapies and elucidating the mechanistic underpinnings of cannabis-based therapeutics.

Case Study on Cannabis Oil Formulation. To validate CANDI's functionality, performed studies utilizing a commercial cannabis oil formulation comprising various composition of cannabinoids and terpenes (Table 2). The formulation's composition, obtained from experimental data, was input into CANDI by modifying the platform's sample CSV file. Upon analysis, CANDI generated results that were presented in two sections: predicted molecular targets and associated pathways.

Table 2 Dataset of the cannabis compounds with the formulation used for the case studies. All values are noted in percentage

Compounds	Cannabis oil (case study 1)	Cannabi-noids (case study 2)	Ter-penes (case study 3)
Cannabichromene	1.55	3.89	0
Cannabidiol	36.56	34.87	0
Cannabidivarin	0	0	0
Cannabidivarinic acid	0	61.09	0
Cannabigerol	35.63	21.31	0
Cannabigerolic acid	0	8.45	0
Cannabicyclolic acid	0	0.13	0
Cannabinol	0	0.15	0
Delta-9-tetrhydrocannabivarin	0	0.19	0
Delta-9-tetrahydrocannabivarinic acid	0	0.34	0
Delta-8-tetrahydrocannabinol	0	0	0
Delta-9-tetrahydrocannabinol	0	2.64	0
Caryophyllene oxide	0	0	0.47
(+) Cedrol	0	0	0.09
α-Bisabolol	0.04	0	1.58
α-Cedrene	0	0	0.07
trans-trans-α-Farnesene	0	0	1.49
Humulene	0.023	0	1.32
α-Phellandrene	0	0	0.04
α-pinene	0	0	0.01
α-Terpineol	0.001	0	0.09
trans-ß-Farnesene	0	0	5.46
ß-Myrcene	0.001	0	0.25
ß-Pinene	0	0	0.02
Camphene	0	0	0.01
cis-Nerolidol	0	0	0.02
cis-ß-ocimene	0	0	0.03
Fenchyl alcohol	0.002	0	0.08
Eucalyptol	0	0	0.01
Fenchone	0	0	0.01
γ-Terpinene	0	0	0.01
Geranyl Acetate	0	0	0.01
Guaiol	0.04	0	0.56
Linalool	0	0	0.07
(-)-Limonene	0.003	0	0.12
Terpinolene	0	0	0.25
ß-Caryophyllene	0.091	0	5.37
(±)-*trans*-Nerolidol	0	0	0.11
Valencene	0	0	0.72

The predicted targets section displayed a ranked list, ordered by their predicted interaction scores, with the highest-scoring targets listed first. For further reference, each target was linked to its corresponding UniProt entry. The analysis revealed that the formulation was predicted to interact with cannabinoid receptors CB1 and CB2, followed by G protein-coupled receptor 55 (GPR55), cytochrome P450 enzymes, and other receptors (Fig. 3A). The

associated pathways section provided a detailed overview of the Reactome pathways linked to the predicted targets. These pathways were ranked based on their predictive score. Among the identified pathways were nuclear receptor transcription, G alpha(i) signaling events, the release of apoptotic factors from mitochondria, and SUMOylation of intracellular receptors all implicated in various physiological processes (Fig. 3B). Hence, the analysis revealed that this formulation could modulate multiple targets and pathways associated with pain management, inflammation, and neurological disorders. The formulation was predicted to interact with the endocannabinoid system, including the CB1 and CB2 receptors. These interactions could contribute to the formulation's potential analgesic, anti-inflammatory, and neuroprotective effects(Donvito et al. 2018; Gonzalo-Consuegra et al. 2024). Furthermore, the analysis identified several relevant pathways related to pain perception and inflammation (Che 2021; Zhao et al. 2020). Hence, CANDI-generated hypothesis is that this formulation shows promise as a potential therapeutic agent for these conditions. Further research, including preclinical and clinical studies, is warranted to validate these findings and explore the full therapeutic potential of this formulation.

