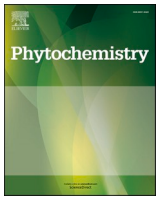
Phytochemistry 205 (2023) 113508

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem



Multivariate analysis of chemical and genetic diversity of wild *Humulus lupulus* L. (hop) collected in situ in northern France

Anne-Sophie Paguet [a], Ali Siah [a], Gabriel Lefèvre [a], Sophie Moureu [a], Thierry Cadalen [a], Jennifer Samaillie [a], Franck Michels [a], Barbara Deracinois [a], Christophe Flahaut [a], Harmony Alves Dos Santos [a], Audrey Etienne-Debaecker [a], Caroline Rambaud [a], Sylvie Chollet [a], Roland Molinié [a], Jean-Xavier Fontaine [a], Christophe Waterlot [b], Marie-Laure Fauconnier [a], Sevser Sahpaz [a], Céline Rivière [a,*]

[a] *Joint Research Unit 1158 BioEcoAgro, Univ. Lille, JUNIA, UPJV, Univ. Liège, INRAE, Univ. Artois, Univ. Littoral Côte D'Opale, ICV—Institut Charles Viollette, F-59650 Villeneuve D'Ascq, France*

[b] *Univ. Lille, Institut Mines-Télécom, Univ. Artois, JUNIA, ULR 4515 – LGCgE, Laboratoire de Génie Civil et Géo-Environnement, F-59000 Lille, France*

ARTICLE INFO

Keywords:
Humulus lupulus L.
Cannabaceae
Wild hops
SSR fingerprinting
Soil analysis
HS-SPME GC-MS
UHPLC-UV-MS
Untargeted metabolomics
Multiple factor analysis

ABSTRACT

The hop plant (*Humulus lupulus* L.) has been exploited for a long time for both its brewing and medicinal uses, due in particular to its specific chemical composition. These last years, hop cultivation that was in decline has been experiencing a renewal for several reasons, such as a craze for strongly hopped aromatic beers. In this context, the present work aims at investigating the genetic and chemical diversity of fifty wild hops collected from different locations in Northern France. These wild hops were compared to ten commercial varieties and three heirloom varieties cultivated in the same sampled geographical area. Genetic analysis relying on genome fingerprinting using 11 microsatellite markers showed a high level of diversity. A total of 56 alleles were determined with an average of 10.9 alleles per locus and assessed a significant population structure (mean pairwise $F_{ST} = 0.29$). Phytochemical characterization of hops was based on volatile compound analysis by HS-SPME GC-MS, quantification of the main prenylated phenolic compounds by UHPLC-UV as well as untargeted metabolomics by UHPLC-HRMS and revealed a high level of chemical diversity among the assessed wild accessions. In particular, analysis of volatile compounds revealed the presence of some minor but original compounds, such as aromadendrene, allo-aromadendrene, isoledene, β-guaiene, α-ylangene and β-pinene in some wild accessions; while analysis of phenolic compounds showed high content of β-acids in these wild accessions, up to 2.37% of colupulone. Genetic diversity of wild hops previously observed was hence supported by their chemical diversity. Sample soil analysis was also performed to get a pedological classification of these different collection sites. Results of the multivariate statistical analysis suggest that wild hops constitute a huge pool of chemical and genetic diversity of this species.

1. Introduction

Humulus lupulus L (hop) is a dioecious (2n = 20) climbing perennial plant belonging to the Cannabaceae family, widely distributed throughout the Northern Hemisphere including Europe, Asia and North America (Small, 1978; Neve, 1991). Traditionally, only female plants of *H. lupulus* are cultivated because of the use of their inflorescences, commonly referred to as hop cones or hops. This part of the plant imparts to the beer several organoleptic and physical qualities such as bitterness, aroma, flavor, foam stability and microbial protection linked to its unique combination of secondary metabolites (Schönberger and Kostelecky, 2011). Actually, from a biogenetic point of view, hop synthetizes original phenolic compounds, including prenylated chalcones as xanthohumol and desmethylxanthohumol and acylphloroglucinol derivatives (α-acids and β-acids) (Bocquet et al., 2018a) (Fig. 1). Hop essential oil is also rich in terpenes and sesquiterpenes (α-humulene,

* Corresponding author. Céline Rivière Joint Research Unit 1158 BioEcoAgro, University of Lille, Faculty of Pharmacy, 3 rue du Professeur Laguesse, BP 83, 59000 Lille, France. celine.riviere@univ-lille.fr

E-mail address: celine.riviere@univ-lille.fr (C. Rivière).

https://doi.org/10.1016/j.phytochem.2022.113508
Received 29 July 2022; Received in revised form 23 October 2022; Accepted 4 November 2022
Available online 10 November 2022

β-myrcene, *trans*-caryophyllene and farnesene) (Fig. 2). These volatile compounds give beer its aromatic qualities, while *α*-acids are isomerized during the boiling of beer wort in iso-*α*-acids, which impart to the beer its bitterness (Kishimoto et al., 2021). As far as *β*-acids are concerned, they are known for their antimicrobial potential (Bocquet et al., 2018b, 2018c, 2019), and seem to have some effect in invertebrate pest control (Naraine and Small, 2017; Korpelainen and Pietiläinen, 2021). This chemical composition is strongly impacted by genetics but also by the terroir (Morcol et al., 2020). Northern France is a lowland region, bordered on the west by the English Channel and in fact the western part is rather exposed to wind. Northern France is a hop-producing region which shares a historical past with Belgium, its border country. Its production of 1220 ha in 1907 has declined to 35 ha but has slowly begun to increase again to around 40 ha (Ducloux et al., 1910; Bart-Haas Group, 2021). Nowadays, Northern France remains the second area for hop production in France, after Alsace in Eastern France. Ten varieties are cultivated by 9 hop growers. In the current context of development of local microbreweries (Statista, 2021) and the search for local ingredients from sustainable agriculture, these surface areas are increasing in order to meet this new demand (Paguet et al., 2022). Brewers and consumers are indeed looking for locally sourced hops that present original and exotic aromatic notes such as citrus or pineapple notes. The current cultivated varieties therefore need to be more adapted to the terroir and more aromatic than bittering in order to be more in line with new market expectations. The varietal development of hops adapted to the pedo-climatic conditions of Northern France, with good aromatic and gustative qualities in brewing could respond to these challenges.

In this context, this project aimed to investigate the chemical and genetic diversity of wild hops collected in the North of France. Indeed, wild diversity may constitute a huge pool of genetic diversity for breeding or other applications (Murakami et al., 2006; Paguet et al., 2022). As mentioned in our previous review (Paguet et al., 2022), several studies have been dedicated to exploring the diversity of hop collections, based on the genetic structure of populations using different types of genetic markers, such as AFLP markers (Solberg et al., 2014), SSR (Mongelli et al., 2016; Rodolfi et al., 2018; Mafakheri et al., 2020; Dabbous-Wach et al., 2021), SNPs (Yamauchi et al., 2014; Zhang et al., 2017; Jiang et al., 2018; Van Holle et al., 2019; Driskill et al., 2022; Machado et al., 2022) or DaRT (Čerenak et al., 2019). Microsatellite DNA fingerprints have a high level of accuracy, co-dominant, multi--allelic and independent from lab and thus became a favored technique for genotyping (Nybom, 2004). Jakse et al. (2002, 2008) and Stajner

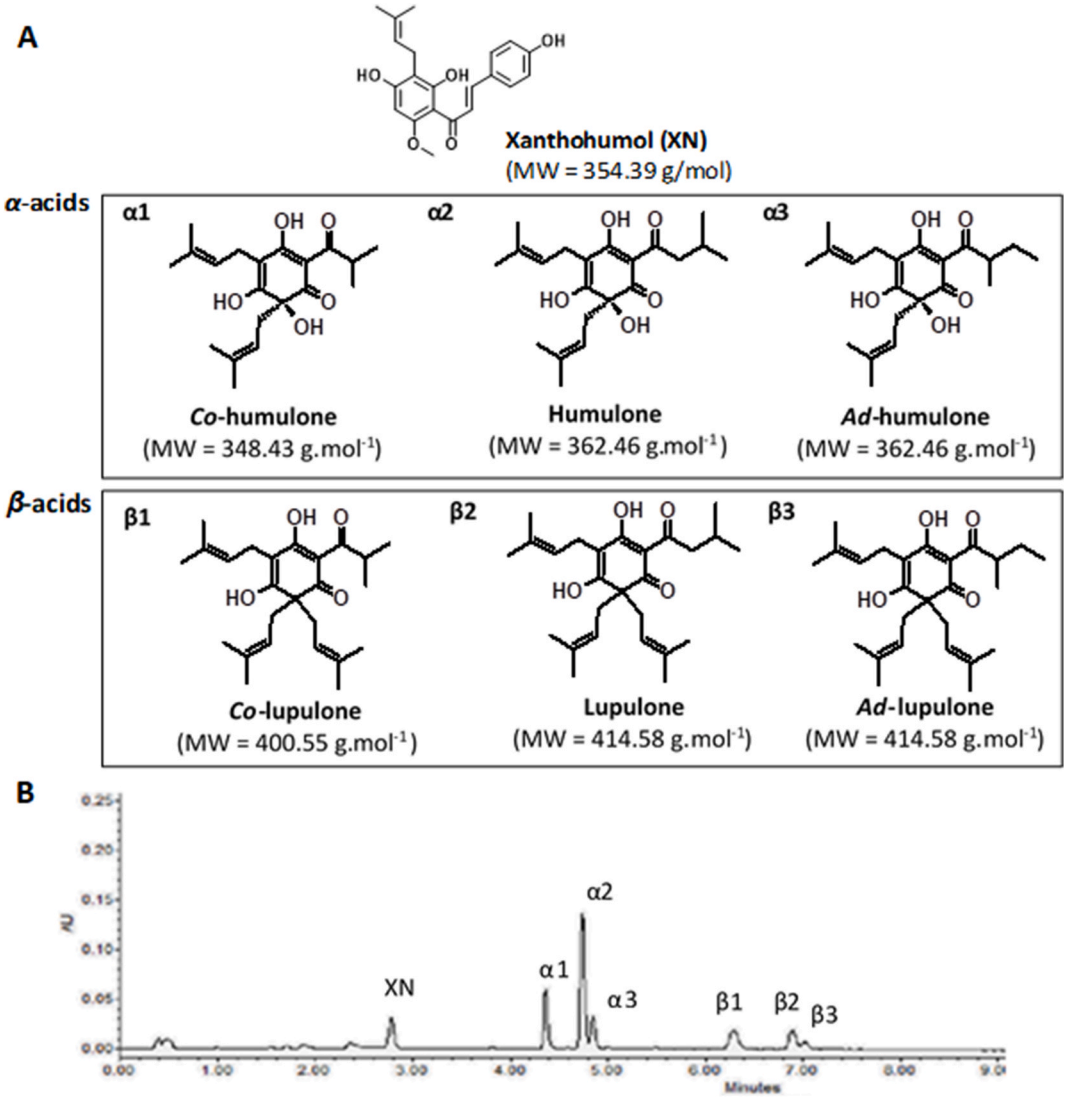


Fig. 1. Main hop prenylated phenolic compounds A. Chemical structure of major chalcones and acylphloroglucinols produced by hops and their molecular weight. **B.** Chromatogram of a crude hydro-ethanolic extract of hops (cultivar Nugget) at 330 nm. XN: xanthohumol, α1: *co*-humulone; α2: humulone; α3: *ad*-humulone; β1: *co*-lupulone; β2: lupulone; β3: *ad*-lupulone.

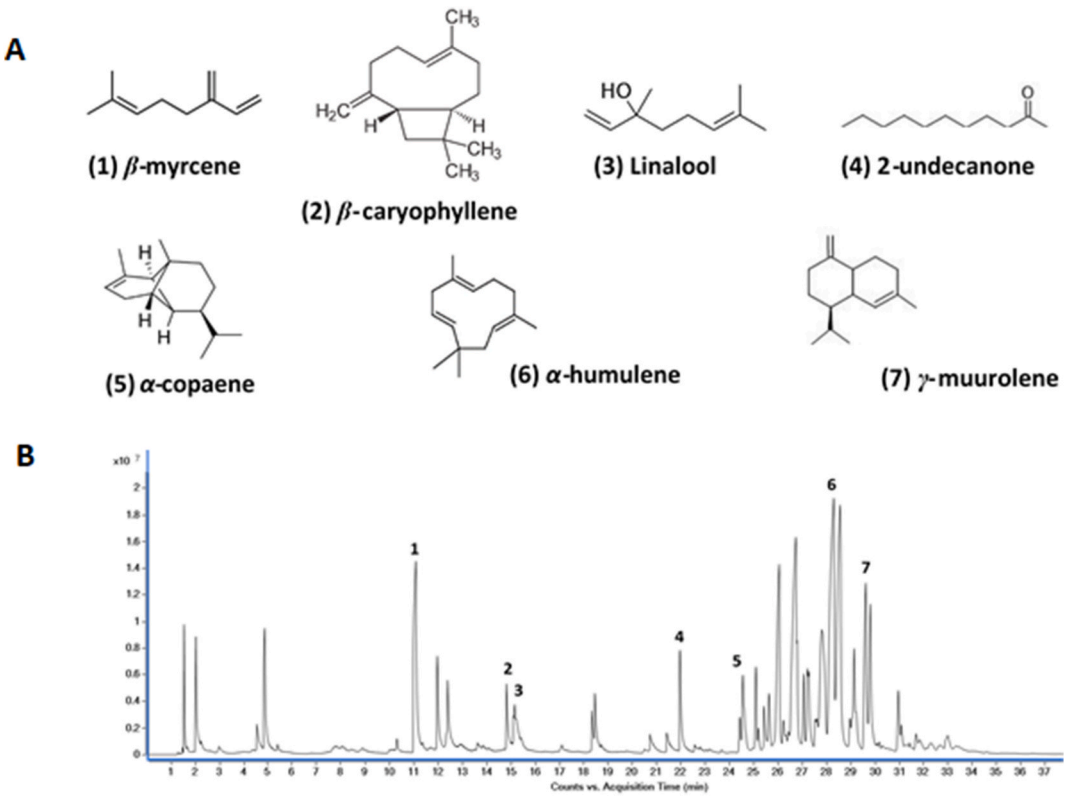


Fig. 2. Analysis of volatile compounds in hop cones by GC-MS. **A.** Chemical structure of main volatile compounds found in hop cones. **B.** GC-MS total ion chromatogram of a hop cone sample (cv. Nugget). Compounds identified correspond to the following compounds: (1) β-myrcene; (2) β-caryophyllene; (3) linalool; (4) 2-undecanone; (5) copaene; (6) α-humulene; (7) γ-muurolene.

et al. (2005) isolated and characterized several microsatellites in the hop genome and submitted their sequences to the NCBI GenBank. The characterization of hops can also be based on morphological analysis (Srečec et al., 2010; Solberg et al., 2014; Mafakheri et al., 2020; Santagostini et al., 2020; Hong et al., 2022), chemical characterization on volatile compound (Solberg et al., 2014; McCallum et al., 2019; Martins et al., 2020; Dabbous-Wach et al., 2021; Hong et al., 2022), and non-volatile compound analysis (Srečec et al., 2010; Mongelli et al., 2016; McCallum et al., 2019; Santagostini et al., 2020; Hong et al., 2022). Most of the time, data are compared by statistical tests such as ANOVA (Analysis of the Variance), Principal Component Analysis (PCA) (Mongelli et al., 2016; Martins et al., 2020) or Partial Least-Square Discriminant Analysis (PLS-DA) (Hong et al., 2022). However, in a multidisciplinary approach, it may be interesting to correlate the different datasets to identify data structuring and to evaluate their interrelationship among the different parameters. To our knowledge, such work has never been done to characterize hop accessions. Multiple Factor Analysis (MFA) makes it possible to analyze several tables of variables simultaneously, and to obtain results, particularly charts, that allow us to study the relationship between observations, variables, and tables (Escofier and Pagès, 2008).

