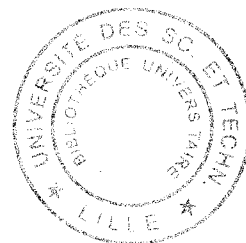




Université des Sciences et Technologies de Lille 1
U.F.R. de Biologie
Ecole doctorale "Sciences de la Matière, du Rayonnement et de l'Environnement"
Laboratoire Génétique et Evolution des Populations Végétales UMR8016

Thèse

Présentée pour l'obtention du grade de
Docteur de l'Université de Lille 1
Discipline: Génétique des Populations



Par Maxime Pauwels

Origine et Evolution de la Tolérance au Zinc chez *Arabidopsis halleri* (Brassicaceae)

Approches phénotypiques et génétiques

Directeur de thèse: Pierre Saumitou-Laprade

Soutenance publique prévue le 12 Mai 2006 devant le jury constitué de:

Pierre MEERTS, Professeur, Université Libre de Bruxelles
Rémy PETIT, Directeur de Recherche INRA, Bordeaux

Rapporteur
Rapporteur

Ophélie RONCE, Chargée de Recherche CNRS, Université Montpellier II
Joël CUGUEN, Professeur, Université de Lille 1
Marcus KOCH, Professor, Heidelberg University
Xavier VEKEMANS, Professeur, Université de Lille 1
Pierre SAUMITOU-LAPRADE, Chargé de Recherche CNRS, Université de Lille 1

Examinatrice
Examineur
Examineur
Examineur
Examineur

"Toute chose étant médiante et immédiate, causée et causante, je tiens impossible de connaître la partie si je ne connais le tout, ni de connaître le tout si je ne connaissais les parties."

Pascal, *Pensées*

"Il faut du temps pour inventer une nouvelle voie et se modifier soi-même. La vitesse ne sait pas qu'elle se répète. La lenteur le sait, elle en éprouve de la confusion et elle se montre circonspecte à l'égard de ce qu'elle croit avoir trouvée."

P. Sansot, *Du bon usage de la lenteur*

Remerciements

L'aventure de la thèse restera dans mes souvenirs comme une autre adolescence. Après que l'enfant est devenu (ou presque) adulte, faible maillon d'une étrange société, l'étudiant anonyme, siégeant aux derniers rangs de l'amphi, et qui passait un temps égal à écouter et noter les cours et à rêver le monde (aidé en cela par le toujours fidèle Lescureux, salut camarade!), voyait son nom s'inscrire sur une porte d'un laboratoire d'abord, sur quelques comptes-rendus de colloques ensuite, sous le titre d'articles enfin, pour finalement s'imaginer membre d'une troublante communauté.

L'aventure a commencé par hasard. Un appel à stagiaire comme il y en a mille à la fin d'un T.D. J'y ai répondu. Nous y avons répondu. Car tu étais déjà là, Stéf. Il n'y a depuis pas un pas que j'ai fait sans qu'il soit doublé par un des tiens. Nos parcours sont plus semblables qu'il n'y paraît. Qu'ils gardent encore longtemps leurs marches parallèles, quelles qu'en soient les formes.

L'appel était double, lancé par un couple improbable, deux chercheurs sans qui les pages qui suivent ne seraient sûrement pas. Pierre Saumitou-Laprade, d'abord. Mais trêve de solennité, c'est Pierre, Pierrot, PSL, Saumitou. La porte et les oreilles toujours ouvertes malgré les files d'attentes à leurs entrées. J'ai très tôt senti une très agréable paternité. Et le respect n'a jamais faibli. En bon adolescent, j'ai lentement déployé mes propres ailes. Mais le souvenir reste. Et restera. Isabelle Bonnin, ensuite. J'ai tremblé (et je tremble encore) avant chaque entretien que j'ai passé avec toi. Et j'ai pris un antalgique après la plupart. Mais, sans masochisme, avec toujours un vrai plaisir à discuter. Car j'ai beaucoup appris. Je te dois une belle part de ma rigueur scientifique. Avec elle sûrement, une grande part aussi de mon parcours à venir.

Ce fût donc la première entrée au GEPV (la première crise d'acné, en quelques sortes, ou le premier pic d'hormones, c'est selon). Qu'à travers son actuel directeur, le GEPV soit remercié pour l'ensemble des moyens dont j'ai pu disposer. A entendre certains des mes pairs, tout le monde n'a pas la chance d'avoir ce qu'il souhaite à la demande. Je veux m'adresser aussi directement à Joël Cuguen. Je t'aurai beaucoup observé diriger ce petit monde, avec un mélange difficile à doser d'autorité, de distance et d'écoute, et un souci régulier, je pense, de rendre disponible pour chacun (du stagiaire de maîtrise au MCF) les facilités du labo.

Pas de travaux sans équipes. Au sein des Métaux Lourds, deux techniciennes qui sauraient se chamailler l'ordre des noms dans ce paragraphe. Je serai chronologique. Anne-Cat' d'abord et mes premiers doigts dans la BM. A coups de "spayceurs" et de gels d'acryl en morceaux. Un travail efficace et dans la bonne humeur, indispensable pour la suite. Cécile ensuite. Le même âge et les expériences qui vont avec aidant, une grande proximité, un rapport au delà du travail. Je t'ai détestée autant qu'admiration, et tu as dû en faire de même. J'ai toujours pensé que nos rapports à fleur de peau étaient signes de profond respect. J'espère avoir toujours une aide technique aussi précise et précieuse que la tienne. Puis il y a Glenda, Glenneke si je peux, l'autre thésarde. Tu m'auras laissé comme un môme à qui on a piqué le sucre d'orge, moi qui aurais souhaité être le premier thésard de Saumitou. Pour enfoncer le clou, tu n'auras jamais partagé ton chocolat, si belge. Allez, sans rancune, parce que nous nous complétons bien, qu'il n'y aura pas eu un orage en trois ans et demi, et qu'avoir cette fraîcheur toute flamande à mes côtés aura été un vrai plaisir. Enfin, il y a Claire-Lise, "la petite" comme Ln et moi disons parfois derrière toi, si affectueusement. Ta place dans ce paragraphe est hautement discutable. Car nos liaisons, malgré des débuts éclatants dont nous farderons les ultimes leçons, sont tout

sauf métalliques. Mais ta place dans notre quotidien est imprenable, fixée par la multitude des atomes (délicatement) crochus qui se sont lentement et sereinement dévoilés, souvent bien loin des attendus.

J'ai une pensée pour tous ceux, éphémères ou permanents, avec qui j'ai partagé le bureau 108. Un chouette bureau. Travailleur sans être pénible: je m'y sens bien, point. Mention spéciale à Vincent. Un jeune MCF comme tout thésard devrait avoir à portée, en exemple réjouissant de ce qu'il pourrait devenir, ou pas. Je m'égare, mais pas loin de toi il y a Xavier. Vous formez une équipe attachante, un pont précieux et solide pour le grand saut dans la recherche scientifique.

Quelque part entre les équipes et les bureaux, symboliquement au milieu des couloirs, il y a vous deux, mes deux pieds sur terre, chaînons manquants entre la vraie vie et les étudiants de la vie. Merci Sandrine (Sandy) et merci Michèle (Salut Michèle, comment vas-tu?), j'ai pris de nombreuses bouffées d'air dans vos bureaux, entre deux plongées en apnée.

Et puis il y a tous les autres dont les noms n'apparaîtront pas. Beaucoup d'entre vous m'ont beaucoup aidé, souvent à leur insu. En formant un patchwork de profils très différents. J'ai pu, au fil des mois, picorer lentement vos commentaires et attitudes, à la découverte de ce que je croyais bon de faire... ou ne pas faire, être... ou ne pas être. Merci.

Aux rangs des "autres", je ne peux pas ne pas mentionner les amis, fugitifs ou pérennes, récents ou de toujours, proches ou lointains, universitaires ou non. Vous êtes l'immense arrière-plan de ce travail, avec lequel vous faites deux vases communicants que je souhaite pouvoir équilibrer.

J'ai une pensée toute particulière pour ma famille, en commençant par mes parents et frangins. Il y a longtemps je pense que vous avez renoncé à comprendre quelle mouche avait piqué cet étudiant-peter-pan dont on ne sait jamais ce qu'il fait, où il est ni même comment il va. L'incompréhension n'a jamais entamé votre admiration. Je le sais, même si les chemins de traverse de son expression me sont aujourd'hui parfois inaccessibles. S'affranchir de ses racines par les études est un parcours douloureux. Mais je suis fier aujourd'hui d'apporter cette branche mutante à notre modeste ramage, qui n'aurait pu forcer sans notre tronc commun, et que j'espère florissante.

Enfin, j'ai gagné une autre famille durant ma thèse. Elle porte sa sympathie en son nom. Vous, les Frérot, m'avaient accueilli à bras ouverts, m'offrant un cocon que je n'aurais pas rêvé. En votre sein, une étoile danseuse, riieuse et séminante. Grâce à toi, j'écris ces lettres bleues sans ridicule, je pleure sans honte, je ris sans gêne, je ferais tout, et son contraire. Le reste est anecdote. Merci.

Sommaire

AVANT PROPOS	1
INTRODUCTION	4
I LA TOLERANCE AUX METAUX LOURDS DES VEGETAUX SUPERIEURS	4
I.1 Définitions	4
I.1.1 Métal, lifère, lophyte, licole : définitions.....	4
I.1.2 La tolérance aux métaux.....	6
I.1.2.1 Deux définitions relatives.....	6
I.1.2.2 La tolérance dans les bacs: mesure et définition de la tolérance en milieu contrôlé.....	8
I.2 Mécanismes physiologiques de la tolérance aux métaux	10
I.2.1 Des mécanismes spécifiques	10
I.2.2 Les grandes stratégies physiologiques de tolérance aux métaux	11
I.2.2.1 L'exclusion	11
I.2.2.2 L'accumulation et l'hyperaccumulation	12
I.2.3 Les processus moléculaires de la tolérance aux métaux	13
I.3 Architecture génétique de la tolérance aux métaux	18
I.3.1 Un déterminisme simple.....	18
I.3.1.1 Un ou quelques gènes à effets majeurs.....	18
I.3.1.2 Des gènes modificateurs à effets mineurs	19
I.3.2 Les interactions géniques: dominance et épistasie.....	20
I.4 Evolution de la tolérance aux métaux lourds	21
I.4.1 Tolérer ou ne pas tolérer: les hasards de l'adaptation.....	22
I.4.1.1 Tolérance et genostasis.....	22
I.4.1.2 Tolérance et phylogénie	23
I.4.1.3 Exclure ou accumuler?	24
I.4.2 Dynamique évolutive de la tolérance aux métaux	25
I.4.2.1 Une adaptation rapide.....	25
I.4.2.2 ...avec (ou sans) bottleneck... ..	26
I.4.2.3 ... peu coûteuse.....	27
I.4.2.4 ...et répétée.	28
I.4.3 Tolérance et isolement reproducteur.....	29
I.4.3.1 Tolérance et isolement reproducteur entre populations	29
I.4.3.2 Tolérance et spéciation écologique	31
I.4.3.3 Isolement reproducteur et distribution de la tolérance.....	32
I.4.4 La tolérance aux métaux: une adaptation locale aux milieux métallifères?	33
II TOLERANCE AUX METAUX ET PHYLOGEOGRAPHIE	35
I.1 La boîte à outils de la phylogéographie	37
II.1.1 Le concept de structure phylogéographique.....	37
II.1.2 La mesure de la structure phylogéographique	38
II.1.3 Le génome chloroplastique	38
II.2 Résultats empiriques en Europe	39

II.2.1	Le contexte paléo-environnemental et les données fossiles	39
II.2.2	La confirmation phylogéographique	40
II.2.3	L'extension aux espèces herbacées.....	42
III	LE MATERIEL BIOLOGIQUE ET LES OBJECTIFS DE L'ETUDE	43
III.1	<i>Arabidopsis halleri</i> (L.): généralités.....	43
III.1.1	Ecologie et distribution.....	44
III.1.2	Position phylogénétique.....	44
III.2	Tolérance aux métaux lourds d' <i>A. halleri</i>	46
III.2.1	Les connaissances actuelles	46
III.2.2	Le statut de modèle.....	46
IV	REFERENCES	48
CHAPITRE 1: POPULATION DIFFERENTIATION FOR ZN-TOLERANCE		61
CHAPITRE 2: ORIGIN OF METALLICOLOUS POPULATIONS.....		82
CHAPITRE 3: POPULATION GENETIC STRUCTURE		95
DISCUSSION GENERALE ET PERSPECTIVES		133
I	AMPLITUDE ECOLOGIQUE D' <i>A. HALLERI</i>	133
II	ORIGINE DE LA TOLERANCE AU ZINC D' <i>A. HALLERI</i>	134
III	EVOLUTION DE LA TOLERANCE AU ZINC CHEZ <i>A. HALLERI</i>	136
IV	REFERENCES	138
ANNEXES.....		142

Avant propos

Depuis soixante-dix ans que l'existence de capacités de tolérance aux métaux lourds a été mise en évidence chez les plantes, l'attention que porte au phénomène une communauté de scientifiques appartenant à une majorité des champs de la biologie (physiologie, écologie, biologie évolutive...) n'a pas diminué. Bien au contraire, les connaissances sur le sujet se sont continuellement accumulées et ont su profiter des évolutions techniques et conceptuelles opérées dans chaque discipline.

Après l'avènement de la biologie moléculaire, la plupart des recherches actuellement effectuées visent, d'une part, à l'identification de mécanismes moléculaires normalement impliqués dans l'homéostasie des métaux, mais pouvant conférer les capacités extrêmes de tolérance tissulaires rencontrées chez les espèces métallophytes, d'autre part, à la détermination directe, chez les métallophytes, de l'architecture génétique de la tolérance aux métaux, soit le nombre de gènes impliqués et la nature des interactions épistatiques entre ces gènes. Ces études engagent des technologies lourdes. Elles nécessitent la focalisation des recherches sur quelques espèces modèles et, au sein de ces espèces, sur quelques accessions soigneusement choisies comme représentatives de l'impact de la pression métallique sur le génome de l'espèce. Se posent à travers elles la question du choix de ces espèces et, au sein des espèces choisies, du choix des accessions analysées.

Initiées durant les années 80, les études de phylogéographie suggèrent que les espèces biologiques ne constituent pas, le plus souvent, un ensemble génétique homogène mais rassemblent, au contraire, des sous-groupes génétiquement différenciés de populations. En ajoutant l'analyse des relations généalogiques (phylogénétiques) entre variants génétiques à l'analyse classique de leurs variations de fréquences entre populations, le champ d'étude de la phylogéographie a révélé que, chez de nombreuses espèces, des phénomènes démographiques passés avait pu fortement influencer la structure en populations actuelle des polymorphismes génétiques, neutres ou sélectionnés. Les oscillations climatiques entre période glaciaires et interglaciaires notamment, ont provoqué des oscillations des aires de distributions des espèces pouvant aboutir à un fractionnement de ces aires, c.-à-d. à l'isolement géographique de populations. Ces populations n'échangeant plus de gènes ont pu diverger. En conséquence, on

observe, au sien de nombreuses espèces, l'existence de groupes de populations dont les patrimoines génétiques sont différents.

Classiquement, on distingue, au sien des espèces végétales tolérantes aux métaux, les populations sur sol pollués, tolérantes, des populations sur sol non pollué, non tolérantes. L'évolution de la tolérance aux métaux lourds, et la différenciation génétique des populations pour le caractère, est en effet associée à l'exposition de certaines de ses populations à des doses importantes de métaux dans leur milieu de vie. Elles sont consécutives à la rencontre de l'espèce et d'un milieu contaminé. La phylogéographie semble enseigner que cette "règle" ne dépend pas que des processus microévolutifs locaux mais est aussi conditionnée par l'histoire des populations. De l'origine des populations et de leur appartenance ou non à des groupes aux patrimoines génétiques différents peut dépendre leurs capacités à acquérir la tolérance aux métaux du sol.

Le travail présenté ici consistait à appliquer l'approche phylogéographique à une espèce végétale présentant des populations sur sites pollués par les métaux lourds afin de mieux comprendre l'histoire évolutive du caractère de tolérance au sein de l'espèce et de resituer cette histoire dans le contexte plus général de l'histoire de l'espèce.

Pour une bonne compréhension des intérêts de cette approche, je synthétiserai successivement dans l'introduction de ce rapport :

- a) afin de définir en détail le caractère qui sera suivi, l'état des connaissances actuelles sur la tolérance aux métaux lourds chez les espèces végétales supérieures,
- b) afin de comprendre les spécificités de l'approche adoptée, les principes fondamentaux de l'analyse phylogéographique, ainsi que des résultats illustratifs,
- c) afin d'identifier le modèle biologique, les connaissances disponibles sur l'espèce utilisée dans cette étude : *Arabidopsis halleri* (Brassicaceae).

Puis seront présentés en trois chapitres et à travers des articles scientifiques écrits en anglais et soit publié, soumis, ou en cours de rédaction, les résultats de l'application de

l'approche phylogéographique dans l'étude de l'évolution de la tolérance au zinc chez *A. halleri*. Le chapitre I rapportera les résultats d'estimation des capacités de tolérance de populations réparties dans toute l'aire de distribution européenne de l'espèce. Une bonne évaluation du polymorphisme pour le caractère et de sa distribution dans les populations de l'espèce était en effet un préalable indispensable à toute tentative de compréhension de son évolution. Le chapitre II présentera l'étude menée, à l'aide de la généalogie de variants chloroplastiques, pour déterminer l'origine de populations métallophiles rencontrées à la limite nord de l'aire de répartition. Le chapitre III élargira cette étude en ouvrant l'échantillonnage à l'ensemble de l'aire de distribution Européenne de l'espèce et en ajoutant, pour la caractérisation de la structure génétique neutre de l'espèce, des marqueurs nucléaires aux marqueurs chloroplastiques. Enfin, une conclusion générale permettra un rappel des résultats majeurs et recensera les perspectives qu'ils ouvrent.

Introduction

I La tolérance aux métaux lourds des végétaux supérieurs

I.1 Définitions

I.1.1 Métal, lifère, lophyte, licole : définitions

La catégorie des métaux couvre 75% des éléments chimiques connus (Figure 1). Les nombreuses fonctions des métaux essentiels dans les systèmes biologiques s'étendent de la régulation de processus métaboliques à la composition des protéines, notamment de leur site actif (Borovik 1990).

Parmi les métaux, la sous-classe des **métaux lourds** est équivoque (Duffus 2002). Sa définition est suffisamment vague et variable d'une discipline scientifique à l'autre pour qu'il soit parfois conseillé de l'éviter (Borovik 1990; Duffus 2002). Le terme est d'ailleurs fréquemment utilisé sans aucune précision et il est difficile de dresser une "liste des métaux lourds". Notons néanmoins parmi les éléments le plus souvent cités par les écotoxicologistes comme appartenant aux métaux lourds: l'arsenic (As), le cadmium (Cd), le chrome (Cr), le cuivre (Cu), le plomb (Pb), le mercure (Hg), le nickel (Ni), le sélénium (Se) et le zinc (Zn). Dans la suite du rapport le terme sera employé pour désigner tout métal présentant une certaine toxicité quand il est présent à faible (ions non essentiels) ou forte (ions essentiels) dose dans les tissus biologiques (Baker & Proctor 1990). Et pour faciliter la lecture, le terme "métal" sera souvent employé comme synonyme de "métal lourd".

On identifie trois sources majeures de métaux lourds dans les sols: la roche-mère, l'atmosphère, la biosphère (Friedland 1990). Le plus souvent, les métaux n'y sont qu'en très faibles teneurs (Bourrelier & Berthelin 1998). Mais on trouve sur 1/10¹⁰ de la surface continentale de la planète des sols dits **métallifères** qui contiennent une quantité anormalement importante d'un ou le plus souvent de plusieurs métaux lourds (Bradshaw *et al.* 1990). Une telle contamination peut-être naturelle et est souvent alors associée à la lente minéralisation, aux niveaux d'affleurements, d'une roche-mère riche en certains métaux

TABLEAU PÉRIODIQUE DES ÉLÉMENTS

<http://www.kipit.kr/periodic/fr/>

PÉRIODE	GROUPE																	
	I A	II A	III B	IV B	V B	VI B	VII B	VIII	IX	X	XI	XII	III A	IV A	V A	VI A	VII A	VIII A
1	1 1.0079 H HYDROGÈNE																	2 4.0026 He HÉLIUM
2	3 6.941 Li LITHIUM	4 9.0122 Be BÉRYLLIUM											5 10.811 B BORE	6 12.011 C CARBONE	7 14.007 N AZOTE	8 15.999 O OXYGÈNE	9 18.998 F FLUOR	10 20.180 Ne NÉON
3	11 22.990 Na SODIUM	12 24.305 Mg MAGNÉSIMUM											13 26.982 Al ALUMINIUM	14 28.086 Si SILICIUM	15 30.974 P PHOSPHORE	16 32.065 S SOUFRE	17 35.453 Cl CHLORE	18 39.948 Ar ARGON
4	19 39.098 K POTASSIUM	20 40.078 Ca CALCIUM	21 44.956 Sc SCANDIUM	22 47.867 Ti TITANE	23 50.942 V VANADIUM	24 51.996 Cr CHROME	25 54.938 Mn MANGANESE	26 55.845 Fe FER	27 58.933 Co COBALT	28 58.693 Ni NICKEL	29 63.546 Cu COUVRE	30 65.39 Zn ZINC	31 69.723 Ga GALLIUM	32 72.64 Ge GERMANIUM	33 74.922 As ARSENIC	34 78.96 Se SÉLÉNIUM	35 79.904 Br BROME	36 83.80 Kr KRYPTON
5	37 85.468 Rb RUBIDIUM	38 87.62 Sr STRONTIUM	39 88.906 Y YTTRIUM	40 91.224 Zr ZIRCONIUM	41 92.906 Nb NIOBIUM	42 95.94 Mo MOYBÈNE	43 (98) Tc TECHNETIUM	44 101.07 Ru RUTHÉNIUM	45 102.91 Rh RHODIUM	46 106.42 Pd PALLADIUM	47 107.87 Ag ARGENT	48 112.41 Cd CADMIUM	49 114.82 In INDIUM	50 118.71 Sn ÉTAIN	51 121.76 Sb ANTIMOINE	52 127.60 Te TELLORE	53 126.90 I IODE	54 131.29 Xe XÉNON
6	55 132.91 Cs CÉSURIUM	56 137.33 Ba BARYUM	57-71 La-Lu Lanthanides	72 178.49 Hf HAFNIUM	73 180.95 Ta TANTALE	74 183.84 W TUNGSTÈNE	75 186.21 Re RHÉNIUM	76 190.23 Os OSMIUM	77 192.22 Ir IRIDIUM	78 195.08 Pt PLATINE	79 196.97 Au OR	80 200.59 Hg MERCURE	81 204.38 Tl THALLIUM	82 207.2 Pb PLOMB	83 208.98 Bi BISMUTH	84 (209) Po POLONIUM	85 (210) At ASTATE	86 (222) Rn RADON
7	87 (223) Fr FRANCIUM	88 (226) Ra RADIUM	89-103 Ac-Lr Actinides	104 (261) Rf RUTHÉRFORIUM	105 (262) Db DUBNIUM	106 (263) Sg SEABORGIUM	107 (264) Bh BOHRNIUM	108 (267) Hs HASSIUM	109 (268) Mt MEITNERIUM	110 (268) Uu UNUNNIUM	111 (272) Uub UNUNBIUM	112 (285) Uuq UNUNQUADIUM						

(*) Pure Appl Chem, 73, No. 4, 667-683 (2001)
La masse atomique relative est donnée avec un seul chiffre significatif. Pour les éléments qui n'ont pas de données établies, la valeur entre parenthèses indique le nombre de masse de l'isotope de l'élément ayant la durée de vie la plus grande.
Tous les éléments Th, Pa et U ont une composition isotopique connue, une masse atomique est indiquée.

Editor: Michel Dula

LANTHANIDES											ACTINIDES																		
57 138.91 La LANTHANE	58 140.12 Ce CÉRIUM	59 140.91 Pr PRASEODYME	60 144.24 Nd NÉODYME	61 (145) Pm PROMÉTHIUM	62 150.36 Sm SAMARIUM	63 151.96 Eu EUROPIUM	64 157.25 Gd GADOLINIUM	65 158.93 Tb TERBIUM	66 162.50 Dy DYSPROBIUM	67 164.93 Ho HOLMIUM	68 167.26 Er ÉRBIUM	69 168.93 Tm THULIUM	70 173.04 Yb YTTÉRIUM	71 174.97 Lu LUTÉTIUM	89 (227) Ac ACTINIUM	90 232.04 Th THORIUM	91 231.04 Pa PROTACTINIUM	92 238.03 U URANIUM	93 (237) Np NEPTUNIUM	94 (244) Pu PLUTONIUM	95 (243) Am AMÉRICIUM	96 (247) Cm CURIUM	97 (247) Bk BERKÉLIUM	98 (251) Cf CALIFORNIUM	99 (252) Es EINSTEINIUM	100 (257) Fm FERMIUM	101 (258) Md MENDELÉVIUM	102 (259) No NOBÉLIUM	103 (262) Lr LAWRENCIUM

Figure 1 Tableau périodique des éléments de Mendeleiev. Les métaux sont les éléments dont la couche électronique la plus élevée n'est que partiellement remplie par des électrons.

lourds. Ainsi les sols serpentiniques, produits par l'érosion de roches ultramafiques sont composés d'au moins 70% de minéraux ferromagnésiens (ou mafiques), et sont riches en fer, nickel, chrome, et cobalt (Brady *et al.* 2005). On constate aussi, dans les régions fortement industrialisées, des pollutions d'origine anthropique très significatives, le plus souvent dues aux activités minières et métallurgiques d'extraction et de transformation de minerais (Antonovics *et al.* 1971; Bourrelier & Berthelin 1998). Les sites dont la pollution est anthropique se distinguent notamment des sites naturellement riches en métaux par leur histoire comparativement récente (Macnair 1987).

Les milieux métallifères présentent des conditions environnementales particulièrement stressantes pour les organismes, non seulement en raison de la présence des métaux à des niveaux toxiques, mais aussi en raison de la combinaison souvent associée de propriétés physico-chimiques extrêmes (mauvaise structure du sol, érosion importante, matière organique peu abondante, entraînant une sécheresse plus forte, manque en oligo-éléments, etc., Macnair 1997). En conséquence se développent sur ces milieux des communautés biologiques singulières (laissons pour le moment de côté la signification de leur présence). Les espèces de plantes vasculaires présentes dans ces communautés sont dites **métallophytes** (Antonovics *et al.* 1971). Elles sont classées comme métallophytes absolues (ou eumétallophytes ou métallophytes strictes) lorsqu'elles sont endémiques des sites pollués et comme pseudométallophytes (ou métallophytes facultatives) lorsqu'elles se rencontrent conjointement sur sites pollués et non pollués (Tableau 1). Au sein des pseudométallophytes, on qualifie de **métallicoles (M)** les populations de l'espèce sur site métallifère et de **non métallicoles (NM)** les populations sur site non métallifère.

1.1.2 La tolérance aux métaux

1.1.2.1 Deux définitions relatives

Dans le cadre des espèces métallophytes, il ne fait plus de doute que la compréhension des processus biologiques conditionnant l'acquisition de la capacité à se développer sur sol métallifère nécessitera l'intégration des approches récentes de biologie moléculaire, aux approches classiques d'écologie (Purugganan & Gibson 2003). Alors que les premières sont intéressées surtout par la mécanique moléculaire des caractères biologiques, les secondes se

Tableau 1 Classification des métallophytes avec exemples pris dans la flore d'Europe (selon Lambinon et Auquier 1964, cité dans Antonovics *et al.* 1971). Depuis, n'ont surtout été retenus que les termes "métallophytes absolues" pour désigner les espèces endémiques des milieux pollués, et "pseudométallophytes" pour désigner les espèces rencontrées à la fois sur sites métallifères et non métallifères.

1. *Metallophytes*—taxa found only on metal contaminated soils.
 - (a) *Absolute metallophytes*—
found only on metal contaminated soil over all their distribution. e.g. *Viola calaminaria*, *Thlaspi alpestre* ssp. *calaminare*, *Minuartia verna* ssp. *hercynica*.
 - (b) *Local metallophytes*—
found only on metal contaminated soil within a given region but occurring also in a phyto-geographically distinct non-contaminated area. e.g. *Armeria maritima*.