Case Study on Cannabinoids. In the second case study, we analyzed a cannabis oil formulation containing only cannabinoids (Table 2). From the analyses we could decipher that the formulation was predicted to interact with cannabinoid receptors CB1 and CB2, DNA polymerase kappa, and G protein-coupled receptor 55 (GPR55), vitamin receptor, and other receptors (Supp. Figure 1A). The associated pathways were G alpha(i) signaling events, Interlukin-4 and Interlukin-13 signaling, and neutrophil degranulation, entirely associated in several biological activities (Supp. Figure 2B). Notably, this formulation was predicted to interact with the well-characterized CB1 and CB2 receptors, which are primary targets in cannabinoid research. These receptors are involved in various physiological processes, including pain modulation and inflammatory responses(Raup-Konsavage et al. 2023; Turcotte et al. 2016). In accord with our findings, the mapped pathways, particularly G alpha(i) signaling events and Interleukin-4 and Interleukin-13 signaling, have been implicated in pain perception and inflammatory processes(Ibsen et al. 2017; Oláh et al. 2017).

Case Study on Terpenes. In the third case study, a formulation containing only terpenes was analyzed (Table 2). Predicted targets and associated pathways were charted. The formulation was predicted to interact with solute carrier organic anion transporter family members 1B1 and 1B3, bile acid receptor FXR, arachidonate 15-lipoxygenase receptors and also cannabinoid CB2 receptor (Supp. Figure 2A). Associated pathways included nuclear receptor transcription, aspirin ADME,
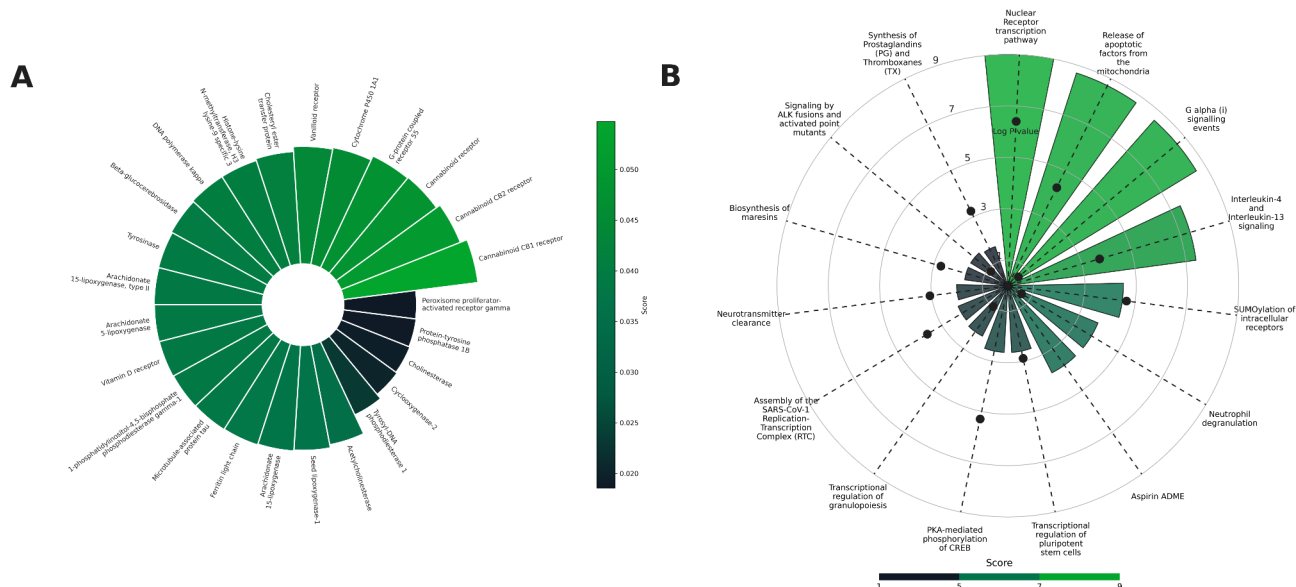
Fig. 3 A case study on cannabis oil from the experimental studies. (**A**) The predicted targets for the cannabis oil formulation were ranked according to the predictive scores suggesting that CB1 and CB2 are top targets for the given formulation (**B**) Mapped pathways for the targets predicted were plotted elucidating the involvement of the targets in nuclear receptor transcription factor and G alpha (i) signalling events pathways