Our study is the first on the diversity of wild hop from the North of France. Actually, hop plants can be found in various locations throughout this territory, without knowing if they are wild or relics from earlier cultivations. The specimens collected could be wild or former cultivars whose identity is lost, local or naturalized clones, and even populations. Since the identity and properties of these accessions were unknown, plants have been collected and maintained in clonal archives for further characterization and evaluation.

Hence, our study was focused on 50 wild hop accessions (Fig. 3 and Table 1) compared to 10 commercial varieties cultivated in Northern France (Brewers Gold, Cascade, Challenger, Fuggle, Goldings, Magnum,

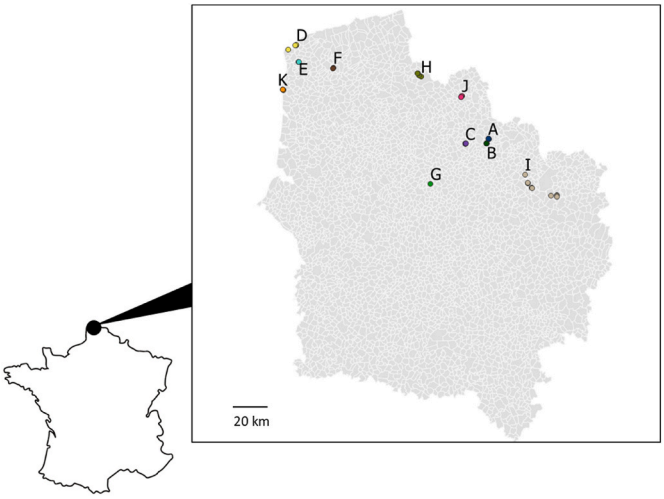


Fig. 3. Geographical repartition of the fifty accessions of wild hop (*Humulus lupulus* L.) collected on the 11 locations A to K in the North of France.

Northern Brewer, Nugget, Strisselspalt and Target) and 3 heirloom varieties formerly grown in the region (Groene Bel, Star and Coigneau). These heirloom varieties were cultivated in the region until the middle of the 20th century. Coigneau, a variety developed in Belgium during the 18th century, was named for its breeder and appreciated for its good yields and resistance to pests. This variety was cultivated until around 1910. At that time, Groene Bel progressively replaced Coigneau (Bart-Haas Report, 1930). Groene Bel is a variety with a light green stem coming from Belgian clonal selection in the late 19th century and known for its high levels in humulene. Star is an old Belgian variety, no longer

Table 1
List of the 50 accessions of wild hops collected in the Northern France (Hauts-de-France region). N: number of samples belonging to a location site. The GPS coordinates corresponds to the barycenter of the location of the N samples collected on the same collection site.

Location code	N = 50	Localities, city (departmental code)	GPS coordinates	Ecological particularities
A	5	Pré des Nonnettes, Scarpe-Escaut Regional Natural Park Marchiennes (59)	N 50°24′59.7419″ E 3°15′18.8099″	Regional nature reserve in the North of France Former estate cultivated by abbey of Marchiennes Wet meadow and megaphorbiaies
B	5	Tourbière de Vred, Scarpe-Escaut Regional Natural Park Vred (59)	N 50° 23′ 37.5″ E 3° 14′ 34.5″	Regional nature reserve in the North of France Alkaline bog Formerly exploited for peat
C	5	Douai (59)	N 50° 23′ 28.793″ E 3° 4′ 9.131″	Hedge of the agricultural domain of the Lycée de Wagnonville Edge of the Wagnonville marsh Regional nature reserve in the North of France
D	4	Cap Blanc Nez, Regional Natural Park of the Caps et Marais d'Opale Wissant (62)	N 50° 53′ 49.973″ E 1° 40′ 53.423″	White dune sand at stream outlet Hedge along a pond
E	2	Marquise (62) Regional Natural Park of the Caps et Marais d'Opale	N 50°48′33.4260″ E 1°42′10.6679″	Hedge Border of pastures
F	2	Audrehem (62) Regional Natural Park of the Caps et Marais d'Opale	N 50° 46′ 53.945″ E 1° 59′ 13.164″	Hedge Megaphorbiaies along a river
G	2	Moyeneville (60)	N 50° 10′ 59.375″ E 2° 47′ 10.631″	Hedge Edge of gardens, village area
H	4	Flanders, Météren (59)	N 50° 44′ 21.521″ E 2° 42′ 22.368″	Hills of Flanders Hedges on the edges of cultivated fields Forest edge
I	10	Regional Natural Park of Avesnois Avesnes-sur-Helpe (59)	N 50°9′24.1826″ E 3°41′27.106″	Hedge at the edge of the meadow Regional natural park of Avesnois
J	6	Citadel of Lille, Lille (59)	N 50°38′2.2140″ E 3°2′4.620″	Urban area Ditches of Vauban citadel of Lille Deûle riversides
K	5	Dunes d'Ecault, Saint-Étienne-au-Mont (62)	N 50°39′51.4679″ E 1°34′37.3260″	Grove, backshore dunes Creek banks

grown commercially, characterized by low vigor and low yield, but excellent storage stability and a continental-like aroma for home-brewing (Damseaux, 1911). As regards wild hops, some accessions were collected on sites which are of undisputed ecological value (seaside, peat bog). Others were taken from sites with ethnobotanical interest, where former hop fields may have been abandoned leaving potentially remnant hops. Hops were gathered in eleven locations presenting the same ecological characteristics (Table 1). The objectives of this research work were to (1) investigate the intraspecific genetic diversity of native *H. lupulus* in Northern France; (2) further explore the secondary metabolite profiles of these accessions; (3) search for any correlation between genetic characteristics and geographical, pedological or chemical data. The analysis of these different datasets provided a solid framework for the investigation of this diversity for diverse applications including varietal improvement research.

2. Results and discussion

2.1. Soil characterization

The analysis of collected soil samples close to root environment of the hop accessions emphasized the wide pedo-geographical diversity of the hop collection sites. These analyses were based on soil pH and conductivity measurements as well as the dosage of the elements Ca, Na, K and Mg. The data analysis through a heatmap associated with hierarchical cluster analysis (N = 8) represented as a dendrogram highlighted the pedological originality of some biotopes (Fig. 4 A and B). For example, Location B formed a very separate cluster (Cluster 1), characterized by high amounts of organic matter, a low ionic content, an acid pH and high conductivity. This can be explained by the fact that Location B is a peatland (the Tourbière de Vred, Vred, 59), a wetland with very high organic matter content. Locations D and K also constituted separate clusters (Clusters 6 and 7 respectively), globally characterized by low ionic contents (except for a high Na content for Cluster 7), low organic matter content, low conductivity but a basic pH. These two locations are both located on the seashore in the dunes (Fig. 3 and Table 1), around the Cap Blanc Nez (Wissant, 62), and in the Ecault Dunes (Saint-Étienne-au-Mont, 62), respectively. Other soil samples do not follow any particular trend. Indeed, the other clusters can include some soil samples from different locations, as is for example the case for the soil samples found in Cluster 8. This strong difference among soil samples was also confirmed by the visual aspect of soils. Actually, soils from locations B and C displayed a sandy texture while the peat soil in location A was rich in organic debris (Fig. 4C).

2.2. Genetic diversity and population structure

2.2.1. Genetic diversity

A collection of 63 hop genotypes including: (i) ten commercial varieties (Nugget, Strisselspalt, Northern Brewer, Target, Fuggle, Golding, Challenger, Magnum, Cascade, and Brewers Gold); (ii) three heirloom varieties (Groene Bel, Star, and Coigneau); (iii) fifty wild hops sampled from eleven different locations in Northern France (Hauts-de-France region) presenting different ecological characteristics (Table 1 and Fig. 3), was assessed for genetic diversity. This genetic characterization and the analysis of the population structure was carried out using eleven microsatellite markers (Supplementary material Table S6). All markers were polymorphic in all examined collections. A total of 56 different alleles was scored over the eleven tested loci. The number of amplified alleles per locus ranged from 5 (Hl-AGA1 locus) to 16 (GT1-K1-4 locus), with an average of 10.9 alleles per locus. The average number of different alleles per locus ranged from 1.636 (Location F) to 4.909

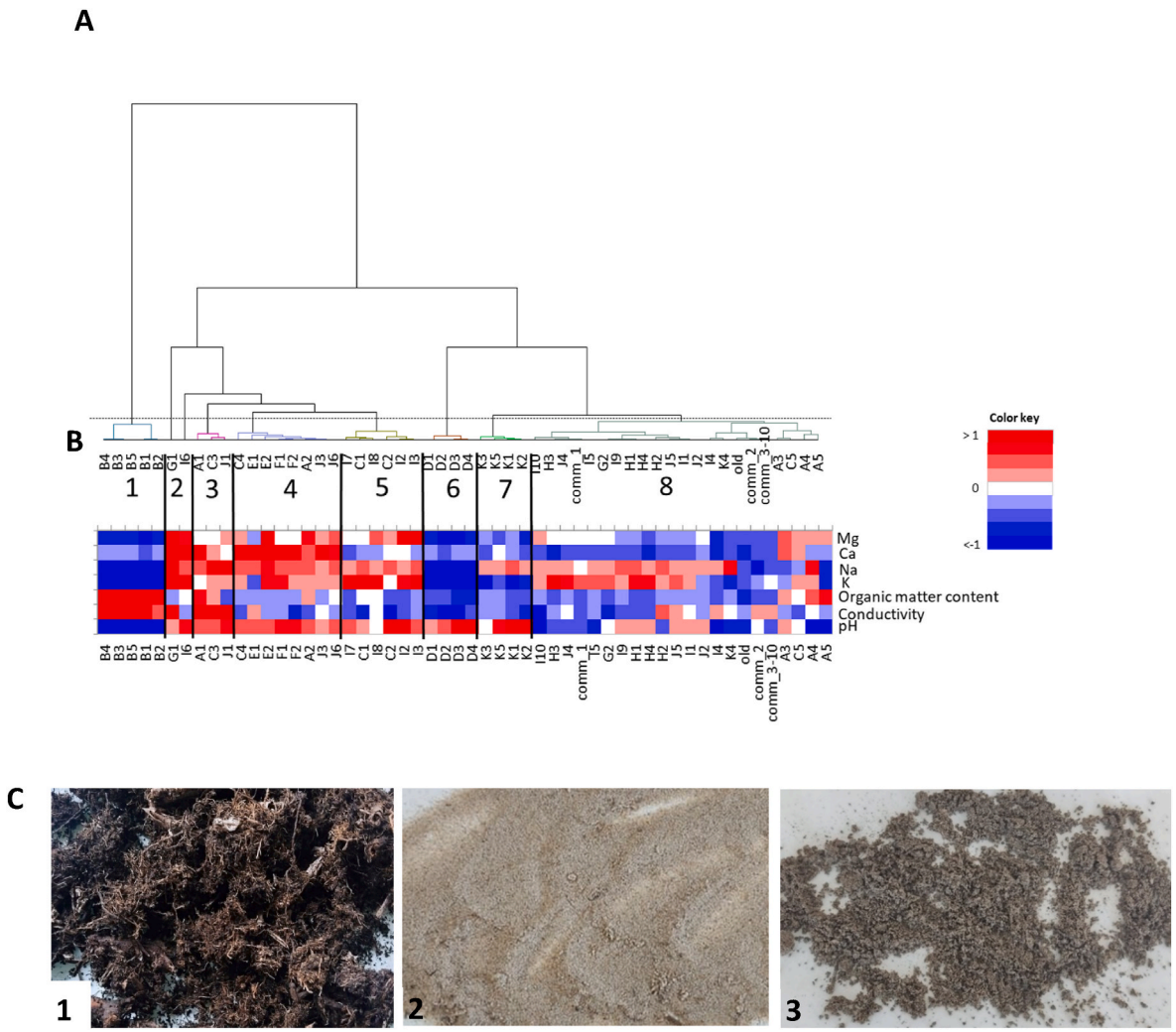


Fig. 4. Results of characterization of soil samples collected close to the root environment of the 63 hops studied. These 63 hops samples are represented by ten commercial varieties: Nugget comes from « comm_1 »; Strisselspalt comes from « comm_2 » and other commercial cultivars come from the same field « comm_3–10 »; three heirloom varieties coming from the same field identified as « old »; and fifty wild hops identified according to Table 1. The characterization of soil samples was based on pH and conductivity measurements, on organic matter content determined by loss of ignition as well as on dosage of the sodium, potassium, calcium and magnesium elements. **A.** Dendrogram including hierarchical cluster analysis (N = 8) among soil samples determined by soil characterization. **B.** Heatmap associated to the dendrogram. The 8 clusters of the hierarchical clustering were reported on the heatmap. **C.** Pictures of soil samples from locations B (Tourbière de Vred, Vred) (1), D (Cap Blanc nez, Wissant) (2) and K (Dunes d'Ecault) (3).

(commercial varieties), with an average of 2.95 alleles per locus across locations (Supplementary material Table S7). The average expected heterozygosity (H_E) and observed heterozygosity (H_0) values were 0.45 and 0.52, respectively (Supplementary material Table S7). Allele diversity analysis also revealed high and close levels of gene diversity among the studied sub-collections. Nei's gene diversity index was 0.53 on average for the whole collection, with very slight variations among each sub-collection (Supplementary material Table S7). The values of unbiased gene diversity index (standardized for sample size) were overall slightly higher than those of Nei's index, but globally displayed a similar extent of gene diversity when compared to the values of Nei's index (Supplementary material Table S7). Surprisingly, despite hop capacity of vegetative propagation and natural and human spread, it appeared from these data that all samples are genetically discrete, absolutely no clones were found. In the studies of Karlsson Strese et al. (2014) and Rodolfi et al. (2018), clonality was found within locations. This lack of clonality could be explained by the fact that we collected the wild accessions far enough apart to avoid sampling the same parent plant twice. Indeed, the accessions collected are at least 300 m from each other, thus underlying that the leaves collected for genotyping come from two different parent plant. The lack of clonality could also be due to the high level of gamete mixing, which may be explained by the strong winds from the ocean front of the region, which easily pushed pollen from one accession to another, thus dispersing the genetic resources and avoiding the identification of clones among the accessions in the collection. However, this hypothesis needs to be confirmed.