2. *Pseudometallophytes*—taxa occurring both on contaminated soils and on normal soils in the same region.
 - (a) *Elective pseudometallophytes*—
abundant and often more vigorous on contaminated soil. e.g. *Agrostis tenuis*, *Campanula rotundifolia*, *Polygala vulgaris*, *Thymus pulegioides*, *Rumex acetosa*.
 - (b) *Indifferent pseudometallophytes*—
live on contaminated soil regularly but show neither abundance nor particular vitality. e.g. *Plantago lanceolata*, *Avena pubescens*, *Genista tinctoria*, *Linum catharticum*.
 - (c) *Accidental metallophytes*—
usually weeds and ruderals appearing sporadically and showing reduced vigour on metal contaminated soils

concentrent sur les forces évolutives agissant sur les organismes et les niveaux d'organisation biologiques supérieurs (populations, espèces...). Aux différences de perspective correspondent des différences de définition, notamment du stress environnemental (Bijlsma & Loeschcke 2005). Les physiologistes moléculaires envisagent surtout le stress comme une perturbation des processus physiologiques nécessaires au bon fonctionnement de l'individu alors que les écologistes y voient un facteur de réduction de la valeur sélective des organismes (Bijlsma & Loeschcke 2005). En conséquence, on trouve dans la bibliographie relative à la tolérance aux métaux deux définitions différentes en fonction de la spécialité des auteurs.

En physiologie moléculaire, la tolérance aux métaux désigne l'ensemble des mécanismes homéostatiques (transport, chélation, séquestration) visant "à maintenir la concentration des métaux essentiels dans les limites physiologiques et à minimiser les effets délétères des métaux non essentiels" (Clemens 2001). Cette tolérance de base, constitutive et ubiquiste (Clemens 2001), est surtout définie par opposition à l'(hyper)sensibilité aux métaux créée par mutagenèse chez certains organismes (Clemens 2001; Hall 2002; Krämer 2005). Elle suggère déjà une capacité (limitée) à gérer des variations de concentrations en métal dans l'environnement et à résister à des expositions excessives (Ernst *et al.* 1992). Mais dans ce cas, la "capacité [exceptionnelle des métallobytes] à tolérer des concentrations [anormalement] élevées dans les cellules racinaires et foliaires" est qualifiée d'"hypertolérance" (Chaney *et al.* 1997; repris par Clemens 2001; Clemens 2006).

En écologie et génétique des populations, le terme "tolérance aux métaux lourds" est employé pour qualifier la "capacité d'une plante à survivre et se reproduire sur des sols toxiques ou défavorables à la plupart des autres organismes en raison d'une contamination métallique" (Turner 1969; Antonovics *et al.* 1971; Macnair & Baker 1994; Macnair *et al.* 2000). Le terme s'applique préférentiellement aux individus d'une espèce plutôt qu'à l'ensemble de l'espèce afin de tenir compte de l'absence de tolérance généralement observée chez les individus NM d'une espèce pseudométallophyte (Antonovics *et al.* 1971; cf. § I.4). Mais on qualifie aussi la tolérance d'une population par la moyenne des capacités des individus qui la composent (Pollard *et al.* 2002).

Au-delà de la différence de "catégorie écologique" auquel s'applique le terme (le règne végétal en physiologie végétale, les métallobytes en écologie végétale), on peut noter des

différences de conceptions du caractère. Les physiologistes envisagent la (hyper)tolérance comme la réponse des mécanismes de l'homéostasie des métaux à la pression métallique, réduisant la possibilité d'autres mécanismes¹ et ignorant les autres contraintes environnementales auxquelles sont généralement exposés les métallobytes. A l'opposé, les généticiens des populations voient dans la tolérance un phénomène global d'adaptation à un environnement multiple extrême et confondent souvent les caractères "tolérant" et "métallicole".

Dans ce rapport, le terme "tolérance" sera toujours employé pour qualifier la capacité héréditaire d'une plante à survivre et se reproduire dans des conditions d'exposition aux métaux lourds normalement toxiques. Cette définition envisage la tolérance comme une capacité adaptative à supporter des concentrations toxiques dans le milieu de vie sans inclure les adaptations conjointes que sous-entend la colonisation des sols métallifères. C'est cette définition qui a été retenue lors des expériences de phénotypage d'individus durant ma thèse.

1.1.2.2 La tolérance dans les bacs: mesure et définition de la tolérance en milieu contrôlé

La tolérance ainsi définie, il suffirait d'observer la survie d'individus exposés à des doses anormalement élevées de métaux pour déduire l'existence de capacités de tolérance chez ces individus. Reste néanmoins, d'une part, à déterminer les stades de développement et les organes à observer ainsi que les variables biologiques à mesurer pour établir la survie et donc la tolérance d'individus à une dose donnée, et, d'autre part, à définir les doses d'exposition qui seront considérées comme anormales.

En conditions expérimentales contrôlées, c.-à-d. une fois la variance environnementale réduite au minimum, les différences génétiquement déterminées de capacités de tolérance sont les seules à expliquer significativement des différences de réponses à des expositions aux métaux. On attribue ainsi généralement la tolérance à certains génotypes de façon relative,

¹ Par exemple, la tolérance à l'As de *Holcus lanatus* semble être en partie conférée par la perte de transporteurs ayant une grande affinité pour le phosphate, antérieurement à l'exposition au métal (Meharg, 1993)

après comparaison en milieu contrôlé avec d'autres génotypes (souvent conspécifiques ou congénères et d'origine non métallicole) ne supportant pas de faibles expositions aux métaux (pour une revue, voir Baker 1987).

Les différences de capacité de tolérance sont généralement si marquées qu'il suffit de tests de survie à cours terme utilisant des plantules ou des propagules pour les détecter (Baker 1987). Les premiers travaux sur la tolérance au zinc d'*Agrostis capillaris* (syn. *A. tenuis*, Bradshaw 1952) ayant montré que la **croissance racinaire** était particulièrement sensible à la présence de métal, une méthode dite d'élongation racinaire a été très tôt développée et utilisée pour quantifier la tolérance au métal d'un grand nombre d'espèce (Wilkins 1978; De Koe *et al.* 1992). Elle consiste à mesurer quantitativement, en **culture hydroponique**, la croissance des racines de jeunes individus ou clones respectivement exposés (solution test) ou non (solution témoin) à une certaine concentration d'un métal et à utiliser le rapport de ces deux mesures pour calculer un indice de tolérance à ce métal (Tolerance Index, $TI = \text{croissance racinaire en solution test} / \text{croissance racinaire en solution témoin}$). La méthode suppose ainsi que la croissance racinaire d'un génotype dans une solution toxique est fonction à la fois du taux intrinsèque de croissance (mesuré en solution témoin) et de son niveau de tolérance. Elle intègre cependant dans le calcul du TI la variance des gènes codant spécifiquement pour la croissance racinaire et l'effet de leur interaction avec l'environnement (Macnair 1990). Les effets de ces gènes en solutions témoin et test sont non proportionnels, ce qui engendre un bruit statistique important diminuant d'autant la précision de l'indice (Macnair 1983; Macnair 1990; Schat & Ten Bookum 1992). On pourrait ajouter que la notion de solution témoin pour des génotypes naturellement exposés à de fortes concentrations en métal n'est pas évidente. Pour ces raisons, (Macnair 1983) a proposé de mesurer plutôt la **capacité des plantules à produire ou non de nouvelles racines**, toujours en hydroponie, dans une solution dont la dose est fixée. La nature purement qualitative du critère proposé, i.e. croissance ou non à une **concentration fixée**, s'est révélée inadéquate pour estimer la variation quantitative du caractère (Schat & Ten Bookum 1992). Pour éviter l'usage de tests à concentrations multiples dans lesquels des clones sont exposés à différentes concentrations, nécessitant beaucoup de temps, d'espace, et la clonabilité de l'espèce étudiée, (Schat & Ten Bookum 1992) a proposé un **test séquentiel**, au cours duquel les mêmes individus sont exposés à des concentrations croissantes de métal dans la solution de culture. Le niveau de tolérance d'un génotype correspond à la plus petite concentration stoppant sa croissance racinaire (EC100, pour

Effective Concentration 100%, Schat & Ten Bookum 1992). Ce test a depuis été largement utilisé dans les études de la distribution en populations et du déterminisme génétique de la tolérance (Bert *et al.* 2000; Assunção *et al.* 2003a; Bert *et al.* 2003; Pauwels *et al.* sous presse; Willems *et al.* soumis)

Il est clairement impossible d'extrapoler sans précaution les niveaux de tolérance à un métal obtenus en hydroponie pour une espèce à des situations naturelles. Le problème le plus évident est lié à la chimie complexe des métaux dans les sols et à la difficulté d'y estimer leur biodisponibilité et leur toxicité (Friedland, 1990). Il est en effet difficile de déterminer si une concentration expérimentée en hydroponie est supérieure ou inférieure à celle rencontrée sur un sol donné et de définir en hydroponie le niveau de tolérance requis pour permettre à une plante de se développer sur un sol particulier (Wilkins 1978). Cependant, la méthode propose de travailler avec des conditions reproductibles dans l'espace ou le temps et permet ainsi la comparaison de résultats.

I.2 Mécanismes physiologiques de la tolérance aux métaux

I.2.1 Des mécanismes spécifiques

Les sols métallifères sont souvent enrichis d'une combinaison de différents métaux. Par conséquent, les métallobytes montrent souvent des capacités de tolérance à plusieurs métaux. La combinaison de ces capacités peut reposer soit sur une combinaison de mécanismes indépendants de tolérance respectivement spécifiques d'un seul métal, on parle de **tolérance multiple**, soit sur des mécanismes moins spécifiques qui confèrent simultanément, par pléiotropie, une tolérance à plusieurs métaux, on parle alors de **cotolérance** (Schat *et al.* 2000).

Les travaux effectués très tôt sur *Agrostis tenuis* (Gregory & Bradshaw 1965) ont montré que la tolérance à un métal présent dans le sol n'était accompagnée de la tolérance à un autre métal presque exclusivement que si ce métal était lui aussi présent dans le sol. Les auteurs ont donc conclu à une grande spécificité des mécanismes de tolérance. De nombreux auteurs ont néanmoins reporté chez certaines plantes l'existence de tolérances à des métaux qui n'étaient pas en concentrations toxiques dans les sols (voir, pour une revue Schat & Vooijs

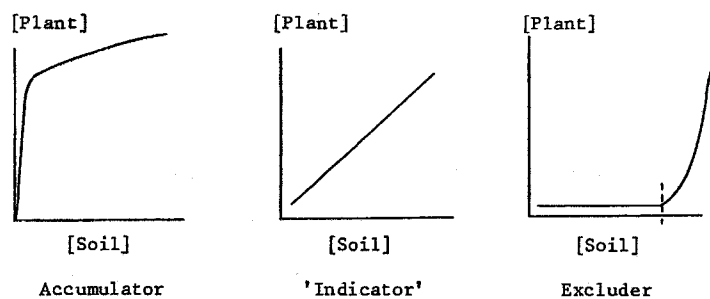


Figure 2 Représentations schématiques des stratégies rencontrées chez les plantes en réponse à une augmentation des concentrations en métal dans le sol (Baker 1981). Trois catégories sont distinguées. La stratégie "indicatrice" (Indicator), peu étudiée, représente la stratégie intermédiaire entre deux autres extrêmes: l'(hyper)accumulation (Accumulator) et l'exclusion (Excluder). Les espèces indicatrices régulent, activement ou non, l'absorption et le transport des métaux de façon à ce que les concentrations dans les organes aériens soient proportionnelles à celles du sol.

1997). Ces tolérances, "non fonctionnelles", ont parfois été interprétées comme le sous-produit, pléiotropique, d'une tolérance "fonctionnelle" à un métal présent dans le sol, favorisant ainsi l'hypothèse d'une coto tolérance (Schat & Vooijs 1997).

Le débat "tolérance multiple versus coto tolérance" est resté longtemps très ouvert, jusqu'à l'apport des analyses génétiques de ségrégation du caractère (Macnair 1993). Il semble à présent, comme il a été montré pour la tolérance au cuivre de *Mimulus guttatus* (Tilstone & Macnair 1997) et les tolérances au Zn, Cd et Ni de *Thlaspi caerulescens* (Assunção *et al.* 2003a), que les mécanismes de tolérance aux métaux lourds soient hautement métal-spécifiques. Il se peut néanmoins que la tolérance à un métal produise un faible niveau de coto tolérance à un autre métal aux propriétés biochimiques proches (Schat *et al.* 2000). Ainsi, par exemple, l'étude combinée du déterminisme génétique des tolérances au Cu, Zn, Ni, Co et Cd chez *Silene vulgaris* montre que la tolérance au Zn peut conférer un certain niveau de tolérance au Ni (Schat & Vooijs 1997).

I.2.2 Les grandes stratégies physiologiques de tolérance aux métaux

C'est en confrontant, pour différents métaux et de nombreuses espèces végétales, des données de concentrations dans les racines et les organes aériens aux données de concentrations dans les sols que Baker (Baker 1981) a suggéré que deux types extrêmes de "stratégies" physiologiques pouvaient rendre compte de la tolérance aux métaux à l'échelle de l'organisme: **l'exclusion et l'accumulation** (Figure 2). Ces stratégies, qui définissent le comportement de différentes espèces face à un même métal, mais aussi parfois, en accord avec la spécificité des mécanismes (voir I.2.1), une même espèce face à différents métaux, se distinguent par l'organe où sont détoxifiés et stockés les métaux en excès. Pour un métal donné, les espèces excluantes (excluders) et accumulatrices (accumulators) sont distinguées par la valeur du rapport "concentration en métal dans les parties aériennes / concentration dans les racines" ([organes aériens] / [racines], respectivement < 1 et > 1).

I.2.2.1 L'exclusion

La stratégie d'exclusion ("excluder strategy", Baker 1981) s'illustre par le maintien de concentrations en métal dans les parties aériennes constantes et relativement faibles sur une large gamme de concentration dans le sol, de telle façon que le rapport [organes aériens] / [racines] est inférieur à 1. Au-delà d'une valeur critique de concentration dans le sol,

Tableau 2 Exemples d'espèces Angiospermes identifiées comme excluantes dans la littérature

Espèce	Famille	Tolérance
<i>Mimulus guttatus</i>	Scrophulariaceae	Cu ^{a,b}
<i>Plantago lanceolata</i>	Plantaginaceae	Zn, Pb ^c
<i>Silene vulgaris</i>	Caryophyllaceae	Zn, Cu, Pb, Cd
<i>Elsholtzia splendens</i>	Labiatae	Cu ^d
<i>Rumex acetosa</i>	Polygonaceae	Al ^e
<i>Festuca rubra</i>	Poaceae	Zn ^f
<i>Eriophorum angustifolium</i>	Cyperaceae	Zn ^f
<i>Carex rostrata</i>	Cyperaceae	Zn ^f
<i>Silene maritima</i>	Caryophyllaceae	Zn ^g
<i>Agrostis capillaris</i> (<i>A. tenuis</i>)	Poaceae	Zn, Cu ^h
<i>Agrostis stolonifera</i>	Poaceae	Zn, Cu ⁱ

^a: (Tilstone & Macnair 1997); ^b: (Macnair 1981); ^c: (Lefèbvre et Denaeyer-De Smet, 1986, cité dans Baker & Proctor 1990); ^d: (Song *et al.* 2004); ^e: (Tolra *et al.* 2005); ^f: (Matthews *et al.* 2005); ^g: (Baker 1981); ^h: (Peterson, 1969, Bradshaw *et al.*, 1965, cités dans Baker & Walker 1990); ⁱ: (Wu & Antonovics 1975)

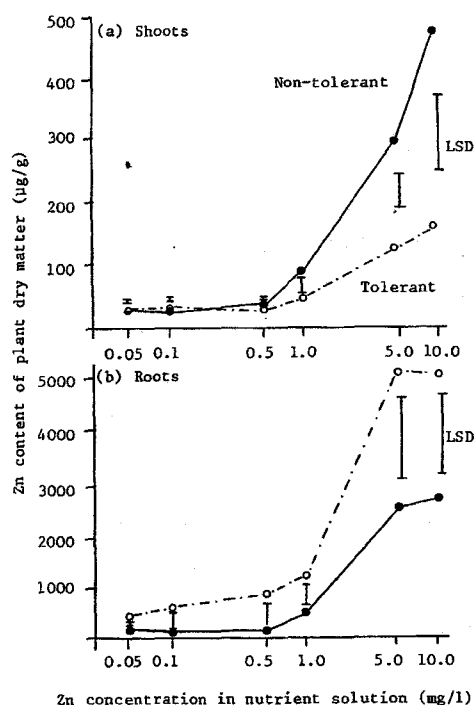


Figure 3 Concentrations en Zn dans les organes aériens (a) et les racines (b) de populations tolérantes (M) et non tolérantes (NM) de *Silene maritima* exposées en hydroponie et pendant 15 jours à différentes concentrations en Zn dans la solution nutritive (Baker 1981). La tolérance au Zn de *Silene maritima* illustre la stratégie d'exclusion, dans laquelle le métal est préférentiellement accumulé dans les racines. Les barres verticales représentent les LSD (Least Significant Difference; P = 0,05) entre les moyennes des populations

néanmoins, le contrôle est perdu, la plante s'intoxique et ce rapport augmente. L'évolution vers une telle stratégie d'exclusion a été mise en évidence pour divers métaux chez de nombreuses espèces appartenant à de nombreuses familles taxonomiques (Tableau 2). L'adaptation apparaît par exemple clairement lorsque l'on compare, au sein d'une espèce pseudométallophyte comme *Silene maritima*, l'évolution des concentrations en Zn dans les organes aériens et dans les racines d'accessions NM (c.-à-d. non tolérante) et M (c.-à-d. tolérante) exposées en hydroponie à des doses variables de Zn (Figure 3).

Au sens le plus large, la stratégie d'exclusion inclut la capacité à empêcher l'absorption des métaux depuis le sol vers les organes de la plante. Dans ce cas précis où les métaux du sols ne pénètrent pas la plante, on parle plus spécifiquement d'**évitement** ("avoidance", Baker & Walker 1990). L'évitement est communément observé chez les bactéries, les champignons et les algues (Baker & Walker 1990). Il reste comparativement rare chez les végétaux vasculaires tolérants (Baker 1981; Baker & Walker 1990). A l'exception de quelques espèces réellement capables d'évitement¹, la plupart des espèces vasculaires excluantes ne peuvent, au mieux, que réduire le transport vers les parties aériennes. La capacité de ces espèces à survivre sur sols toxiques correspond ainsi à une tolérance vraie de concentrations excessives de métaux dans les tissus, et notamment racinaires. En résumé, la **stratégie d'exclusion correspond donc surtout à une forte réduction du transport des métaux vers les parties aériennes et à une détoxification et une séquestration dans les tissus racinaires.**

1.2.2.2 L'accumulation et l'hyperaccumulation

Les espèces accumulatrices ("accumulators") sont définies comme celles dont les concentrations en métaux dans les parties aériennes sont supérieures à celles du sol, ou comme celles dont les concentrations en métaux dans les parties aériennes sont, sur un sol donné, supérieures aux valeurs normales de concentrations rencontrées chez les espèces végétales (Peterson, 1983, cité dans Baker & Proctor 1990). De façon générale, les espèces accumulatrices se caractérisent par des concentrations fortes en métaux dans les tissus aériens,

¹ Par différents mécanismes comme la complexation à des molécules organiques dans le sol, l'association avec des symbiotes, etc. (pour des revues, lire Meharg & Hartley-Whitaker 2002; Christie *et al.* 2004; Gaur & Adholeya 2004)

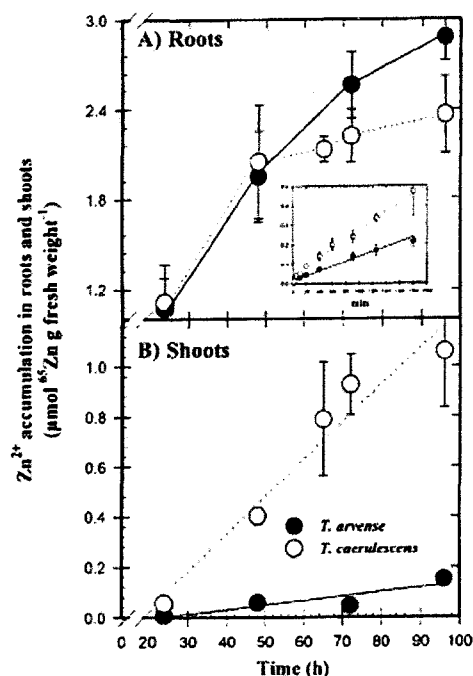


Figure 4 Evolution temporelle du Zn^{2+} dans les racines (A) et les organes aériens (B) de deux espèces de *Thlaspi* cultivées en hydroponie avec une solution de culture contenant $10\mu M$ de Zn^{2+} (Lasat *et al.* 2000). Les points et barres d'erreur représentent les moyennes et écart-types obtenus sur quatre réplicats. *T. caerulescens* est hyperaccumulatrice de Zn, *T. arvense* ne l'est pas. En comparaison, le transfert de Zn^{2+} vers les organes aériens est nettement accru chez *T. caerulescens*.

Tableau 3 Nombre d'hyperaccumulateurs connus et familles dans lesquels ils sont le plus souvent trouvés, pour différents métaux. Noter le nombre généralement croissant de taxa identifiés entre 1998 et 2000, symptomatique de l'intérêt porté sur les espèces hyperaccumulatrices.

Elément	Nb. de taxa* (**)	Nb. de familles**	Familles prédominantes*
Cadmium (Cd)	1 (1)	1	Brassicaceae
Cobalt (Co)	26 (28)	11	Lamiaceae, Scrophulariaceae
Cuivre (Cu)	24 (37)	15	Cyperaceae, Lamiaceae, Poaceae, Scrophulariaceae
Plomb (Pb)	- (14)	14	
Manganse (Mn)	11 (9)	5	Apocynaceae, Cunoniaceae, Proteaceae
Nickel (Ni)	290 (317)	37	Brassicaceae, Cunoniaceae, Euphorbiaceae, Flacourtiaceae, Violaceae
Selenium (Se)	19 (-)	-	Fabaceae
Thallium	1 (-)	-	Brassicaceae,
Zinc	16 (11)	5	Brassicaceae, Violaceae

* in (Brooks 1998c); ** in (Baker *et al.* 2000)

de tel sorte que le rapport [organes aériens] / [racines] est supérieur à 1. A l'opposé de l'exclusion, l'accumulation ("accumulator strategy") consiste ainsi à concentrer le métal dans les tissus aériens (Figure 4). L'accumulation suppose **un important transport actif des métaux depuis les racines, où sont absorbés les métaux du sol, vers les organes chlorophylliens, qui sont les organes de détoxification et de stockage.**

De grandes variations de niveaux d'accumulation ont été observées entre espèces, depuis une concentration à peine supérieure à celle du sol jusqu'à des niveaux représentant un pourcentage significatif du poids sec de la plante (Baker *et al.* 2000). Ainsi, si l'accumulation est déjà, en soi, une forme spectaculaire de réponse physiologique (Baker 1981), on distingue encore, au sein des accumulateurs, les hyperaccumulateurs, découverts relativement récemment (Brooks 1998b), dont la concentration en métal dans les organes aériens excède un seuil allant, en fonction du métal, de 0,01% à 1% du poids sec des feuilles (Ginocchio & Baker, 2004).

Bien que représentée dans de nombreuses familles taxonomiques (Tableau 3), **l'hyperaccumulation reste une stratégie de réponse relativement peu commune** [418 espèces (Prasad & Freitas 2003), soit 0,2% des Angiospermes (Baker *et al.* 2000)]. Néanmoins, en tant que type extrême de réponse et pour son intérêt économique dans de potentielles applications en phytoremédiation (Lasat 2002; Prasad & Freitas 2003; Ghosh & Singh 2005), l'hyperaccumulation a récemment considérablement attiré l'intérêt des scientifiques et a fait l'objet de plusieurs monographies (Brooks 1998b; Terry & Bañuelos 2000). Les hyperaccumulateurs sont aujourd'hui les espèces métallophytes les plus intensément étudiées. Les deux espèces majoritairement proposées comme modèle pour l'étude de la tolérance aux métaux lourds à l'aide des techniques de génomiques modernes sont d'ailleurs les hyperaccumulateurs *Thlaspi caerulescens* (**Brassicaceae**) et *Arabidopsis halleri* (**Brassicaceae**) (Assunção *et al.* 2003b; Krämer 2003).

I.2.3 Les processus moléculaires de la tolérance aux métaux

Aux stratégies globales de réponse à la pression métallique, observées à l'échelle de l'organisme, correspondent inmanquablement des mécanismes moléculaires aux niveaux cellulaire et tissulaire. Puisque, dans les études à ces niveaux, la tolérance est surtout envisagée comme une adaptation des mécanismes physiologiques de transport, de détoxification et de stockage des métaux (voir ci-dessus), de gros efforts sont aujourd'hui

consacrés à une meilleure compréhension des mécanismes moléculaires de base de l'homéostasie des métaux (présents chez toutes les espèces végétales) et de leurs modifications chez les espèces métallophytes (Clemens 2001; Hall 2002).

Cinq niveaux dans la voie métabolique suivie par les métaux depuis le sol d'où ils sont absorbés jusqu'à l'organe dans lequel ils sont utilisés ou stockés ont été caractérisés et plus particulièrement étudiés (Figure 5):

- **la mobilisation depuis le sol et l'absorption racinaire.** Les mécanismes actifs de mobilisation des métaux lourds par les racines des plantes supérieures sont encore peu connus. Cependant, alors que peu d'études ont su démontrer des différences d'absorption racinaire entre accessions M et NM de pseudometallophytes (voir la mise en évidence de modifications des flux transmembranaires de Cu chez des accessions tolérantes de *S. vulgaris*, van Hoof *et al.* 2001), l'hypothèse d'une régulation de l'absorption racinaire est rarement formulée pour expliquer l'existence de capacités de la tolérance aux métaux [voir néanmoins la résistance à l'As de quelques graminées (Meharg & Hartley-Whitaker 2002) et les tolérances de *T. caerulea* (Lasat *et al.* 2000; Pence *et al.* 2000)

- **la chélation et la séquestration dans les racines.** Chez toutes les espèces végétales supérieures, les métaux, séquestrés surtout dans les vacuoles des cellules, sont associés à des agents chélateurs (Callahan *et al.* 2006). Au niveau racinaire, les phytochélatines (PCs) et les métallothionéines (MTs), protéines riches en S (Cobbett 2000; Cobbett & Goldsbrough 2002) et dont la synthèse est activée par l'exposition aux métaux (Zenk 1996; Cobbett 2000), semblent généralement jouer un rôle important dans la détoxification des métaux, notamment du Cd (Clemens *et al.* 2002). Elles ont donc logiquement été supposées très impliquées dans la tolérance aux métaux des plantes. Après étude, ces protéines ne sont cependant pas apparues comme un mécanisme majeur de tolérance, notamment chez les espèces excluantes (qui accumulent pourtant les métaux dans les racines). Pour exemple, les tolérances au Cu, Cd, Zn, Co et Ni de *S. vulgaris* ne semblent pas dépendre pas d'une séquestration fondée sur la production de PCs (Leopold *et al.* 1999; Schat *et al.* 2002). La tolérance au Zn de *S. vulgaris* serait plutôt associée à un transport important de l'ion non chélaté vers la vacuole des cellules racinaires (Chardonens *et al.* 1999).

- **le transport via le xylème.** Dans le xylème des plantes supérieures, l'histidine (His), la nicotianamine et le citrate (Callahan *et al.* 2006) sont des chélateurs importants des métaux (Clemens *et al.* 2002). L'exposition d'*Alyssum lesbiacum* (hyperaccumulateur de Ni) au Ni augmente la quantité d'His libre et l'apport d'His améliore les capacités de tolérance et de

transport vers les feuilles de l'espèce congénère non accumulatrice *Alyssum montanum* (Krämer *et al.* 1996). Ces deux résultats suggèrent un rôle de l'His dans la tolérance à des métaux comme le Ni. Ce rôle de l'His semble cependant varier entre espèces hyperaccumulatrices de Ni (Callahan *et al.* 2006). Ainsi, la concentration en Zn dans le xylème est plus forte chez *T. caerulescens* que chez *T. arvense*, espèce congénère non tolérante (Lasat *et al.* 1998; Lasat *et al.* 2000).