SUMOylation of intracellular receptors, Interleukin-4 and Interleukin-13 signaling, and G alpha (i) signaling events all implicated in various biological processes (Supp. Figure 2B). These findings suggest that this terpene formulation may contribute to modulating diverse physiological functions through its interactions with these targets and pathways. In accord with our findings, the identified targets and pathways are commonly involved in inflammatory bowel disorders, various inflammatory diseases, and metabolic disorders(Del Prado-Audelo et al. 2021; Kim et al. 2020; LaVigne et al. 2021).

Conclusion

The development of cannabis-based therapeutics holds significant potential for treating diverse medical conditions. However, this potential is constrained by the intricacy of the cannabis plant and the current lack of standardized, targeted therapies. The study exemplifies a noteworthy improvement in overcoming these challenges by leveraging computational approaches, specifically deep learning algorithms. CANDI facilitates the identification of molecular targets and associated pathways for specific combinations of cannabis-derived compounds, addressing research gaps related to the entourage effect. Additionally, the user-friendly interface allows researchers to investigate the complex interplay between these compounds and their potential therapeutic targets. By integrating information on compound-target interactions and relevant biological pathways, CANDI facilitates a comprehensive analysis of the molecular mechanisms underlying the therapeutic effects of cannabis formulations and offers a plausible hypothesis on health

outcomes of such compounds and formulations. Hence, the study contributes to the advancement of drug discovery efforts aimed at harnessing the therapeutic potential of cannabis compounds for the effective treatment of various disorders.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42238-025-00268-w.

Supplementary Material 1

Author contributions
SE, JW and NVD contributed to the design of the study, conceptualization, methodology, and data analysis. SE and JW led the interpretation of the final data and the writing of original draft of the manuscript. All authors approved the final version of the study and collectively edited the manuscript.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethical approval
Not applicable.

Competing interests
The authors declare no competing interests.

References

Abd-Nikfarjam B, Dolati-Somarin A, Baradaran Rahimi V, et al. Cannabinoids in neuroinflammatory disorders: focusing on multiple sclerosis, parkinsons, and alzheimers diseases. BioFactors. 2023;49(3):560–83. https://doi.org/10.1002/biof.1936.

Agamah FE, Mazandu GK, Hassan R, et al. Computational/in silico methods in drug target and lead prediction. Brief Bioinform. 2019;21(5):1663–75. https://doi.org/10.1093/bib/bbz103.

Alves P, Amaral C, Teixeira N, et al. Cannabis sativa: much more beyond Δ9-tetrahydrocannabinol. Pharmacol Res. 2020;157:104822. https://doi.org/10.1016/j.phrs.2020.104822.

Anand U, Pacchetti B, Anand P, et al. Cannabis-based Medicines and Pain: a review of potential synergistic and Entourage effects. Pain Manage. 2021;11(4):395–403. https://doi.org/10.2217/pmt-2020-0110.

Askr H, Elgeldawi E, Aboul Ella H, et al. Deep learning in drug discovery: an integrative review and future challenges. Artif Intell Rev. 2023;56(7):5975–6037. https://doi.org/10.1007/s10462-022-10306-1.

Badowski ME. A review of oral cannabinoids and medical marijuana for the treatment of chemotherapy-induced nausea and vomiting: a focus on pharmacokinetic variability and pharmacodynamics. Cancer Chemother Pharmacol. 2017;80(3):441–9. https://doi.org/10.1007/s00280-017-3387-5.

Bonn-Miller MO, ElSohly M, Loflin MJE, et al. Cannabis and Cannabinoid Drug Development: evaluating Botanical Versus single molecule approaches. Int Rev Psychiatry. 2018;30(3):277–84. https://doi.org/10.1080/09540261.2018.1474730.