2.2.2. *Population structure*

F_{ST} calculated was equal to 0.29, which attested a significant structuration of the collection (>0.25). Comparison of F_{ST} and Φ_{ST} pairwise revealed significant genetic differentiation between all sampled locations, with F_{ST} and Φ_{ST} values among locations ranging from 0.06 to 0.28 and from 0.002 to 0.19, respectively (Supplementary material Table S8). The value range of both fixation indexes was overall similar among all locations and all F_{ST} and Φ_{PT} comparisons were significant at $P = 0.01$, except for commercial and heirloom varieties, Location B/Locations A, C, D, I, J and K, Location F/Location G, H, I, Location I/Locations J, K and Location J/Location K (Supplementary material Table S8). Some locations, such as E, F or H, seem to be very apart from the others, presenting for the whole of the pairwise F_{ST}, a value higher than

5%. Nonetheless for the commercial lines, the only pairwise F_{ST} value lower than 5% is for the cross with "heirloom" lines ($F_{ST} = 0.015$). This result hence revealed the strong link between "Commercial" and "Heirloom" which may assert the phylogenetic link between former and current varieties, because current cultivars derive from heirloom varieties (Darby, 2005). This hypothesis was also confirmed by the not significantly different Φ_{PT}. The pairwise F_{ST} higher than 5% for whole of the cross of commercial and heirloom lines with wild population suggested that the wild accessions are genuine wild genotypes and not remnant hops derived from commercial or heirloom lines. Bayesian analysis implemented within STRUCTURE version 2.3.4 supported the occurrence of genetic differentiation among locations and showed that several haplotypes had mixed origins (Fig. 5A). The measures of ΔK calculated from lnP(D) revealed that K = 13 was the most likely number of clusters (highest ΔK value = 5.87). Nevertheless, in Fig. 5A, we preferred to analyze the structure of the population with K = 6 clusters, corresponding to a plateau, in order to better highlight the structural differences of the genotypes within a location. The six identified clusters revealed a high genetic structuration among studied hops, notably in Fig. 5B. Strong variations were hence visible in the genetic structuration of some populations, such as that of Location H, mainly dominated by the yellow, or Locations F and G mainly dominated by turquoise (Fig. 5B). Furthermore, structuration of commercial and ancient lines was very close, dominated by the color redand widely different from the

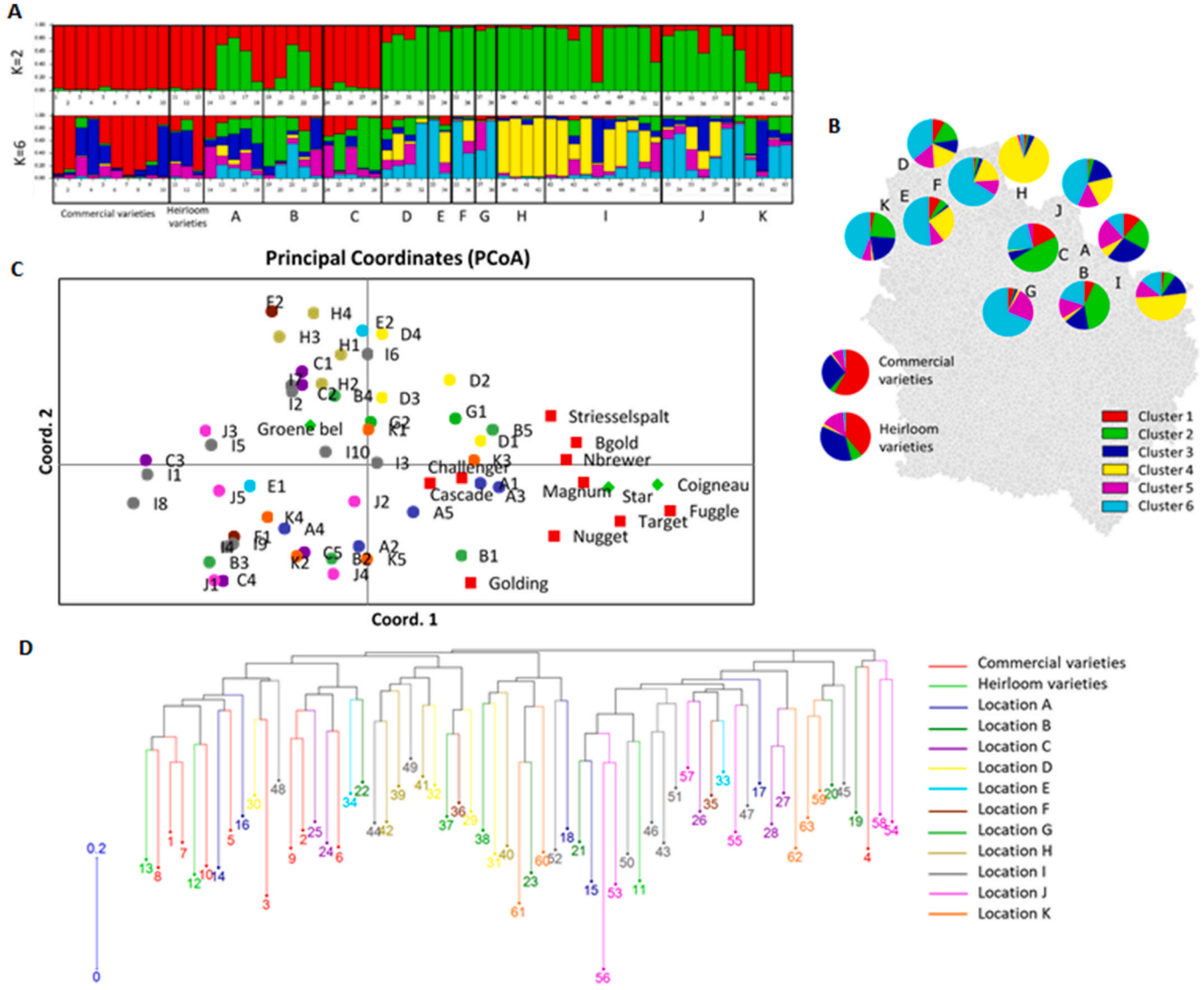


Fig. 5. Identification of population genetic structure of the 63 accessions of *Humulus lupulus* L. sampled in Northern France (Hauts-de-France region) using 11 microsatellites. **A.** Bar plot showing the distribution of individual assignations estimated for K = 2 and K = 6 clusters, from Bayesian inference cluster analysis performed with the 53 *Humulus lupulus* haplotypes sampled from the 14 locations (from A to K). Each vertical line represents an individual and the length of each colored line corresponds to the membership coefficient (scale at the left of the bar plot) for each cluster. Individuals are grouped according to their sampling locations. **B.** Frequencies of the 6 clusters (represented by colors) within each sampled location. Colors are same than on Fig. 5A. **C.** Principle Component Analysis (PCoA) based on genetic distances between each accession. Individuals were colored according to their sample site collection. **D.** Dendrogram underlying genetic clustering of the 63 hop accessions, including 10 commercial varieties (samples 1 to 10), 3 heirloom varieties (samples 11 to 13) and 50 wild sampled from 11 geographical locations (cf Table 1). 1: Nugget, 2: Strisselspalt, 3: Golding, 4: Challenger, 5: Brewers Gold, 6: Cascade, 7: Magnum, 8: Northern Brewer, 9: Target, 10: Fuggle, 11: Groene Bel, 12: Star, 13: Coigneau, Location A: 14 to 18; Location B: 19 to 23; Location C: 24 to 28; Location D: 29 to 32; Location E: 33 and 34; Location F: 34 and 35; Location G: 37 and 38; Location H: 39 to 42; Location I: 43 to 52; Location J: 53 to 58; Location K: 59 to 63. The tree was constructed using the unweighted neighbor-joining method based on genetic dissimilarity among the haplotypes according to microsatellite markers. Each branch corresponds to a hop genotype and the colors of branches indicate locations from which the genotypes were sampled. The color code is the same as the one on Fig. 5C.

wild lines. This was also simply confirmed by the bar plot with K = 2 clusters (Fig. 5A), on which commercial and heirloom lines were only composed of the red cluster, while wild lines were mostly composed of the green cluster. This observation hence confirmed the previous observations based on pairwise F_{ST}, according to which commercial and ancient lines may have a common background. Interestingly, the dendrogram constructed using the unweighted neighbor-joining method (Fig. 5D) clearly showed a greater diversity among wild hops compared to commercial varieties (in red) or heirloom varieties (in bright green). Only Challenger and Groene Bel (respectively accessions numbers 4 and 11) seem to be more genetically distant from other commercial and heirloom varieties. Overall, hops collected from the same area (characterized by the same color) were globally genetically close to each other. The analysis by principal components (PCoA) (Fig. 5C) supported this analysis underlying the global proximity of commercial lines (in red), compared to the wide diversity of wild accessions. Haplotype clustering on the dendrogram was consistent with genetic differentiation highlighted by the *STRUCTURE* program, with several common cluster features (Fig. 5A and D). Locations H and K, identified previously with a dominating color (green and yellow respectively) hence generated separate clusters with these two different approaches. These different features seem to assert the close parentage of commercial varieties, which may contain genetic material of ancient varieties, except for the ancient variety Groene Bel, which appears to be genetically closer to the wild accessions. Overall, these observations converged to show the proximity of the commercial accessions to each other and indeed to the diversity of the wild accessions in our collection. This observation may be linked to the relative proximity of our commercial lines that contain large amounts of North American genetic material. For instance, Northern Brewer, Magnum, Nugget and Target are progeny of Brewers Gold, itself containing half of north American *H. lupulus* ssp. *Lupuloides* (Turner et al., 2011).

Location H corresponds to the current location of hop fields where commercial varieties are currently cultivated in the North of France, in particular in the Monts de Flandres, near Bailleul and Meteren, (59, Hauts-de-France). We also previously noted the original genetic structure of hops collected in Location H, with insignificant pairwise F_{ST}, which suggests that wild hops from Location H were real wild hops and not remnant hops or scattered plants. AMOVA (Supplementary material Table S14) revealed that 5% of the genetic variation could be explained by differences among sub-collections (locations) whereas 95% was detected within sub-collections, although variation among populations was highly significant ($P < 0.001$).

Our results provided a first investigation of the genetic diversity of hops collected in Northern France (Hauts-de-France region) with a high level of polymorphism. In other parts of the world, for example in Iran or in Italy, previous studies conducted on the identification of genetic structure of wild hops have already underlined a high level of diversity (Rodolfi et al., 2018; Mafakheri et al., 2020). In these two studies, genetic analysis with STRUCTURE software revealed a clusterization by collection site, which may be explained by the great distance between them. In our study, our observations were less contrasted among wild accessions but it must be put in perspective with the relatively small collection area (12,414 km^2 for the Hauts-de-France region). Our results

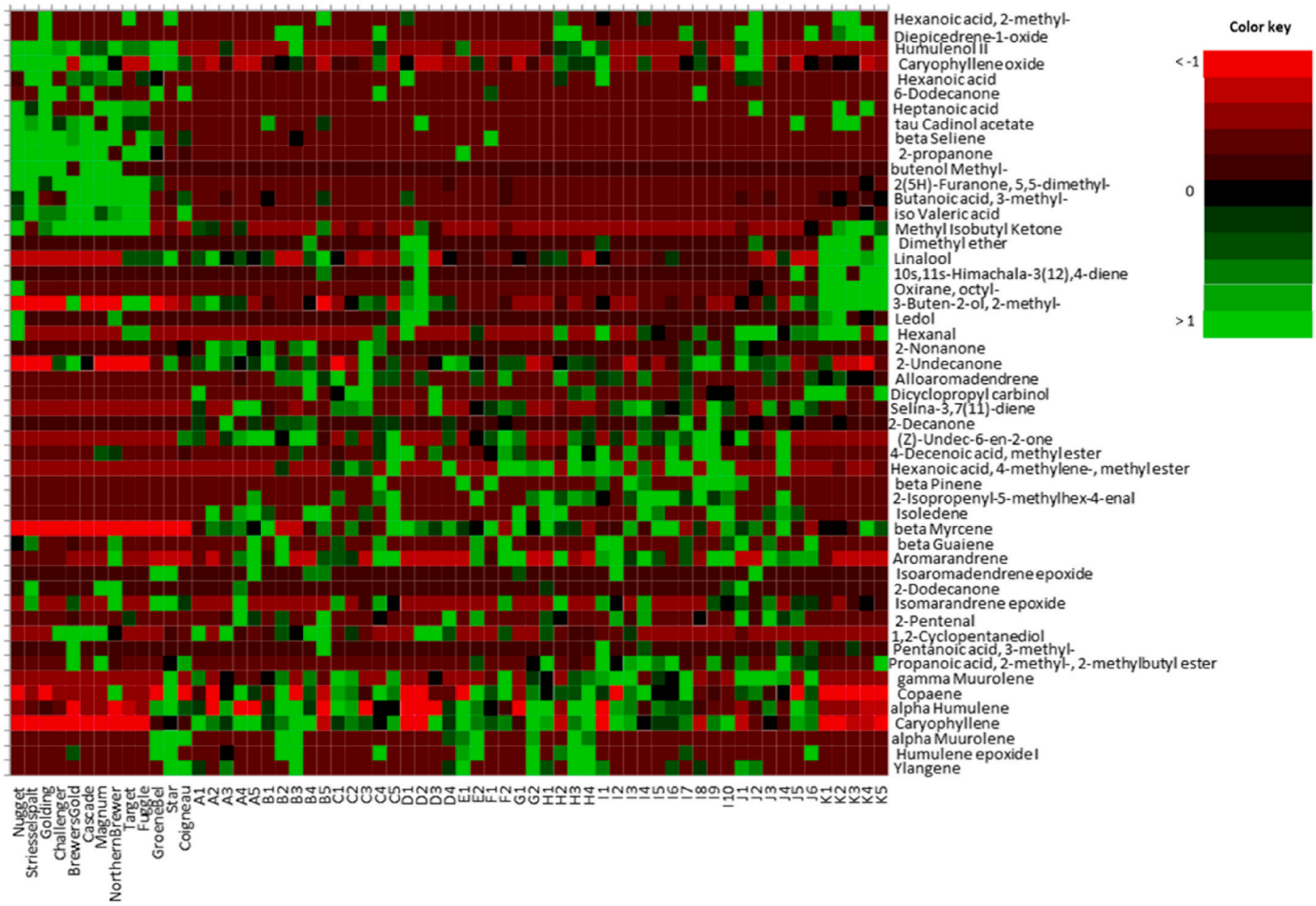


Fig. 6. Heatmap highlighting variation of volatile compounds across the 63 hop accessions from Northern France. This heatmap has been generated with normalized data for the top 51 molecules responsible for differences between the chemical profiles. Red and green colors indicate lowest and highest performance of the traits, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