- **la distribution dans les tissus aériens.** Dans les tissus aériens, les métaux sont généralement surtout stockés dans les vacuoles des cellules foliaires, même si les niveaux d'accumulation des différents types cellulaires varient selon les espèces. Chez *T. caerulescens*, l'accumulation du Zn est 5 à 6,5 fois supérieure dans les cellules épidermiques que dans les cellules mésophylliennes (Küpper *et al.* 1999). Ce type de distribution est aussi retrouvé chez de nombreux hyperaccumulateurs de Ni (revus dans Baker *et al.* 2000). A l'inverse, le Zn et le Cd sont majoritairement accumulés dans le mésophylle après exposition au métal chez *A. halleri* (Küpper *et al.* 2000). Des quantités importantes de Ni et de Zn ont aussi été détectées dans les trichomes respectifs d'*Alyssum lesbiacum* (Krämer *et al.* 1997) et d'*A. halleri* (Küpper *et al.* 2000; Sarret *et al.* 2002).

- **la chélation et la séquestration dans les feuilles.** Alors que le transport intracellulaire des métaux essentiels jusqu'aux organelles fait, chez les végétaux supérieurs, communément intervenir des protéines appelées (métallo)chaperones (Himmelblau & Amasino 2000; O'Halloran & Culotta 2000), le stockage des mêmes métaux dans les vacuoles des cellules photosynthétiques est surtout assuré par des acides organiques ou carboxyliques auxquels les métaux sont chélatés (Smith *et al.*, 1996, cité dans Baker *et al.* 2000). Leur rôle des tels acides dans l'hyperaccumulation a été évoqué dans de nombreuses études, même s'il semble dépendre de l'âge des feuilles et varie entre espèce (pour un revue, voir Callahan *et al.* 2006). Par exemple, le Ni est respectivement trouvé complexé au citrate et à l'His chez *Thlaspi goesingense* (hyperaccumulatrice de Ni) et *T. arvense* (non hyperaccumulatrice Krämer *et al.* 2000)); le Zn est respectivement complexé au malate et à des groupes carboxylés et hydroxylés dans le mésophylle et les trichomes d'*A. halleri* (Sarret *et al.* 2002).

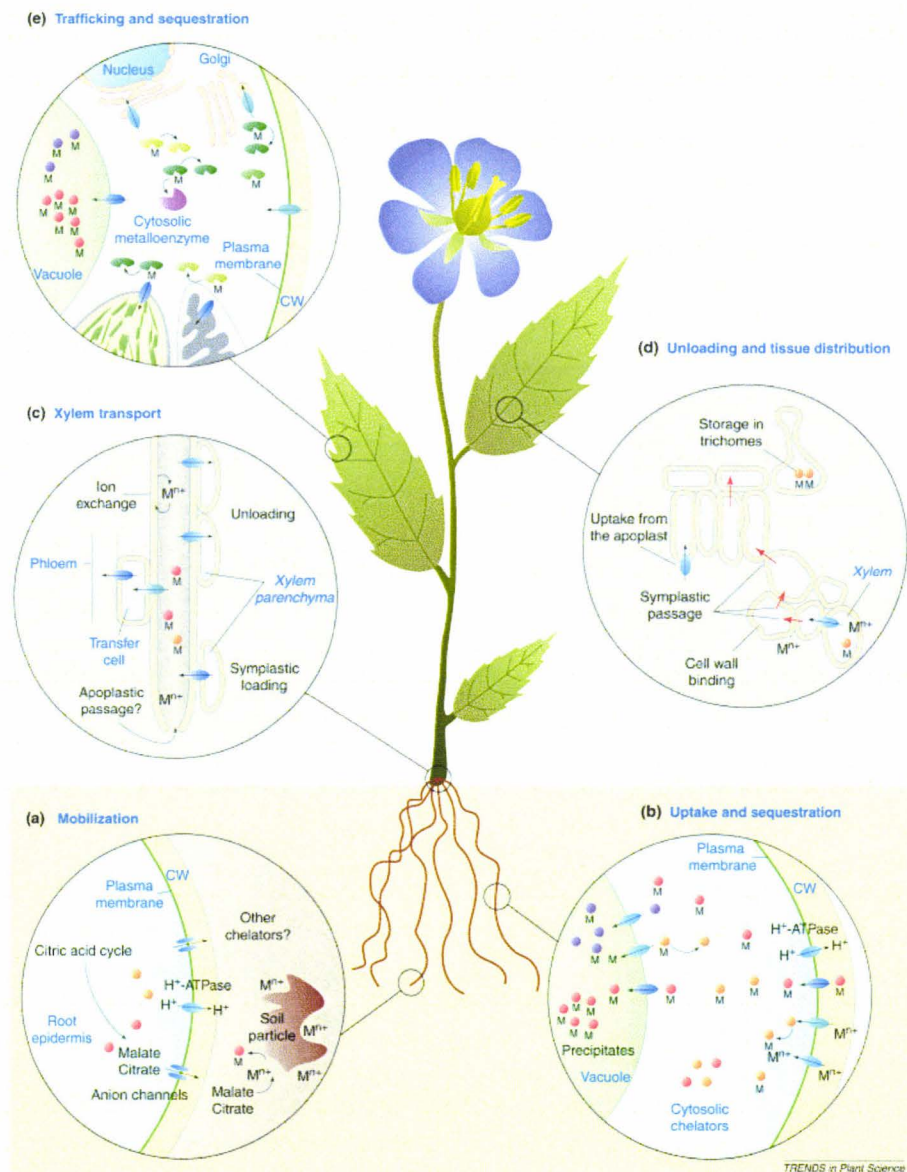


Figure 5 Mécanismes moléculaires supposés impliqués dans l'homéostasie des métaux chez les plantes et potentiellement responsables du développement de tolérances aux métaux (Clemens *et al.* 2002). Cinq niveaux majeurs d'actions sont distingués (voir texte): (a) la mobilisation depuis le sol et l'absorption racinaire (les ions métalliques sont mobilisés par sécrétion de chélateurs et acidification de la rhizosphère); (b) la chélation* et la séquestration* dans les racines, dont la régulation est potentiellement impliquée dans la stratégie d'exclusion (l'absorption dans la cellule est assurée par des transporteurs transmembranaires. Les métaux sont chélatés et les métaux en excès sont transportés et séquestrés dans la vacuole); (c) le transport via le xylème, (d) la distribution dans les tissus aériens et (e) la chélation et la séquestration dans les feuilles, dont les régulations sont potentiellement impliquées dans la stratégie d'hyperaccumulation (les métaux sont capturés par différents types de cellules foliaires et transportés de cellule en cellule à travers les membranes. La distribution intracellulaires des métaux est assurée par les métallochaperones et les transporteurs localisés dans les endomembranes).

CW: paroi cellulaire (Cell Wall); M: métal; ●: métal; ○: transporteur; ◐: métallochaperones

Comme le montre la figure 5, deux grands types "fonctionnels" de molécules interviennent dans l'homéostasie des métaux: les chélateurs, auxquels sont fixés les métaux transportés ou stockés, et les transporteurs, qui, à chacun des niveaux présentés, assurent le passage des ions métalliques entre les tissus et leur entrée dans les différents compartiments cellulaires. La nature, la biochimie et la localisation des chélateurs engagés dans les processus homéostatiques de base et, potentiellement, dans la tolérance des (pseudo)métallophytes, sont aujourd'hui de mieux en mieux connues (Hall 2002; Callahan *et al.* 2006; Krämer & Clemens 2006). Ces chélateurs sont cependant peu spécifiques et ne peuvent expliquer la grande spécificité des phénomènes de tolérance observés (Callahan *et al.* 2006). Cette spécificité reposerait surtout **sur la nature, la fonction et la régulation de transporteurs transmembranaires** (Williams *et al.* 2000).

De nombreuses études sont donc aujourd'hui consacrées à **l'identification chez *A. thaliana* des gènes codant pour ces transporteurs** (Clemens 2001; Clemens *et al.* 2002; Krämer & Clemens 2006). Identifier, parmi ces gènes ceux potentiellement impliqués dans la tolérance des métallophytes résulte ensuite soit (1) d'approches de génétique écologique ("ecological genetics", Purugganan & Gibson 2003), au cours desquelles des différences d'expression des gènes entre espèces tolérantes et non tolérantes sont recherchées en conditions environnementales contrôlées (Assunção *et al.* 2001; Becher *et al.* 2004; Dräger *et al.* 2004; Weber *et al.* 2004); soit (2) d'études d'association ("association studies", Purugganan & Gibson 2003) au cours desquels des correspondances sont recherchées entre la ségrégation de gènes candidats et la ségrégation du caractère de tolérance dans des descendances de croisement hybride entre génotypes tolérants et non tolérants (Willems 2006; Assunção *et al.* sous presse).

Les gènes candidats identifiés appartiennent tous aux quatre familles suivantes de protéines de transport transmembranaires des cations: les facilitateurs de la diffusion des cations (Cation Diffusion Facilitator, CDF, Krämer 2005), les transporteurs régulés par le Zn et/ou le Fe (Zinc-regulated transporter (ZRT)-Iron-regulated transporter (IRT) like Protein, ZIP, Guerinot 2000), les ATPases de type P (famille des HMA heavy-metal P-type ATPases), qui utilisent l'ATP pour transporter des ions à travers les membranes biologiques et sont caractérisées par la formation d'un composé phosphorylé (Williams *et al.* 2000), les protéines

Metal	Species	Number of genes	Reference
Aluminium	Barley	1 major gene	Reid (1971)
	<i>Zea mays</i>	1 major gene, with multiple alleles	Rhue <i>et al.</i> (1978)
	<i>Nicotiana glauca</i>	1 major gene	Connor & Meredith (1985)
	Wheat (cv. Chinese Spring)	1 major gene + at least 2 modifiers	Aniol (1990)
	Wheat (cv. BH1146)	3 major genes, + some minor genes	Aniol (1991)
Arsenic	Sorghum	1 major gene	Furlani & Bastos (1990)
	<i>Holcus lanatus</i>	1 major gene + modifiers	Macnair, Curnea & Meharg (1992)
Boron	<i>Agrostis capillaris</i>	1 major gene + modifiers	Watkins & Macnair (1991)
	Wheat	3 major genes	Paul, Rathgen & Cartwright (1991)
Cadmium	<i>Chlamydomonas reinhardtii</i>	2 major genes	Collard & Matagne (1990)
Copper	<i>Mimulus guttatus</i>	1 major gene + modifiers	Macnair (1983)
	<i>Silene vulgaris</i>	1 major gene + modifiers	Schat & ten Bookum (1992a)
Manganese	<i>Silene vulgaris</i>	2 major genes + modifiers	Schat <i>et al.</i> (1993)
	Wheat	2 major genes	Foy, Scott & Fisher (1988)

Note: In all cases, the major genes detected could be a number of tightly linked genes.

Tableau 4 Etudes rapportant l'existence d'un ou de quelques gènes majeurs à l'origine de la tolérance chez les végétaux supérieurs (d'après Macnair 1993)

Nramp (natural resistance-associated macrophage proteins) impliqués dans le transport des métaux chez de nombreux organismes (Thomine *et al.* 2000).

I.3 Architecture génétique de la tolérance aux métaux

I.3.1 Un déterminisme simple

I.3.1.1 Un ou quelques gènes à effets majeurs

Un grand nombre de caractères, notamment ceux associés à la croissance végétative, sont gouvernés par des systèmes polygéniques dans lesquels de nombreux gènes interagissent pour produire de la variation quantitative (Hartl & Clark 1997).

Les premières études génétiques de la tolérance aux métaux, et notamment l'analyse de la distribution des phénotypes tolérants dans des populations naturelles ou dans des descendance de croisements entre accessions tolérantes et non tolérantes (F2 et backcrosses), ont suggéré que la tolérance aux métaux lourds chez les plantes devait aussi être gouvernée par des systèmes polygéniques (Antonovics *et al.* 1971). Il est apparu cependant que les techniques de mesures de la tolérance d'alors, trop imprécises, ne pouvaient qu'aboutir à la description d'un caractère à variation quantitative, c.-à-d. déterminé par un nombre important de gènes. L'indice de tolérance, notamment, en raison des inconvénients évoqués plus hauts (voir I.1.2.2) ne permettait d'observer que des distributions phénotypiques continues, soutenant l'hypothèse d'un déterminisme polygénique (Macnair 1990, 1993).

Une revue analytique des résultats d'études obtenus sur des espèces sauvages et cultivées a mené Macnair (1990, 1993) à conclure que la tolérance aux métaux lourds pourrait, au contraire, n'être généralement déterminée que par **un petit nombre de gènes** (Tableau 4). Simultanément, l'analyse de la ségrégation du caractère dans des croisements hybrides entre populations NM et M de *M. guttatus* et *S. vulgaris*, soit deux espèces pseudométallophytes parmi les plus intensément étudiées, a mené à l'élaboration de modèles génétiques beaucoup plus simples: des gènes uniques, à effet majeur, contrôlant respectivement la tolérance au cuivre de *Mimulus guttatus* (Macnair 1983; Smith & Macnair 1998) et les tolérances au Cu et au Zn de *Silene vulgaris* (Schat & Ten Bookum 1992; Schat

et al. 1993; Schat *et al.* 1996). Chez *S. vulgaris*, néanmoins, les niveaux de tolérance les plus élevés engageraient un deuxième gène majeur (Schat & Ten Bookum 1992; Schat *et al.* 1993).

Réciproquement aux analyses de ségrégation, la constatation, chez *S. vulgaris*, que des populations M ayant évolué indépendamment vers la tolérance partagent un même déterminisme génétique, a soutenu l'hypothèse d'un déterminisme génétique simple (Schat *et al.* 2000). L'hypothèse est que la variabilité génétique nécessaire à l'évolution locale de la tolérance chez cette espèce pourrait n'être présente qu'à quelques gènes spécifiques, que donc le nombre de mécanismes "sélectionnables" est réduit, ce qui explique que les mêmes mécanismes soient sélectionnés dans des situations indépendantes. En conséquence, le nombre de gènes impliqués est lui-aussi limité.

L'hypothèse d'un **nombre limité de gènes impliqués dans la tolérance aux métaux**, élaborée au début des années 90 est devenue depuis **le nouveau paradigme** (Macnair *et al.* 2000; Schat *et al.* 2000). Elle contraste avec la multitude des voies métaboliques de l'homéostasie des métaux mises en évidence par la physiologie moléculaire et qui sont envisagées comme autant de mécanismes potentiels de tolérance (voir I.2.3). Elle est néanmoins fortement soutenue par les applications récentes des techniques nouvelles de biologie moléculaire. Des analyses QTL (Quantitative Trait Loci), notamment, qui couplent les analyses classiques de ségrégation de caractères quantitatifs à l'étude de la ségrégation de marqueurs moléculaires (Barton & Keightley 2002), ont ainsi attesté du rôle déterminant d'un seul locus QTL dans la tolérance à l'aluminium de la luzerne *Medicago sativa* subsp. *coerulea* (Sledge *et al.* 2002) et du blé *Triticum aestivum* (Ma *et al.* 2005), et de trois loci QTL dans la tolérance au zinc du riz *Oryza sativa* (Dong *et al.* 2006). Il reste néanmoins à déterminer combien de gènes sont révélés par un seul QTL. Ceci est rendu possible, par exemple chez *A. halleri*, par des études de synténie avec le génome séquencé d'*A. thaliana* et l'identification des gènes présents dans la région QTL.

1.3.1.2 Des gènes modificateurs à effets mineurs

Les gènes majeurs semblent cependant être plus responsables de l'existence de capacités de tolérance au sein d'un individu, en comparaison avec un individu non tolérant,

que des variations observées entre individus. Le degré d'expression du phénotype "tolérant" serait plutôt sous l'influence de gènes modificateurs (Macnair 1993; Macnair *et al.* 2000). A nouveau, les hypothèses faites sur l'existence de ces gènes modificateurs reposent surtout sur l'étude de la ségrégation du caractère dans des descendances de croisements entre individus de niveaux de tolérance variables chez *S. vulgaris* (Macnair 1993; Smith & Macnair 1998) et *M. guttatus* (Schat *et al.* 1993; Schat & Vooijs 1997).

I.3.2 Les interactions géniques: dominance et épistasie

Il existe deux grands types d'interactions géniques: les interactions intragéniques ou interalléliques (dominance/récessivité) et les interactions intergéniques (épistasie).

L'insertion, par croisements répétés, du gène de la tolérance au Cu d'une accession M dans le génome non tolérant d'une accession NM chez *M. guttatus* aboutit à un phénotype tolérant et met en évidence la dominance du gène de tolérance (Strange & Macnair 1991). De même, les deux gènes responsables de la tolérance au Cu de *S. vulgaris* sont à effets dominants et additifs (Schat *et al.* 1993). De façon générale, la majorité des tolérances aux métaux étudiées chez les plantes s'est révélée être contrôlée par **un ou quelques gènes à effets dominants et additifs** (Macnair 1993). Cette observation, en écho à la tendance générale à la récessivité des mutations entraînant des pertes de fonctions, suggère que la tolérance aux métaux est liée à la modification positive d'une fonction (comme l'amélioration d'une activité ou d'une spécificité, Macnair 1993).

Un **modèle épistatique** est requis pour expliquer les variations observées entre génotypes tolérants (Macnair 1983; Schat & Ten Bookum 1992): les gènes modificateurs ne peuvent pas conférer seuls la tolérance aux métaux et leur l'impact sur le phénotype global est tributaire de la présence des gènes majeurs (ils sont donc hypostatiques Macnair 1993; Smith & Macnair 1998).

Il existe une différence fondamentale entre un déterminisme polygénique et un **déterminisme multigénique** comprenant un ou plusieurs gènes majeurs dominants associés à des gènes modificateurs hypostatiques. Elle réside dans l'importance relative des effets des différents gènes. Alors que dans le premier modèle, les différents gènes sont codominants et

leurs rôles équivalents, dans le second, le(s) gène(s) majeur(s) est(ont) à la fois nécessaire(s) et suffisant(s) pour conférer la tolérance.

I.4 Evolution de la tolérance aux métaux lourds

La théorie originale, proposée par Darwin, qui explique l'évolution biologique par l'action de la sélection naturelle sur une variabilité héritable des propriétés individuelles reste, au-delà des précisions fondamentales apportées depuis (e.g. l'hérédité mendélienne, le néodarwinisme, la théorie de l'évolution neutre, *etc.*) l'hypothèse la plus acceptable pour rendre compte de l'évolution des organismes. Dans le cadre de la théorie, les **caractères adaptatifs** sont donc des traits **génétiquement contrôlés et héréditaires**, acquis par les organismes lorsqu'ils possèdent les **ressources génétiques** nécessaires pour répondre à une/des pression(s) de **sélection** exercée(s) par le milieu (en l'absence de contraintes ou d'autres forces antagonistes), et qui confèrent un **avantage sélectif** dans les conditions environnementales locales (Kawecki & Ebert 2004).

Dans ce contexte, la tolérance aux métaux lourds, parce qu'elle est surtout observée en milieux métallifères, a, très tôt, logiquement été envisagée comme la **réponse adaptative** des végétaux au stress engendré par les concentrations toxiques de métaux dans les sols sur lesquels ils se développent. Le contrôle génétique et l'**hérédité** du caractère sont admis depuis les travaux originels de Prat (Prat 1934) qui, semant des graines (vierges de toute expérience environnementale) de *S. vulgaris* provenant de populations NM et M sur un sol contaminé par le Cu, a montré que celles des populations M se développaient mieux (c.-à-d. toléraient la contamination). L'hérédité apparaît aussi dans la relation forte entre les indices de tolérance d'individus ou de populations d'*Agrostis tenuis* et ceux des graines issues de ces individus ou populations (McNeilly & Bradshaw 1968)¹. Enfin, elle est un postulat évident des études de ségrégation du caractère (voir I.3). Mais le mode d'évolution du caractère en milieu métallifère est peu compris. Il faudrait expliquer pourquoi certaines espèces acquièrent une ou des tolérances aux métaux quand d'autres en sont incapables dans les mêmes conditions, quels sont les mécanismes microévolutifs conditionnant la dynamique évolutive

¹ De rares cas de tolérance induites par des expositions antérieures ont cependant été rapportés (Macnair 1993).

Tableau 5 Pourcentages d'individus tolérants au Cu dans des populations NM d'espèces de Graminées, en relation avec la présence ou l'absence de l'espèce sur des sites pollués par le cuivre ou à proximité de ces sites (d'après Ingram, 1988, cité dans Baker & Proctor 1990). Les pourcentages sont exprimés après évaluation de 20 000 graines par espèce en culture hydroponique.

Species	% occurrence of tolerant individuals	Species presence or absence	
		on mines	at margins
<i>Holcus lanatus</i>	0.16	+	+
<i>Agrostis capillaris</i>	0.13	+	+
<i>Festuca ovina</i>	0.07	-	+
<i>Dactylis glomerata</i>	0.05	+	+
<i>Deschampsia flexuosa</i>	0.03	+	+
<i>Anthoxanthum odoratum</i>	0.02	-	+
<i>Festuca rubra</i>	0.01	+	+
<i>Lolium perenne</i>	0.005	-	+
<i>Poa pratensis</i>	0	-	+
<i>Poa trivialis</i>	0	-	+
<i>Phleum pratense</i>	0	-	+
<i>Cynosurus cristatus</i>	0	-	+
<i>Alopecurus pratensis</i>	0	-	+
<i>Bromus</i> spp.	0	-	+
<i>Arrhenatherum elatius</i>	0	-	+

de la tolérance, quelle niveau de divergence génétique entre accessions tolérantes et non tolérantes engendre l'évolution du caractère (observe-t-on véritablement des différenciations adaptatives? S'accompagnent-elles d'isolement reproducteur?), mais aussi, parce qu'il reste parfois peu évident, quel est la nature du rôle de l'exposition aux métaux dans l'évolution du caractère.

I.4.1 Tolérer ou ne pas tolérer: les hasards de l'adaptation

Seule une minorité d'espèces a évolué vers la tolérance, souvent à un ou seulement quelques métaux, et est rencontrée sur les sols métallifères. L'existence au sein d'une espèce d'une diversité génétique appropriée est probablement l'élément fondamental conditionnant d'une part sa capacité à coloniser un milieu pollué et, d'autre part le type de stratégie de tolérance vers laquelle elle évoluera.

I.4.1.1 Tolérance et genostasis

En testant la présence d'individus tolérants au Cu dans des populations NM de plusieurs espèces de Graminées, Ingram (1988, cité dans Baker & Proctor 1990) a montré que des individus tolérants étaient présents, en faible fréquence, dans les populations de certaines espèces (Tableau 5). Les populations NM d'espèces pseudométallophytes par ailleurs rencontrées sur sol contaminé par le Cu comprennent clairement des individus tolérants au Cu, "pré-adaptés" aux milieux métallifères. Inversement, les individus des populations NM d'espèces non métallophytes sont tous non tolérants. Ce constat a mené Bradshaw (1991) à soutenir que la tolérance aux métaux ne peut pas évoluer chez certaines espèces (par conséquent non métallophytes) en raison de l'absence, dans le génome de ces espèces, de la variabilité génétique nécessaire à l'évolution de la tolérance (*genostasis*).

Dans l'hypothèse où la tolérance est contrôlée par quelques gènes majeurs (voir I.3.1.1), la faible fréquence des individus tolérants dans les populations NM des espèces pseudométallophytes suggère une faible fréquence des gènes de tolérance dans les populations NM (Macnair 1997). En population NM, en l'absence de sélection (voir I.4.2.1), et, éventuellement, en l'absence de contre-sélection ou coût (voir I.4.2.2), l'évolution de la fréquence de ses gènes est fort probablement soumise à des forces stochastiques. Des variations de fréquences sont attendues entre populations, avec certaines populations montrant des fréquences nulles, et la variabilité nécessaire à l'évolution d'une tolérance peut-être

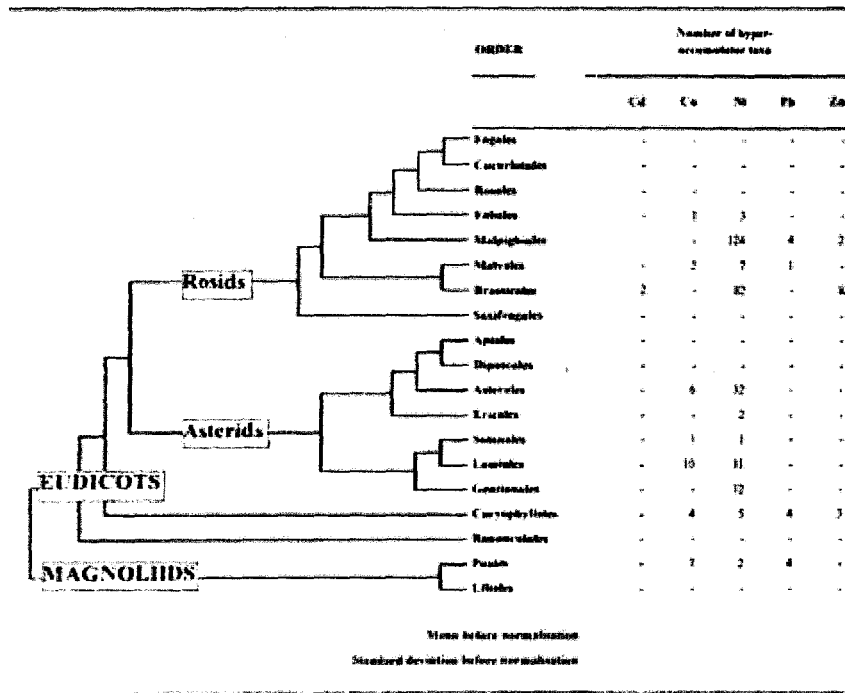


Figure 6 Distribution des espèces hyperaccumultrices dans les différents ordres d'Angiospermes (d'après Broadley *et al.* 2001). On observe la concentration importante des hyperaccumulateurs (tous métaux confondus) dans certains ordres phylogénétiquement éloignés

présente dans le patrimoine génétique globale d'une espèce pseudométallophyte mais localement absente dans une ou plusieurs de ses populations NM (Macnair 1997). C'est ainsi qu'en comparant les situations répétées de pollution du sol par l'installation de pylônes électriques galvanisés (Zn), Al-Hiyaly *et al.* (1988; 1993) ont montré que les populations d'*A. tenuis* avoisinant ces pylônes n'étaient pas toutes capables de fonder des populations tolérantes. L'examen détaillé de ces populations suivant la méthode d'Ingram a révélé que certaines d'entre elles ne comprenaient aucun individus tolérants (Al-Hiyaly *et al.* 1993).

Enfin, on remarque dans le Tableau 5 que certaines espèces, bien que possédant la variabilité nécessaire et au contact de sites pollués, ne fondent pas localement de population M (*e.g. Festuca ovina*). Il est probable que ces espèces soient exclues des sites pollués par une incapacité (variable entre populations) à tolérer les autres contraintes environnementales de ces sites inhospitaliers (Macnair 1997). **La tolérance physiologique des métaux pourrait n'être pas suffisante pour coloniser un site métallifère et fonder une population M.**

1.4.1.2 Tolérance et phylogénie

Les premières synthèses sur la tolérance aux métaux chez les plantes supérieures ont suggéré que le phénomène était surtout **particulier à quelques genres et espèces** dans les familles Caryophyllaceae, Lamiaceae et Poaceae (Baker 1987). Pareillement, Baker & Proctor (1990) rapportent que la majorité des espèces largement représentées sur les sols pollués de Grande Bretagne appartiennent aux familles Poaceae, Brassicaceae et Caryophyllaceae et que d'autres grandes familles sont étrangement absentes (Cyperaceae, Asteraceae, Rosaceae et Apiaceae). **Il est pourtant clair que des tolérances aux métaux sont apparues indépendamment dans de nombreuses familles** (Tableaux 2 et 3). Et l'absence de certaines familles, notamment d'espèces ligneuses, pourrait n'être que le résultat d'un biais lié aux difficultés rencontrées dans l'étude, en milieu contrôlé, de ces espèces, en comparaison des espèces de graminées et d'herbacées beaucoup plus abordables (Macnair 1997).

Si la variabilité génétique nécessaire à l'évolution de tolérances aux métaux est effectivement présente dans de nombreux taxa, il reste qu'elle semble inégalement distribuée entre taxa. Ainsi, une méta-analyse de données de concentrations en métaux dans les organes aériens d'Angiospermes provenant de 44 études montre par exemple une **influence significative de la phylogénie** sur les capacités d'(hyper)accumulation du Cd, Cr, Cu, Ni, Pb (Figure 6, Broadley *et al.* 2001). Parmi les 27 familles d'Ericales définies, 18 ne comportent

aucun accumulateurs de Al lorsque les 9 autres sont soit partiellement soit presque totalement constituées d'accumulateurs de Al (Jansen *et al.* 2004). De même, dans le genre *Alyssum* (Brassicaceae), qui comportent 48 des 317 espèces hyperaccumulatrices recensées (Pollard *et al.* 2002), le caractère semble avoir été acquis et perdu plusieurs fois indépendamment (Mengoni *et al.* 2003).