Carkaci-Salli N, Raup-Konsavage WM, Karelia D, et al. Cannabinoids as potential Cancer therapeutics: the Concentration Conundrum. Cannabis Cannabinoid Res. 2023. https://doi.org/10.1089/can.2022.0344.

Cassano T, Villani R, Pace L, et al. From Cannabis sativa to Cannabidiol: promising therapeutic candidate for the treatment of neurodegenerative diseases. Front Pharmacol. 2020;11. https://doi.org/10.3389/fphar.2020.00124.

Chacon FT, Raup-Konsavage WM, Vrana KE, et al. Secondary terpenes in Cannabis sativa L. Synthesis Synergy Biomedicines. 2022;10(12):3142. https://doi.org/10.3390/biomedicines10123142.

Che T. Advances in the treatment of Chronic Pain by targeting GPCRs. Biochemistry. 2021;60(18):1401–12. https://doi.org/10.1021/acs.biochem.0c00644.

Chen H, King FJ, Zhou B, et al. Drug target prediction through deep learning functional representation of gene signatures. Nat Commun. 2024;15(1):1853. https://doi.org/10.1038/s41467-024-46089-y.

Cheng T, Li Q, Wang Y, et al. Identifying compound-target associations by combining Bioactivity Profile Similarity Search and Public Databases Mining. J Chem Inf Model. 2011;51(9):2440–8. https://doi.org/10.1021/ci200192v.

Chirasani VR, Wang J, Sha C, et al. Whole proteome mapping of compound-protein interactions. Curr Res Chem Biology. 2022;2:100035. https://doi.org/10.1016/j.crchbi.2022.100035.

Dalli M, Azizi S, Azghar A, et al. Cannabis sativa L.: a comprehensive review on legislation, decriminalization, phytochemistry, antimicrobial activity, and safety. J Food Drug Anal. n.d.;31(3):408–35. https://doi.org/10.38212/2224-6614.3471.

Del Prado-Audelo ML, Cortés H, Caballero-Florán IH, et al. Therapeutic applications of terpenes on inflammatory diseases. Front Pharmacol. 2021;12:704197. https://doi.org/10.3389/fphar.2021.704197.

Donvito G, Nass SR, Wilkerson JL et al. The Endogenous Cannabinoid System: A Budding Source of Targets for Treating Inflammatory and Neuropathic Pain. Neuropsychopharmacol. 2018;43(1):52–79. https://doi.org/10.1038/npp.2017.204

Ezzat A, Wu M, Li X-L, et al. Computational prediction of drug-target interactions using chemogenomic approaches: an empirical survey. Brief Bioinform. 2019;20(4):1337–57. https://doi.org/10.1093/bib/bby002.

Fadaka AO, Taiwo OA, Dosumu OA, et al. Computational prediction of potential drug-like compounds from Cannabis sativa leaf extracts targeted towards Alzheimer therapy. J Mol Liq. 2022;360:119393. https://doi.org/10.1016/j.molliq.2022.119393.

Farrelly AM, Vlachou S, Grintzalis K. Efficacy of Phytocannabinoids in Epilepsy Treatment: Novel approaches and recent advances. Int J Environ Res Public Health. 2021;18(8):3993. https://doi.org/10.3390/ijerph18083993.

Ferber SG, Namdar D, Hen-Shoval D, et al. The entourage effect: terpenes coupled with cannabinoids for the treatment of Mood disorders and anxiety disorders. Curr Neuropharmacol. 2020;18(2):87–96. https://doi.org/10.2174/1570159X17666190903103923.

Fordjour E, Manful CF, Sey AA, et al. Cannabis: a multifaceted plant with endless potentials. Front Pharmacol. 2023;14. https://doi.org/10.3389/fphar.2023.1200269.