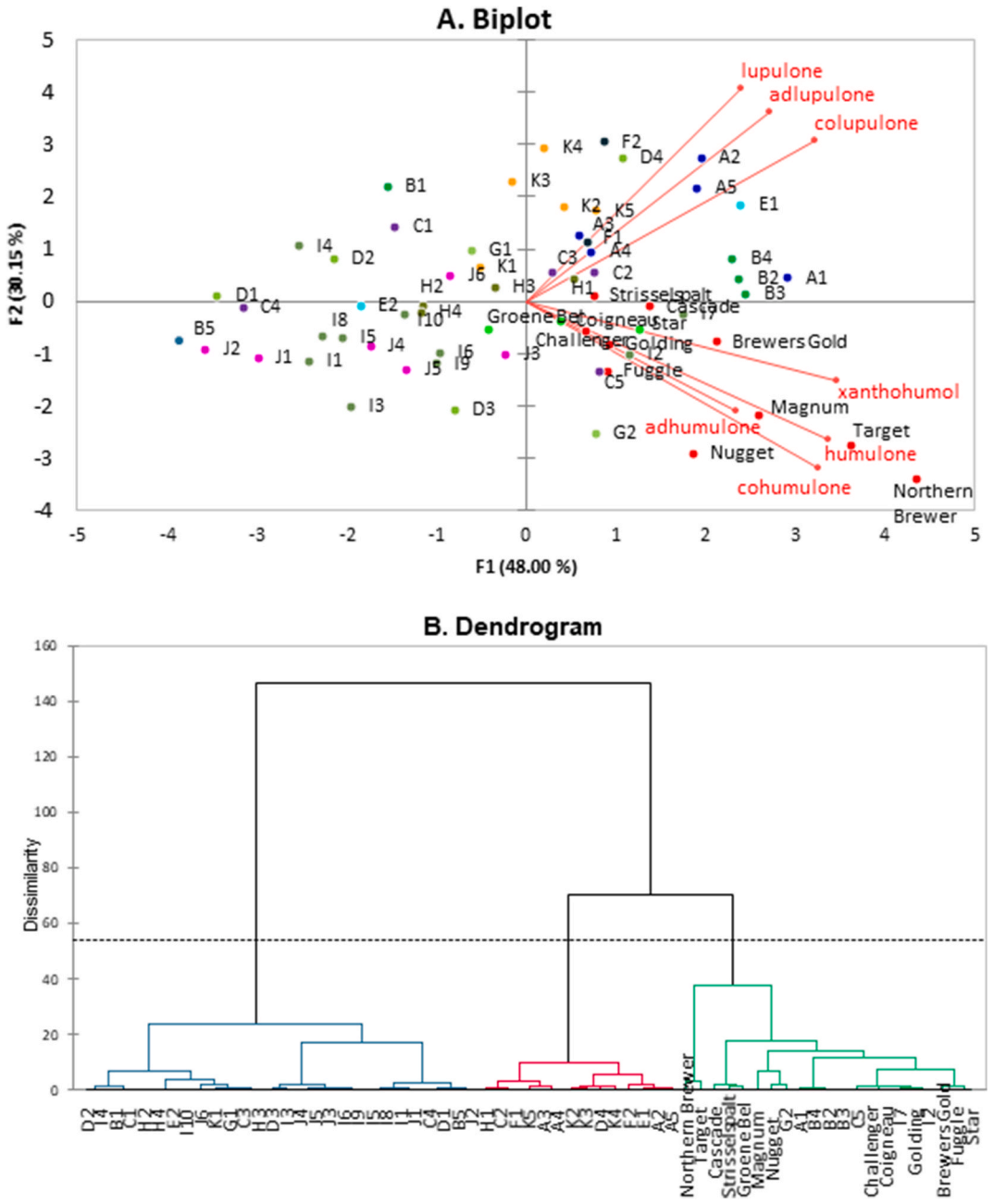


Fig. 7. Results of the statistical treatment of data for the quantitation of xanthohumol, *co-, n-, ad*-humulone and *co-, n-, ad*-lupulone. This quantitation has been performed on the 63 crude hydro-ethanolic extracts of hop cone powder from Northern France, including 10 commercial varieties, 3 heirloom varieties and 50 wild hops (Fig. 1, Table 2). **A.** PCA biplot of quantitation data with score plot and loading plot of variables. Individuals were colored by collection site for a given observation. Variable contribution to component was represented by arrows length. **B.** Dendrogram of the hierarchical cluster analysis among the 63 hops based on the quantitation similarity (Ward's method, distance scale) (N = 3).

also underlined that our wild samples, by their genetic diversity compared to commercial varieties, could be accessions resulting from migrations through Europe, and not from human behavior. These results confirmed the hypotheses of Murakami et al. (2006).

2.3. Phytochemical characterization

Phytochemical characterization analyses were done on the powder of cones collected *in-situ* of the same accessions as previously genetically characterized.

2.3.1. Volatile compounds analysis

The identification of volatile compounds by HS-SPME GC-MS revealed the presence of 101 different compounds, present in at least 1% of relative content (relative peak area), in all of the studied hops (Fig. 2 and Supplementary material Table S15). The profiles were composed of a large diversity of oxygenated (alcohols, aldehydes, ketones, carboxylic acids) and non-oxygenated terpenes and sesquiterpenes. Hop samples contained a greater or lesser diversity of volatile compounds. The sample with the greater diversity of volatile compounds was accession I1 (44 volatile compounds identified), while the sample with the lowest number of volatile compounds was accession E1 (13 volatile compounds

identified). Among these 101 volatile compounds, some were identified more frequently than others. The heatmap (Fig. 6) summarizes the relative abundance of the 51 volatile compounds more frequently identified in the samples, that is to say, identified in at least 15 hop samples on the 63 studied. Hence, the main volatile compounds of hops, including non-oxygenated monoterpenes (*β*-myrcene), non-oxygenated sesquiterpenes (*β*-caryophyllene, *α*-humulene, *α*-copaene, *γ*-muurolene), oxygenated monoterpenes (linalool) and 2-undecanone (methyl nonyl ketone) were found in all hops studied (Fig. 2). These compounds are known as the main volatile compounds found in hop cones and therefore they do not constitute a particular chemical originality' (Vázquez-Araújo et al., 2013). They are known for their woody,

Table 2
Data of the quantitation of the main phenolic compounds (xanthohumol, *co*-,*n*-, *ad*-humulone and *co*-, *n*-, *ad*-lupulone) in the 10 commercial varieties, the 3 old varieties and the 50 wild accessions (Locations A to K), expressed in % of drdry weight, quantified by UHPLC-UV.

	Xanthohumol		Cohumulone		Humulone		Adhumulone		Colupulone		Lupulone		Adlupulone	
Nugget	0.58%	ab	2.35%	a-d	6.78%	ab	0.39%	b-g	1.21%	a-e	1.04%	c-f	0.30%	c-p
Strisselspalt	0.27%	b-d	0.74%	d-f	1.86%	b-g	0.58%	ab	1.75%	a-e	1.86%	a-f	0.19%	d-p
Golding	0.34%	b-d	1.03%	b-f	3.23%	b-g	0.24%	b-g	1.25%	a-e	1.53%	a-f	0.19%	c-p
Challenger	0.38%	b-d	0.97%	c-f	2.94%	b-g	0.44%	a-e	1.63%	a-e	1.76%	a-f	0.26%	c-p
Brewers Gold	0.41%	bc	2.08%	a-c	2.86%	b-g	0.20%	c-g	1.43%	ab	1.03%	b-f	0.26%	b-p
Cascade	0.25%	b-d	0.73%	b-f	1.30%	b-g	0.28%	a-d	1.45%	a-e	1.28%	a-f	0.13%	d-p
Magnum	0.27%	b-d	1.71%	a-e	5.23%	a-c	0.68%	a	1.60%	a-e	1.90%	a-f	0.19%	d-p
Northern Brewer	0.76%	a	2.27%	ab	6.43%	a	0.33%	a-g	1.51%	a-e	1.42%	a-f	0.19%	d-p
Target	0.73%	a	2.71%	a	4.30%	a-e	0.10%	d-g	1.78%	a-e	0.81%	c-f	0.24%	c-p
Fuggle	0.38%	ab	1.09%	b-f	2.56%	b-g	0.15%	c-g	1.00%	a-e	0.87%	b-f	0.19%	c-p
Groene Bel	0.17%	b-d	0.32%	d-f	0.66%	d-g	0.32%	a-c	1.09%	a-e	0.84%	a-f	0.07%	k-p
Star	0.31%	ab	0.94%	b-f	2.02%	b-g	0.11%	c-g	1.07%	a-e	0.88%	a-f	0.21%	a-l
Coigneau	0.20%	b-d	0.50%	c-f	1.82%	b-g	0.20%	b-g	0.94%	a-e	0.83%	a-f	0.18%	a-p
A1	0.16%	b-d	0.66%	b-f	2.58%	a-c	0.17%	a-g	1.06%	a-e	1.16%	a-f	0.21%	a-c
A2	0.16%	b-d	0.38%	d-f	1.31%	b-g	0.11%	c-g	2.60%	a	1.42%	a-d	0.25%	ab
A3	0.13%	b-d	0.31%	d-f	1.55%	b-g	0.08%	d-g	0.95%	a-e	1.20%	a-f	0.14%	a-o
A4	0.15%	b-d	0.50%	d-f	2.79%	b-g	0.14%	c-g	1.08%	a-e	1.82%	a-e	0.22%	a-l
A5	0.21%	b-d	0.42%	d-f	1.92%	b-g	0.12%	c-g	1.27%	a-e	1.84%	ab	0.26%	a-c
B1	0.06%	cd	0.10%	f	0.41%	fg	0.02%	fg	0.92%	a-e	1.17%	a-f	0.17%	a-m
B2	0.13%	b-d	0.39%	d-f	1.40%	a-g	0.10%	b-g	0.67%	a-e	0.82%	a-f	0.11%	a-j
B3	0.16%	b-d	0.65%	d-f	2.18%	a-f	0.15%	b-g	0.97%	a-e	1.07%	a-f	0.17%	a-i
B4	0.21%	b-d	0.83%	d-f	1.99%	b-g	0.17%	b-g	1.31%	a-e	1.39%	a-f	0.22%	a-d
B5	0.03%	d	0.16%	ef	0.18%	g	0.01%	g	0.17%	e	0.19%	f	0.03%	op
C1	0.09%	b-d	0.17%	f	0.63%	d-g	0.06%	d-g	0.62%	a-e	1.14%	a-f	0.18%	a-m
C2	0.15%	b-d	0.50%	c-f	1.32%	b-g	0.11%	b-g	1.11%	a-e	0.96%	a-f	0.12%	a-p
C3	0.17%	b-d	0.49%	c-f	1.88%	b-g	0.14%	c-g	1.18%	a-e	1.19%	a-f	0.19%	a-p
C4	0.16%	b-d	0.15%	f	0.84%	e-g	0.05%	efg	0.92%	a-e	1.08%	c-f	0.11%	m-p
C5	0.22%	b-d	1.16%	b-f	3.38%	b-g	0.29%	a-g	1.01%	a-e	0.90%	b-f	0.25%	a-n
D1	0.07%	cd	0.08%	f	0.44%	g	0.01%	g	0.59%	b-e	0.75%	c-f	0.08%	l-p
D2	0.06%	b-d	0.08%	f	0.47%	d-g	0.03%	d-g	0.42%	a-e	0.64%	a-f	0.08%	d-p
D3	0.19%	b-d	1.01%	b-f	3.71%	b-g	0.30%	b-g	0.80%	a-e	0.76%	d-f	0.13%	j-p
D4	0.29%	b-d	0.46%	ef	1.43%	c-g	0.16%	c-g	2.37%	a-c	2.08%	a-f	0.47%	a
E1	0.08%	b-d	0.21%	b-f	0.56%	b-g	0.04%	c-g	0.51%	ab	0.54%	a-c	0.07%	a-f
E2	0.28%	b-d	0.67%	ef	2.44%	c-g	0.20%	d-g	1.59%	a-e	1.70%	b-f	0.30%	d-p
F1	0.22%	b-d	0.82%	c-f	2.46%	b-g	0.20%	c-g	2.33%	a-d	2.08%	a-f	0.25%	c-p
F2	0.12%	b-d	0.25%	ef	1.22%	b-g	0.07%	d-g	1.28%	a-e	2.45%	a	0.21%	a-e
G1	0.14%	b-d	0.35%	ef	1.24%	b-g	0.14%	c-g	1.07%	a-e	1.29%	a-f	0.21%	a-m
G2	0.27%	b-d	1.36%	b-f	4.51%	a-d	0.35%	a-f	0.89%	a-e	1.16%	b-f	0.12%	j-p
H1	0.18%	b-d	0.71%	c-f	2.24%	b-g	0.18%	b-g	1.57%	a-e	1.77%	a-f	0.12%	g-p
H2	0.26%	b-d	0.36%	ef	1.42%	b-g	0.11%	d-g	0.83%	a-e	1.02%	b-f	0.18%	c-p
H3	0.33%	b-d	0.74%	d-f	2.94%	b-g	0.29%	b-g	1.87%	a-e	2.29%	a-f	0.29%	c-p
H4	0.21%	b-d	0.45%	d-f	1.20%	c-g	0.13%	c-g	0.97%	a-e	0.92%	b-f	0.13%	d-p
I1	0.20%	b-d	0.39%	ef	1.57%	b-g	0.12%	d-g	0.58%	c-e	0.72%	d-f	0.08%	n-p
I2	0.25%	b-d	0.78%	b-f	2.53%	b-g	0.25%	a-g	0.87%	a-e	1.07%	a-f	0.17%	c-p
I3	0.14%	b-d	0.86%	c-f	2.90%	b-g	0.19%	c-g	0.51%	de	0.67%	def	0.07%	p
I4	0.18%	b-d	0.10%	f	0.56%	g	0.02%	g	1.09%	a-e	1.77%	a-f	0.23%	d-p
I5	0.12%	b-d	0.37%	d-f	1.41%	b-g	0.08%	d-g	0.50%	c-e	0.68%	c-f	0.10%	h-p
I6	0.20%	b-d	0.77%	c-f	2.24%	b-g	0.19%	b-g	0.98%	a-e	0.88%	c-f	0.15%	e-p
I7	0.29%	b-d	1.16%	b-f	3.40%	b-g	0.34%	a-g	1.95%	a-e	1.76%	a-f	0.24%	b-p
I8	0.23%	b-d	0.71%	d-f	1.96%	b-g	0.16%	d-g	1.20%	a-e	1.09%	c-f	0.18%	i-p
I9	0.26%	b-d	0.83%	d-f	3.65%	b-g	0.32%	b-g	0.98%	a-e	1.39%	b-f	0.19%	f-p
I10	0.18%	b-d	0.57%	d-f	2.03%	b-g	0.15%	c-g	1.01%	a-e	1.27%	a-f	0.17%	d-p
J1	0.09%	bcd	0.21%	ef	0.59%	d-g	0.06%	d-g	0.35%	c-e	0.29%	f	0.04%	p
J2	0.04%	cd	0.11%	f	0.30%	e-g	0.03%	d-g	0.19%	e	0.18%	f	0.03%	p
J3	0.15%	b-d	0.98%	b-f	2.71%	b-g	0.27%	b-g	1.12%	a-e	0.97%	b-f	0.19%	c-p
J4	0.19%	b-d	0.78%	d-f	2.05%	b-g	0.25%	b-g	1.05%	a-e	0.89%	d-f	0.20%	c-p
J5	0.12%	b-d	0.71%	c-f	1.60%	b-g	0.18%	b-g	0.70%	a-e	0.48%	ef	0.13%	c-p
J6	0.13%	b-d	0.23%	ef	0.96%	b-g	0.14%	b-g	0.63%	a-e	0.72%	b-f	0.20%	a-i
K1	0.12%	b-d	0.27%	d-f	1.02%	b-g	0.06%	d-g	0.76%	a-e	0.93%	a-f	0.10%	d-p
K2	0.17%	b-d	0.35%	d-f	1.29%	b-g	0.12%	c-g	1.28%	a-e	1.50%	a-f	0.24%	a-g
K3	0.14%	b-d	0.07%	f	0.40%	e-g	0.08%	c-g	1.02%	a-e	1.14%	a-f	0.16%	a-k
K4	0.13%	b-d	0.06%	f	0.28%	e-g	0.02%	e-g	0.90%	a-c	0.85%	a-f	0.14%	a-d
K5	0.22%	b-d	0.48%	d-f	1.90%	b-g	0.04%	e-g	1.35%	a-e	1.63%	a-f	0.25%	a-h

The value reported is the mean of the technical triplicate.
Different letters indicate significant differences between treatments according to ANOVA followed by Tukey's test at $p \leq 0.05$.

herbaceous or spicy aromatic potential after brewing (The Chemical Sources Association, 2010). Other terpenes such as aromadendrene, allo-aromadendrene, isoledene, *β*-guaiene, *α*-ylangene and *β*-pinene were also putatively identified to a lesser extent in some samples (in samples 44, 25, 18, 17, 17 and 10 respectively) and constituted a chemical originality in the composition of volatile compounds of these wild hop samples. Nevertheless, these volatile compounds in hop samples have already been identified in some hop samples, as reported in different publications (Vázquez-Araújo et al., 2013). Lastly, some scarcer volatile compounds, and especially some terpenes, were putatively identified in some hop samples, and constituted an originality in the chemical fingerprint of some wild hops (Supplementary material Table S15). For example, *β*-panasinene and aristolene were only identified in 2 samples. *α*-copaene and *γ*-cadinene were even more rarely identified in hop samples (Bocquet et al., 2018a), just like *α*-bergamotene, a compound with aromatic tea and woody characteristics (The Chemical Sources Association, 2010). *β*-panasinene has already been identified as a major volatile compound in the Perle hop variety (Tofana et al., 2009). To our knowledge, aristolen had never been found in hop samples, but is not known to have particular aromatic characteristics. In some cases, we also noticed the presence of oxidized compounds, such as caryophyllene oxide, humulene I epoxide, humulenol II or humulene oxide II. This strong presence of oxidized compounds, especially in some wild hop samples, was probably due to the ripening of cones of some wild accessions at the time of collection. Indeed, for logistical reasons, the collection of wild hops was spread over the month of September 2019. Thus, some cones of wild hops had already reached an advanced phenotypical stage, which may explain the presence of these degradation compounds (Uemoto et al., 2022). This phenomenon may disappear under *ex-situ* conditions in experimental hop fields. The dendrogram associated with the heatmap (Fig. 6) grouped together the commercial varieties. In the same way, it also underlined a relative proximity in volatile compound composition between the 3 heirloom varieties Groene Bel, Star and Coigneau. As was stated previously regarding the genetic features, commercial lines share an important part of their genetic heritage. Furthermore, they also come from the same localities and have been exposed to the same biotic and abiotic pressures. As chemotype can be considered as the result of the influence of the environment and the genotype, this proximity of commercial and heirloom varieties was not really surprising. Wild hop samples collected at Location K were also clustered together, the singularity of their volatile compounds could be the consequence of the influence of their environment, in this case very sandy and highly exposed to westerly winds.