1.4.1.3 Exclure ou accumuler?

Les stratégies de tolérance (exclusion-hyperaccumulation) diffèrent entre espèces et, au sein d'une espèce, entre métaux. **La stratégie adoptée par une espèce pour un métal donné pourrait dépendre de la position systématique de l'espèce mais aussi du métal toléré.**

Les résultats d'analyses phylogénétiques (voir 1.4.1.2) suggèrent un biais de la phylogénie dans la distribution des stratégies adoptées, en soulignant que l'hyperaccumulation est plus répandue que l'exclusion chez certains taxa. On observe par exemple une **surreprésentation des métallophytes hyperaccumulatrices chez les Brassicaceae**, avec 87 des 418 espèces hyperaccumulatrices recensées (Prasad & Freitas 2003), et notamment 82 des 317 espèces accumulatrices de Ni. Inversement, Reilly et Reilly (1973, cités dans Baker & Walker 1990) ont montré, en calculant le rapport [organes aériens] / [racines] pour le cuivre chez un grand nombre d'espèces, non seulement que ce rapport variait entre espèces (avec des valeurs allant de 0,003 à 2,61), mais surtout que les valeurs les plus faibles (i.e. exclusion) étaient typiques des graminées alors que les valeurs les plus élevées (i.e. (hyper)accumulation) caractérisaient plutôt les autres espèces herbacées et les ligneuses, ce qui suggérait à nouveau **un biais phylogénétique dans la distribution des stratégies de tolérance**. Si les deux stratégies ont évolué dans de nombreuses familles de plantes vasculaires (Tableaux 1 et 3), il semble donc que la nature des ressources génétiques disponibles diffèrent entre taxa et que la stratégie de tolérance adoptée dépende de la position taxonomiques des espèces.

Il ressort enfin de la bibliographie que, même si l'hyperaccumulation est une stratégie de tolérance à de nombreux métaux (Tableau 3), **une majorité des espèces hyperaccumulatrices sont hyperaccumulatrices de Ni**, suggérant que la stratégie de tolérance dépende aussi du métal toléré. Cette observation doit néanmoins être pondérée par

le focus important porté sur les espèces serpentiniques (Brooks 1987; Baker *et al.* 1992; Reeves *et al.* 1999; Brady *et al.* 2005), par définition tolérantes au Ni.

Enfin, sans qu'aucune explication soit avancée, il est souvent admis que **la majorité des pseudométallophytes sont excluantes lorsque l'hyperaccumulation est surtout rencontrée chez les métallophytes absolues** (Baker & Proctor 1990; Pollard *et al.* 2002). Cette "règle" souffre bien évidemment d'exceptions : les capacités d'exclusion du cuivre d'*Elsholtzia haichowensis*, pourtant endémique des sols contaminés par le Cu (Lou *et al.* 2004); plus connues, les capacités d'hyperaccumulation d'*Arabidopsis halleri* (Bert *et al.* 2000) et de *Thlaspi caerulescens* (Meerts & van Isacker 1997), pourtant pseudométallophytes.

I.4.2 Dynamique évolutive de la tolérance aux métaux

I.4.2.1 Une adaptation rapide

L'existence des sites contaminés d'origine anthropique, et dont l'histoire est relativement récente, a permis de suivre en détail l'établissement de populations tolérantes et ainsi la dynamique de l'adaptation à des concentrations toxiques aux métaux. Il est apparu que **l'évolution vers la tolérance était rapide dans les populations M**. Ainsi, des populations M d'*Agrostis stolonifera* âgées de 4 à 14 ans montrent déjà des capacités de tolérances élevées alors que celles des populations NM sont faibles (Wu *et al.* 1975). Mais l'un des exemples les plus illustratifs concerne l'évolution de la tolérance au Zn chez *A. tenuis*. Al-Hiyaly *et al.* (1988) ont montré que des populations exposées au Zn à la suite de l'installation, 20 à 30 ans auparavant, de pylônes galvanisés possédaient des capacités de tolérance parfois comparables à celles de populations installées sur des sites miniers beaucoup plus anciens.

Dans les deux exemples cités, les niveaux de tolérance varient entre populations d'âges différents. Il apparaît, notamment chez *A. stolonifera*, que, à pollution égale, des populations plus anciennes (70 ans) montrent des niveaux de tolérance nettement plus élevés (Wu *et al.* 1975). De façon générale, la **sélection secondaire** vers une tolérance accrue serait un phénomène commun dans les populations M d'espèces métallophytes (Macnair 1993).

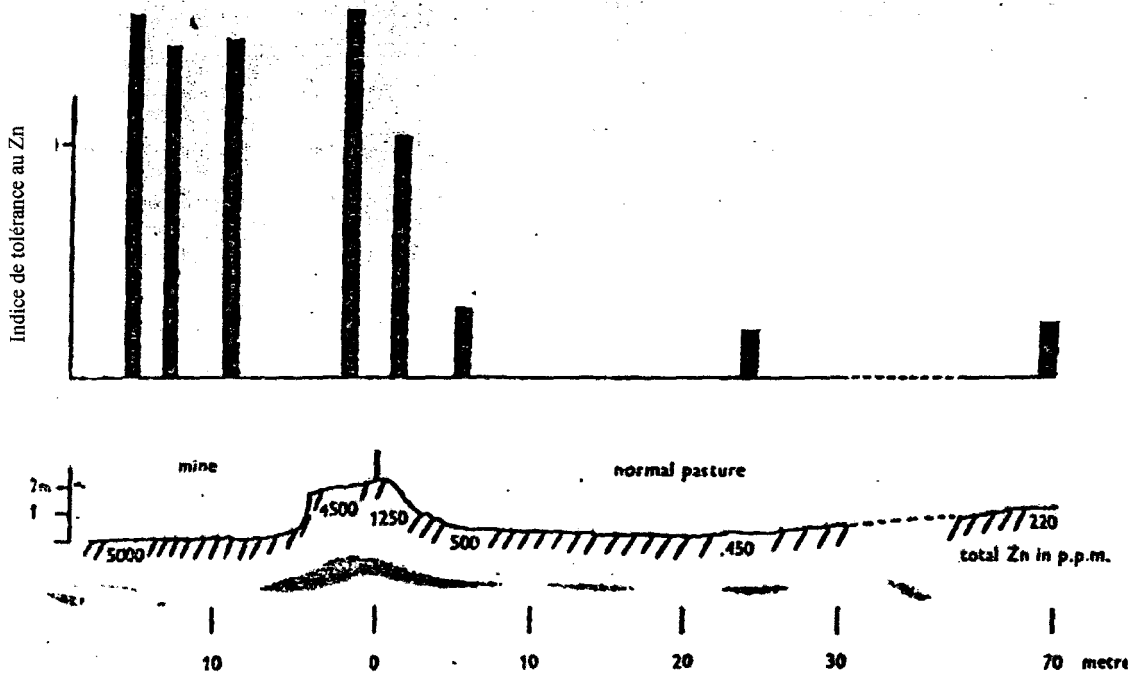


Figure 7 Indices de tolérance au Zn de populations d'*Anthoxanthum odoratum* autour d'une mine de Zn (d'après Jain & Bradshaw 1966). On observe une transition rapide dans ces capacités de tolérance de part et d'autre de la frontière séparant un sol minier et une prairie peu contaminée.

A partir de ce double constat (évolution rapide de capacités basiques de tolérance et sélection secondaire de niveau de tolérance plus forts), et en s'appuyant sur le déterminisme génétique probable de la tolérance et de ses variations quantitatives (voir I.3.1), Macnair (1997) a suggéré un **scénario pour l'évolution de la tolérance dans les populations M**. Selon lui, la sélection rapide, en milieu métallifère, du ou des gène(s) majeur(s) codant pour le caractère suffirait à la colonisation de ces milieux; dans un second temps, une adaptation plus fine, par exemple vers plus de tolérance sur les sols les plus pollués, engagerait la sélection des gènes modificateurs hypostatiques.

La dynamique de la tolérance aux métaux est probablement conditionnée par les effets antagonistes de la sélection et les flux de gènes entre population M et NM et par la nature du déterminisme génétique pour le caractère (Macnair 1993). Il a été montré que la sélection exercée sur site pollué contre les génotypes non tolérants était très forte (voir pour des estimations Jain & Bradshaw 1966). L'intensité de la sélection apparaît aussi dans l'expérience de (Wu *et al.* 1975). Après avoir semé, sur un sol contaminé par une raffinerie de Cu, des graines (2000 par population) d'*A. stolonifera* provenant de 5 populations installées sur des sols non contaminés par le Cu, les auteurs ont au mieux comptabilisé 0,1% de survivants au bout de 9 mois de culture. Dans cet exemple, les génotypes non tolérants sont éliminés en une seule génération. Une telle **intensité de la sélection** et la **dominance des gènes majeurs** de tolérance (voir I.3.2) expliquent que, **même en présence de flux de gènes depuis des populations NM adjacentes**, l'évolution vers un premier niveau de tolérance se produise rapidement (Lefèbvre & Vernet 1990). Elles expliquent aussi que le passage vers de capacités de tolérance importantes soit brutal lorsque l'on passe d'une population NM à la population M adjacente (Figure 7).

1.4.2.2 ...avec (ou sans) bottleneck...

La sélection forte exercée pour la tolérance devrait faire subir à la population M un goulot d'étranglement (bottleneck) fort qui, ajouté à de potentiels effets de fondation, résulterait en une diminution importante de la diversité génétique intra-population. Une **diversité génétique réduite est donc attendue dans les populations M** des espèces pseudométallophytes, en comparaison de leurs populations NM (Lefèbvre & Vernet 1990;

Wu 1990). Mais si l'**attendu théorique** est fort, il est **peu confirmé par les données empiriques** de génotypage nucléaires obtenues jusqu'ici. Avec 1 ou 2 marqueurs allozymiques, Wu *et al.* (1975) et Ducouso *et al.* (1990) ne mettent pas en évidence de différence de diversité génétique entre populations M et NM chez *A. stolonifera* et *A. elatius*. Avec un nombre plus important de loci allozymiques utilisés, Bush and Barrett (1993), Vekemans et Lefèbvre (1997) et Nordal *et al.* (1999) détectent des différences chez *Deschampsia cespitosa*, *Armeria maritima* et *Lychnis alpina* mais insuffisantes pour attester la réalité d'effets de fondation importants.

L'existence d'un **flux de gènes nucléaires depuis les populations fondatrices** est souvent invoquée pour expliquer l'absence de réduction de la diversité génétique. Un tel flux est supposé restaurer la diversité génétique des populations M après leur fondation (Wu *et al.* 1975; Ducouso *et al.* 1990; Mengoni *et al.* 2001). Cet effet devrait ne concerner que la diversité aux marqueurs nucléaires et rester sans conséquence sur les loci cytoplasmiques. L'hypothèse est appuyée par l'observation, chez *Silene paradoxa*, d'une réduction de la diversité chloroplastique dans une population M alors qu'aucune différence n'était observée sur des marqueurs RAPD nucléaires (Mengoni *et al.* 2001).

I.4.2.3 ... peu coûteuse...

Plusieurs observations ont mené à penser que les plantes tolérantes pouvaient être désavantagées par rapport à des plantes conspécifiques sur sol non contaminé, autrement dit que la tolérance pouvait comporter un **coût**. Tout d'abord, la faible fréquence des individus tolérants dans les populations NM de la plupart des pseudométallophytes et les clines importants observés à la frontière de sites miniers, suggèrent l'existence d'un tel coût (Figure 7). Ensuite, les comparaisons des taux de croissance et de compétitivité (revues dans Macnair *et al.* 2000) montrent souvent des capacités moindres chez les plantes tolérantes¹. Enfin, il semblerait aussi que ces dernières aient des besoins essentiels en métaux plus

¹ Les modifications physiologiques que demande la tolérance aux métaux lourds (voir I.2.3) engendreraient l'existence de dépenses métaboliques additionnelles qui réduiraient la disponibilité des ressources pour les autres fonctions biologiques.

importants que les plantes non tolérantes, potentiellement insatisfaits en milieu non métallifère (Macnair *et al.* 2000).

Ces observations et hypothèses suggèrent un vrai désavantage des plantes M sur sol non pollué. Mais elles ne montrent pas que la tolérance elle-même est coûteuse. Les études s'attachant à ne regarder que le coût de la tolérance aux métaux (entre autres caractères liés à l'adaptation aux sites métallifères) sont moins concluantes. En conditions contrôlées, la comparaison de lignées isogéniques de *Mimulus guttatus* ne différant que par la présence de gène(s) de tolérance au Cu ne permet pas de mettre en évidence un coût à cette tolérance en conditions non polluées (Harper *et al.* 1997). La comparaison d'accessions d'*Agrostis tenuis* tolérantes et non tolérantes mais sélectionnées dans une même population NM (et donc présentant une probable proximité génétique globale) échoue à montrer l'existence d'un coût (Nicholls & McNeilly 1985). En résumé, il n'existe aujourd'hui **aucune évidence indiscutable d'un coût direct à la présence de gènes de tolérance aux métaux en milieu non métallifère.**

Les sols miniers sont généralement pauvres en nutriments, sujets à la sécheresse en raison du manque de matière organique et d'une texture plus grossière: des adaptations à ces conditions environnementales particulières sont attendues (Macnair 1987). Ainsi, en plus de la tolérance aux métaux, des différences morphologiques (tailles des plantes, des fleurs, des tiges...) et physiologiques (tolérance au soufre, aux carences en calcium et phosphate) aux métaux ont souvent été rapportées entre accessions M et NM, même si elles n'ont que rarement été mises en évidence en conditions contrôlées (Bradshaw 1970; Antonovics *et al.* 1971). Si les populations M montrent un désavantage en milieu non pollué, celui-ci serait ainsi plus dû au **complexe d'adaptations** requis par la colonisation des sols métallifères.

Enfin, si les gènes majeurs de la tolérance ne montrent **pas de coût évident**, l'observation d'une disparition plus rapide des niveaux élevés de la tolérance que de la tolérance elle-même dans une population de *M. guttatus* à la frontière d'une mine de Cu a mené Macnair (1993) à suggérer qu'**une contre-sélection pourrait agir sur les gènes modificateurs**. Un tel coût n'a cependant pas pour le moment pu être mis en évidence expérimentalement (Harper *et al.* 1997; Harper *et al.* 1998).

1.4.2.4 ...et répétée.

La distribution des sites pollués d'origine anthropique est caractérisée par une grande **discontinuité géographique** (Bradshaw 1970; Buse *et al.* 2003) et la distance entre chaque site est souvent trop importante pour être couverte par les flux de gènes (Schat *et al.* 1996). Deux hypothèses simples ont été confrontées pour expliquer la colonisation, par des espèces pseudométallophytes, de sites pollués géographiquement distant. Soit la tolérance n'a évolué qu'une fois durant la colonisation d'un site métallifère puis s'est répandu entre les sites; soit elle a, chaque fois, évolué à partir de populations NM différentes, et de façon indépendante. La majorité des études menées avec l'objectif de tester les deux hypothèses, dont celles sur *Agrostis capillaris* (Al-Hiyaly *et al.* 1988), *Armeria maritima* (Vekemans & Lefèbvre 1997), *Deschampsia cespitosa* (Bush & Barrett 1993), *Silene dioica* (Westerbergh & Saura 1992) ou *Silene paradoxa* (Mengoni *et al.* 2001) ont conclu à l'**indépendance des origines des populations M**.

I.4.3 Tolérance et isolement reproducteur

En théorie, des populations de plantes se développant sur des types de sols très contrastés doivent subir des sélections divergentes suffisamment fortes pour que la différenciation des populations s'accompagne d'isolement reproducteur, et que, dans les cas extrêmes, celui-ci résulte en spéciation écologique (Schluter 2001).

I.4.3.1 Tolérance et isolement reproducteur entre populations

L'isolement reproducteur est un pré-requis à l'adaptation locale des populations sur les sites contaminés (Kawecki & Ebert 2004). Il permet la transmission du caractère adaptatif à la descendance et la survie de la population. Chez diverses espèces pseudométallophytes, des différences morphologiques ou physiologiques (génétiquement déterminées) entre populations M et NM ont été observées (Antonovics *et al.* 1971). Par exemple, des populations d'*Anthoxanthum odoratum* de part et d'autre d'un mine de Zn et de Pb différent, malgré les probables flux géniques, par la taille des plantes, la taille des feuilles ou encore le nombre de hampes florales (Antonovics & Bradshaw 1970)¹.

¹ Les auteurs ne mettent pas en évidence de corrélations entre les variations pour les caractères morphologiques et la tolérance. Les différences sont attribuées aux nombreuses pressions de sélection s'exerçant sur sites pollués (voir I.1.1), en parallèle et indépendamment de la sélection pour la tolérance aux métaux.

The Drws-y-coed transect (*Agrostis*)

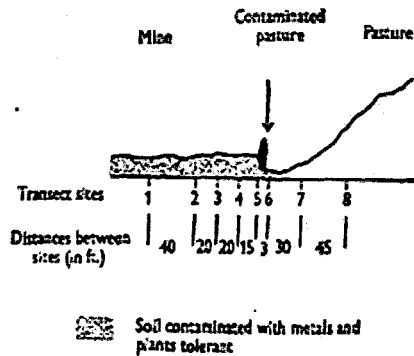
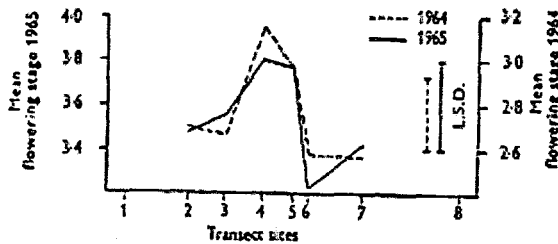


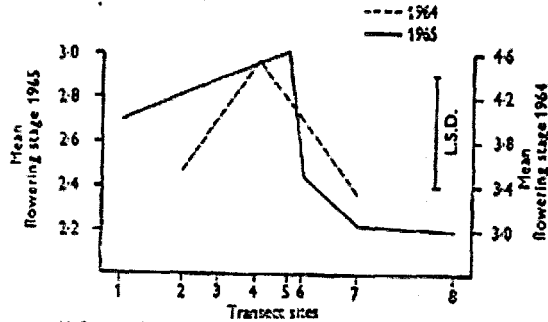
Figure 8 Décalage de floraison entre populations d'*Agrostis tenuis* le long d'un transect (I) traversant la frontière entre un site minier et la pelouse non polluée voisine (d'après McNeilly & Antonovics 1968). Les niveaux de floraison sont exprimés pour une date fixe sur une échelle de 1 à 6, en situation naturelle (IIa) ou en conditions contrôlées (IIb). Enfin, le parallèle est fait avec la sécheresse des sols (IIc). On observe un décalage fort au niveau de la frontière entre la mine et la pelouse, soit entre les sites 5 et 6

I —Transects across the boundaries between mine and pasture showing positions of sites from which plants were sampled.

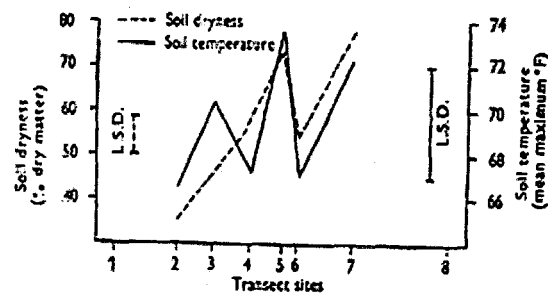
(a) Flowering time in field.



(b) Flowering time in greenhouse.



(c) Ecological conditions



II —Flowering time and ecological conditions along the Drws-y-Coed (*Agrostis*) transect.

Mais l'existence de variations morphologiques n'est pas en soit indicateur de l'établissement d'un isolement reproducteur entre les populations. Chez les espèces allofécondantes, la possibilité d'un isolement reproducteur entre populations doit être recherchée par la mise en évidence de mécanismes d'isolement, soit **directs**, comme la mise en place de barrières à la reproduction entre phénotypes dissemblables (*e.g.* décalage des phénologies de floraison, évolution vers l'autofécondation...), soit **indirects**, par effet pléiotropique du gène d'adaptation ou par liaison de ce gène avec un gène d'incompatibilité.

Des divergences dans les périodes de floraison sont souvent observées entre populations M et NM d'espèces pseudométallophytes (Lefèbvre & Vernet 1990). Un exemple évocateur concerne des divergences des périodes de floraison entre populations M et NM voisines chez *Agrostis capillaris* et *Anthoxanthum odoratum* (McNeilly & Antonovics 1968). Chez ces espèces, les divergences observées entre accessions étaient d'autant plus fortes que les observations étaient proches de la frontière entre les populations (Figure 8), ce qui suggèrerait que la sélection divergente était d'autant plus forte que le flux de gènes entre phénotypes dissemblables était probable (McNeilly & Antonovics 1968). Mais un net décalage de floraison n'est pas toujours observé et des situations équivoques sont aussi rencontrées (Lefèbvre & Vernet 1990). Par exemple, une forte diversité de types phénologiques a été enregistrée dans des populations M d'*Arrhenatherum elatius*, avec des phénotypes parfois peu divergents des phénotypes NM voisins (Ducousso *et al.* 1990). Enfin, lorsqu'il existe, un décalage de floraison peut avoir évolué sur site métallifère pour des raisons autre que la présence de métaux, comme la sécheresse du sol (McNeilly & Antonovics 1968). Il ne participe pas moins à l'adaptation globale aux sites métallifères.

Des taux d'autofécondation plus importants dans les populations M que dans les populations NM ont été observés chez *Anthoxanthum odoratum* et *Agrostis tenuis* (Antonovics 1968; McNeilly & Antonovics 1968), *Armeria maritima* (Lefebvre 1970) et *Arrhenatherum elatius* (Cuguen *et al.* 1989). Mais à l'opposé, les taux d'autofécondation sont par exemple plus importants dans les populations NM de *T. caerulescens* que dans ses populations M (Dubois *et al.* 2003). Il semblerait au final que les différences de régime observées, parfois d'ailleurs dans des populations M et NM géographiquement trop distantes pour que l'évolution de barrières aux flux de gènes puissent constituer une hypothèse solide, soient plus attribuables à des différences de densité de populations (Lefebvre 1970; Ducousso

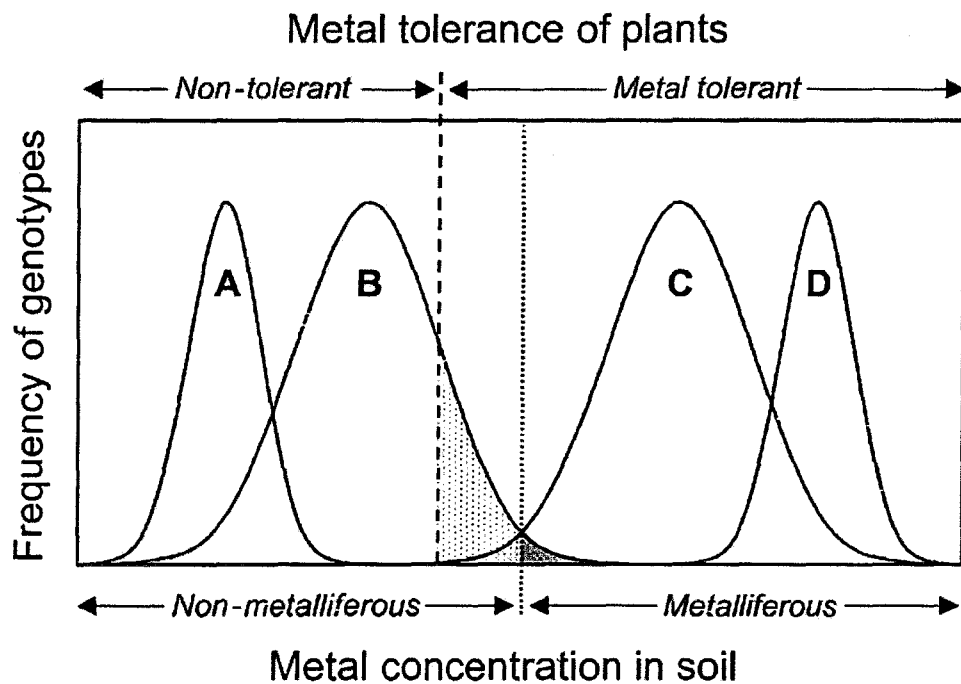


Figure 9 Distributions écologiques des espèces végétales et fréquences des génotypes non tolérants/ tolérants illustrées pour quatre espèces théoriques (d'après Pollard *et al.* 2002). L'espèce A représente les non métalrophytes absolues, chez lesquelles on ne rencontre aucun génotype tolérant. L'espèce B représente les pseudométalrophytes présentant un polymorphisme qualitatif pour la tolérance. Les génotypes tolérants sont rares dans les populations NM de l'espèce mais fixés dans les populations M. En moyenne, la tolérance est donc caractéristique des populations M, même si celle si sont rares. On parle de tolérance écotypique ou populationnelle. L'espèce C représente les pseudométalrophytes dont la tolérance est constitutive. On ne rencontre pas chez cette espèce d'individus non tolérants. Cette espèce est surtout rencontrée sur sol pollué et les populations NM son rares. L'espèce D représente les métalrophytes absolues dont la tolérance est constitutive.

et al. 1990; Dubois *et al.* 2003). Le niveau d'autofécondation plus important dans les populations M témoignerait alors simplement de l'avantage de ce régime lors de la colonisation de nouveaux milieux (pollués), lorsque les individus se trouvent en faible densité (Lefebvre 1970).

En 1983, Macnair et Christie (1983) décrivait chez *Mimulus guttatus* un système génétique simple produisant un isolement reproducteur entre populations de *M. guttatus* (une population était tolérante, deux ne l'étaient pas). Les populations incompatibles s'opposaient surtout par leur tolérance au Cu. En associant tolérance et incompatibilité, les auteurs ont supposé que l'isolement reproducteur entre populations tolérantes et non tolérantes devait faire intervenir **un gène d'incompatibilité en liaison étroite avec, ou pléiotropique de, un gène de tolérance au cuivre**, concluant à la génération possible d'un isolement reproducteur à travers la sélection d'un gène d'adaptation pour la tolérance.

1.4.3.2 Tolérance et spéciation écologique

Si l'étude des pseudométallophytes ne montre pas de tendance générale, le rôle de la colonisation des sites pollués dans la formation de nouvelles espèces est particulièrement évoqué chez les espèces endémiques des milieux métallifères (métalloendémiques). La marche vers la formation des espèces endémiques des sols serpentiniques a notamment été modélisé par (Kruckeberg 1986; Kruckeberg & Kruckeberg 1990). Kruckeberg propose une série d'étapes depuis la colonisation d'un sol serpentinique par une population et la formation d'un écotype serpentinique, jusqu'à un isolement reproducteur total avec la population d'origine et la formation consécutive d'une espèce endémique (Tableau 6). Dans ce scénario, l'isolement reproducteur est la conséquence secondaire de la différenciation génétique due à la **sélection divergente** exercée sur les traits adaptatifs selon la nature du sol. Le modèle de Kruckeberg correspond au modèle le plus simple de spéciation écologique (Schluter 2001).