Gonzalo-Consuegra C, Santos-García I, García-Toscano L, et al. Involvement of CB1 and CB2 receptors in neuroprotective effects of cannabinoids in experimental TDP-43 related frontotemporal dementia using male mice. Biomed Pharmacother. 2024;174:116473. https://doi.org/10.1016/j.biopha.2024.116473.

Helcman M, Šmejkal K. Biological activity of Cannabis compounds: a modern approach to the therapy of multiple diseases. Phytochem Rev. 2022;21(2):429–70. https://doi.org/10.1007/s11101-021-09777-x.

Hryhorowicz S, Kaczmarek-Ryś M, Zielińska A, et al. Endocannabinoid System as a Promising Therapeutic Target in Inflammatory Bowel Disease – a systematic review. Front Immunol. 2021;12. https://doi.org/10.3389/fimmu.2021.790803.

Ibsen MS, Connor M, Glass M. Cannabinoid CB1 and CB2 receptor signaling and Bias. Cannabis Cannabinoid Res. 2017;2(1):48–60. https://doi.org/10.1089/can.2016.0037.

Kim Y-A, Wuchty S, Przytycka TM. Identifying causal genes and dysregulated pathways in complex diseases. PLoS Comput Biol. 2011;7(3):e1001095. https://doi.org/10.1371/journal.pcbi.1001095.

Kim T, Song B, Cho KS, et al. Therapeutic potential of volatile terpenes and terpenoids from forests for inflammatory diseases. Int J Mol Sci. 2020;21(6):2187. https://doi.org/10.3390/ijms21062187.

Koltai H, Namdar D. Cannabis Phytomolecule entourage: from domestication to Medical Use. Trends Plant Sci. 2020;25(10):976–84. https://doi.org/10.1016/j.tplants.2020.04.007.

LaVigne JE, Hecksel R, Keresztes A, et al. Cannabis sativa terpenes are cannabimimetic and selectively enhance cannabinoid activity. Sci Rep. 2021;11(1):8232. https://doi.org/10.1038/s41598-021-87740-8.

Li H-L. An Archaeological and historical account of Cannabis in China. Econ Bot. 1974;28(4):437–48.

Li G, Peng X, Guo Y, et al. Currently available strategies for Target Identification of Bioactive Natural products. Front Chem. 2021;9:761609. https://doi.org/10.3389/fchem.2021.761609.

Liu Y, Chance MR. Pathway analyses and understanding disease associations. Curr Genet Med Rep. 2013;1(4). https://doi.org/10.1007/s40142-013-0025-3.

Milacic M, Beavers D, Conley P, et al. The Reactome Pathway Knowledgebase 2024. Nucleic Acids Res. 2024;52(D1):D672–8. https://doi.org/10.1093/nar/gkad1025.

Namdar D, Anis O, Poulin P, et al. Chronological Review and rational and future prospects of Cannabis-based Drug Development. Molecules. 2020;25(20):4821. https://doi.org/10.3390/molecules25204821.

Newman DJ, Cragg GM. Natural products as sources of new drugs from 1981 to 2014. J Nat Prod. 2016;79(3):629–61. https://doi.org/10.1021/acs.jnatprod.5b01055.

Ng T, Gupta V, Keshock MC. Tetrahydrocannabinol (THC). In: StatPearls StatPearls Publishing: Treasure Island (FL);; 2024.

Nogueira MS, Koch O. The Development of Target-Specific Machine Learning Models as Scoring functions for Docking-based target prediction. J Chem Inf Model. 2019;59(3):1238–52. https://doi.org/10.1021/acs.jcim.8b00773.

Oláh A, Szekanecz Z, Bíró T. Targeting Cannabinoid Signaling in the Immune System: high-ly exciting questions, possibilities, and challenges. Front Immunol. 2017;8. https://doi.org/10.3389/fimmu.2017.01487.

Pacher P, Bátkai S, Kunos G. The endocannabinoid system as an emerging target of pharmacotherapy. Pharmacol Rev. 2006;58(3):389–462. https://doi.org/10.1124/pr.58.3.2.