2.3.2. Main phenolic compound composition

UHPLC analysis revealed the presence of 7 main peaks at retention times 2.76, 4.35, 4.73, 4.83, 6.26, 6.87 and 7.00 min (Fig. 1) attributed to xanthohumol (XN), α-acids *co*, *n*-, *ad*-humulone (α1, α2, α3) and β-acids *co*-, *n*-, *ad*-lupulone (β1, β2, β3) respectively. This attribution was based both on their retention times compared with those of the standards purified in the laboratory and on their mass spectra (Bocquet et al., 2019). Xanthohumol and *co*-, *n*-and *ad*-humulone, and *co*-, *n*- and *ad*-lupulone were then quantified in the 63 crude extracts of cone samples using the quantitation method set up by UHPLC-UV. Acceptable linearity was observed for each compound over the concentration range used for calibration. Evaluation of the recovery data of the quantitation method showed acceptable intra and inter-day precisions for xanthohumol (RSD % = 16.9, 15.78), *co*-humulone (RSD % = 17.91, 19.59), humulone (RSD % = 15.55, 16.75), *co*-lupulone (RSD % = 14.12, 12.51) and lupulone (RSD % = 13.95, 12.39). Quantitation data for *ad*-humulone and *ad*-lupulone were obtained by extrapolation of dosing equation of humulone and lupulone respectively. Results were expressed in μg. mL^{-1} Statistical analysis with Tukey's test revealed significant differences among hop samples in term of phenolic compound composition (Table 2). The higher content in xanthohumol, *co*-, *n*-, *ad*-humulone and *co*-, *n*-, *ad*-lupulone were obtained for the Northern Brewer (0.76%), Target (2.71%), Northern Brewer (6.43%), Magnum (0.68%), A2 (2.60%), F2 (2.45%) and D4 (0.47%) samples respectively. While the lower content measured in xanthohumol, *co*-, *n*-, *ad*-humulone and *co*-, *n*-, *ad*-lupulone were obtained for the B5 (0.03%), K4 (0.06%), B5 (0.18%), B5/D1 (0.01%), B5 (0.17%), J2 (0.18%) and B5/J2 (0.03%) samples respectively. The output process of the quantification in the form of PCA (Fig. 7A) revealed a good explanation for the variable's distribution, with 83.46%, and a contribution of PC1 and PC2 equal to 48.00% and 30.15% respectively. Moreover, Fig. 7A clearly revealed a good separation between *α*-acids contents on one side and *β*-acids on the

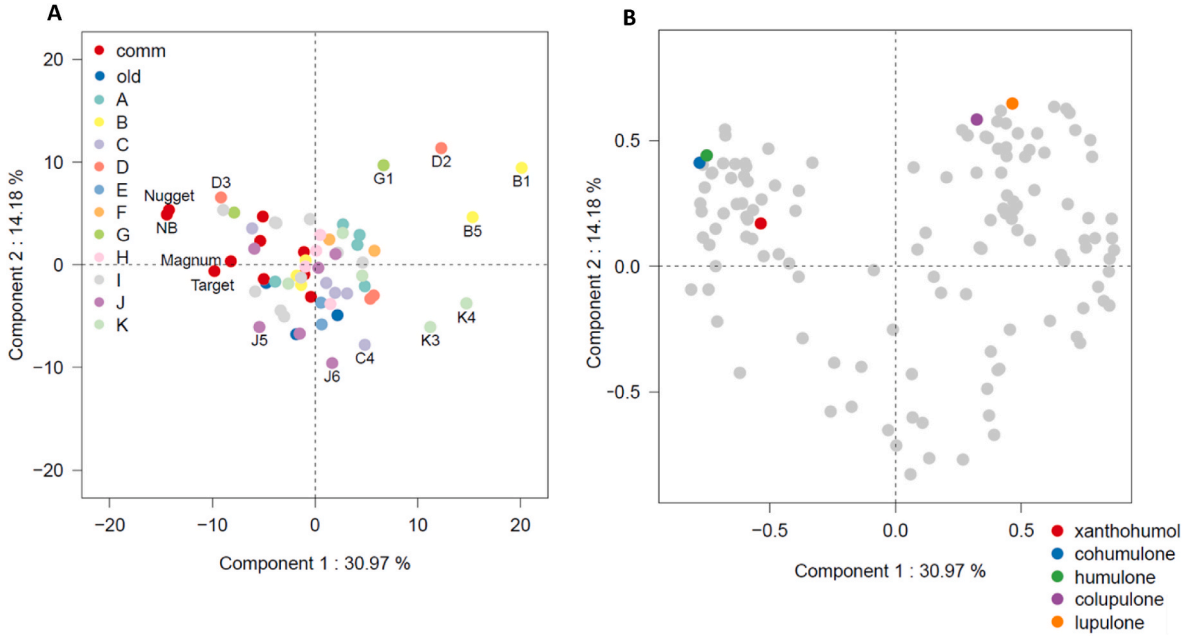


Fig. 8. Untargeted metabolomic analysis **A.** Principle component analysis of the 63 chemotypes of hop studied. Each symbol represents a single plant from the different accessions. Commercial varieties (10 accessions), heirloom varieties (3 accessions), wild hops collected on different locations (50 accessions, Fig. 3). **B.** Principle component analysis of the chemical markers.

other side. Hops are traditionally classified according to their content in α-acids (Palmer, 2006). Depending on the biplot (Fig. 7A), commercial varieties (in red) had globally higher content of α-acids (including *co*-, *n*- and *ad*-humulone) and of xanthohumol. On the contrary, wild hops (Locations A to K) had higher β-acid content (*co*-, *n*- and *ad*-lupulone). Fig. 7A thus revealed a relatively low bitterness potential of wild hops compared to commercial varieties. However, some exceptions were noted for the wild hops G2 (1.36%), I7 (1.16%), C5 (1.16%), D3 (1.01%) and J3 (0.98%) that showed *co*-humulone content relatively close to that of dual-purpose Challenger hops (0.97%). These wild hops can further provide a softer bitterness due to a slightly lower *co*-humulone content by comparison with Target (2.71%) or Nugget (2.35%) (Table 2). The volatile compounds analysis of these four accessions revealed a high content of the most common volatile compounds of hop (α-humulene, β-myrcene, β-caryophyllene, α-copaene, γ-muurolene) (Fig. 6) and highlighted the presence of some original volatile compounds such as ylangene (A1, B3, G2) or aromarandrene (A1, B2). The B2, B3 and G2 accessions stand out for their originalities in volatile compounds, with the presence of β-guaiene, longifolene, α-muurolene, α-cubebene in the B2 accession, α-copaene, α-muurolene, β-selinene in the B3 accession, and α-muurolene and γ-muurolene in G2. These volatile compounds may offer some potential for brewing. Actually, allo-aromarandrene, α-copaene, β-guaiene, longifolene are for example known for their woody aromatic potential (The Chemical Sources Association, 2010). These results also underlined that wild hops were richer in β-acids known for their properties in brewing, in particular as a bitterness stabilizer and preservation enhancer, as well as their biological activities, particularly sought after by pharmaceutical or agro-food companies (Bocquet et al., 2018a). This high content of β-acids in wild accessions has already been reported by (McCallum et al. (2019) on samples collected in Canada. On the biplot (Fig. 7A), wild hops seemed to be globally unified by collection site (by color). This observation was partially confirmed by the dendrogram of the hierarchical ascending clustering (HAC, Euclidean distance, Ward's agglomeration, dissimilarity, N = 3) (Fig. 7B), where wild hops collected on the same site nearly belong to the same branch. For example, wild hops from Location I were in the green cluster, whereas hops from Location J (except J6), were in the pink cluster. Furthermore, this dendrogram clearly confirmed the opposition between commercial lines (containing North American germplasm) and heirloom varieties, mostly gathered in the three same clusters (in purple), and wild accessions dispersed through other clusters. Wild accessions lying within this pink cluster (G2, A1, B4, B2, B3, C5, I7, I2) may contain North American genetic material, due to outcrossing events.

In particular, we have already showed in Fig. 5 the genetic proximity of accessions I2 and I7 with commercial genotypes, which is also in line with this hypothesis. Actually, hop pollen can travel more than 10 km on the wind, and North American hops may possess chemicals helpful in resisting against local pests and pathogens.

These results were then very informative about the bitterness potential of wild hops from Northern France, and their aromatic potential if we consider the results previously detailed. Nevertheless, as we mentioned the degradation of volatile compounds, α- and β-acids are very prone to oxidation (Taniguchi et al., 2013). Even if we had paid particular attention during the storage of the powder and the realization of the samples, it remains nevertheless a parameter to be kept in mind in these quantification results.

2.3.3. Untargeted metabolomics analysis

From the Progenesis matrix where 270 signals were detected, the cleaning allowed only 132 compounds to be kept. The data of this final matrix were processed by PCA (Fig. 8). The first two principle components explained 45.15% of the total variance (30.97% for the PC1 and 14.18% for the PC2). Among the 132 compounds we were able to identify the five markers quantified previously, namely xanthohumol, *co*- and *n*-humulone, *co*- and *n*-lupulone (Fig. 8A). It may be very interesting to further this observation by identifying markers that explain the distribution of the individuals regarding this feature. The untargeted metabolomics analysis underlined a great diversity among wild hops compared to commercial varieties (Fig. 8B). These results confirmed the results obtained with the quantification of xanthohumol, α- and β-acids. This strong difference in the phytochemical composition between the 50 wild hops and the 13 varieties (including heirloom varieties and commercial varieties), may be explained by genetic difference or terroir influence. Nevertheless, it may also be influenced by the treatment of the samples post-harvest.

2.4. Comparison of the different data sets by multifactorial analysis

The five datasets previously obtained were compiled to obtain a global matrix of 393 variables, distributed in the five tables as follows: soil characterization: 5 variables; genetic characterization: 63 variables; volatile compounds analysis: 188 variables; quantification analysis: 5 variables and untargeted metabolomics: 132 variables. As noted later in the materials and methods section, all data were considered quantitative variables, except data of volatile compound analysis which were considered qualitative variables (presence/absence of the compound). To identify specific patterns in these five data matrices, MFA was applied (Fig. 9 and Supplementary material Table S17). MFA performed individual PCAs on the data matrices obtained from the characterizations. Fig. 9a presents the scores plot of the MFA containing the results of all the datasets. The first and the second principal components explained 10.02% and 7.51% of the variation, respectively. It was not a very high percentage of explained variance, but it was satisfactory, considering the large number of variables (393). Fig. 9a confirms the wider diversity of wild hops compared to commercial varieties from a genetic and chemical point of view. Fig. 9b presents the connections among the different data sets. It shows that the first dimension separates on the one hand the soil characterization and on the other hand the chemical and the genetic characterization. The second dimension separates the genetic characterization, and the analysis of volatile compounds. Fig. 9c illustrates in the form of dendrogram the result of the HAC (Ward's method, N=7) on the results of the MFA. It was interesting to observe the distribution of hop accessions among the seven clusters. Varieties (commercial and heirloom) belonged mostly to the same cluster (in red, 2). Only "Groene Bel" and "Challenger" were included in two other clusters (Clusters 5 and 6, respectively). The genotype Groene Bel was already separated from other accessions on the dendrogram processed according to the genetic results (Fig. 5D). Furthermore, within the same cluster, we globally found wild hops coming from the same location. For example, Cluster 1 was globally dominated by wild hops from Location B (3 hops B2, B3, B4); and the remaining two hops from Location B (B1 and B5), belonged to the same green cluster (3). Cluster 4 (pink) contained three of the four wild hops from Location D, and as well as two wild hops from Location K. Cluster 6 contained wild hops from six locations (of eleven) including four wild hops from Location C (C1, C2, C3, C4). It also contained the two wild hops from Location F (F1 and F2), three wild hops from Location K (K1, K2, K5), as well as some wild hops from Location A (A3, A4), Location I (I4 and I6) and Location J (J2 and J6). Lastly, Cluster 7 contained seven wild hops from Location I (I1, I2, I5, I7, I8, I9, and I10) out of 10 total; it also had three of the four wild hops from Location H (H1, H3, H4). This "terroir effect" is well known and well documented (Moore et al., 2014). However, even if some authors mention and discuss a possible terroir effect on the hops and their chemical composition (Morcol et al., 2020; Dabbous-Wach et al., 2021), no demonstration has yet been made to our knowledge. Lg coefficient measured the extent to which the tables are linked two by two (Supplementary Material Table S17). RV coefficients corresponded to a normalization of the Lg coefficients; their values were then between 0 and 1. The stronger the linkage, the more all the variables in one table are linked to the second. These results highlight that the strongest link was observed between the datasets of the quantification of phenolic