Pour tester le modèle de Kruckeberg, Macnair et Gardner (1998) ont étudié en détail le niveau d'isolement reproducteur entre taxa du complexe d'espèces autour de *Mimulus guttatus*. Ce complexe comprend: des populations *M. guttatus* NM et non tolérantes aux métaux, des écotypes *M. guttatus* tolérants au Cu et ne présentant qu'un faible isolement reproducteur avec les populations NM, des espèces endémiques des sols serpentiniques (*M. pardalis* et *M. nudatus*) et une espèce endémique de sols cuprifères d'origine anthropique (*M. cupriphilus*). Les auteurs ont rapproché, dans ce complexe, les événements de spéciation (et

Tableau 6 Différentes étapes menant à la formation d'une espèce endémique édaphique (D'après Kruckeberg 1986)

Etape	Evénement
0.	Un certain niveau de préadaptation existe pour la tolérance aux sols serpentiniques dans les populations normales d'une espèce. Il offre les conditions qui permettront l'évolution d'un écotype tolérant aux sols serpentiniques
1.	La sélection divergente provoque la séparation de l'espèce en pools génétiques tolérant et non tolérant. Il y a formation d'un écotype serpentinique
2.	Une différenciation de traits structuraux et fonctionnels s'effectue au sien de l'écotype
3.	L'isolement reproducteur entre les deux écotypes est total, de telle sorte qu'ils ne peuvent plus échanger de gènes
4.	La divergence entre l'espèce en formation et l'espèce ancestrale s'accroît. Il y a formation d'une espèce endémique.

notamment la formation d'espèces endémiques à partir de l'espèce ancestrale probable *M. guttatus*) de la colonisation et l'adaptation aux sols métallifères, et conclu que **la colonisation des sols métallifères pouvait engendrer un isolement reproducteur fort allant jusqu'à la différenciation en espèces**. L'existence d'écotypes *M. guttatus* M, interfécond avec les populations NM de l'espèce, suggérait cependant que cette liaison n'était pas systématique. **Dans les cas où l'adaptation n'engendre qu'un faible isolement reproducteur, la spéciation n'a jamais lieu et la différenciation reste écotypique.**

La comparaison par Macnair et Gardner (1998) d'espèces endémiques de sols serpentiniques et d'un écotype minier est limitée par les échelles de temps respectives sur lesquelles l'isolement reproducteur a pu s'établir. L'**âge des sols** naturellement enrichis en métal sur lesquels les espèces serpentiniques sont rencontrées est bien plus ancien que les quelques centaines d'années ou moins du sol d'origine anthropique colonisé par l'écotype. (Brooks 1987). C'est pourquoi on suppose parfois que le faible isolement entre populations M et NM d'espèces pseudométallophytes est lié à la courte histoire des populations M, trop courte pour permettre l'évolution de mécanismes d'isolement (Ducouso *et al.* 1990). Néanmoins, le cas de l'espèce *M. cupriphilus*, endémique d'un sol d'origine anthropique, suggère que le temps seul n'est pas une raison suffisante pour qu'une population passe du stade d'écotype à celui d'espèce à part entière (Macnair & Gardner 1998). L'intensité de la **réduction du flux de gènes** engendrée par l'adaptation aux sols métallifères semble au final bien être le facteur majeur conditionnant le devenir (écotype ou nouvelle espèce) des populations M (Macnair & Gardner 1998).

1.4.3.3 Isolement reproducteur et distribution de la tolérance

En fonction du niveau d'isolement reproducteur entre des accessions M et NM, la distribution de la tolérance au sein d'une espèce varie. **Chez les espèces pseudométallophytes, la tolérance n'est le plus souvent qu'"écotypique"**, et ne concerne que les populations / écotypes de l'espèce présents sur sol pollué. Chez les **métalloendémiques**, au contraire, la tolérance, probablement impliquée dans la formation de l'espèce (voir 1.4.3.2), est **constitutive**, et concerne toutes les populations de l'espèce. En conséquence, dans les schémas généraux, l'amplitude écologique d'une espèce, c.-à-d. sa distribution sur différents types édaphiques, est souvent étroitement mise en correspondance

avec la distribution de la tolérance dans ses populations (Figure 9). Et chez les espèces pseudométallophytes, les caractères "métallicole" et "tolérant" sont souvent confondus.

I.4.4 La tolérance aux métaux: une adaptation locale aux milieux métallifères?

Les pseudométallophytes sont surtout étudiées sur les sites dont la pollution est d'origine anthropique (sites miniers ou industriels). Ils sont pour cela parfois appelées "mine taxa" (Brooks 1987). Chez ces espèces, la tolérance aux métaux apparaît clairement comme le **résultat de la sélection naturelle** agissant en raison d'une pression forte exercée par des concentrations toxiques de métaux (Turner 1969; Antonovics *et al.* 1971; Macnair 1987). Par exemple, à l'échelle micro-géographique, les relations entre indice de tolérance au Zn de populations d'*Anthoxanthum odoratum* distantes d'à peine quelques mètres et quantités de Zn dans le sol sur lesquels se développent ces populations, soulignent très fortement le caractère adaptatif de la tolérance (Figure 7). Plus indirectement, la grande spécificité des tolérances à un métal (voir I.2.1) qui, chez les espèces pseudométallophytes, ne sont presque exclusivement trouvées que dans les populations exposées à ce métal en milieu naturel, suggère encore une tolérance adaptée à la pression métallique (Schat & Vooijs 1997).

La tolérance des espèces pseudométallophytes est souvent associée à une stratégie d'exclusion (voir I.4.1.3). Chez ces espèces, **des différences de capacités d'exclusion sont mises en évidence entre accessions M et NM (Figure 3) et apparaissent clairement comme une réponse adaptative à l'exposition aux métaux.** Récemment cependant, quelques cas de tolérance par exclusion constitutive ont été rapportés chez des espèces pseudométallophytes de zones humides (Matthews *et al.* 2005), pouvant suggérer une origine de la tolérance indépendante de l'exposition aux métaux¹.

¹ L'existence de capacités de tolérance dans les populations NM de ces espèces serait expliquée par la biogéochimie de la rhizosphère. La nature des sols de zones humides faciliterait la mobilisation des métaux et un enrichissement des zones en contact immédiat avec les racines (Matthews *et al.* 2005). Là toujours donc, les capacités d'exclusion des métaux seraient une adaptation à l'exposition aux métaux.

Le rôle fondamental joué par la pression de sélection métallique dans l'évolution de la tolérance chez les métallobytes hyperaccumulatrices est plus discuté. Parce que les espèces hyperaccumulatrices sont surtout des métallobytes absolues (voir I.4.1.3), il est impossible de lier solidement l'acquisition de la tolérance à l'exposition aux métaux, comme cela est fait après comparaison des capacités de tolérance d'accessions M et NM d'espèces pseudométallobytes excluantes. De plus, chez les deux espèces pseudométallobytes hyperaccumulatrices les plus étudiées (parmi seulement quelques-unes connues, appartenant presque toutes au genre *Thlaspi*, Pollard *et al.* 2002), soit *Thlaspi caerulescens* et *Arabidopsis halleri*, la tolérance aux métaux, notamment au Zn, est constitutive. Elle est rencontrée dans toutes les populations de ces espèces, même si une tolérance plus forte des populations M est observée (Meerts & van Isacker 1997; Bert *et al.* 2000). Il est donc envisageable que la tolérance de ces espèces ait d'abord évolué en l'absence de pression métallique, en milieu non métallifère, où l'espèce est majoritairement présente. C'est pourquoi des hypothèses nouvelles sont élaborées, qui font de l'hyperaccumulation un caractère dont l'évolution serait conditionnée par **des pressions évolutives indépendantes de la pression métallique** (Boyd & Martens 1998; Pollard 2000; Boyd 2004). Dans un tel scénario, l'acquisition de capacités de tolérance aux métaux serait d'abord la conséquence de l'augmentation des capacités de séquestration du métal dans les feuilles, potentiellement acquises en milieu non contaminé. Plusieurs pressions de sélection peuvent rendre aboutir à la séquestration du métal dans les feuilles. L'hypothèse la plus étudiée implique la **pression exercée par l'herbivorie** et attribue à l'hyperaccumulation un rôle dans la défense contre les herbivores et les pathogènes (Pollard & Baker 1997; Jhee *et al.* 1999; Davis & Boyd 2000; Huitson & Macnair 2003). Elle reste très discutée (Davis *et al.* 2001; Noret *et al.* 2005).

L'hypothèse d'un **déterminisme génétique de l'hyperaccumulation indépendant de celui de la tolérance** (faut-il comprendre de l'exclusion?) a aussi été plusieurs fois avancée (Ingrouille & Smirnoff 1986; Macnair *et al.* 1999), même si des résultats récents obtenus chez *Thlaspi caerulescens* montrent des corrélations négatives entre les deux caractères (Assunção *et al.* 2003c; Frérot *et al.* 2005). C'est pourquoi l'hyperaccumulation, autrefois considérée comme une stratégie de tolérance, est aujourd'hui parfois aujourd'hui étudiée *per se*, en parallèle ou indépendamment des études de tolérance (voir Macnair, 2000, où sont synthétisés, dans un même chapitre de livre, les résultats disponibles sur les deux caractères,

mais dans deux parties successives, et sans que la relation entre les deux soit vraiment discutée).

II Tolérance aux métaux et phylogéographie

Avec le développement des techniques de biologiques moléculaires, l'identification des bases moléculaires de la tolérance aux métaux est devenue l'enjeu majeur des recherches menées sur les espèces métallobytes (voir I.2.3). Le déploiement des efforts consacrés en physiologie rend probable une identification prochaine de gènes candidats. De nouvelles études pourront alors être menées. Elles chercheront à définir les conséquences de la sélection de la tolérance sur l'évolution interspécifique des gènes impliqués dans le caractère et, à l'échelle intraspécifique, l'impact de la colonisation des sols métallifères sur la structure en populations de la variabilité génétique présentée par ces gènes.

La variation des fréquences de gènes dans et entre populations est principalement conditionnée par les pressions évolutives suivantes: mutation, migration, dérive génétique et sélection naturelle.

- On considère comme **mutation** toute variation héréditaire de l'information génétique. La mutation est donc la source ultime de la variation génétique. Les taux de mutation sont généralement faibles, de l'ordre de l'ordre de 10^{-6} par gène et par génération. C'est à travers la mutation qu'ont pu apparaître les gènes conférant la tolérance.

- La **migration**, communément évoquée sous l'expression "flux de gènes", fait référence au mouvement de gènes entre populations assuré par la dispersion de pollen ou de graines. La migration tend à homogénéiser la distribution de la variation génétique entre les populations d'une espèce et agit donc contre les forces favorisant la différenciation des populations. C'est par la dispersion de graines que sont fondées les populations M d'une espèce pseudométallophyte et c'est par l'existence de flux de gènes qu'on explique par exemple que la divergence génétique entre populations ou écotypes M et NM n'évoluent pas vers une spéciation écologique.

- La transmission des gamètes d'une génération à la suivante résulte d'un échantillonnage aléatoire des allèles présents dans la population parentale. Cet échantillonnage engendre au sein des populations une évolution aléatoire de la variation génétique appelée **dérive génétique**. L'intensité de la dérive est inversement proportionnelle à la taille de la population. En l'absence de migration, des populations isolées peuvent diverger sous la seule pression de la dérive. La

dérive génétique explique par exemple que certaines populations NM d'espèces pseudométallophytes ont perdu la variation nécessaire à la fondation de populations M.

- La **sélection naturelle** est la seule force évolutive à l'origine de l'adaptation. Elle agit sur les génotypes et leurs avatars, les phénotypes. Dans un environnement donné, certains génotypes, mieux adaptés, ont des taux de reproduction et de survie supérieurs aux autres génotypes. Leurs gènes sont plus transmis à la génération suivante et la fréquence de ces gènes augmente. La sélection naturelle répartie la variation génétique en fonction des pressions de sélection locales. La forte fréquence des génotypes tolérants sur site métallifère est une conséquence directe de la sélection naturelle.

A petite échelle géographique, ces pressions suffiraient à rendre compte de différenciation génétique entre populations à des locus neutres ou sélectionnés. La sélection explique par exemple les différences de capacités de tolérance observées de part et d'autre de la frontière séparant un sol minier et une prairie peu contaminée (Figure 7 et 8).

Néanmoins, à plus grande échelle, et en particulier lorsqu'elle est décrite sur l'ensemble de l'aire de distribution d'une espèce, la distribution actuelle entre populations de la variation génétique intraspécifique aux locus neutres ou sélectionnés, peut avoir aussi été très influencée par des **processus démographiques historiques**, réalisés sur de longues périodes de temps. L'avènement, au cours des vingt dernières années, de la phylogéographie, a en effet révélé que l'ensemble des populations d'une espèce ne constituaient le plus souvent pas un ensemble génétique homogène (Avice 2000). Au contraire, elles sont rassemblées en sous-groupes génétiques distincts dont les histoires diffèrent. En raison d'histoires biogéographiques différentes (et indépendamment des quatre forces évolutives évoquées ci-dessus), ces sous-groupes peuvent, pour un gène ou un locus donné: (1) différer par la **nature des variants génétiques ou allèles** pour le gène qui les composent; (2) différer par le **niveau de la diversité génétique** qu'ils présentent. Parce que leurs patrimoines génétiques respectifs peuvent varier, ces sous-groupes peuvent différer dans leur réponse à une variable environnementale locale. Mieux comprendre les effets locaux d'un stress comme le stress métallique sur l'évolution d'une population et les différences de réponses de populations distinctes à un même stress nécessite de pouvoir **replacer la ou les population(s) dans la phylogéographie générale de l'espèce à laquelle elle(s) appartient(en)t** (Staton *et al.* 2001).

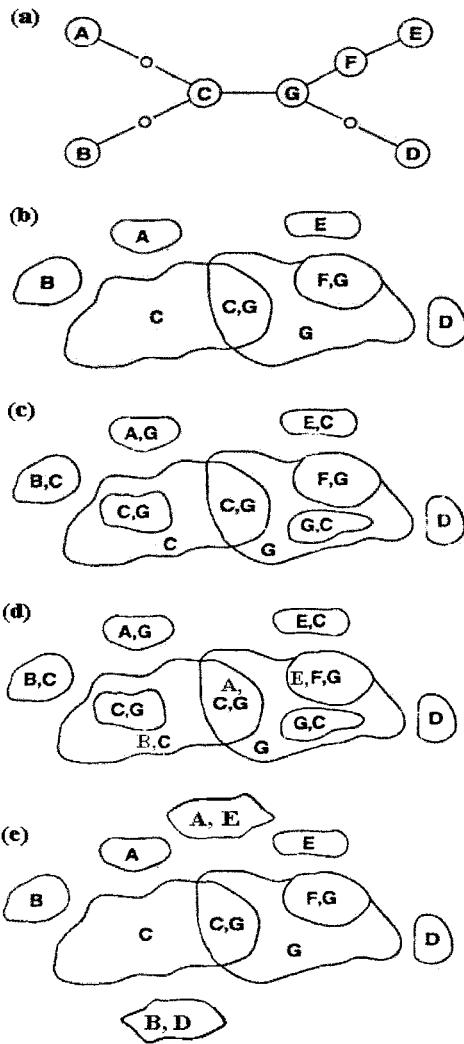


Figure 10: Arbre phylogénétique hypothétique de variants génétiques (A-G) et exemples de distributions géographiques de ces variants dans des groupes de populations (d'après Schaal *et al.* 1998).

(a): Arbre phylogénétique des variants génétiques. Chaque branche représente un pas de mutation. Les variants internes sont supposés ancestraux, variants externes dérivés par mutation.

(b): Correspondance entre phylogénie des variants et distribution géographique en populations.

(c): Non correspondance entre phylogénie des variants et distribution géographique. Les variants ancestraux restent présents et sont différemment éliminés dans les populations périphériques.

(d): Non correspondance entre phylogénie des variants et distribution géographique. Les variants dérivés (A, B, D, E) sont présents dans les populations d'origine.

(e): Non correspondance entre phylogénie des variants et distribution géographique. Des variants dérivés éloignés (A,E ou B,D) sont présents dans les mêmes populations.

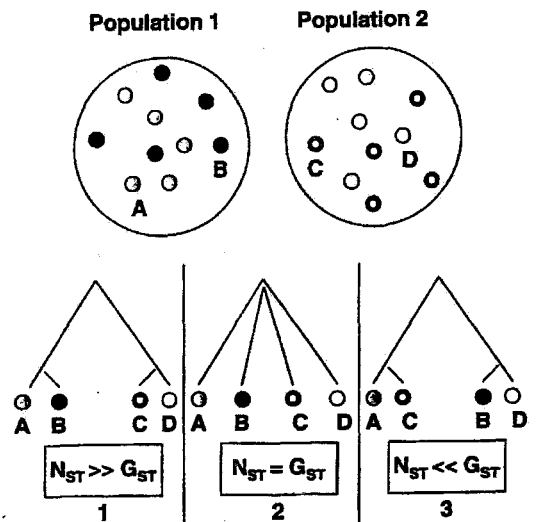
Figure 11: Estimation de la correspondance entre phylogénies d'allèles et distributions géographiques de ces allèles (d'après Pons & Petit 1996)

L'indice G_{ST} est un indice de différenciation ne tenant compte que des fréquences alléliques (Annexe XXX); l'indice N_{ST} est un indice de différenciation tenant compte en plus des distances génétiques entre haplotypes.

(1) Lorsqu'il y a correspondance entre phylogénie des allèles et distribution géographique (i.e. structure phylogéographique), l'indice N_{ST} est supérieur au G_{ST} . Dans le cas illustré, les allèles sont différents entre deux populations et deux allèles d'une même population divergent moins que deux allèles de populations différentes. La différenciation mesurée en tenant compte des distances phylogénétiques est plus forte entre deux populations: $N_{ST} > G_{ST}$.

(2) Lorsqu'il y a égale distance phylogénétique entre tous les allèles d'une espèce, l'information phylogénétique n'apporte rien au calcul de la différenciation des populations: $N_{ST} = G_{ST}$.

(3) Lorsque deux allèles à l'intérieur d'une même population sont phylogénétiquement plus éloignés que deux allèles de deux populations différentes, alors la différenciation tenant compte des distances moléculaires entre haplotypes est supérieure à celle n'en tenant pas compte: $N_{ST} < G_{ST}$.



I.1 La boîte à outils de la phylogéographie

II.1.1 Le concept de structure phylogéographique

Le terme de phylogéographie a été introduit par *Avise et al.* en 1987 (*Avise et al.* 1987) pour décrire le "champ d'étude concerné par les principes et processus gouvernant les distributions géographiques de lignées généalogiques, principalement au niveau intraspécifique" (*Avise* 2000).

La phylogéographie a apporté un moyen de détecter l'empreinte de flux de gènes anciens ou continus entre populations et de faire la différence, parmi les facteurs rendant compte de la structure génétique actuelle d'une espèce, **entre l'existence de flux de gènes entre populations et le partage d'un polymorphisme ancestral** (*Schaal et al.* 1998; *Avise* 2000) et, corollairement, de discuter les flux de gènes ou mouvements migratoires passés. Ceci est rendu possible par l'intégration, dans l'analyse de la distribution géographique des allèles à un locus, des relations généalogiques entre ces allèles et par la recherche de correspondance ou d'absence de correspondance entre la distribution géographique d'haplotypes et leurs relations généalogiques (Figure 10).

La correspondance traduit l'existence d'une **structure phylogéographique** dans laquelle les distributions d'allèles phylogénétiquement proches sont géographiquement limitées et proches (Figure 10b). L'existence d'une structure phylogéographique révèle une **situation durable de flux de gènes réduits**. En conséquence, les nouvelles mutations restent localisées dans la région géographique où elles se produisent. Les populations isolées divergent: c'est la **vicariance** (*Avise* 2000). Dans cette situation, l'analyse phylogéographique met en évidence des histoires évolutives différentes entre groupes de populations génétiquement isolés et qui constituent finalement des ensembles génétiques distincts (*Avise* 1989).

Plusieurs scénarios peuvent rendre compte d'une **absence de structure phylogéographique** (*Avise et al.* 1987; *Schaal et al.* 1998). La position phylogénétique des allèles à l'origine de l'absence de structure phylogéographique doit permettre de distinguer les scénarios. La **présence d'allèles ancestraux dans les populations divergentes** (Figure 10c) traduit plutôt la rémanence d'un polymorphisme ancestral dans ces populations que l'existence d'un actuel flux de gènes entre populations. A l'opposé, la distribution large des allèles, et

Encadré

Statistiques utilisées en génétiques des populations pour la mesure de la diversité et de la structure génétique

Mesure de la diversité génétique

A: Richesse allélique (allelic richness). Mesure le nombre d'allèles rencontrés dans une population S (A_S) ou sur l'ensemble des populations échantillonnées (A_T). Les valeurs sont souvent standardisées à une taille particulière de population (méthode de raréfaction) pour permettre des comparaisons entre populations

H_O: Hétérozygotie observée (observed heterozygosity), autrement dit la proportion d'hétérozygotes dans la population échantillonnée. Varie de 0 (tous les individus sont homozygotes) à 1 (tous les individus sont hétérozygotes). Nécessite l'utilisation de marqueurs codominants sur des locus diploïdes.

H_E: diversité génique (gene diversity) ou hétérozygotie attendue sous l'hypothèse de régime panmictique (locus diploïdes). C'est la probabilité de tirer deux allèles différents lors de tirages aléatoires avec remise dans une population. Varie de 0 à 1. Notée H_S lorsqu'elle est mesurée à l'échelle de la population, H_T lorsqu'elle est mesurée sur l'ensemble des populations échantillonnées.

Soit la population S, x_{is} la fréquence de l'allèle i dans S et x_i la fréquence de l'allèle dans l'ensemble des populations:

$$H_S = 1 - \sum_i x_{is}^2 \qquad H_T = 1 - \sum_i x_i^2$$

Mesure de la structure génétique

F_{IS}: coefficient de consanguinité. Varie de -1 à 1. Une valeur positive traduit l'existence de croisements consanguins, une valeur proche de 1 est caractéristique des espèces autogames.

$$F_{IS} = \frac{H_E - H_O}{H_E}$$

F_{ST}: indice de fixation entre populations. Mesure la proportion de la diversité totale distribuée entre les populations. Varie de 0 (pas de différenciation entre populations, la diversité totale est représentée dans chacune des populations) à 1 (populations complètement différentes)

$$F_{ST} = \frac{H_T - \overline{H_S}}{H_T}$$

G_{ST}: Généralisation du F_{ST} à des locus multialléliques

R_{ST}: Comparable au F_{ST} . Développé pour évaluer la contribution des différences de tailles à des locus microsatellites dans la différenciation des populations.

N_{ST}: Comparable au F_{ST} . Développé dans le contexte phylogéographique pour évaluer la contribution des distances phylogénétiques entre allèles dans la différenciation des populations.

notamment des allèles dérivés (Figure 10d), est plutôt révélatrice de l'existence de **flux de gènes entre populations** qui ne se différencient pas. Enfin, la rencontre d'allèles phylogénétiquement très distants dans une même population (Figure 10e) signale des **zones dites hybrides ou de suture** dans lesquelles des lignées alléliques ayant divergé dans des populations différenciées auparavant isolées sont mises en contact après migration.

II.1.2 La mesure de la structure phylogéographique

La structure phylogéographique est une structure génétique particulière dans laquelle on observe un certain niveau de différenciation entre les populations d'une espèce. La différence avec les indices classiques de mesure de la différenciation génétique entre populations, fondés sur des différences de fréquences alléliques entre populations (cf. Encadré), se situe dans l'attention portée sur les **distances phylogénétiques entre les allèles** présents dans les populations. Une façon simple d'identifier l'existence d'une structure phylogéographique consiste ainsi à inclure dans les calculs des indices de différenciation un facteur mathématique rendant compte de la distance phylogénétique entre allèles (Figure 11). Une valeur de cet indice transformé significativement supérieure à la valeur de l'indice classique traduit une distance plus grande entre allèles présent dans des populations différentes (géographiquement séparées) qu'entre allèles présents au sein d'une même population, soit un certain niveau de correspondance entre la distribution géographique des allèles et leurs relations généalogiques, soit une structure phylogéographique. CQFD.

II.1.3 Le génome chloroplastique

Le génome chloroplastique (ADNcp), s'est rapidement imposé, en phylogéographie des végétaux, comme un marqueur approprié pour l'obtention de phylogénies de gènes au niveau intraspécifique.

- Bien qu'en plusieurs copies dans chaque cellule, le génome chloroplastique est **unique pour un individu donné**.
- L'ADNcp est haploïde et, chez la majorité des Angiospermes, transmis uniquement par voie maternelle (Birky 1995). Cette **transmission uniparentale** réduit la taille efficace des populations et accentue les effets de la dérive génétique. En conséquence, le polymorphisme chloroplastique est plus structuré que le polymorphisme nucléaire (Petit *et al.* 1993; McCauley 1995). De plus, la dispersion par les graines est souvent limitée, ce qui préserve les patrons de structuration établis (Crandall & Templeton 1996; Templeton 1998)
- L'ADNcp est **non recombinant** : tous les gènes présents sur une même molécule possèdent la même histoire évolutive (Clegg & Zurawski 1992), ce qui facilite la reconstruction

de relation généalogiques (Cruzan & Templeton 2000). En conséquence de la liaison des gènes chloroplastiques, les allèles observés à différents locus sur une même molécule peuvent être combinés pour former des **haplotypes** ou **chlorotypes**.

- le taux de mutation de l'ADNcp est **conservatif**. Il est suffisamment fort pour détecter du polymorphisme intraspécifique et suffisamment faible pour éviter l'apparition de mutations homoplasmes (Soltis *et al.* 1992; Clegg *et al.* 1994).
- la **séquence** du génome chloroplastique est connue pour un nombre croissant d'espèces : huit génomes chloroplastiques sont aujourd'hui complètement séquencés chez les plantes vasculaires (une Gymnosperme, cinq Dicotylédones et deux Monocotylédones, Grivet 2001). Etant donnée la conservation du génome, les outils moléculaires mis au point à partir des séquences connues sont facilement applicables à d'autres espèces (Taberlet *et al.* 1991; Demesure *et al.* 1995; Dumolin-Lapegue *et al.* 1997).

II.2 Résultats empiriques en Europe

II.2.1 Le contexte paléo-environnemental et les données fossiles

Le climat du **Pléistocène** (de -1,7 millions d'années à -10 000 ans) a été dominé par de longues périodes glaciaires, chacune durant environ 100 000 ans, interrompues par des périodes chaudes et humides plus courtes (10 000 à 20 000 ans, Webb & Bartlein 1992). Au cours des périodes glaciaires, l'avancée des glaciers et des zones de permafrost ainsi que la chute des températures ont rendu inhabitables la plupart des régions situées à moyenne ou haute latitude. L'**aire de distribution des espèces s'en est trouvée modifiée**: elles se sont déplacées; les espaces disponibles se réduisant, elles se sont réduites; des fragmentations ont pu avoir lieu (Willis & Niklas 2004).

Au cours du siècle dernier, les enregistrements des distributions de pollens fossiles, en particulier d'espèces ligneuses, sur de longues périodes géologiques ont permis de suivre dans le temps l'évolution des aires de distributions d'espèces végétales (Willis 1996). Il est apparu qu'une majorité des espèces tempérées Européennes s'étaient retrouvées, pendant les épisodes glaciaires, et notamment pendant le dernier maximum glaciaire (Last Glacial Maximum, LGM, approximativement -115 000 à -10 000 ans), isolées dans de petites régions aux conditions environnementales favorables, les **refuges glaciaires**, localisées dans les trois péninsules sud-Européennes: les péninsules ibérique, italienne et balkanique (Willis 1996; Willis *et al.* 2000; Willis & van Andel 2004). Des séquences chronologiques de relevés polliniques répartis en Europe ont aussi attesté d'une **recolonisation plutôt rapide** de

l'Europe centrale depuis les refuges glaciaires à la fin du LGM, (avec parfois une progression de plusieurs centaines de mètres par an, Brewer *et al.* 2002). Le modèle proposé pour rendre compte de la recolonisation postglaciaire est celui d'une expansion des populations les plus au nord des refuges par des événements rares de dispersion à longue distance, entraînant une perte progressive de la diversité génétique conservée dans les refuges ("leading edge dispersal model" Hewitt 1996).

II.2.2 La confirmation phylogéographique

Chez la plupart des espèces végétales supposées recluses dans les refuges glaciaires, les études phylogéographiques ont montré (1) l'existence d'une forte structuration phylogéographique de la diversité, (2) une distribution latitudinale de lignées haplotypiques différenciées (Figure 12a et 12b), (3) la présence de la majorité des haplotypes d'Europe centrale dans une ou plusieurs des trois péninsules, alors que la réciproque n'est pas vraie (Grivet & Petit 2002; Petit *et al.* 2002; Grivet & Petit 2003; Rendell & Ennos 2003). Ces données attestent une réduction passée des flux de gènes entre des populations qui ont divergé dans des zones géographiques distinctes (Avisé *et al.* 1987). Elles permettent une relocalisation, indépendante des relevés polliniques, de ces zones (les refuges glaciaires) dans les péninsules du sud et des descriptions précises des voies de colonisation postglaciaires à partir des refuges (Figure 12c).