Raup-Konsavage WM, Carkaci-Salli N, Greenland K, et al. Cannabidiol (CBD) Oil does not display an entourage effect in reducing Cancer Cell viability in vitro. Med Cannabis Cannabinoids. 2020;3(2):95–102. https://doi.org/10.1159/000510256.

Raup-Konsavage WM, Sepulveda DE, Wang J, et al. Antinociceptive effects of Cannabichromene (CBC) in mice: insights from Von Frey, Tail-Flick, Formalin, and

acetone tests. Biomedicines. 2023;12(1):83. https://doi.org/10.3390/biomedicines12010083.

Ren G, Zhang X, Li Y, et al. Large-scale whole-genome resequencing unravels the domestication history of Cannabis sativa. Sci Adv. 2021;7(29):eabg2286. https://doi.org/10.1126/sciadv.abg2286.

Scherma M, Muntoni AL, Riedel G, et al. Cannabinoids and their therapeutic applications in mental disorders. Dialogues Clin Neurosci. 2020;22(3):271–9. https://doi.org/10.31887/DCNS.2020.22.3/pfadda.

Sepulveda DE, Vrana KE, Graziane NM, et al. Combinations of Cannabidiol and Δ9-Tetrahydrocannabinol in reducing Chemotherapeutic Induced Neuropathic Pain. Biomedicines. 2022;10(10):2548. https://doi.org/10.3390/biomedicines10102548.

Simiyu DC, Jang JH, Lee OR. Understanding Cannabis sativa L.: current status of propagation, Use, legalization, and haploid-inducer-mediated genetic Engineering. Plants (Basel). 2022;11(9):1236. https://doi.org/10.3390/plants11091236.

The UniProt Consortium. UniProt: the Universal protein knowledgebase in 2023. Nucleic Acids Res. 2023;51(D1):D523–31. https://doi.org/10.1093/nar/gkac1052.

Turcotte C, Blanchet M-R, Laviolette M, et al. The CB2 receptor and its role as a regulator of inflammation. Cell Mol Life Sci. 2016;73(23):4449–70. https://doi.org/10.1007/s00018-016-2300-4.

Varshney K, Patel A, Ansari S, et al. Cannabinoids in treating Parkinson's disease symptoms: a systematic review of Clinical studies. Cannabis Cannabinoid Res. 2023;8(5):716–30. https://doi.org/10.1089/can.2023.0023.

Wang J, Dokholyan NV, Yuel. Improving the generalizability of structure-free compound-protein Interaction Prediction. J Chem Inf Model. 2022;62(3):463–71. https://doi.org/10.1021/acs.jcim.1c01531.

Wishart DS, Hiebert-Giesbrecht M, Inchehborouni G, et al. Chemical composition of Commercial Cannabis. J Agric Food Chem. 2024;72(25):14099–113. https://doi.org/10.1021/acs.jafc.3c06616.

Zeng X, Zhu S, Lu W, et al. Target identification among known drugs by deep learning from heterogeneous networks. Chem Sci. n.d.;11(7):1775–97. https://doi.org/10.1039/c9sc04336e.

Zhang X, Wu F, Yang N, et al. In silico methods for identification of potential therapeutic targets. Interdiscip Sci Comput Life Sci. 2022;14(2):285–310. https://doi.org/10.1007/s12539-021-00491-y.

Zhao X, Xia B, Cheng J, et al. PKCε SUMOylation is required for mediating the Nociceptive Signaling of Inflammatory Pain. Cell Rep. 2020;33(1):108191. https://doi.org/10.1016/j.celrep.2020.108191.

Zhou L, Wang Y, Peng L, et al. Identifying potential drug-target interactions based on ensemble deep learning. Front Aging Neurosci. 2023;15:1176400. https://doi.org/10.3389/fnagi.2023.1176400.

Zou S, Kumar U. Cannabinoid receptors and the Endocannabinoid System: signaling and function in the Central Nervous System. Int J Mol Sci. 2018;19(3):833. https://doi.org/10.3390/ijms19030833.

Publisher's note