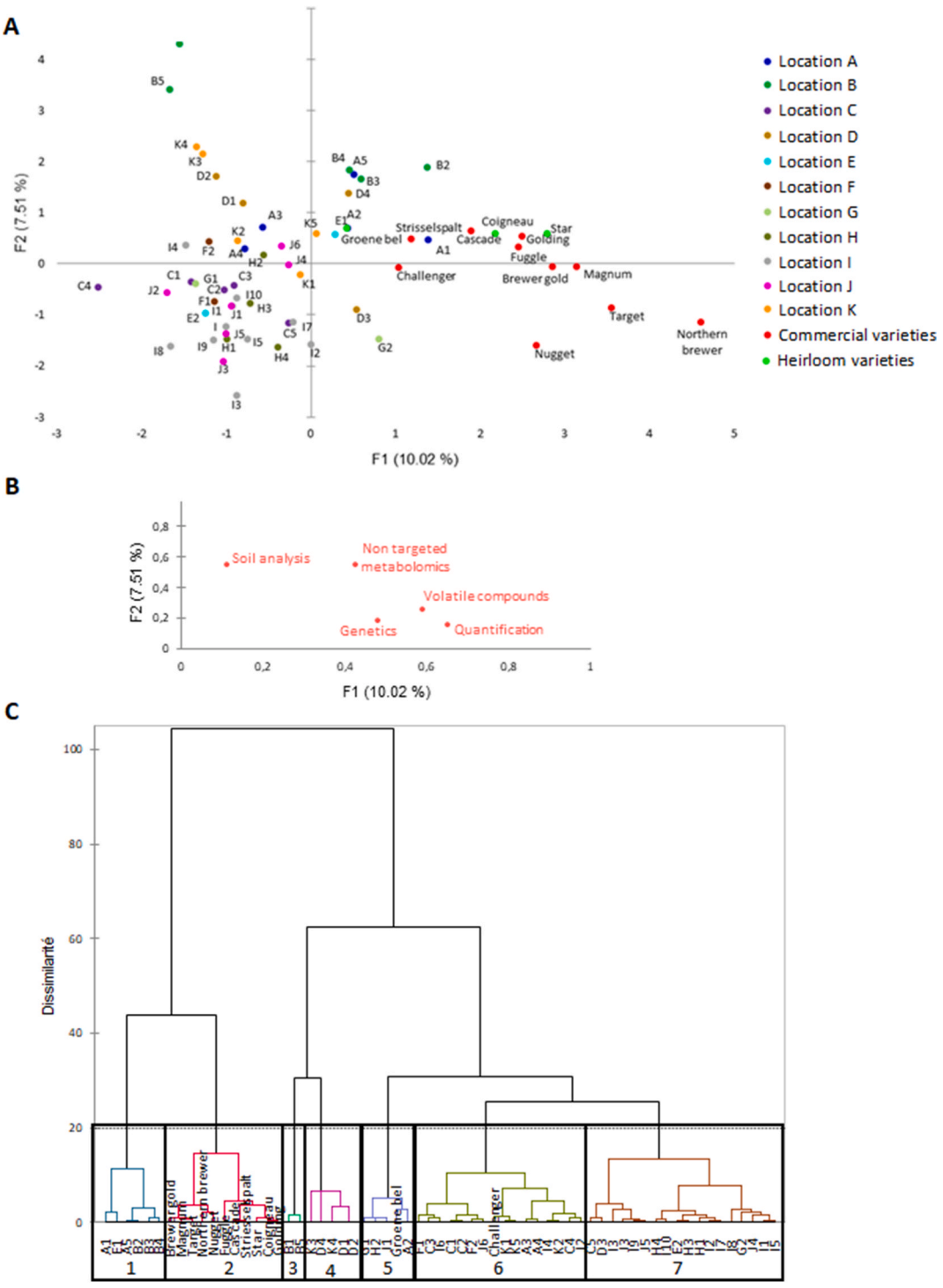


Fig. 9. Identification of the common structures of the variable blocks by multifactorial analysis. This multifactorial analysis has been done on the five datasets of this study: the soil analysis, the genetic characterization and the phytochemical analysis (volatile compounds analysis, quantitation of main prenylated phenolic compounds and untargeted metabolomic analysis) of the 63 hops samples (10 commercial varieties, 3 heirloom varieties and 50 wild hops). **A.** Multiple factor analysis scores plot **B.** Table coordinate plot **C.** Hierarchical clustering Ward's method (distance scale) (N = 7).

compounds and that of untargeted metabolomics ($RV = 0.382$, $p = 0.95$). By contrast the lower linkage was underlined between the datasets of the soil characterization and the quantification ($RV = 0.047$, $p = 0.95$). Untargeted metabolomics was the closest dataset of the global configuration ($RV = 0.76$, $p = 0.95$), while soil characterization was the most distant dataset ($RV = 0.454$, $p = 0.95$). It could have been interesting to modulate the weighting of the different tables used to process this AFM in order to ponder each dataset. Actually, in our study, all the datasets were considered with the same weighting, but we can suggest that the weighting of the genetic characterization could be stronger than those of the phytochemical characterization. This last parameter can effectively contain different analytical and technical biases. Furthermore, the biotope characterization was only partial; we could have taken into account the meteorological factors or the analysis of the microfauna of the soil. On this point, meteorological data (like sunshine, rainfall, average temperatures …) during the cone-ripening period could

have constituted a supplementary data set to integrate into the MFA in order to better characterize the biotopes. However, this dataset could not be constituted because sample locations of wild hops were close, and it was not possible to get one weather record per location.

3. Conclusions

This study constitutes the first identification and characterization of wild hop diversity in Northern France using molecular methods and phytochemical tools. Different datasets were acquired and revealed the great chemical and genetic diversity of the 50 collected wild hops. Genetic statistical analysis and Bayesian methods highlighted a clear distinction between the wild hop populations collected in Northern France and hop commercial varieties. These data suggest that the accessions identified as wild are not remnant hops. Moreover, the diversity among wild hops seems to be more related to the terroir than to their genotype. Nonetheless, the underlying philosophy of terroir encompasses the interaction of the soil, micro, and mesoclimate with the genetics of a variety, combining to create unique organoleptic quality. Since our wild accessions are not genetically identical, multivariate analysis does not allow us to determine which factors (genetic vs. environmental) are at play to explain the chemical diversity of our collection. Furthermore, data from several growing seasons are generally necessary to be able to make robust conclusions about terroir effects. Hence, these first *in-situ* investigations have to be consolidated with *ex-situ* analysis in order to disregard the observed potential terroir effect. This will be also the opportunity to carry out crop monitoring and to evaluate the agronomic performances of these accessions in order to select the best genotypes to be utilized as a base for breeding programs.

4. Materials and methods

4.1. Hop genotypes and sampling

4.1.1. Materials collection

A collection of 63 female *Humulus lupulus* L. (Cannabaceae) (hop) accessions was obtained, divided as follows: (1) *10 commercial varieties.* They were selected among the ten most relevant varieties, including so-called “top” hops used by craft brewers, such as Brewers Gold, Cascade, Challenger, Fuggle, Goldings, Magnum, Northern Brewer, Nugget, Strisselspalt and Target. Commercial hops were provided by the Northern France Hops Cooperative, the CoopHounord. (2) *3 heirloom varieties*. Three genebank clones representing former Flanders cultivars were included: Groene Bell, Coigneau, and Star, from Belgium (J. Cambié, Poperinge). (3) *50 wild hops*. These wild hops were collected by Drs C. Rivière and G. Lefèvre on various sites in Northern France (Hauts-de-France region) in collaboration with several natural site managers (department of Nord, EDEN 62, regional natural parks) or on private lands in accordance with their owners (Table 1). Wild hops were collected according to ecological or ethnobotanical criteria. A location gathered several sample sites with the same ecological characteristic (Table 1). In Table 1, the GPS coordinates of the barycenter of each location are defined as the gravity center of the sample points of this location, and calculated as the mean of the different sample site coordinates of the same location.

These samples were collected in accordance with the rules of the Nagoya Protocol and the French biodiversity law of 2017 (decision of June 9, 2020 issued by the Ministry of Ecological and Inclusive Transition; NOR: TREL2002508 S/284). Voucher specimens were deposited in the herbarium of the Faculty of Pharmacy (University of Lille). Global Positioning System (GPS) coordinates were used on the QGIS software to generate a map of these wild hop collection sites (Fig. 1). Cones were sampled during the month of September 2019 at their maturity and used for phytochemical analysis. Rhizomes of these wild accessions were collected and replanted into an experimental hop collection at the Douai-Wagnonville High School (France) to have a conservatory for these genetic resources. Samples from younger leaves and stems were used for genetic and plant biotechnology studies the following year in *ex-situ* conditions. Finally, soil samples were taken close to the root environment of the hop accessions in the A horizon, and then air-dried and sieved to 250 μm. Different types of analyses were performed on these samples to characterize the original biotope of hops studied.

4.1.2. Vitroplants

As the genotyping analyses were more efficient on young leaves whereas the collection had taken place in September in order to obtain mature cones, the majority of the DNA extractions were performed on the leaves once rhizome fragments had been transplanted to our experimental hop field at the Douai-Wagnonville High School (France). Nevertheless, some accessions did not recover from this transplant and in these cases, we used leaves from our collection of vitroplants. *In vitro* cultures of hop were established from plants from the experimental hop collection. Herbaceous stems of hop were surface sterilized with 80 $g.L^{-1}$ calcium hypochloride (60% active chlorine) for 10 min, followed by three washes with sterile water, then stems were cut into nodal fragments and placed in 160 × 24 mm glass tubes with transparent plastic covers, containing a half-strength medium (MS medium) (Murashige and Skoog 1962) supplemented with 100 $mg.L^{-1}$ myo-inositol, 30 $g.L^{-1}$ glucose and 6 $g.L^{-1}$ agar, pH 5.6. Nodal fragments develop to give vitroplants which are maintained by regular transfers (every 2 months) of stem cuttings (nodal segments) in fresh media. The vitroplant cultures were grown at 22 °C under a 16 h light photoperiod provided by cool-white fluorescent lamps (40 $\mu mol.m^{-2}.s^{-1}$).

4.1.3. Characterization of the biotopes and soil analysis

(1) Measure of dry weight and organic matter content. Approximately 3 g of sieved soil sample, weighed with a Denver Instrument balance scale were taken in a tared crucible and placed in an oven at 105 °C for 16 h. A weighing was carried out, to measure the moisture content of the samples. Calcination of the soil sample at high temperature (>400 °C) resulted in the calcination of all the organic matter that it contained. The difference in mass between this remaining mineral fraction and that of the initial sample is the loss on ignition (LOI). It can be equated to the amount of organic matter contained in the sample when the sample is low in carbonate (Heiri et al., 2001). The LOI protocol was then carried out as follows: (i) 1 h of temperature rise to 650 °C, (ii) 3 h at 650 °C, (iii) once cooled, crucibles were weighed, and organic matter content was then calculated by comparison with the dry weight. *(2) Measure of* pH. A volume of 6 mL of soil was placed in lidded jars with 30 mL of distilled water, shaken for 1 h at 750 rpm with a magnet bar on the shaking table and then left to stand for 2 h. The pH was measured directly with a calibrated pH meter (pHmeter Knick™, Berlin, Germany). *(3) Measure of conductivity*. The electrical conductivity of the soil was measured by mixing 4 g of soil with 20 mL of bidistilled water in Falcon tubes. The mixture was shaken for 2 h on a rotary shaker at 180 rpm, before being centrifuged at 1000 G for 10 min. Conductivity was measured with a calibrated pH/conductometer (Metrohm, Herisau Switzerland). *(4) Sample mineralization and dosage of Ca, Na, K and Mg.* 300 mg of soil sample were added to a mix containing 4.5 mL of 37% HCl and 1.5 mL of 70% HNO_3. The resulting suspension was heated at 95 °C for 90 min in a digester (HotBlock™ Environmental Express®SC100, Charleston, USA). Ultra-pure water was then added to reach a total volume of 25 mL. The solution was then filtered over an acetate Millipore membrane (0.45 μm porosity). Before analysis, 1 mL of mineralized soil solution was diluted 10 times in lanthanum chloride ($LaCl_3$). The concentrations of calcium (Ca), sodium (Na), potassium (K) and magnesium (Mg) elements were measured by flame atomic absorption spectrometry (FAAS, AA-6800 Shimadzu, Tokyo, Japan) following recommendations described in the literature (Waterlot and Hechelski, 2019). The calibration curves were established for each mineral to be dosed in the linearity range of the apparatus to avoid potential spectral interferences. Details on the characteristics of the light source and the method of quantification were

given in Waterlot and Hechelski (2019). A lanthanum chloride blank and technical blank were also injected. Data obtained from this assay, expressed as $mg.L^{-1}$, were first converted into $mg.kg^{-1}$, based on the mass of soil analyzed. Data were then expressed as $mg.kg^{-1}$ dry weight based on the moisture content of the sample, measured previously. Data of soil analysis were summarized into an Excel file and analyzed with a heat map and an HAC (N = 8) generated with XLSTAT® software version 2020- 3.1. HAC combined similar individuals or variables into clusters and arranged these clusters into a hierarchy.

4.2. DNA extraction and microsatellite-based DNA fingerprinting

4.2.1. Plant material and DNA isolation

DNA was extracted from young leaves collected in early spring for the 63 genotypes studied. These samples were collected directly from the hop growers for the 10 commercial varieties (cooperative Coophounord) and the 3 heirloom varieties (farm of Joris Cambié). For wild hops, samples were taken directly in the Douai high school experimental hop field or on vitroplants. The leaves, after grinding in liquid nitrogen, were stored at - 80 °C until use. DNA extraction was performed following the NucleoSpin Plant II protocol mini kit for DNA from plants (Macherey-Nagel, Germany). The concentration and the quality of each DNA sample were assessed by NanoDrop measure (NanoDrop One, ThermoScientific, Wilmington, USA) and 1% agarose gel electrophoresis.

4.2.2. Microsatellite assay

All isolates were fingerprinted using eleven microsatellite markers: Hl-ACA3, Hl-GA27, Hl-AGA1, Hl-GT14, Hl-GT16, Hl-AGA35, Hl-GA36 (Stajner et al., 2005) and GT1-K1-4, GA4-P11–9, GA8-K15–4 and GA7-I6-16 (Jakse et al., 2008) (Supplementary material Table S6). These markers were chosen due to their high allelic polymorphism during the preliminary experiments. PCR reactions were performed in a 50 µL reaction volume containing 8.0 µL of 25 mM MgCl2, 5.0 µL of 10X buffer B, 4.0 µL of 2.5 mM dNTP (Thermo Fischer Scientific, USA), 1 µL of each primer (20 µM) (Thermo Fischer Scientific, USA), 1 U of Taq DNA polymerase (Thermo Fischer Scientific, USA) and 5 µL of hop DNA (1 $ng.\mu L^{-1}$ final DNA concentration). The amplifications were carried out on a thermal cycler (C1000 Touch, Bio-Rad, Hercules, CA) under the following PCR cycling conditions: initial denaturation at 94 °C for 10 min; followed by 40 cycles of 94 °C for 1 min, 30 s at the annealing temperature (depending on the primer pair used), and 72 °C for 1 min; with a final extension at 72 °C for 10 min. After the final step, samples were stored at 4 °C. The separation was performed using capillary array electrophoresis in multiplex after pooling the amplicons into five pools of two amplicons per pool and one pool of one amplicon (Supplementary material Table S6). Within each pool, the primer pair used to amplify each marker was labeled with a fluorescent dye (Applied Biosystem, CA), with either blue (6-FAM) or green (HEX) fluorochrome, for better resolution of the capillary system. Mixtures were performed by using 1 µL of each pool-PCR product mixed with 10 µL of formamide (Sigma-Aldrich, France) and 0.15 µL of GeneScan 500 LIZ Size standard (Applied Biosystem, CA). After a denaturation step at 94 °C for 10 min, the pooled-PCR products were analyzed on a 3130 xl Genetic analyzer DNA sequencer (Applied Biosystem, CA). Fluorescent peaks were standardized using the Genescan Analysis software v. 3.7.1 Genotyper (Applied Biosystem, CA). The conversion into alleles and the determination of allele size were performed using the software Genotyper (Applied Biosystem, CA). Alleles scored with each microsatellite marker were summarized in an Excel sheet before undergoing statistical analysis.