Les études phylogéographiques ont aussi révélé les conséquences des épisodes démographiques passés sur la distribution de la diversité génétique intraspécifique en identifiant souvent des patrons de structuration de diversité génétique associés à ces épisodes. Des divergences importantes sont attendues entre les haplotypes de populations ayant survécu isolément dans des refuges glaciaires différents (Lumaret *et al.* 2002). Le niveau de différenciation entre les populations de refuges est généralement plus important que le niveau de différenciation moyen entre populations (Petit *et al.* 2003). Enfin, la migration depuis les zones refuges est supposée avoir été accompagnée d'événements répétés de fondation. En conséquence, une perte progressive de la diversité génétique est attendue au cours de la colonisation, résultant en un **gradient géographique de la diversité génétique**, décroissant depuis les populations sources dans les refuges vers les populations du nord (Hewitt 1996). Une diversité génétique plus importante a effectivement été reportée dans de nombreuses

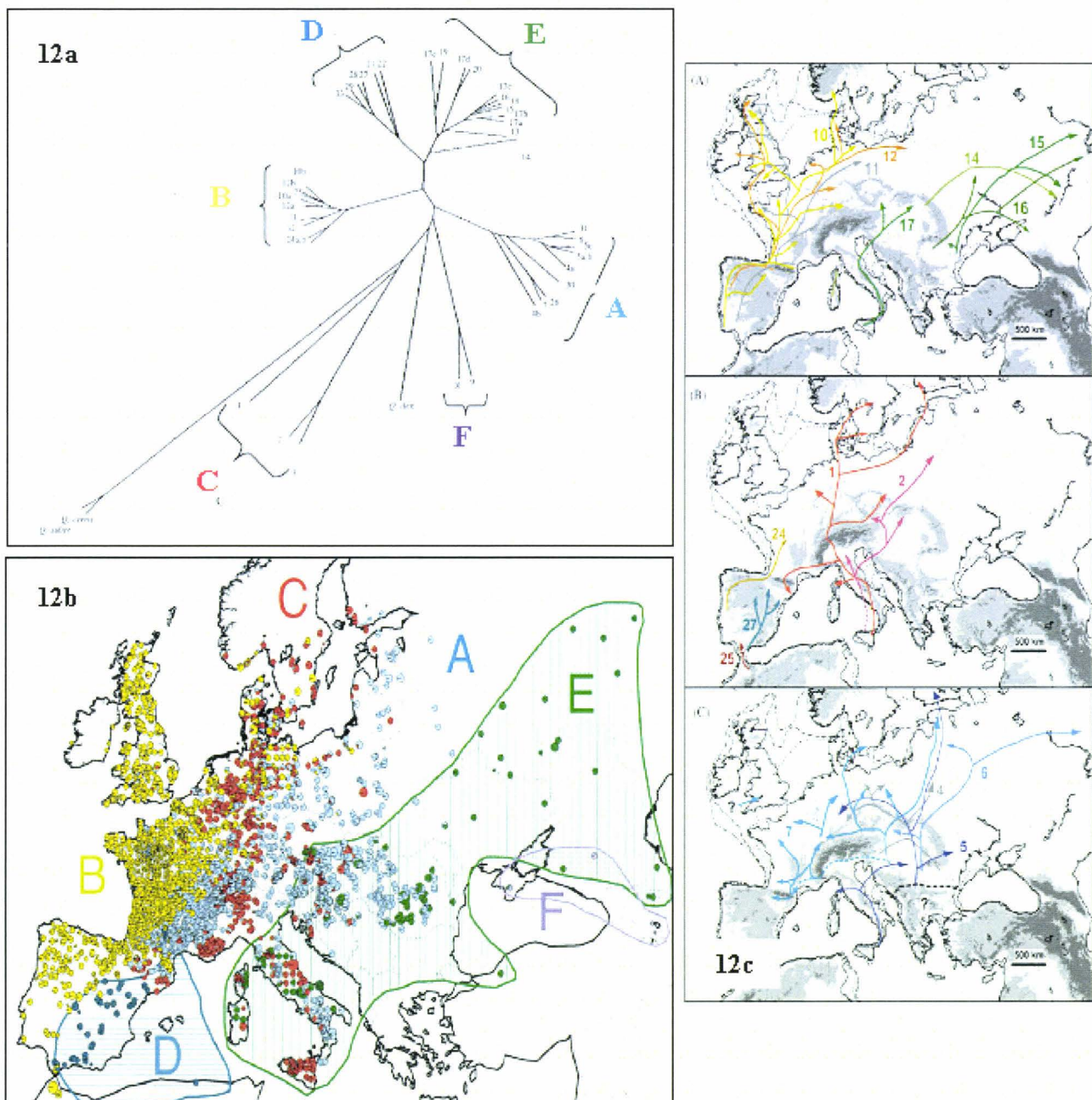


Figure 12 a: Arbre phylogénétique des lignées d'haplotypes identifiées dans les populations Européennes de chênes. 42 haplotypes (1-42) ont été identifiés durant l'étude de 2613 populations (soit 12 214 individus). Ces haplotypes se regroupent en 6 lignées (A-F). **b:** distribution géographiques des lignées haplotypiques. Chaque lignée montre une distribution distincte. **c:** Voies de migrations suivies par certains haplotypes depuis les refuges glaciaires jusqu'à leur distribution actuelle (Petit *et al.* 2002a; Petit *et al.* 2002b)

études phylogéographiques fondées sur des marqueurs allozymiques ou chloroplastiques (revues dans Comes & Kadereit 1998).

Néanmoins, les observations plus récentes nuancent ce patron général. Chez le hêtre commun (*Fagus sylvatica* L.), la richesse allélique (A_S), mesurée à l'aide 12 marqueurs isozymiques, est effectivement plus importante dans les populations des zones refuges, mais la diversité génique (H_S) est plus forte dans les régions plus récemment colonisées (Comps *et al.* 2001). Une étude sur 22 espèces ligneuses tempérées montre que le nombre moyen d'haplotypes chloroplastiques et la diversité génique sont plus importants dans les peuplements situés aux latitudes moyennes, au centre des aires de distribution (Petit *et al.* 2003). Dans les deux cas, les résultats ne contredisent cependant pas l'existence d'événements de fondation associés aux recolonisation postglaciaires (Comps *et al.* 2001; Petit *et al.* 2003).

Une diversité génétique forte aux latitudes moyennes pourrait traduire la **rencontre de voies de colonisation** et le mélange subséquent d'haplotypes divergents provenant de refuges distincts (Petit *et al.* 2003).

II.2.3 L'extension aux espèces herbacées

Bien qu'à aborder avec précaution, surtout en l'absence de relevés polliniques précis permettant d'inférer des scénarios démographiques (Comps *et al.* 2001; Petit *et al.* 2003; Lascoux *et al.* 2004), la correspondance observée chez les espèces ligneuses des régions tempérées entre événements démographiques passés (fragmentation de l'aire de distribution, isolement, expansion postglaciaire) et patrons de structure génétique (divergences entre populations, gradient de diversité) est à présent utilisée pour la phylogéographie d'espèces herbacées pour lesquelles les données polliniques sont rares, voire absentes.

Chez les espèces alpines, les analyses phylogéographiques ont surtout reposé sur l'analyse de la **distribution de la diversité génétique neutre** (Tribsch & Schonswetter 2003; Schonswetter *et al.* 2005). Chez nombre d'espèces, l'observation de groupes de populations géographiquement proches, divergents les uns des autres, et entre lesquels une part importante de la diversité génétique intraspécifique est répartie, ont permis de soutenir les hypothèses antagonistes de l'existence ou de l'absence de refuges glaciaires isolés (Stehlik *et al.* 2001; Holderegger *et al.* 2002; Kropf *et al.* 2002; Schonswetter *et al.* 2002). La position géographique de ces refuges a été déduite de la présence de populations présentant des

niveaux importants de diversité génétique sur des sites qui, à partir de données géologiques et paléoenvironnementales, sont supposés être restés à l'abri des glaces durant le Pléistocène et avoir pu servir de refuges (Tribsch & Schonswetter 2003; Schonswetter *et al.* 2005). L'observation de gradients géographiques de diversité génétique a été associée à des événements de fondation successifs et a été utilisée pour retracer des voies de migration postglaciaires (Tribsch *et al.* 2002; Schonswetter *et al.* 2003; Schonswetter *et al.* 2006).

Ces études sont dites **phylogéographiques au sens large**. Par manque des outils moléculaires nécessaires ou par souci d'identifier des niveaux de diversité génétique suffisants, elles reposent le plus souvent sur des marqueurs RAPD (Random Amplified Polymorphic DNAs), AFLP (Amplified Fragment Length Polymorphism) ou microsatellites (Tribsch & Schonswetter 2003; Schonswetter *et al.* 2005) ne permettant pas la reconstruction de l'histoire des allèles. Elles ne disposent donc **pas de phylogénie de gènes** mais tirent avantage des variations génétiques entre populations pour identifier des frontières entre populations et décrire des relations historiques ou des processus démographiques (Goldstein *et al.* 1999). La pertinence de ces approches est encore discutée. Il semble néanmoins par exemple, que les microsatellites puissent aider à déterminer justement des relations généalogiques entre populations (Schlotterer 2001).

III Le matériel biologique et les objectifs de l'étude

III.1 *Arabidopsis halleri* (L.): généralités

L'arabette de Haller *Arabidopsis halleri* (syn. *Cardaminopsis halleri*, Brassicaceae) est une espèce diploïde, pérenne, formant des rosettes et capable de se reproduire végétativement par clonage via la formation de stolons (Figure 13). Le régime de reproduction sexuée est auto-incompatible (Bechsgaard *et al.* soumis) et la pollinisation est surtout assurée par les insectes (van Rossum *et al.* 2004). Les graines sont nombreuses et petites, ne présentent pas de mécanismes de dispersion particulier et leur capacité de dispersion est limitée (van Rossum *et al.* 2004).

III.1.1 Ecologie et distribution

A. halleri est assez largement distribuée en Europe continentale (Figure 14) où on la rencontre dans des zones de montagne et de plaine, sur des sols plutôt acides, humides et oligotrophiques (Clapham & Akeroyd 1993; Al-Shehbaz & O'Kane 2002).

Deux sous-espèces sont actuellement décrites en Europe. *A. h. subsp. halleri* est présente dans presque toute l'Europe continentale, à l'exception des régions les plus à l'Ouest, l'extrême Nord et la majeure partie de l'Italie (O' Kane & Al-Shehbaz 1997). On la rencontre sur les versants des montagnes et les lisières de forêt, à des altitudes comprises entre 66 et 2200 mètres (O' Kane & Al-Shehbaz 1997; Al-Shehbaz & O'Kane 2002). *A. h. subsp. ovirensis* est distribuée sur les versants rocheux d'Europe de l'Est et du Sud, entre 1200 et 2400 mètres d'altitude (O' Kane & Al-Shehbaz 1997; Al-Shehbaz & O'Kane 2002).

La sous-espèce *A. h. subsp. halleri* est mentionnée dans **trois régions fortement industrialisées et polluées** par les métaux dans le nord de l'Europe: le Nord de la France, le Harz en Allemagne et la Silésie en Pologne (Berton 1946; Fabiszewski 1986; Ernst 1990; Ernst *et al.* 2004). La pollution du Harz est surtout dû aux activités minières rapportées dans la région depuis le Moyen Age (Liessmann 1997; Brooks 1998a) alors que la pollution en France et en Pologne est liée aux activités industrielles développées avec la révolution industrielle aux XVIIIe et XIXe siècle (Figure 15, Berton 1946; Grodzinska *et al.* 2000).

III.1.2 Position phylogénétique

Arabidopsis halleri a récemment été placée dans le genre *Arabidopsis* sur la base de critères morphologiques (Al-Shehbaz & O'Kane 2002). Les études moléculaires récentes, fondées sur des marqueurs nucléaires et cytoplasmiques s'accordent sur la **proximité phylogénétique d'*A. halleri* avec *A. thaliana*** (Figure 16, Koch *et al.* 2000; Koch *et al.* 2001; O'Kane & Al-Shehbaz 2003). L'événement de spéciation séparant *A. thaliana* des autres *Arabidopsis* est daté d'environ 5 millions d'années (Koch *et al.* 2000).



Figure 13 *Arabidopsis halleri* en milieu naturel. A droite: rosettes isolées, on peut voir un stolon s'allonger vers le haut. A gauche: tapis de rosettes. En l'absence de données génétiques, il est difficile d'identifier des individus (genets) distincts.

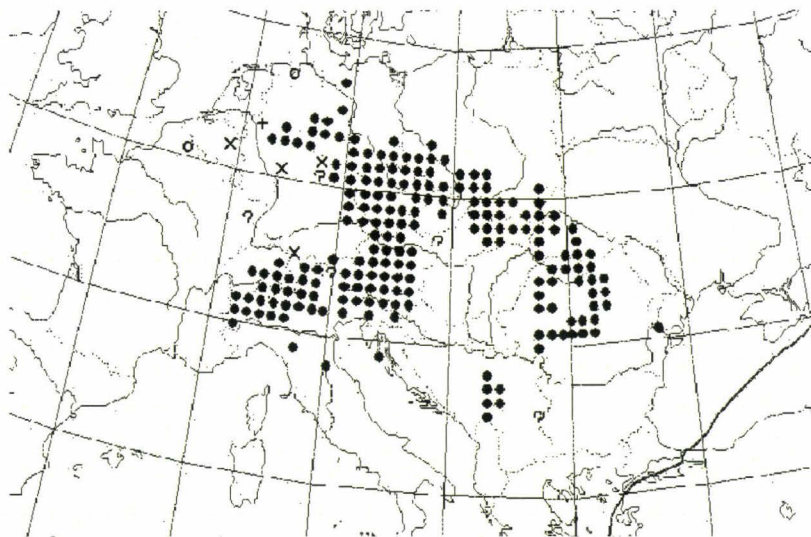


Figure 14 Carte de distribution d'*A. halleri* en Europe (*Atlas Flora Europea*, 1999, Botanical Museum, Finnish Museum of Natural History, <http://www.helsinki.fi/kmus/afe.html>)



Figure 15 Site métallifère sur lequel se développe *A. halleri* subsp. *halleri* en Silésie, Pologne.

III.2 Tolérance aux métaux lourds d'*A. halleri*

III.2.1 Les connaissances actuelles

Une première étude menée en hydroponie (Bert *et al.* 2000) a montré des capacités de tolérance et d'hyperaccumulation du Zn élevées dans deux populations métallophiles du Nord de la France. La même étude a aussi montré des capacités de tolérance et d'hyperaccumulation du Zn dans trois populations non métallophiles de Sumava, en République Tchèque. Les auteurs ont suggéré l'existence de **propriétés constitutives** chez *A. halleri*. Des variations de tolérance ont été notées avec des capacités significativement plus fortes dans les populations M, suggérant que, même constitutive, la tolérance continue à évoluer sur sol pollué (Bert *et al.* 2000). Depuis, une accumulation de Zn, Pb et Cd a été reportée dans les tissus aériens de plantes récoltées dans 33 populations M et NM réparties au nord de l'aire de distribution de l'espèce (Bert *et al.* 2002). *A. h. halleri* est aussi connue comme hyperaccumulatrice de Cd (Brooks 1998a; Küpper *et al.* 2000).

III.2.2 Le statut de modèle

La proximité d'*A. halleri* avec *A. thaliana* et la forte identité des régions codantes des génomes entre les deux espèces (90 à 95% Assunção *et al.* 2003b; Weber *et al.* 2004) rendent la majorité des **outils moléculaires développés chez *A. thaliana***, grâce à la connaissance de son génome, facilement **transférables à *A. halleri*** (Feder & Mitchell-Olds 2003). En conséquence, *A. halleri* est aujourd'hui proposé comme espèce modèle pour l'étude de la tolérance aux métaux lourds chez les plantes (Mitchell-Olds 2001; Lexer & Fay 2005). Elle est déjà intensément étudiée dans des études de génomique, transcriptomique, protéomique et métabolomique (Becher *et al.* 2004; Dräger *et al.* 2004; Weber *et al.* 2004; Willems 2006).

La comparaison, par la technique des micro-puces, de profils d'expression de plus de 8000 gènes entre *A. halleri* et *A. thaliana*, a permis de mettre en évidence une expression plus forte de deux gènes dans les racines d'*A. halleri* (Weber *et al.* 2004). L'un d'entre eux code pour une enzyme responsable de la synthèse d'un chélateur organique: la nicotianamine. La même technique appliquée aux gènes s'exprimant dans les feuilles a révélé une expression accrue de **plusieurs gènes impliqués dans l'homéostasie des métaux** (Becher *et al.* 2004).

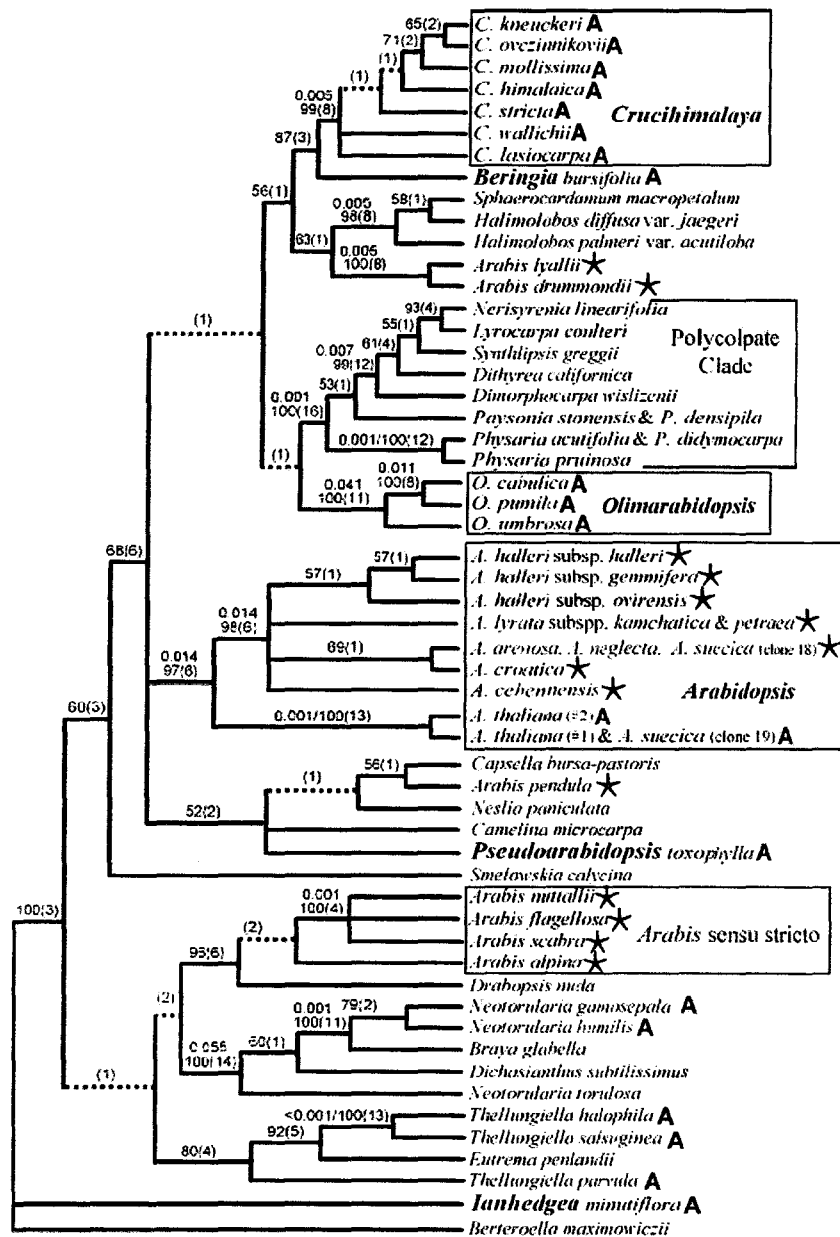


Figure 16 Arbre phylogénétique d'espèces de Brassicaceae, obtenu à partir de données de séquences ITS (O'Kane & Al-Shehbaz 2003). Arbre consensus obtenu à partir de 24 arbres de parsimonie. Les limites du genre *Arabidopsis*, incluant neuf espèces, sont indiquées. A: espèces récemment ou plus anciennement classées dans le genre *Arabidopsis*; *: espèces traditionnellement placées dans le genre *Arabis*.

L'analyse QTL de la tolérance au Zn sur une descendance de backcross d'un croisement interspécifique *A. halleri* x *A. lyrata* a permis d'identifier trois régions génomiques expliquant 29% de la variance phénotypique et 42% de la variance génotypique observées pour le trait dans la population de backcross (Willems *et al.* soumis). Ces régions comprennent des transporteurs potentiellement impliqués dans la séquestration vacuolaire du métal et une ATPase. L'analyse QTL de la tolérance au Cd sur la même population met en évidence trois régions QTL dont une explique seule 42% de la variance génétique (Willems 2006). Cette région co-localise avec la région comprenant l'ATPase identifiée dans l'analyse de la tolérance au Zn.

La connaissance des mécanismes moléculaires de la tolérance au métal chez *A. halleri* progresse rapidement. Elle permettra bientôt l'identification de gènes candidats qui pourront constituer la cible d'études d'évolution moléculaire. Malgré ce contexte dynamique, les connaissances nécessaires à la confirmation de la sélection de ces gènes en populations naturelles, sont encore déficientes. C'est pourquoi, l'objectif de la thèse présentée ici est triple:

1. Obtenir une description précise de la distribution des capacités de tolérance au zinc d'*A. halleri* à l'échelle de l'aire de distribution européenne de l'espèce. C'est aussi tester la constitutivité du caractère de tolérance au Zn chez *A. halleri*, évaluer le niveau de polymorphisme présent dans l'espèce et la distribution des capacités entre populations M et NM (Chapitre 1).

2. Utiliser la phylogénie d'haplotypes chloroplastiques pour déterminer l'origine des populations métallicoles situées au nord de l'aire de distribution de l'espèce. Tester deux hypothèses alternatives : une origine unique à toutes les populations ou des origines indépendantes pour les populations géographiquement isolées. Définir si les phénotypes observés sur des sites métallifères distants résultent d'une adaptation unique ou s'ils correspondent à des évolutions indépendantes. Rechercher l'existence d'événements de fondation dans ces populations (Chapitre 2).

3. Mettre en évidence en employant les outils de la phylogéographie, le patron d'organisation de la diversité génétique neutre et les facteurs majeurs responsables de ce patron. Poser implicitement la question du rôle de l'exposition aux métaux dans la structure génétique neutre de l'espèce (Chapitre 3).

IV Références

- Al-Hiyaly SA, McNeilly T, Bradshaw AD (1988) The effect of zinc contamination from electricity pylons - evolution in a replicated situation. *New Phytologist*, **110**, 571-580.
- Al-Hiyaly SA, McNeilly T, Bradshaw AD, Mortimer AM (1993) The effect of zinc contamination from electricity pylons. Genetics constraints on selection for zinc tolerance. *Heredity*, **70**, 22-32.
- Al-Shehbaz IA, O'Kane SLJ (2002) Taxonomy and Phylogeny of *Arabidopsis* (Brassicaceae). In: *The Arabidopsis Book* (eds. Somerville CR, Meyerowitz EM). American Society of Plant Biologist, Rockville.
- Antonovics J (1968) Evolution in closely adjacent plant populations. V. The evolution of self-fertility. *Heredity*, **23**, 219-238.
- Antonovics J, Bradshaw AD (1970) Evolution in closely adjacent plant populations. VIII. Clinal patterns at a mine boundary. *Heredity*, **25**, 349-362.
- Antonovics J, Bradshaw AD, Turner RG (1971) Heavy metal tolerance in plants. *Advances in Ecological Research*, **7**, 1-85.
- Assunção AGL, Bookum WM, Nelissen HJM, *et al.* (2003a) Differential metal-specific tolerance and accumulation patterns among *Thlaspi caerulescens* populations originating from different soil types. *New Phytologist*, **159**, 411-419.
- Assunção AGL, da Costa Martins P, de Folter S, *et al.* (2001) Elevated expression of metal transporter genes in three accessions of the metal hyperaccumulator *Thlaspi caerulescens*. *Plant, Cell and Environment*, **24**, 217-226.
- Assunção AGL, Pieper B, Vromans J, *et al.* (in press) Construction of a genetic linkage map of *Thlaspi caerulescens* and quantitative trait loci analysis of zinc accumulation. *New Phytologist*, **0**, ???-???
- Assunção AGL, Schat H, Aarts MGM (2003b) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. *New Phytologist*, **159**, 351-360.
- Assunção AGL, Ten Bookum WM, Nelissen HJM, *et al.* (2003c) A cosegregation analysis of zinc (Zn) accumulation and Zn tolerance in the Zn hyperaccumulator *Thlaspi caerulescens*. *New Phytologist*, **159**, 383-390.
- Avise JC (1989) Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution Int J Org Evolution*, **43**, 1192-1208.
- Avise JC (2000) *Phylogeography: the history and formation of species* Harvard University Press, Cambridge, Massachusetts and London, England.
- Avise JC, Arnold J, Ball RMJ, *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **7**, 489-522.
- Baker AJM (1981) Accumulators and Excluders - Strategies in the response of Plants to heavy metals. *Journal of Plant Nutrition*, **3**, 643-654.
- Baker AJM (1987) Metal tolerance. *New Phytologist*, **106(suppl.)**, 93-111.

- Baker AJM, McGrath SP, Reeves DR, Smith J (2000) Metal hyperaccumulator plants: a review of the ecology and physiology of a biological resource for phytoremediation of metal-polluted soils. In: *Phytoremediation of contaminated soils and water* (ed. Terry N, Banuelos, G. eds), pp. 171-188. CRC press, Boca Raton, FL, USA.
- Baker AJM, Proctor J (1990) The influence of cadmium, copper, lead, and zinc on the distribution and evolution of metallophytes in the British Isles. *Plant Systematics and Evolution*, **173**, 91-108.
- Baker AJM, Proctor J, Reeves DR (1992) The Ecology of Ultramafic (Serpentine) Soils. Intercept Ltd., Andover.
- Baker AJM, Walker PL (1990) Ecophysiology of metal uptake by tolerant plants. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. Shaw AJ), pp. 155-177. CRC Press, Boca Raton, Florida.
- Barton NH, Keightley PD (2002) Understanding quantitative genetic variation. *Nat Rev Genet*, **3**, 11-21.
- Becher M, Talke IN, Krall L, Kramer U (2004) Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. *The Plant Journal*, **37**, 251-268.
- Bechsgaard JS, Castric V, Charlesworth D, Vekemans X, Schierup MH (submitted) The transition to self-compatibility in *Arabidopsis thaliana* and evolution within S-haplotypes over 10 million years. *Molecular Biology and Evolution*.
- Bert V, Bonnin I, Saumitou-Laprade P, de Laguerie P, Petit D (2002) Do *Arabidopsis halleri* from nonmetallicolous populations accumulate zinc and cadmium more effectively than those from metallicolous populations? *New Phytologist*, **155**, 47-57.
- Bert V, Macnair MR, de Laguerie P, Saumitou-Laprade P, Petit D (2000) Zinc tolerance and accumulation in metallicolous and non-metallicolous populations of *Arabidopsis halleri* (Brassicaceae). *New Phytologist*, **146**, 225-233.
- Bert V, Meerts P, Saumitou-Laprade P, *et al.* (2003) Genetic basis of Cd tolerance and hyperaccumulation in *Arabidopsis halleri*. *Plant and Soil*, **249**, 9-18.
- Berton A (1946) Presentation de plantes *Arabis Halleri*, *Armeria elongata*, *Oenanthe fluviatilis*, *Galinsoga parviflora discoidea*. *Bulletin de la Société de Botanique de France*, **93**, 139-145.
- Bijlsma R, Loescheke V (2005) Environmental stress, adaptation and evolution: an overview. *Journal of Evolutionary Biology*, **18**, 744-749.
- Birky C, Jr (1995) Uniparental Inheritance of Mitochondrial and Chloroplast Genes: Mechanisms and Evolution. *PNAS*, **92**, 11331-11338.
- Borovik AS (1990) Characterization of metal ions in biological systems. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. Shaw AJ), pp. 3-5. CRC Press, Boca Raton, Florida.
- Bourelle P-H, Berthelin J (1998) Contamination des sols par les éléments en traces : les risques et leur gestion. In: *Rapport de l'Académie des sciences*. Lavoisier, Paris.
- Boyd RS (2004) Ecology of metal hyperaccumulation. *New Phytologist*, **162**, 563-567.

- Boyd RS, Martens SN (1998) The significance of metal hyperaccumulation for biotic interactions. *Chemoecology*, **8**, 1-7.
- Bradshaw AD (1952) Populations of *Agrostis tenuis* resistant to lead and zinc poisoning. *Nature*, **169**, 1098.
- Bradshaw AD (1970) Plants and industrial waste. *Trans. Bot. Soc. Edinb.*, **41**, 71-64.
- Bradshaw AD (1991) Genostasis and the limits to evolution. *Philos Trans R Soc Lond B Biol Sci*, **333**, 289-305.
- Bradshaw AD, McNeilly T, Putwain PD (1990) The essential qualities. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. Shaw AJ), pp. 323-334. CRC Press, Boca Raton, Florida.
- Brady KU, Kruckeberg AR, Bradshaw Jr. HD (2005) EVOLUTIONARY ECOLOGY OF PLANT ADAPTATION TO SERPENTINE SOILS. *Annual Review of Ecology, Evolution, and Systematics*, **36**, 243-266.
- Brewer S, Cheddadi R, de Beaulieu JL, Reille M (2002) The spread of deciduous *Quercus* throughout Europe since the last glacial period. *Forest Ecology and Management*, **156**, 27-48.
- Broadley MR, Willey NJ, Wilkins JC, *et al.* (2001) Phylogenetic variation in heavy metal accumulation in angiosperms. *New Phytologist*, **152**, 9-27.
- Brooks R (1998a) Phytoarcheology and hyperaccumulators. In: *Plants That Hyperaccumulate Heavy Metals* (ed. Brooks R), pp. 181-202. CAB International, Oxon, U.K.
- Brooks RR (1987) *Serpentine and its Vegetation: A Multidisciplinary Approach* Dioscorides Press, Portland, Oregon.
- Brooks RR (1998b) Plants that hyperaccumulate heavy metals. Their role in phytoremediation, microbiology, mineral exploration and phytomining, p. 380. CAB International, Oxon, U.K.
- Brooks RR (1998c) General introduction. In: *Plants that hyperaccumulate heavy metals* (ed. Brooks RR), pp. 1-14. CAB international, Oxford.
- Buse A, Norris D, Harmens H, *et al.* (2003) Heavy metals in European mosses: 2000/2001 survey. Centre for Ecology and Hydrology, University of Wales Bangor.
- Bush EJ, Barrett SCH (1993) Genetics of mine invasions by *Deschampsia cespitosa* (Poaceae). *Canadian Journal of Botany*, **71**, 1336-1347.
- Callahan D, Baker AJM, Kolev S, Wedd A (2006) Metal ion ligands in hyperaccumulating plants. *Journal of Biological Inorganic Chemistry*, **11**, 2-12.
- Chaney RL, Malik M, Li YM, *et al.* (1997) Phytoremediation of soil metals. *Curr Opin Biotechnol*, **8**, 279-284.
- Christie P, Li X, Chen B (2004) Arbuscular mycorrhiza can depress translocation of zinc to shoots of host plants in soils moderately polluted with zinc. *Plant and Soil*, **261**, 209-217.
- Clapham AR, Akeroyd JR (1993) *Cardaminopsis*. In: *Flora Europea* (eds. Tutin TG, Heywood VH, Burges NA, *et al.*), p. 290. Cambridge University Press, Cambridge.

- Clegg MT, Gaut BS, Learn GH, Jr., Morton BR (1994) Rates and patterns of chloroplast DNA evolution. *Proceeding of the National Academy of Science USA*, **91**, 6795-6801.
- Clegg MT, Zurawski G (1992) Chloroplast DNA and the study of Plant phylogeny: present status and future prospects. In: *Molecular systematics of Plant* (eds. Soltis PS, Soltis DE, Doyle JJ), pp. 1-13. Chapman and Hall, New York.
- Clemens S (2001) Molecular mechanisms of plant metal tolerance and homeostasis. *Planta*, **212**, 475-486.
- Clemens S (2006) Evolution and function of phytochelatase synthases. *Journal of Plant Physiology*, **163**, 319-332.
- Clemens S, Palmgren MG, Krämer U (2002) A long way ahead: understanding and engineering plant metal accumulation. *TRENDS in Plant Science*, **7**, 309-315.
- Cobbett C, Goldsbrough P (2002) Phytochelatins and metallothioneins: Roles in Heavy Metal Detoxification and Homeostasis. *Annu. Rev. Plant Biol.*, **53**, 159-182.
- Cobbett CS (2000) Phytochelatase biosynthesis and function in heavy-metal detoxification. *Current Opinion in Plant Biology*, **3**, 211-216.
- Comes HP, Kadereit JW (1998) The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, **3**, 432-438.
- Comps B, Gomory D, Letouzey J, Thiebaut B, Petit RJ (2001) Diverging Trends Between Heterozygosity and Allelic Richness During Postglacial Colonization in the European Beech. *Genetics*, **157**, 389-397.
- Crandall KA, Templeton AR (1996) Applications of intraspecific phylogenetics. In: *New Uses for New Phylogenies* (eds. Harvey PH, Brown AJL, Maynard Smith J, Nee S). Oxford University Press, Oxford, U. K.
- Cruzan MB, Templeton AR (2000) Paleoecology and coalescence: phylogeographic analysis of hypotheses from the fossil record. *Trends in Ecology and Evolution*, **15**, 491-496.
- Cuguen J, Acheroy M, Loutfi AL, Vernet P (1989) Breeding system differentiation in *Arrhenatherum elatius* populations: evolution toward selfing? *Evolutionary Trends in Plants*, **2**, 17-24.
- Davis MA, Boyd RS (2000) Dynamics of Ni-based defence and organic defences in the Ni hyperaccumulator, *Streptanthus polygaloides* (Brassicaceae). *New Phytologist*, **146**, 211-217.
- Davis MA, Murphy JF, Boyd RS (2001) Nickel Increases Susceptibility of a Nickel Hyperaccumulator to Turnip mosaic virus. *Journal of Environmental Quality*, **30**, 85-90.
- De Koe T, Geldmeyer K, Jaques NMM (1992) Measuring maximum root growth instead of longest root elongation in metal tolerance tests for grasses (*Agrostis tenuis*, *Agrostis capillaris* and *Agrostis castellana*). *Plant and Soil*, **144**, 305-308.
- Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*, **4**, 129-131.
- Dong Y, Ogawa T, Lin D, *et al.* (2006) Molecular mapping of quantitative trait loci for zinc toxicity tolerance in rice seedling (*Oryza sativa* L.). *Field Crops Research*, **95**, 420-425.

- Dräger DB, Desbrosses-Fonrouge A-G, Krach C, *et al.* (2004) Two genes encoding *Arabidopsis halleri* MTP1 metal transport proteins co-segregate with zinc tolerance and account for high *MTP1* transcript levels. *The Plant Journal*, **39**, 425-439.
- Dubois S, Cheptou P-O, Petit C, *et al.* (2003) Genetic structure and mating systems of metallicolous and nonmetallicolous populations of *Thlaspi caerulescens*. *New Phytologist*, **157**, 633-641.
- Ducouso A, Petit D, Valero M, Vernet P (1990) Genetic variation between and within populations of a perennial grass: *Arrhenatherum elatius*. *Heredity*, **65**, 179-188.
- Duffus JH (2002) "Heavy metals" a meaningless term? *Pure and Applied Chemistry*, **74**, 793-807.
- Dumolin-Lapegue S, Pemonge MH, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology*, **6**, 393-397.
- Ernst WHO (1990) Mine vegetation in Europe. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. Shaw AJ), pp. 21-37. CRC Press, Boca Raton, Florida.
- Ernst WHO, Knolle F, Kratz S, Schnug E (2004) Aspects of ecotoxicology of heavy metals in the Harz region – a guided excursion. *Landbauforschung Völkenrode*, **54**, 53-71.
- Ernst WHO, Verkleij JAC, Schat H (1992) Metal tolerance in plants. *Acta Botanica Neerlandica*, **41**, 229-248.
- Fabiszewski J (1986) Heavy metal tolerance of *Cardaminopsis halleri* (L.) Hayek populations in the Polish Tatra Mts. *Acta Societatis Botanicorum Poloniae*, **55**, 421-428.
- Feder ME, Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Nature Reviews Genetics*, **4**, 649-655.
- Frérot H, Lefebvre C, Petit C, *et al.* (2005) Zinc tolerance and hyperaccumulation in F1 and F2 offspring from intra and interecotype crosses of *Thlaspi caerulescens*. *New Phytologist*, **165**, 111-119.
- Friedland AJ (1990) The movement of metals through soils and ecosystems. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. Shaw AJ), pp. 7-19. CRC Press, Boca Raton, Florida.
- Gaur A, Adholeya A (2004) Prospects of arbuscular mycorrhizal fungi in phytoremediation of heavy metal contaminated soils. *Current Science*, **86**, 528-534.
- Ghosh M, Singh SP (2005) A review on phytoremediation of heavy metals and utilization of its byproducts. *Applied Ecology and Environmental Research*, **3**, 1-18.
- Ginocchio R, Baker AJM (2004) Metalófitas en América Latina: un recurso biológico y genético único poco conocido y estudiado en la región. *Revista chilena de historia natural*, **77**, 185-194.
- Goldstein DB, Roemer GW, Smith DA, *et al.* (1999) The Use of Microsatellite Variation to Infer Population Structure and Demographic History in a Natural Model System. *Genetics*, **151**, 797-801.
- Gregory RPG, Bradshaw AD (1965) Heavy metal tolerance in populations of *Agrostis tenuis* Sibth. and other grasses. *New Phytologist*, **64**, 131-143.

- Grivet D (2001) *Phylogéographie et évolution moléculaire comparée d'arbres forestiers à l'aide des marqueurs chloroplastiques*.
- Grivet D, Petit RJ (2002) Phylogeography of the common ivy (*Hedera* sp.) in Europe: genetic differentiation through space and time. *Molecular Ecology*, **11**, 1351-1362.
- Grivet D, Petit RmJ (2003) Chloroplast DNA phylogeography of the hornbeam in Europe: Evidence for a bottleneck at the outset of postglacial colonization. *Conservation Genetics*, **4**, 47-56.
- Grodzinska K, Korzeniak U, Szarek-Lukaszewska G, Godzik B (2000) Colonization of zinc mine spoils in southern Poland - preliminary studies on vegetation, seed rain and seed bank. *Fragm. Flor. Geobot.*, **45**, 123-145.
- Guerinot ML (2000) The ZIP family of metal transporters. *Biochimica et Biophysica Acta*, **1465**, 190-198.
- Hall JL (2002) Cellular mechanisms for heavy metal detoxification and tolerance. *J. Exp. Bot.*, **53**, 1-11.
- Harper FA, Smith SE, Macnair MR (1997) Can an increased copper requirement in copper-tolerant *Mimulus guttatus* explain the cost of tolerance ? I. Vegetative growth. *New Phytologist*, **136**, 455-467.
- Harper FA, Smith SE, Macnair MR (1998) Can an increased copper requirement in copper-tolerant *Mimulus guttatus* explain the cost of tolerance? II. Reproductive phase. *New Phytologist*, **140**, 637-654.
- Hartl DL, Clark AG (1997) *Principles of Population Genetics*, Third edition edn. Sinauer Associates, Inc., Sunderland, Massachusetts.
- Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276.
- Himelblau E, Amasino RM (2000) Delivering copper within plant cells. *Curr Opin Plant Biol*, **3**, 205-210.
- Holderegger R, Stehlik I, Abbott RJ (2002) Molecular analysis of the Pleistocene history of *Saxifraga oppositifolia* in the Alps. *Molecular Ecology*, **11**, 1409-1418.
- Huitson SB, Macnair MR (2003) Does zinc protect the zinc hyperaccumulator *Arabidopsis halleri* from herbivory by snails? *New Phytologist*, **159**, 453-459.
- Ingrouille MJ, Smirnoff N (1986) *Thlaspi caerulescens* J. & C. presl. (T. alpestre) in Britain. *New Phytologist*, **102**, 219-233.
- Jain SK, Bradshaw AD (1966) Evolutionary divergence among adjacent plant populations. I. The evidence and its theoretical analysis. *Heredity*, **20**, 407-441.
- Jansen S, Watanabe T, Caris P, *et al.* (2004) The distribution and phylogeny of aluminium accumulating plants in the Ericales. *Plant Biology (Stuttg)*, **6**, 498-505.
- Jhee EM, Dandridge KL, Christy Jr. AM, Pollard AJ (1999) Selective herbivory on low-zinc phenotypes of the hyperaccumulator *Thlaspi caerulescens* (Brassicaceae). *Chemoecology*, **9**, 93-95.
- Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225-1241.

- Koch M, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of Chalcon Synthase and Alcohol Dehydrogenase loci in *Arabidopsis*, *Arabis* and related genera (Brassicaceae). *Molecular Biology and Evolution*, **17**, 1483-1498.
- Koch M, Haubold B, Mitchell-Olds T (2001) Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. *American Journal of Botany*, **88**, 534-544.
- Krämer U (2003) Phytoremediation to phytochelatin - plant trace metal homeostasis. *New Phytologist*, **158**, 4-6.
- Krämer U (2005) MTP1 mops up excess zinc in *Arabidopsis* cells. *Trends in Plant Science*, **10**, 313-315.
- Krämer U, Clemens S (2006) Functions and homeostasis of zinc, copper, and nickel in plants. *Topics in Current Genetics*, **XXX**, 215-271.
- Krämer U, Cotter-Howells JD, Charnock JM, Baker AJM, Smith JA (1996) Free histidine as a metal chelator in plants that accumulate nickel. *Nature*, **379**, 635-638.
- Krämer U, Pickering IJ, Prince RC, Raskin I, Salt DE (2000) Subcellular Localization and Speciation of Nickel in Hyperaccumulator and Non-Accumulator *Thlaspi* Species. *Plant Physiology*, **122**, 1343-1354.
- Krämer U, Smith RD, Wenzel WW, Raskin I, Salt DE (1997) The Role of Metal Transport and Tolerance in Nickel Hyperaccumulation by *Thlaspi goesingense* Halacsy. *Plant Physiology*, **115**, 1641-1650.
- Kropf M, Kadereit JW, Comes HP (2002) Late Quaternary distributional stasis in the submediterranean mountain plant *Anthyllis montana* L. (Fabaceae) inferred from ITS sequences and amplified fragment length polymorphism markers. *Molecular Ecology*, **11**, 447-463.
- Kruckeberg AR (1986) An essay: the stimulus of unusual geologies for plant speciation. *Systematic Botany*, **11**, 455-463.
- Kruckeberg AR, Kruckeberg AL (1990) Endemic metallophytes: their taxonomic, genetic, and evolutionary attributes. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. Shaw AJ), pp. 301-312. CRC Press, Boca Raton, Florida.
- Küpper H, Lombi E, Zhao FJ, McGrath SP (2000) Cellular compartmentation of cadmium and zinc in relation to other elements in the hyperaccumulator *Arabidopsis halleri*. *Planta*, **212**, 75-84.
- Küpper H, Zhao FJ, McGrath SP (1999) Cellular compartmentation of zinc in the leaves of the hyperaccumulator *Thlaspi caerulescens*. *Plant Physiology*, **212**, 75-84.
- Lasat MM (2002) Phytoextraction of toxic Metals: A Review of Biological Mechanisms. *Journal of Environmental Quality*, **31**, 109-120.
- Lasat MM, Baker AJM, Kochian LV (1998) Altered Zn Compartmentation in the Root Symplasm and Stimulated Zn Absorption into the Leaf as Mechanisms Involved in Zn Hyperaccumulation in *Thlaspi caerulescens*. *Plant Physiology*, **118**, 875-883.
- Lasat MM, Pence NS, Garvin DF, Ebbs SD, Kochian LV (2000) Molecular physiology of zinc transport in the Zn hyperaccumulator *Thlaspi caerulescens*. *Journal of Experimental Botany*, **51**, 71-79.

- Lascoux M, Palme AE, Cheddadi R, Latta RG (2004) Impact of Ice Ages on the genetic structure of trees and shrubs. *Philosophical Transactions of the Royal Society of London Serie B*, **359**, 197-207.
- Lefebvre C (1970) Self-fertility in maritime and zinc mine populations of *Armeria maritima* (Mill.) Willd. *Evolution*, **24**, 571-577.
- Lefebvre C, Vernet P (1990) Microevolutionary processes on contaminated deposits. In: *Heavy metal tolerance in plants: evolutionary aspects* (ed. Shaw J), pp. 285-300. CRC Press Inc, Boca Raton.
- Leopold I, Gunther D, Schmidt J, Neumann D (1999) Phytochelatins and heavy metal tolerance. *Phytochemistry*, **50**, 1323-1328.
- Lexer C, Fay MF (2005) Adaptation to environmental stress: a rare or frequent driver of speciation? *Journal of Evolutionary Biology*, **18**, 893-900.
- Liessmann W (1997) *Historischer Bergbau Im Harz* Springer-Verlag, Berlin Heidelberg New York.
- Lou L-Q, Shen Z-G, Li X-D (2004) The copper tolerance mechanisms of *Elsholtzia haichowensis*, a plant from copper-enriched soils. *Environmental and Experimental Botany*, **51**, 111-120.
- Lumaret R, Mir C, Michaud H, Raynal V (2002) Phylogeographical variation of chloroplast DNA in holm oak (*Quercus ilex* L.). *Molecular Ecology*, **11**, 2327-2336.
- Ma H-X, Bai G-H, Carver B, Zhou L-L (2005) Molecular mapping of a quantitative trait locus for aluminum tolerance in wheat cultivar Atlas 66. *Theoretical and Applied Genetics*, **112**, 51-57.
- Macnair M, Gardner M (1998) The Evolution of Edaphic Endemics. In: *Endless Forms, Species and Speciation* (eds. Howard DJ, Berlocher SH), pp. 157-171. Oxford University Press, New York.
- Macnair M (1981) The uptake of copper plants of *Mimulus guttatus* differing in genotype primarily at a single major copper tolerance locus. *New Phytologist*, **88**, 723-730.
- Macnair MR (1983) The genetic control of copper tolerance in the yellow monkey flower, *Mimulus guttatus*. *Heredity*, **50**, 283-293.
- Macnair MR (1987) Heavy metal tolerance in plants: a model evolutionary system. *Trends in Ecology and Evolution*, **2**, 354.
- Macnair MR (1990) The genetics of metal tolerance in natural populations. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. Shaw AJ), pp. 235-253. CRC Press, Boca Raton, Florida.
- Macnair MR (1993) Tansley review N°49: the genetic of metal tolerance in vascular plants. *New Phytologist*, 541-559.
- Macnair MR (1997) The evolution of plants in metal-contaminated environments. In: *Environmental stress, Adaptation and Evolution* (eds. Bijlsma R, Loeschke V), pp. 1-24. Birkhäuser Verlag Basel.
- Macnair MR, Baker JM (1994) Metal-tolerant plant: An evolutionary perspective. In: *Plants and the chemical elements, biochemistry, uptake, Tolerance and toxicity* (ed. Farago ME), pp. 68-83. VCH, NY.

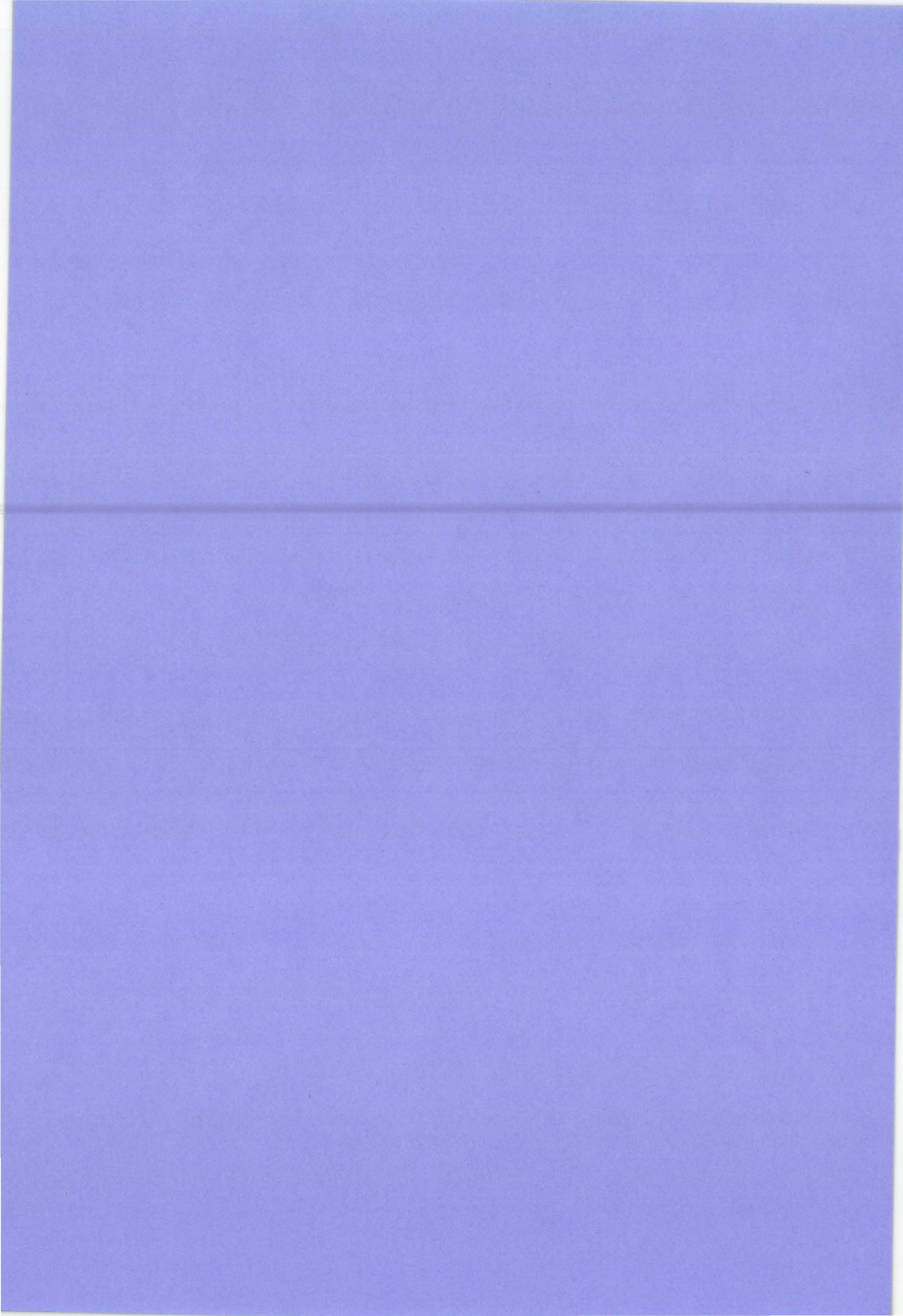
- Macnair MR, Bert V, Huitson SB, Saumitou-Laprade P, Petit D (1999) Zinc tolerance and hyperaccumulation are genetically independent characters. *Proceeding of the Royal Society of London, Biological Serie*, **266**, 2175-2179.
- Macnair MR, Christie P (1983) Reproductive isolation as a pleiotropic effect of copper tolerance in *Mimulus guttatus*? *Heredity*, **50**, 295-302.
- Macnair MR, Tilstone GH, Smith SE (2000) The Genetics of Metal Tolerance and Accumulation in Higher Plants. In: *Phytoremediation of Contaminated Soil and Water* (eds. Terry N, Banuelos G, Vangronsveld J), pp. 235-250. CRC Press, Boca Raton.
- Matthews DJ, Moran BM, Otte ML (2005) Screening the wetland plant species *Alisma plantago-aquatica*, *Carex rostrata* and *Phalaris arundinacea* for innate tolerance to zinc and comparison with *Eriophorum angustifolium* and *Festuca rubra* Merlin. *Environmental Pollution*, **134**, 343-351.
- McCauley DE (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology & Evolution*, **10**, 198-202.
- McNeilly T, Antonovics J (1968) Evolution in closely adjacent plant population. IV. Barriers to gene flow. *Heredity*, **23**, 205-218.
- McNeilly T, Bradshaw AD (1968) Evolutionary processes in populations of copper tolerant *Agrostis tenuis* Sibth. *Evolution*, **22**, 108-118.
- Meerts P, van Isacker N (1997) Heavy metal tolerance and accumulation in metallicolous and nonmetallicolous populations of *Thlaspi caerulescens* from continental Europe. *Plant Ecology*, **133**, 221-231.
- Meharg AA, Hartley-Whitaker J (2002) Arsenic uptake and metabolism in arsenic resistant and nonresistant plant species. *New Phytologist*, **154**, 29-43.
- Mengoni A, Barabesi C, Gonnelli C, *et al.* (2001) Genetic diversity of heavy metal-tolerant populations in *Silene paradoxa* L. (Caryophyllaceae): a chloroplast microsatellite analysis. *Molecular Ecology*, **10**, 1909-1916.
- Mengoni A, Gonnelli C, Brocchini E, *et al.* (2003) Chloroplast genetic diversity and biogeography in the serpentine endemic Ni-hyperaccumulator *Alyssum bertolonii*. *New Phytologist*, **157**, 349-356.
- Mitchell-Olds T (2001) *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends in Ecology & Evolution*, **16**, 693-700.
- Nicholls MK, McNeilly T (1985) The performance of *Agrostis capillaris* L. genotypes, differing in copper tolerance, in ryegrass swards on normal soils. *New Phytologist*, **101**, 207-217.
- Nordal I, Haraldsen KB, Ergon A, Eriksen AB (1999) Copper resistance and genetic diversity in *Lychnis alpina* (Caryophyllaceae) populations on mining sites. *Folia Geobotanica*, **34**, 471-481.
- Noret N, Meerts P, Tolra R, *et al.* (2005) Palatability of *Thlaspi caerulescens* for snails: influence of zinc and glucosinolates. *New Phytologist*, **165**, 763-772.
- O' Kane SL, Al-Shehbaz IA (1997) A synopsis of *Arabidopsis* (Brassicaceae). *Novon*, **7**, 323-327.

- O'Halloran TV, Culotta VC (2000) Metallochaperones, an Intracellular Shuttle Service for Metal Ions
10.1074/jbc.R000006200. *Journal of Biological Chemistry*, **275**, 25057-25060.
- O'Kane SL, Al-Shehbaz IA (2003) Phylogenetic position and generic limits of *Arabidopsis* (Brassicaceae) based on sequences of nuclear ribosomal DNA. *Annals of the Missouri Botanical Garden*, **90**, 603-612.
- Pauwels M, Frérot H, Bonnin I, Saumitou-Laprade P (submitted) A broad-scale analysis of population differentiation for Zn-tolerance in an emerging model species for tolerance study: *A. halleri* (Brassicaceae). *Journal of Evolutionary Biology*.
- Pence NS, Larsen PB, Ebbs SD, *et al.* (2000) The molecular physiology of heavy metal transport in the Zn/Cd hyperaccumulator *Thlaspi caerulescens*. *Proceeding of the National Academy of Science USA*, **97**, 4956-4960.
- Petit RJ, Aguinagalde I, de Beaulieu J-L, *et al.* (2003) Glacial Refugia: Hotspots But Not Melting Pots of Genetic Diversity. *Science*, **300**, 1563-1565.
- Petit RJ, Brewer S, Bordacs S, *et al.* (2002a) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, **156**, 49-74.
- Petit RJ, Csaikl UM, Bordacs S, *et al.* (2002b) Chloroplast DNA variation in European white oaks: Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management*, **156**, 5-26.
- Petit RJ, Csaikl UM, Bordacs S, *et al.* (2002) Chloroplast DNA variation in European white oaks: Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management*, **156**, 5-26.
- Petit RJ, Kremer A, Wagner DB (1993) Finite island model for organelle and nuclear genes in plant. *Heredity*, **71**, 630-641.
- Pollard AJ (2000) Metal hyperaccumulation: a model system for coevolutionary studies. *New Phytologist*, **146**, 179-181.
- Pollard AJ, Baker AJM (1997) Deterrence of herbivory by zinc hyperaccumulation in *Thlaspi caerulescens* (Brassicaceae). *New Phytologist*, **135**, 655-658.
- Pollard AJ, Powell KD, Harper FA, Smith JAC (2002) The genetic basis of metal hyperaccumulation in plants. *Critical Reviews in Plant Sciences*, **21**, 539-566.
- Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered *versus* unordered alleles. *Genetics*, 1237-1245.
- Prasad MNV, Freitas HMO (2003) Metal hyperaccumulation in plants - Biodiversity prospecting for phytoremediation technology. *Electronic Journal of Biotechnology [online]*, **6**, 285-321.
- Prat S (1934) Die Erbllichkeit der Resistenz gene Kupfer. *Berichte der Deutschen Botanischen Gesellschaft*, **52**, 65-78.
- Purugganan M, Gibson G (2003) Merging ecology, molecular evolution, and functional genetics. *Molecular Ecology*, **12**, 1109-1112.
- Reeves RD, Baker AJM, Borhidi A, Berazain R (1999) Nickel Hyperaccumulation in the Serpentine Flora of Cuba. *Annals of Botany*, **83**, 29-38.