4.2.3. Statistical analysis

Indices per locus, such as the number of alleles (N_a), expected (H_e) and observed (H_o) heterozygosity and the inbreeding coefficient (F_{is}) were calculated utilizing the GenAlEx v. 6.501 (Peakall and Smouse, 2012). Genetic structure was tested according to different approaches. (1) First, the magnitude of genetic differentiation among sub-collections, according to each locus, was assessed based on several fixation indices: a) Wright's F index (F_{ST}) implemented in GenAlEx version 6.501, b) Nei's G_{ST} index, a multiallelic analog of FST that is suitable for haploids as well as diploids (Nei, 1973), implemented in POPGENE version 1.32.3 (Yeh et al., 2000)) new standardized estimators of genetic structure, including Hedrick's standardized fixation index G''_{ST} (Meirmans and Hedrick, 2011) and Jost's standardized fixation index Jost's Dest (Jost, 2008), implemented in GenAlEx version 6.501. These different indices were calculated because the literature mentions and discusses the utility of these different statistic tools for measuring genetic differentiation between or among populations (Jost, 2008). (2) In addition, the program GenAlEx version 6.501 (Peakall and Smouse, 2012) was run to calculate F_{ST} and Φ_{ST} indices to highlight the amount of population differentiation among pairs of sub-collections (locations). Sources of genetic structure in each collection were investigated using hierarchical analysis of molecular variance (AMOVA) implemented in GenAlEx version 6.501, by the estimation of the degree of genetic differentiation within and among sub-collections using 1000 permutations. This software also allowed the distance matrix between the 63 accessions to be obtained, then processed by PCoA. (3) Hop population structure was also analyzed using *STRUCTURE* version 2.3.4, a model-based clustering program that uses a Bayesian approach to estimate the number of clusters and to assign individuals probabilistically to populations (Pritchard et al., 2000; Falush et al., 2003). The analysis was performed without prior information on the sub-collections to which the individuals belonged. 10 replicated runs, obtained with 50,000 iterations of the Markov Chain Monte Carlo method (MCMC) used as 'burn-in' and then followed by 500,000 MCMC iterations were produced to compare the log probability Pr(K), of each model, where K is the number of clusters. The data were analyzed with K ranging from 1 to 20, with 10 repeat runs for each K. Values were computed and plotted by the software Structure Harvester (Earl and vonHoldt, 2012). The best estimate of K was based on ΔK values calculated from ln P(D), as described in Evanno et al. (2005). (4) Furthermore, a dendrogram estimating genetic clustering of the haplotypes was produced using the unweighted neighbor-joining method based on the dissimilarity matrix (10,000 bootstraps), as implemented in the DARwin version 6.0.014 software (Perrier and Jacqemoud-Collet, 2006). (5) In addition, directional relative migration on network among the 11 locations and the 13 varieties of the region scale was determined using the divMigrate-online software using N_m as a measure of genetic differentiation (Sundqvist et al., 2016). This approach provides network plots, which facilitates the viewing of patterns of directional relative migration between locations.

4.3. Phytochemical characterization

All the hop cones used for this phytochemical characterization come from the collection gathered in September 2019. Cones were dried in an oven at 35 °C, ground and then stored at −20 °C until experimentation.

4.3.1. Volatile compound analysis quantification with HS-SPME GC-MS

Volatile compounds analyses were performed using HS-SPME GC-MS. The SPME fiber divinylbenzene/carboxen/polydimethylsiloxane (DVB, CAR, PDMS, 50/30 µm) (Supelco, Darmstadt, Germany) was selected because it was the most suitable to analyze terpenes. Prior to use, the fiber was conditioned 10 min at injection temperature according to the recommendations of the manufacturer. Extractions were performed on 1 g of hop cone powder placed in a 5 mL glass vial. For each sample, the incubation was initiated with a temperature of 45 °C for 5 min, then the extraction lasted 30 min at the same temperature. The analyses were performed on an Agilent 7890 A gas chromatograph coupled to an Agilent 5975C mass spectrometer (Agilent Technologies, Santa Clara, CA, USA) equipped with an MPS auto-sampler. Then, injections were performed in splitless mode at 280 °C. An HP-5 MS

capillary column (30 m × 250 μm × 0.25 μm, Agilent Technologies, Santa Clara, CA, USA) was used, with helium employed as the carrier gas at a flow rate of 1.2 mL.min^{-1}. The separation conditions were as follows: initial column temperature of 40 °C for 2 min; then it was increased by 4 °C.min^{-1} to 200 °C; then increased by 20 °C.min^{-1} to 300 °C, where it was maintained for 5 min. The mass spectrometer was set to have a temperature of the ion source at 230 °C and was programmed with SCAN acquisition mode. The mass spectra were acquired within a range of 35–500 Da m/z (mass-to-charge ratio). All main peaks corresponding to volatile compounds were integrated using the software Agilent MassHunter Unknowns Analysis software and general volatile compound profiles were established through a chromatographic deconvolution process (Fig. 2). Putative identification of volatile compounds was performed by comparing their mass spectra to those of commercial databases National Institute of Standards and Technology (NIST17) and Wiley 7. Additionally, experimental retention index (RI) of the compounds was calculated following the injection of a mixture of n-alkanes C_8–C_{20} (Sigma Aldrich, Darmstadt, Germany) and compared with those of compounds referenced in the NIST online database (https://webbook.nist.gov/chemistry/cas-ser.html, accessed on November 5, 2021). Data were analyzed in Excel to provide a heatmap and an HAC generated by a general analysis on XLSTAT® software version 2020- 3.1.

4.3.2. Non-volatile compounds analysis

4.3.2.1. Metabolite extraction and sample preparation for UHPLC-DAD-MS and UHPLC-hrms-qtof analysis. Hydro-ethanolic extracts of the 63 studied hops were performed in triplicate using 50 mg of dried cone powder with 1 mL of ethanol-water mixture (9:1, v/v). A 1-h maceration session was carried out in an ultrasonic bath. Afterwards, tubes were centrifuged at 4000 rpm at 20 °C for 5 min. Supernatant was then transferred into a tared tube. The exhausted matrix was then re-extracted two further times following the same protocol, and the supernatants were pooled with the first extract. At the end of the three maceration sessions, the solvent extraction was then evaporated in a Genevac™ centrifugal concentrator to obtain a crude hydro-ethanolic extract.

4.3.2.2. UHPLC-UV-MS analysis and quantification of main phenolic compounds. Quantification of xanthohumol, *co-, n-* and *ad*-humulone and *co-, n- and ad*-lupulone in each hydro-ethanolic extract was performed on an Acquity UPLC®H-Class Waters® system (Waters, Guyancourt, France) coupled to a Diode Array Detector (DAD) and a QDan ESI-Quadrupole Mass Spectrometer (Fig. 1). Separation was achieved using a Waters® Acquity BEH C18 column (pore size 300 Å, particle size 1.7 μm, 2.1 × 150mm, Waters, Milford MA) connected to a 0.2 μm in-line filter. Solvent A (water with 0.1% formic acid (*v/v*)) and solvent B (acetonitrile with 0.1% formic acid (*v/v*)) were used as the mobile phase. Compounds were eluted using the following chromatographic conditions: the flow rate was 0.3 mL.min^{-1}; the column temperature was set at 40 °C; the injection volume was 2 μL. The gradient elution (total analysis time of 13 min) was performed using eluent A and eluent B: initial condition in 0–1 min isocratic mode at 50% B, 1–3 min linear gradient to reach 75% B, 3–5 min isocratic mode at 75% B, 5–7 min linear gradient to reach 100% B, 7–9 min isocratic mode at 100% B, 9–13 min return to 50% B to rebalance the column. Chromatograms were acquired at 330 and 370 nm. Parameters for mass spectrometry were as follows: negative ionization mode, cone voltage at 10 V, capillary voltage at 0.8 kV, probe temperature at 600 °C, MS-Scan mode from 100 to 1000 Da. Xanthohumol and acylphloroglucinol derivatives were quantified in crude extracts according to the International Conference for Harmonisation (ICH) of Technical Requirements for Pharmaceuticals for Human Use guideline Q2-R1 (ICH, 2005). Quantification was carried out in UV at 370 nm for xanthohumol and 330 nm for acylphloroglucinols. The quantitation method was set up using standards purified in the laboratory according to the protocol detailed by Bocquet et al. (2019). Stock solutions of xanthohumol, *co-* and *n*-humulone and *co-* and *n*-lupulone were elaborated at a concentration of 1 mg.mL^{-1} in methanol in triplicate and stored at −20 °C prior to use for quantification. Fifteen working solutions (from 100 μg.mL^{-1} to 2.5 ng.mL^{-1}) were then prepared from these stock solutions. Calibration curves were established by plotting peak area (y) as a function of the nominal concentration for each calibration level (x) and then fitted by weighted (1/x) least square linear regression. Linearity and sensitivity of the method were determined and are reported in the Supplementary material section (Table S18). For these reference compounds, the limit of detection (LOD) and the limit of quantification (LOQ) were experimentally determined as the lowest concentration with signal-to-noise ratio of 3 and 10, respectively. The precision of the chromatographic system was tested by performing intra- and inter-day multiple injections of the hop extract and then checking the RSD % of retention times and peak areas. Five injections were performed each day for three consecutive days (n = 3, k = 3). Xanthohumol and α- and β-acids were identified based on the retention time of purified standards and their mass spectra and were quantified using the quantitation method previously set up on the Empower software (Bocquet et al., 2019). Data were exported from Empower to Excel. Data were converted from μg.ml^{-1} to % of dry weight using the total extract volume and the dry weight, and assessed by analysis of variance (ANOVA) considering the 63 hop accessions as factors and xanthohumol, α-acids, β-acids as dependent variables. Post-hoc pairwise testing among all accessions or between types was performed using Tukey's test (with a $P < 0.05$) on R using the Agricolae package (Mediburu, 2021). Data were then processed on a PCA, associated to an HAC on XLSTAT® software version 2020- 3.1. PCA is a technique used to reduce dimensionality of the data by finding linear combinations.

4.3.2.3. Untargeted metabolomics using UHPLC-hrms-qtof. Analyses for untargeted metabolomics were performed on the hydro-ethanolic crude extracts, diluted at 1 mg.mL^{-1} using methanol (Carlo Erba) and chromatographically separated on an ACQUITY UHPLC system (Waters, Guyancourt, France). The column was a Waters® Acquity BEH C18 column (2.1 × 50 mm, 1.7 μm) connected to a 0.2 μm in-line filter. The mobile phase was composed of water (A) and acetonitrile (B) both acidified with 0.1% of formic acid. Elution was performed at a flow rate of 300 mL.min^{-1} with the following gradient program (total analysis time: 10 min): initial condition in 0–0.2 min isocratic mode at 10% B, 0.2–7.5 min linear gradient to reach 100% B, 7.5–8 min isocratic mode at 100% B, 8.1–10 min linear gradient to rebalance the column at 10% B. The column temperature was set at 40 °C and 10 μL of hop extract was injected. The UHPLC eluent was then directly electrosprayed at the end of the column at a voltage of 3 kV, using a desolvation gas (N) at a flow of 600 L.h^{-1}, a nebulizer gas flow of 6.5 bar, and a temperature of 300 °C. The UHPLC was coupled to a SYNAPT-G2-Si mass spectrometer (Waters, Guyancourt, France) previously calibrated using a sodium formate solution. Mass spectrometry measurements were made in high resolution negative-mode using MassLynx software (version 4.1, Waters). Raw data from MassLynx was imported into Progenesis where a list of compound peaks (retention time, m/z molecular ion and peak intensity) contained in the extract was generated. A matrix was then created to obtain the samples per row and the different variables per column for statistical processing and was finally exported. As the samples were analyzed in triplicate, the average of each intensity for the three repetitions was considered. Ions with a maximum intensity lower than 1000 were eliminated. PCA were obtained using R software with the package FactoMineR (R Core Team, 2021).

4.4. Multivariate analysis of metabolites, genetic traits, and soil characteristics

Hop data of the different data sets including soil analysis, genetic characterization using the distance matrix used to construct PCoA and phytochemical analysis (volatile compounds analysis, quantitation of main prenylated compounds and non-targeted metabolomics) were compiled into an input file on an Excel spreadsheet. An MFA, as described by Escofier and Pagès (2008), followed by an HAC were performed to compare the different datasets obtained as mentioned before. Matrices of soil analysis, quantification of main phenolic compounds, untargeted metabolomics and genetic characterization were kept as quantitative datasets, while volatile compound results were considered a qualitative dataset. If variables are structured in different datasets with the same observation number, MFA may be considered to be an extension of the PCA for quantitative variables or an extension of the multiple correspondence analysis for qualitative variables. MFA allows the analysis and visualization of multiple data matrices as a whole. MFA was run using XLSTAT® software version 2020- 3.1 as a nonsupervised method to identify key traits with the largest effect on the overall variability and to evaluate the effect of genetic background on the chemical profiles among different accessions. HAC was performed with Ward's method, and were visualized as a dendrogram (N = 7).

Fundings

This work has been carried out in the framework of the FEADER project 2021–2024 (action 16.02.01 "assistance with pilot projects and the development of new products, practices, processes and technologies" n°35/2020) which is financed by the European Union and the Region of Hauts-de-France (France). The authors gratefully thank the Region of Hauts-de-France and the European Union (Feder), the University of Lille (I-Site) and the "Fondation de la brasserie et de la malterie françaises" for the funding of PhD fellow Anne-Sophie Paguet.

Author contributions

ASP: Investigation, Methodology, Formal analysis, Writing – original draft; AS: Methodology, Supervision (genetical part), Project administration, Formal analysis, Writing – review & editing; GL: Conceptualization, Harvests, Writing; SM, JS, TC, FM, BD: Methodology and Investigation (chemistry and genetics); CF: Responsible for mass spectrometry platform; CRa: Supervision (vitro-plants); AHDS, AED: Methodology (vitroplants); SC: Formal analysis (statistical analysis); RM, JXF: Formal analysis (untargeted metabolomics); MLF: Supervision (volatile compound analysis); CW: Supervision (soil characterization); SS: Project administration; CRi: Conceptualization, Harvests, Methodology, Validation, Supervision, Project administration, Funding acquisition, Investigation (NMR), Formal analysis, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors wish to thank the hop growers of CoopHounord (Berthen, France), in particular Yvon Pruvot, Paul Pruvot, Christophe Gombert, Nicolas Daneels, Riquier Thévenin, Julien Hennon, Denis Beck, Thierry Beck, as well as Joris Cambie (Poperinge, Belgium) for supplying commercial and heirloom varieties of hop. We also warmly thank the natural park managers for their help in identifying collection sites of wild hops. We thank in particular Hubert Brabant and Dominique Derout (Eden 62), Sébastien Mézière and François Charlet (Parc Naturel Régional des Caps et Marais d'Opale), Yann Dulondel and Gérald Duhayon (Parc Naturel Régional Scarpe-Escaut), Yvon Brunelle and Stéphane Marache (Parc Naturel Régional de l'Avesnois) and the Nord department. Our thanks also go to those who advised us or gave their authorization to collect on their private lands: Claude Delasalle and his family, Christophe Noyon, David Lutun, Norman Degardin, Riquier Thévenin. Our great thanks to David Lutun and Fabrice Henry at the Douai-Wagnonville high school for the preservation of our collection of wild hops and to the Hauts-de-France regional chamber of agriculture for their support in this project, in particular Anne-Hélène Martin and Aurélien Honoré. We also thank Richard Boucherie (CRRG) for some historical data on hop production in the North of France and Edouard Rousset, organic hop grower, for our discussions on hops. Many thanks also to Charles Dermont for the initiation of volatile compound analysis, Jean-Louis Hilbert for his great support, as well as Cassandra Harvey for English grammar corrections. Finally, we thank the BRGM (Bureau de Recherches Géologiques et Minières) for their kind permission to use the geological maps. Authors are also grateful to LARMN and to REALCAT platforms for access to equipment.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phytochem.2022.113508.

References

Bart-Haas Report, 1930. Hop Report for 1929-30. Nuremberg, Bavaria, July 18th 1930. https://www.barthhaas.com/fileadmin/user_upload/downloads/barth-berichte-broschueren/barth-berichte/englisch/1920-1930/barth-report-1929-1930.pdf.

Bart-Haas Group, 2021. In: Bart-Haas Report.

Bocquet, L., Sahpaz, S., Hilbert, J.L., Rambaud, C., Rivière, C., 2018a. *Humulus lupulus* L., a very popular beer ingredient and medicinal plant: overview of its phytochemistry, its bioactivity, and its biotechnology. Phytochemistry Rev. 17, 1047–1090. https://doi.org/10.1007/s11101-018-9584-y.

Bocquet, L., Sahpaz, S., Rivière, C., 2018b. An Overview of the Antimicrobial Properties of Hop. Springer International Publishing AG, Sustainable Development and Biodiversity Book Series, pp. 31–54.

Bocquet, L., Rivière, C., Dermont, C., Samaillie, J., Hilbert, J.-L., Halama, P., Siah, A., Sahpaz, S., 2018c. Antifungal activity of hop extracts and compounds against the wheat pathogen *Zymoseptoria tritici*. Ind. Crop. Prod. 122, 290–297. https://doi.org/10.1016/j.indcrop.2018.05.061.

Bocquet, L., Sahpaz, S., Bonneau, N., Beaufay, C., Mahieux, S., Samaillie, J., Roumy, V., Jacquin, J., Bordage, S., Hennebelle, T., Chai, F., Quetin-Leclercq, J., Neut, C., Rivière, C., 2019. Phenolic compounds from *Humulus lupulus* as natural antimicrobial products: new weapons in the fight against methicillin resistant *Staphylococcus aureus, Leishmania mexicana* and *Trypanosoma brucei* strains. Molecules 24 (6), E1024 pii.

Čerenak, A., Kolenc, Z., Sehur, P., Whittock, S.P., Koutoulis, A., Beatson, R., Buck, E., Javornik, B., Škof, S., Jakše, J., 2019. New male specific markers for hop and application in breeding program. Sci. Rep. 9, 14223 https://doi.org/10.1038/s41598-019-50400-z.

Dabbous-Wach, A., Rodolfi, M., Paolini, J., Costa, J., Ganino, T., 2021. Characterization of wild Corsican hops and assessment of the performances of German hops in Corsican environmental conditions through a multidisciplinary approach. Appl. Sci. 11, 3756. https://doi.org/10.3390/app11093756.

Damseaux, A., 1911. Plantes de la grande culture, Troisième édition. Amat, Namur, p. 491.

Darby, P., 2005. Brewery History [WWW Document]. Brewery history. URL. http://www.breweryhistory.com/journal/archive/121/bh-121-094.htm. accessed 6.2.21.

Driskill, Mandie, Pardee, Katie, Hummer E, Kim, Zurn D, Jason, Amundsen, Keenan, Wiles, Annette, Wiedow, Claudia, Patzak, Josef, Henning A, John, Bassil V, Nahla, 2022. Two fingerprinting sets for *Humulus lupulus* based on KASP and microsatellite markers. PLoS ONE 17 (4). https://doi.org/10.1371/journal.pone.0257746.

Ducloux, A., Hédiard, L., Vallez, T., 1910. In: Baillière, J.B., éditeurs, fils (Eds.), Les productions agricoles du Nord de la France (Paris, France).

Earl, D.A., vonHoldt, B.M., 2012. Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4, 359–361. https://doi.org/10.1007/s12686-011-9548-7.

Escofier, B., Pagès, J., 2008. Analyses factorielles simples et multiples. In: Objectifs méthodes et interprétation, Sciences Sup. Dunod.

Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies - PubMed. Genetics 164, 1567–1587.

Heiri, O., Lotter, A.F., Lemcke, G., 2001. Loss on ignition as a method for estimating organic and carbonate content in sediments: reproducibility and comparability of results. J. Paleolimnol. 25, 101–110. https://doi.org/10.1023/A:1008119611481.

Hong, K., Xu, Z., Wang, L., Johnpaul, A., Cheng, Y., Lv, C., Ma, C., 2022. Varietal differences in the phytochemical components' accumulation and aroma profile of three *Humulus lupulus* cultivars. Food Control 132, 108499. https://doi.org/10.1016/j.foodcont.2021.108499.

ICH, 2005. International conference on harmonisation of technical Requirements for registration of pharmaceuticals for human use. In: Handbook of Transnational Economic Governance Regimes. Brill, Nijhoff, pp. 1041–1053. https://doi.org/10.1163/ej.9789004163300.i-1081.897.

Jakše, J., Bandelj, D., Javornik, B., 2002. Eleven new microsatellites for hop (*Humulus lupulus* L.). Mol. Ecol. Notes 2, 544–546. https://doi.org/10.1046/j.1471-8286.2002.00309.x.

Jakse, J., Luthar, Z., Javornik, B., 2008. New polymorphic dinucleotide and trinucleotide microsatellite loci for hop *Humulus lupulus* L. Mol Ecol Resour 8, 769–772. https://doi.org/10.1111/j.1755-0998.2007.02053.x.

Jiang, P., Song, T., Jiang, W., Wang, D., Pu, B., Luan, C., 2018. SNP-based kompetitive allele specific PCR (KASPTM) method for the qualification and quantification of hop varieties. J. Am. Soc. Brew. Chem. 76 (3), 185–189. https://doi.org/10.1080/03610470.2018.1483702.

Jost, L., 2008. GST and its relatives do not measure differentiation. Mol. Ecol. 17, 4015–4026. https://doi.org/10.1111/j.1365-294X.2008.03887.x.

Karlsson Strese, E.M., Lundström, M., Hagenblad, J., Leino, M., 2014. Genetic diversity in remnant Swedish hop (*Humulus lupulus* L.) yards from the 15th to 18th century. Econ. Bot. 68, 231–245. https://doi.org/10.1007/s12231-014-9273-8.

Kishimoto, T., Teramoto, S., Fujita, A., Yamada, O., 2021. Evaluation of components contributing to the international bitterness unit of wort and beer. J. Am. Soc. Brew. Chem. https://doi.org/10.1080/03610470.2021.1878684.

Korpelainen, H., Pietiläinen, M., 2021. Hop (*Humulus lupulus* L.): traditional and present use, and future potential. Econ. Bot. 75 (3), 302–322. https://doi.org/10.1007/s12231-021-09528-1.

Machado, J., Faria, M., Barata, A.M., da Silva, I.G., Cerenak, A., Ferreira, I., 2022. Portuguese wild hop diversity assessment by fast SNP genotyping using high-resolution melting. J. Appl. Genet. 63 (1), 104–114. https://doi.org/10.1007/s13353-021-00668-7.

Mafakheri, M., Kordrostami, M., Rahimi, M., Matthews, P.D., 2020. Evaluating genetic diversity and structure of a wild hop (*Humulus lupulus* L.) germplasm using morphological and molecular characteristics. Euphytica 216, 58. https://doi.org/10.1007/s10681-020-02592-z.

Martins, Z.E., Machado, J.C., Cunha, S.C., Barata, A.M., Ferreira, I.M., 2020. A chemometric approach to compare Portuguese native hops with worldwide commercial varieties. J. Chemom. https://doi.org/10.1002/cem.3285.

McCallum, J.L., Nabuurs, M.H., Gallant, S.T., Kirby, C.W., Mills, A.A.S., 2019. Phytochemical characterization of wild hops (*Humulus lupulus* ssp. *lupuloides*) germplasm resources from the maritimes region of Canada. Front. Plant Sci. 10 https://doi.org/10.3389/fpls.2019.01438.

Mediburu, F. de, 2021. Agricolae: Statistical Procedures for Agricultural Research. R Package Version 1, pp. 3–5.

Meirmans, P.G., Hedrick, P.W., 2011. Assessing population structure: F(ST) and related measures. Mol. Ecol. Resour. 11, 5–18. https://doi.org/10.1111/j.1755-0998.2010.02927.x.

Mongelli, A., Rodolfi, M., Ganino, T., Marieschi, M., Dall'Asta, C., Bruni, R., 2016. Italian hop germplasm: characterization of wild *Humulus lupulus* L. genotypes from Northern Italy by means of phytochemical, morphological traits and multivariate data analysis. Ind. Crop. Prod. 70, 16–27. https://doi.org/10.1016/j.indcrop.2015.02.036.

Moore, B., Andrew, R., Külheim, C., Foley, W., 2014. Explaining intraspecific diversity in plant secondary metabolites in an ecological context. New Phytol. 201, 733–750. https://doi/10.1111/nph.12526.

Morcol, T.B., Negrin, A., Matthews, P.D., Kennelly, E.J., 2020. Hop (*Humulus lupulus* L.) terroir has large effect on a glycosylated green leaf volatile but not on other aroma glycosides. Food Chem. 321, 126644 https://doi.org/10.1016/j.foodchem.2020.126644.

Murakami, A., Darby, P., Javornik, B., Pais, M.S.S., Seigner, E., Lutz, A., Svoboda, P., 2006. Molecular phylogeny of wild Hops, *Humulus lupulus* L. Heredity 97, 66–74. https://doi.org/10.1038/sj.hdy.6800839.

Murashige, T., Skoog, F., 1962. A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol. Plantarum 473–497. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x.

Naraine, S., Small, E., 2017. Germplasm sources of protective glandular leaf trichomes of hop (*Humulus lupulus*). Genet. Resour. Crop Evol. 64, 1491–1497. https://doi.org/10.1007/s10722-017-0540-2.

Nei, M., 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U.S.A. 70, 3321–3323. https://doi.org/10.1073/pnas.70.12.3321.

Neve, R.A., 1991. Hops. Chapman and Hall.

Nybom, H., 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol. Ecol. 13, 1143–1155. https://doi.org/10.1111/j.1365-294X.2004.02141.x.

Paguet, A.S., Siah, A., Lefèvre, G., Sahpaz, S., Rivière, C., 2022. Agronomic, genetic and chemical tools for hop cultivation and breeding. Phytochemistry Rev. https://doi.org/10.1007/s11101-022-09813-4.

Palmer, J.J., 2006. How to Brew: Everything You Need to Know to Brew Beer Right for the First Time. Brewers Publications, Boulder, Colo.

Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28, 2537–2539. https://doi.org/10.1093/bioinformatics/bts460.

Perrier, X., Jacqemoud-Collet, J.P., 2006. DARwin Software.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959. https://doi.org/10.1093/genetics/155.2.945.

R Core Team, 2021. R: A Language and Environment for Statistical Computing [WWW Document]. R Foundation for Statistical Computing. Vienna, Austria. URL. https://www.R-project.org/.

Rodolfi, M., Silvanini, A., Chiancone, B., Marieschi, M., Fabbri, A., Bruni, R., Ganino, T., 2018. Identification and genetic structure of wild Italian *Humulus lupulus* L. and comparison with European and American hop cultivars using nuclear microsatellite markers. Genet. Resour. Crop Evol. 65, 1405–1422. https://doi.org/10.1007/s10722-018-0622-9.

Santagostini, L., Caporali, E., Giuliani, C., Bottoni, M., Ascrizzi, R., Araneo, S.R., Papini, A., Flamini, G., Fico, G., 2020. *Humulus lupulus* L. cv. Cascade grown in Northern Italy: morphological and phytochemical characterization. Plant Biosyst. 154, 316–325. https://doi.org/10.1080/11263504.2019.1610111.

Schönberger, C., Kostelecky, T., 2011. 125th anniversary review: the role of hops in brewing. J. Inst. Brew. 117, 259–267. https://doi.org/10.1002/j.2050-0416.2011.tb00471.x.

Small, E., 1978. A numerical and nomenclatural analysis of morpho-geographic taxa of Humulus. Syst. Bot. 3, 37. https://doi.org/10.2307/2418532.

Solberg, S.Ø., Brantestam, A.K., Kylin, M., Bjørn, G.K., Thomsen J, M.G., 2014. Genetic variation in Danish and Norwegian germplasm collections of hops. Biochem. Systemat. Ecol. 52, 53–59. https://doi.org/10.1016/j.bse.2013.12.014.

Srečec, S., Zechner-Krpan, V., Petravić-Tominac, V., Cerenak, A., Liber, Z., Šatović, Z., 2010. Phenotypic and alpha-acid content diversity of wild hop populations in Croatia. Plant Soil Environ. 37–42. https://doi.org/10.17221/130/2009-PSE.

Stajner, N., Jakse, J., Kozjak, P., Javornik, B., 2005. The isolation and characterisation of microsatellites in hop (*Humulus lupulus* L.). Plant Sci. 168, 213–221. https://doi.org/10.1016/j.plantsci.2004.07.031.

Statista, 2021. Craft Beer: Production Volume by Brewery Type U.S. 2020 [WWW Document]. Statista. URL. http://www.statista.com/statistics/732171/craft-beer-production-by-brewery/. accessed 5.22.21.

Sundqvist, L., Keenan, K., Zackrisson, M., Prodöhl, P., Kleinhans, D., 2016. Directional genetic differentiation and relative migration. Ecol. Evol. 6, 3461–3475. https://doi.org/10.1002/ece3.2096.

Taniguchi, Y., Matsukura, Y., Ozaki, H., Nishimura, K., Shindo, K., 2013. Identification and quantification of the oxidation products derived from α-acids and β-acids during storage of hops (*Humulus lupulus* L.). J. Agric. Food Chem. 61, 3121–3130. https://doi.org/10.1021/jf3047187.

The Chemical Sources Association, 2010. The Good Scents Company ([WWW Document]).

Tofana, M., Socaci, S., Socaciu, C., Mihaiescu, D., Semeniuc, C., Truta, D., 2009. Optimization of HS/GC-MS method for the determination of volatile compounds from some indigenous hop varieties. Bulletin UASVM Agriculture 66, 500–505.

Turner, S., Benedict, C., Darby, H., Hoagland, L., Simonson, P., Sirrine, R., Murphy, K., 2011. Challenges and Opportunities for Organic Hop Production in the United States.

Uemoto, M., Takoi, K., Tanigawa, A., Takazumi, K., Ogushi, K., Koie, K., Suda, N., 2022. Effect of "late harvest" of hops (*Humulus lupulus* L.) on the contents of volatile thiols in furano beauty, furano magical, and cascade varieties. J. Agric. Food Chem. 70, 607–614. https://doi.org/10.1021/acs.jafc.1c08210.

Van Holle, A., Muylle, H., Ruttink, T., Van Landschoot, A., Haesaert, G., Naudts, D., De Keukeleire, D., Roldán-Ruiz, I., 2019. Single nucleotide polymorphisms and biochemical markers as complementary tools to characterize hops (*Humulus lupulus* L.) in brewing practice. J. Agric. Food Chem. 67, 3761–3771. https://doi.org/10.1021/acs.jafc.9b00816.

Vázquez-Araújo, L., Rodríguez-Solana, R., Cortés-Diéguez, S.M., Domínguez, J.M., 2013. Use of hydrodistillation and headspace solid-phase microextraction to characterize the volatile composition of different hop cultivars. J. Sci. Food Agric. 93, 2568–2574. https://doi.org/10.1002/jsfa.6078.

Waterlot, C., Hechelski, M., 2019. Benefits of ryegrass on multicontaminated soils Part 1: effects of fertilizers on bioavailability and accumulation of metals. Sustain. Times 11, 5093. https://doi.org/10.3390/su11185093.

Yamauchi, H., Mukouzaka, Y., Taniguchi, T., Nakashima, K., Furukubo, S., Harada, M., 2014. Newly developed SNP-based identification method of hop varieties. J. Am. Soc. Brew. Chem. 72 (4), 239–245. https://doi.org/10.1094/ASBCJ-2014-1006-01.

Yeh, F.C., Yang, R., Boyle, T., Ye, Z.H., Mao, J., 2000. POPGENE, version 1.32: the user friendly software for population genetic analysis – ScienceOpen. In: Molecular Biology and Biotechnology Center, University of Alberta, and CIFOR (Canada).

Zhang, D., Easterling, K., Pitra, N., Coles, M., Buckler, E., Bass, H., Matthews, P., 2017. Non-Mendelian single-nucleotide polymorphism inheritance and atypical meiotic configurations are prevalent in hop. Plant Genome 10 (3). https://doi.org/10.3835/plantgenome2017.04.0032.