- Rendell S, Ennos RA (2003) Chloroplast DNA diversity of the dioecious European tree *Ilex aquifolium* L. (English holly). *Molecular Ecology*, **12**, 2681-2688.
- Sarret G, Saumitou-Laprade P, Bert V, *et al.* (2002) Forms of zinc accumulated in the hyperaccumulator *Arabidopsis halleri*. *Plant Physiol*, **130**, 1815-1826.
- Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, **7**, 465-474.
- Schat H, Kuiper E, Ten Bookum WM, Vooijs R (1993) A general model for the genetic control of copper tolerance in *Silene vulgaris*: evidence from crosses between plants from different tolerant populations. *Heredity*, **70**, 142-147.
- Schat H, Llugany M, Bernhard R (2000) Metal-specificity patterns of tolerance, uptake, and transport of heavy metals in hyperaccumulating and nonhyperaccumulating metallophytes. In: *Phytoremediation of contaminated soil and water* (eds. Terry N, Bañuelos G), pp. 171-188. Lewis Publishers, Boca Raton, FL, USA.
- Schat H, Llugany M, Vooijs R, Hartley-Whitaker J, Bleeker PM (2002) The role of phytochelatins in constitutive and adaptive heavy metal tolerances in hyperaccumulator and non-hyperaccumulator metallophytes. *Journal of Experimental Botany*, **53**, 2381-2392.
- Schat H, Ten Bookum WM (1992) Genetic control of copper tolerance in *Silene Vulgaris*. *Heredity*, **68**, 219 - 229.
- Schat H, Vooijs R (1997) Multiple tolerance and co-tolerance to heavy metals in *Silene vulgaris*: a co-segregation analysis. *New Phytologist*, **136**, 489-496.
- Schat H, Vooijs R, Kuiper E (1996) Identical major gene loci for heavy metal tolerances that have independently evolved in different local populations and subspecies of *Silene vulgaris*. *Evolution*, **50**, 1888-1895.
- Schlotterer C (2001) Genealogical inference of closely related species based on microsatellites. *Genetical Research*, **78**, 209-212.
- Schluter D (2001) Ecology and the origin of species. *Trends in Ecology & Evolution*, **16**, 372-380.
- Schonswetter P, Paun O, Tribsch A, Niklfeld H (2003) Out of the Alps: colonization of Northern Europe by East Alpine populations of the Glacier Buttercup *Ranunculus glacialis* L. (Ranunculaceae). *Molecular Ecology*, **12**, 3373-3381.
- Schonswetter P, Popp M, Brochmann C (2006) Rare arctic-alpine plants of the European Alps have different immigration histories: the snow bed species *Minuartia biflora* and *Ranunculus pygmaeus*. *Molecular Ecology*, **15**, 709-720.
- Schonswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, **14**, 3547-3555.
- Schonswetter P, Tribsch A, Barfuss M, Niklfeld H (2002) Several Pleistocene refugia detected in the high alpine plant *Phyteuma globulariifolium* Sternb. & Hoppe (Campanulaceae) in the European Alps. *Molecular Ecology*, **11**, 2637-2647.
- Sledge MK, Bouton JH, Dall'Agnoll M, Parrott WA, Kochert G (2002) Identification and Confirmation of Aluminum Tolerance QTL in Diploid *Medicago sativa* subsp. *coerulea*. *Crop Science*, **42**, 1121-1128.

- Smith SE, Macnair M (1998) Hypostatic modifiers cause variation in degree of copper tolerance in *Mimulus guttatus*. *Heredity*, **80**, 760-768.
- Soltis DE, Soltis PS, Milligan BG (1992) Intraspecific Chloroplast DNA Variation: Systematic and Phylogenetic Implications. In: *Molecular systematics of plants* (eds. Soltis PS, Soltis DE, Doyle JJ), pp. 117-150. Chapman and Hall, New York.
- Song J, Zhao F-J, Luo Y-M, McGrath SP, Zhang H (2004) Copper uptake by *Elsholtzia splendens* and *Silene vulgaris* and assessment of copper phytoavailability in contaminated soils. *Environmental Pollution*, **128**, 307-315.
- Staton JL, Schizas NV, Chandler GT, Coull BC, Quattro JM (2001) Ecotoxicology and population genetics: the emergence of "phylogeographic and evolutionary ecotoxicology". *Ecotoxicology*, **10**, 217-222.
- Stehlik I, Schneller JJ, Bachmann K (2001) Resistance or emigration: response of the high-alpine plant *Eritrichium nanum* (L.) Gaudin to the ice age within the Central Alps. *Molecular Ecology*, **10**, 357-370.
- Strange J, Macnair M (1991) Evidence for a role for the cell membrane in copper tolerance of *Mimulus guttatus* Fischer ex DC. *New Phytologist*, **119**, 383-388.
- Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol Biol*, **17**, 1105-1109.
- Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381-397.
- Terry N, Bañuelos G (2000) Phytoremediation of contaminated soil and water, p. 389. CRC press, Boca Raton, FL, USA.
- Thomine S, Wang R, Ward JM, Crawford NM, Schroeder JI (2000) Cadmium and iron transport by members of a plant metal transporter family in *Arabidopsis* with homology to Nramp genes. *Proceeding of the National Academy of Science USA*, **97**, 4991-4996.
- Tilstone GH, Macnair MR (1997) Nickel tolerance and copper - nickel co - tolerance in *Mimulus guttatus* from copper mine and serpentine habitats. *Plant and Soil*, **191**, 173-180.
- Tolra RP, Poschenrieder C, Luppi B, Barcelo J (2005) Aluminium-induced changes in the profiles of both organic acids and phenolic substances underlie Al tolerance in *Rumex acetosa* L. *Environmental and Experimental Botany*, **54**, 231-238.
- Tribsch A, Schonswetter P (2003) Patterns of endemism and comparative phylogeography confirm palaeoenvironmental evidence for Pleistocene refugia in the Eastern Alps. *Taxon*, **52**, 477-497.
- Tribsch A, Schonswetter P, Stuessy TF (2002) *Saponaria pumila* (Caryophyllaceae) and the Ice Age in the European Alps. *American Journal of Botany*, **89**, 2024-2033.
- Turner RG (1969) Heavy metal tolerance in plants. In: *Ecological Aspects of the Mineral Nutrition of Plants* (ed. Rorison IH), pp. 399-410. Blackwell Scientific Publications, Oxford.
- van Rossum F, Bonnin I, Fenart S, et al. (2004) Spatial genetic structure within a metal-tolerant population of *Arabidopsis halleri*, a clonal, self-incompatible and heavy-metal-tolerant species. *Molecular Ecology*, **13**, 2959-2967.

- Vekemans X, Lefèbvre C (1997) On the evolution of heavy-metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. *Journal of Evolutionary Biology*, **10**, 175-191.
- Webb T, Bartlein PJ (1992) Global changes during the last 3 million years: climatic controls and biotic response. *Annual Review of Ecology and Systematics*, **23**, 141-173.
- Weber M, Harada E, Vess C, Roepenack-Lahaye E, Clemens S (2004) Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identifies nicotianamine synthase, a ZIP transporter and other genes as potential metal hyperaccumulation factors. *The Plant Journal*, **37**, 269-281.
- Westerbergh A, Saura A (1992) The effect of serpentine on the population structure of *Silene dioica* (Caryophyllaceae). *Evolution*, **46**, 1537-1548.
- Wilkins DA (1978) The measurement of tolerance to edaphic factors by means of root growth. *New Phytologist*, **80**, 623-633.
- Willems G (2006) *Characterisation of zinc and cadmium tolerance in Arabidopsis halleri*, Université de Sciences et Technologies de Lille 1 - Université Libre de Belgique.
- Willems G, Godé C, Dräger DB, *et al.* (submitted) Quantitative Trait Loci mapping of zinc tolerance in the metallophyte *Arabidopsis halleri* ssp. *halleri*. *Genetics*.
- Williams LE, Pittman JK, Hall JL (2000) Emerging mechanisms for heavy metal transport in plants. *Biochimica et Biophysica Acta*, **1465**, 104-126.
- Willis KJ (1996) Where did all the flowers go? The fate of temperate European flora during glacial periods. *Endeavour*, **20**, 110-114.
- Willis KJ, Niklas KJ (2004) The role of Quaternary environmental change in plant macroevolution: the exception or the rule? *Philos Trans R Soc Lond B Biol Sci*, **359**, 159-172; discussion 172.
- Willis KJ, Rudner E, Sumegi P (2000) The Full-Glacial Forests of Central and Southeastern Europe. *Quaternary Research*, **53**, 203-213.
- Willis KJ, van Andel TH (2004) Trees or no trees? The environments of central and eastern Europe during the Last Glaciation. *Quaternary Science Reviews*, **23**, 2369-2387.
- Wu L, Bradshaw AD, Thurman DA (1975) The potential for evolution of heavy metal tolerance in plants. III. The rapid evolution of copper tolerance in *Agrostis stolonifera*. *Heredity*, **34**, 165-187.
- Zenk MH (1996) Heavy metal detoxification in higher plants - a review. *Gene*, **179**, 21-30.



Chapitre 1: A broad-scale analysis of population differentiation for Zn-tolerance in an emerging model species for tolerance study: *A. halleri* (Brassicaceae). Accepté dans JEB.

Maxime Pauwels, H el ene Fr erot, Isabelle Bonnin et Pierre Saumitou-Laprade

Abstract Although current knowledge about the overall distribution of zinc tolerance in *Arabidopsis halleri* populations is scarce, the species is an emerging model for the study of heavy metal tolerance in plants. We attempted to improve this knowledge by testing the Zn tolerance of scattered European metallicolous (M) and non-metallicolous (NM) populations of *A. h.* subsp. *halleri* and *A. h.* subsp. *ovirensis* in hydroponic culture. The occurrence of constitutive tolerance was unconditionally established in *A. h. halleri* and tolerance was extended to the subspecies *ovirensis*. M populations were the most tolerant but there was continuous range of variation in tolerance from NM to M populations. Finally, relatively high levels of tolerance were detected in some NM populations, suggesting that enhanced tolerance could be present at high frequency in populations that have not experienced metal exposure. We used our results to argue the evolutionary dynamics and origin of Zn tolerance in *A. halleri*.

Keywords: *Arabidopsis halleri halleri*, model species, pseudometallophyte, metallicolous and non-metallicolous populations, hydroponics, metal tolerance evolution, *A. halleri ovirensis*

Introduction

Heavy metal tolerance in plants is commonly defined as the ability to survive on soils which would prove toxic to most living things because the soils contain (very) high levels of one or more metals (Antonovics *et al.* 1971, Macnair and Baker 1994). It was first recognised by (Prat 1934) who showed that a copper mine population of *Silene dioica* (syn. *Melandrium silvestre*) was able to grow far better in copper-contaminated soils than populations from uncontaminated areas. From this initial work, tolerances to a wide range of heavy metals have been encountered in many plant species belonging to widely separated families (Antonovics *et al.* 1971). These taxa have been classified as either absolute (strict or eu-) or facultative (pseudo-) metallophytes, depending on whether they occur on contaminated sites only or on metalliferous as well as non-metalliferous soils (Lambinon and Auquier 1964).

Metal tolerance has been considered to be “an example of more powerful evolution in action than industrial melanism in moths” (Antonovics *et al.* 1971). It has therefore been the focus of many studies and the evolutionary framework of heavy metal tolerance in plants is nowadays relatively well documented. Evolutionary studies have argued that metal tolerance could evolve rapidly following exposure to heavy metal stress (Wu *et al.* 1975, Al-Hiyaly *et al.* 1988) and that it could have evolved independently in geographically distant conspecific populations (Westerbergh and Saura 1992, Schat *et al.* 1996, Vekemans and Lefèbvre 1997, Koch *et al.* 1998, Mengoni *et al.* 2001, Pauwels *et al.* 2005). Recent genetics studies have demonstrated that, in most cases, tolerance is governed by a few major genes (for reviews, see Macnair 1993, Macnair *et al.* 2000). Quantitative polymorphism is nevertheless usually observed and implies the existence of hypostatic “modifier” genes that influence the degree of expression of the tolerance phenotype (Schat and Ten Bookum 1992, Schat *et al.* 1993, Smith and Macnair 1998, van Hoof *et al.* 2001). In particular, modifiers are supposed to be responsible for the continued evolution toward a high level of average tolerance in metalicolous populations, in response to high levels of metal exposure (Schat *et al.* 1993, Smith and Macnair 1998).

In recent years, substantial efforts have been expended to gain further insights in the genetic mechanisms of metal tolerance in plants, *i.e.* identifying candidate genes and characterizing physiological mechanisms. Thus, many studies have involved high throughput

genomics, transcriptomics, proteomics and metabolomics technologies. Genomic and post-genomic technologies are necessarily organized around only a few model species (Feder and Mitchell-Olds 2003) which makes it crucial to search for the appropriate model. Whereas classical models (e.g. *Arabidopsis thaliana*) were selected by scientists for their particular genetic and development features, the study of ecologically important traits is motivated by additional criteria (Jackson *et al.* 2002). In the context of heavy metal tolerance in plants, two pseudometallophyte species that both tolerate and hyperaccumulate zinc (Zn) and cadmium (Cd) have recently been proposed: *Arabidopsis halleri* (L.) (O’Kane & Al-Shehbaz) (syn. *Cardaminopsis halleri* (L.) Hayek) and *Thlaspi caerulescens* J. & C. Presl., both Brassicaceae (Assunção *et al.* 2003b).

Because it is the closest metal tolerant relative of the pre-eminent model system *A. thaliana* (Koch *et al.* 2001, Al-Shehbaz and O’Kane 2002), *Arabidopsis halleri* doubtlessly presents major advantages for the study of tolerance in plants. Indeed, the fact that many molecular tools developed in *A. thaliana* can readily be transferred to *A. halleri* (Mitchell-Olds 2001, Lexer and Fay 2005) offers the unparalleled opportunity of an integrative study of metal tolerance. Topics as diverse as the nature, the number and the regulation of genes conferring metal tolerance, the epistatic relationships between them, their origin and evolution of tolerance genes, the impact of selection on these genes can now be addressed. Thus, *A. halleri* is already included in –omics studies, and its efficacy in identifying candidate genes involved in metal tolerance and hyperaccumulation has been demonstrated (Becher *et al.* 2004, Dräger *et al.* 2004, Weber *et al.* 2004).

However, the best molecular technologies will be of limited interest if they are not sustained by a good understanding of the ecological range of the model species (*i.e.* the relative abundance of populations on and off polluted sites) as well as the distribution of tolerance abilities in populations (is this distribution related to the local amount of metals in soils?). So far, in contrast to *T. caerulescens* which has been studied for many decades (Koch *et al.* 1998, Assunção *et al.* 2003b), this knowledge remains sparse in *A. halleri* which was sometimes said to be typically found on metalliferous soils (Dahmani-Muller *et al.* 2000, Pollard *et al.* 2002) although numerous NM populations has been mentioned (Bert *et al.* 2000, Bert *et al.* 2002). Zn tolerance was assumed to occur throughout the species in *A. halleri* (Bert *et al.* 2000), but only five populations located in two regions of northern Europe and belonging to the single subspecies *A. h. halleri* were studied in controlled conditions.

Meanwhile, the species is widely distributed in Europe and at least two other subspecies are recognized (Al-Shehbaz and O'Kane 2002). Although it is assumed to be constitutive, Zn tolerance is generally expected to have evolved towards increased levels in M populations of *A. halleri*, in response to high selective pressure caused by high level of metals in soils (Bert *et al.* 2000). Recent phylogeographic study using cpDNA (Pauwels *et al.* 2005) showed that *A. h. halleri* population structure was related to geographic isolation rather than to Zn exposure in northern Europe. This implies that geographically close non-metallicolous (NM) and metallicolous (M) populations (e.g. from southern and northern Germany) are more genetically similar than distant M populations (e.g. from Poland and northern Germany), and that distant M populations have evolved independently. Consequently, if enhanced tolerance was confirmed to be a general feature of M populations, it should have evolved several times in parallel. Moreover, a study of population genetic structure using 5 nuclear microsatellites over a 500 m long transect characterized by a gradient of heavy metal concentrations found no evidence of genetic divergence due to spatial heavy metal heterogeneity in these neutral markers, suggesting that long-distance pollen dispersal and extensive pollen flow could attenuate the local effect of metal exposure on differentiation by neutral markers (van Rossum *et al.* 2004). Altogether, population studies based on neutral genetic markers suggest that the relation between Zn exposure in the wild and the distribution of inherited Zn tolerance abilities in populations might not be as straightforward as usually thought.

The present study attempted to provide a better description and understanding of the distribution of Zn tolerance in *A. halleri* populations, in particular in relation to the available Zn content in soils, and discusses the origin and the evolutionary dynamics of Zn tolerance in *A. halleri* in Europe. To assess accurately the overall level of quantitative polymorphism for Zn tolerance, 28 widely distributed populations of *A. h. subsp. halleri*, the only subspecies which has been recorded on metalliferous soils in Europe, were sampled. To be representative of the ecological range of the subspecies, populations were collected in both polluted and non-polluted areas. In addition, three populations of *A. h. subsp. ovirensis* were sampled in the Carpathians Mountains (Romania). To identify the genetic component of population differentiation for Zn tolerance, the tested populations were grown under uniform environmental conditions using hydroponic culture in controlled growth chambers.

Table 1 Names and locations of population samples of *A. halleri*. When shared, the names correspond to those used in Pauwels *et al.* (2005). A: Austria; B: Belgium; CH: Switzerland; D: Germany; F: France; I: Italy; PL: Poland; RO: Romania; SK: Slovakia; M: metallicolous populations; NM: non-metallicolous populations; NMP: non-polluted populations; ovirensis: *A. h. ovirensis* populations. N_i est.: approximate population size; n_s: number of soil samples; n: sample size for sequential test; n.a.: not available.

population	subspecies	type	Origin	GPS coordinates		altitude (m)	ecological background	Ni est.	extractable Zn in soil (µM) ± SE (n _s)	year of test	n
				North	East						
A05	<i>A. h. halleri</i>	NM	Tyrol	47°13'778	11°22'779	807	garden lawn under conifers	<50	7 ± 3 (3)	2004	20
A08	<i>A. h. halleri</i>	NM	Tyrol	47°25'177	11°51'957	522	meadow	100-500	177 ± 123 (4)	2004	20
B01	<i>A. h. halleri</i>	NM	Hautes Fagnes	50°29'63	06°40'	160	underwood	500-1000	n.a.	2000	16
CH01	<i>A. h. halleri</i>	NM	Tessin	45°59'233	08°50'387	340	on a shaded wall, around a private garden	>100	5 (2)	2004	20
CH03	<i>A. h. halleri</i>	NM	Tessin	46°10'632	08°43'182	290	meadow at a wood skirt, along a roadway	<50	17 ± 6 (3)	2004	18
D01 (G1*)	<i>A. h. halleri</i>	NM	Bohemian forest	49°10'64	12°09'88	340	nitrogenous Regen river bank	<50	3 ± 2*	2000	10
D08 (G8*)	<i>A. h. halleri</i>	NMp	Harz	51°53'79	10°29'04	190	old mine (19th century)	<50	123 ± 9*	2000	13
D09 (G9*)	<i>A. h. halleri</i>	NMp	Harz	51°53'46	10°25'16	190	Roadside	<50	52*	2000	15
D11 (G11*)	<i>A. h. halleri</i>	NMp	Harz	51°51'27	10°21'95	674	Underwood	<50	7*	2000	11
D12 (G12*)	<i>A. h. halleri</i>	M	Harz	51°51'91	10°17'90	673	mine rubble	>500	1272 ± 1047*	2000	16
D13 (G13*)	<i>A. h. halleri</i>	M	Harz	51°55'22	10°18'50	237	roadside and lawn	50-500	405 ± 207*	2000	21
F01 (Auby†)	<i>A. h. halleri</i>	M	North	50°25'	03°03'	20	wood near a smelter plant	>1000	5260 ± 280†	2000/2004	14/20
I01	<i>A. h. halleri</i>	NM	Trentin	46°52'783	11°24'308	967	small meadow along a roadway	<100	13 ± 7 (3)	2004	5
I02	<i>A. h. halleri</i>	NM	Trentin	46°49'920	11°43'823	840	meadow along a footpath	100-500	4 ± 4 (2)	2004	20
I07	<i>A. h. halleri</i>	NM	Trentin	46°29'623	10°53'349	1388	Meadow	>1000	8 ± 5 (3)	2004	19
I08	<i>A. h. halleri</i>	NM	Trentin	46°29'731	10°53'739	1363	Meadow	500-1000	17 ± 10 (2)	2004	20
I09	<i>A. h. halleri</i>	NM	Trentin	46°43'882	11°25'958	1530	Meadow	>5000	11 ± 8 (2)	2004	20
PL01 (P1*)	<i>A. h. halleri</i>	NMp	Silesia	50°14'80	18°57'04	200	wood in Katowice suburbs	>500	22 ± 1*	2000	20
PL04 (P4*)	<i>A. h. halleri</i>	M	Silesia	50°29'98	18°55'79	269	metallurgical factory	>1000	481*	2000	10
PL06 (P5*)	<i>A. h. halleri</i>	M	Silesia	50°16'95	19°01'52	306	old mine	>1000	2490*	2000	10
PL07	<i>A. h. halleri</i>	M	Silesia	50°17'014	19°29'564	335	meadow on a ZN/Pb spoil	n.a.	n.a.	2004	20
PL08	<i>A. h. halleri</i>	NMp	Silesia	50°06'243	20°21'569	190	oak-hornbeam forest	n.a.	n.a.	2004	20
RO05	<i>A. h. halleri</i>	NM	Carpathians	46°43'305	23°02'637	990	footpath along a forest skirt, near a brook	<50	3 ± 3 (2)	2004	20
RO06	<i>A. h. halleri</i>	NM	Carpathians	46°39'083	23°02'364	1154	<i>Picea abies</i> forest skirt	<50	14 (1)	2004	20
RO08	<i>A. h. halleri</i>	NM	Carpathians	46°50'688	22°48'082	900	river bank	<10	3 (1)	2004	20
RO09	<i>A. h. halleri</i>	NM	Carpathians	46°49'475	22°45'635	666	meadow beside a road	<50	9 ± 8 (3)	2004	16
RO14	<i>A. h. ovirensis</i>	NM	Fagaras Mountains	45°36'139	24°37'064	2050	alpine lawn	50-500	n.a.	2004	20
RO15	<i>A. h. ovirensis</i>	NM	Fagaras Mountains	45°36'139	24°37'064	2050	alpine lawn	50-500	n.a.	2004	20
RO18	<i>A. h. ovirensis</i>	NM	Fagaras Mountains	45°36'163	24°37'336	2230	alpine lawn	<50	21 ± 25 (2)	2004	18
SK02 (S12*)	<i>A. h. halleri</i>	NM	Tatras	48°46'17	21°07'81	690	shady meadow	>500	1*	2000	15
SK05 (S14*)	<i>A. h. halleri</i>	NM	High Tatras	49°16'98	20°09'24	1027	Tatransla javorina (Nature reserve)	>500	1*	2000/2004	21/8

* in Bert *et al.* (2002); † in Bert *et al.* (2000)

Material and methods

Sampling

In the summers of 1999 and 2003, respectively 12 and 16 widely distributed populations of *A. halleri* subsp. *halleri* were sampled so as to cover the range of the subspecies in continental Europe (Table 1). At each study site, viable mature seeds were collected as maternal progenies from distinct individuals in order to obtain living plants and to perform a tolerance test in hydroponic culture. Soil samples were also collected at each study site in order to estimate the concentrations of extractable Zn. Many results from soil sample analysis have already been published in Bert *et al.* (2002). Additional samples were analysed using the protocol described in Bert *et al.* (2002) and all the results are reported in Table 1.

Most populations (17) were scattered in the European mountain ranges (Figure 1), at middle to high altitudes (mean = 822 m), far from any known metal pollution source, in non-contaminated environments (NM populations, Table 1). They were located in the Tatras and High Tatras mountains (Slovakia), in the Apuseni Mountains (Romania), in the Bohemian Forest (Germany) and on northern and southern slopes of the Alps (Austrian Tyrol; Ticino, Switzerland and Trentino, Italy). Eleven populations (D08-09, D11:13, F01, PL01, PL04, PL06:08, Figure 1) were sampled outside the mountain ranges, at low to moderate altitudes (mean = 314 m), in the three known polluted areas the species has colonized, i.e. in northern France, Silesia (Poland) and Harz (Germany). In these disturbed regions, *A. h. halleri* populations have been shown to locally cluster into homogeneous genetic assemblages (Pauwels *et al.* 2005). Since the extension of the *A. h. halleri* distribution area in initially unsuitable disturbed habitats has to be related to anthropogenic metal pollution via industrial or mining activities (Berton 1946, Fabiszewski 1986, Ernst 1990), all populations were classified as M populations and most of them (D12-13, F01, PL04-06-07) indeed occurred on highly polluted soils (Table 1). However, a detailed study of Zn concentration in the soils of each population revealed that the species also colonized slightly polluted habitats (< 300 µg.g⁻¹, see Bert *et al.* 2002) where selection pressure towards enhanced tolerance should be strongly reduced (such as in NM populations, see Table 1). In industrial areas, "non-M" populations (D08-09-11 and PL01-08) may nevertheless differed from true NM ones by their abilities to exchange genes with the geographically proximate M populations (van Rossum *et al.* 2004, Pauwels *et al.* 2005). Both ecological and genetic considerations (Pauwels *et al.*



Fig. 1: Distribution map of sampled *Arabidopsis halleri* populations in Europe. NM, NMp and M populations are respectively indicated by white, grey and black circles. Populations of *A. h. ovirensis* are indicated by white triangles.

2005) further suggests that M populations might have settled first in these areas and that populations only subsequently colonised slightly polluted habitats. Populations D08-09-11 and PL01-08 were therefore considered to be in a particular category and qualified as NMp (for Non-Metallicolous in a polluted area).

In addition to *A. h. halleri* populations, mature seeds from maternal progenies and soil samples from 3 populations of *A. h. subsp. ovirensis* were collected at high altitudes (mean = 2110 m) in the Fagaras Mountains (Southern Carpathians, Romania).

Zinc tolerance experiment

Seeds sampled from each population were first sown in sand filled pots in a greenhouse. After three weeks of growth, a maximum of 20 seedlings per population (depending on the germination rate) were transferred into hydroponic conditions in a controlled-growth chamber (temperature 20°C day: 15°C night; light 14 h day: 10 h night). Ten-litre containers containing perforated polystyrene trays in which the seedlings were planted into nylon wool were used. In order to avoid local environmental effects, plants were randomly distributed in containers so that each population was represented by at least one individual in each container. Moreover, containers were randomly distributed in the growth chamber and moved around once a week. The composition of the nutrient solution in containers was: 0.2 mM MgSO₄, 0.5 mM Ca(NO₃)₂, 0.5 mM KNO₃, 0.1 mM K₂HPO₄, 0.2 μM CuSO₄, 2 μM MnCl₂, 10 μM H₃BO₃, 0.1 μM MoO₃, 10 μM FeEDDHA and 1 μM ZnSO₄. The nutrient solution was changed once a week. After a three-week period allowing plants to acclimate to hydroponic conditions, the sequential tolerance test started following the method of Schat and Ten Bookum (Schat and Ten Bookum 1992). It involves the qualitative measurement of root growth (growth vs no growth) at sequential concentrations. In our experiment, Zn concentrations varied from 1 μM during the first week to 2000 μM during the 9th and last week of treatment, increasing each week by a constant step of 250 μM. At the start of the sequential test, the roots of all plants were blackened by dipping them in a suspension of activated charcoal, and rinsed in deionized water to eliminate excess of charcoal. Plants were returned to 1 μM Zn for a further week when the presence of any new roots (uncoated) visible beyond the charcoal-coated roots was recorded. Roots were then reblackened and transferred to a fresh test solution with an increased Zn concentration for an additional week. The recording and blackening of new roots were repeated for each plant at each concentration

tested. In such a test, the presence of white roots at a given Zn concentration indicates growth. The lowest concentration at which root growth definitively stopped (EC100 in Schat and Ten Bookum 1992) was determined for each individual. EC100 was interpreted as the lowest concentration that was not tolerated by the individual in question. To check this interpretation, we transplanted “non-tolerant” plants to a fresh solution containing a Zn dose corresponding to their respective EC100s. Mortality usually occurred after 2 weeks of exposure at such a constant dose, validating the use of “presence of root growth or not” as an indicator of plant tolerance and, by definition, survival. Plants for which no new roots were observed were removed from the experiment.

Populations sampled in 1999 and 2003 were tested for Zn tolerance separately in time (in spring 2000 and spring 2004, respectively, see Table 1). However, in order to be able to combine survival data from both experiments into a single data set, the same protocol was used, except for the plant density in containers (28 plants per container in 2000, 24 in 2004). Populations F01 (M) and SK05 (NM) were also included in each experiment to allow comparison and to ensure of the homogeneity of data sets. In 2000, 12 individuals of *A. thaliana* were randomly distributed in containers so as to be represented 1 to 3 times in each container; in 2004, 20 individuals of *A. lyrata petrea* were included and similarly randomized. Both these non-metallophyte relatives were used as non-tolerant control species.

Statistical analysis

At each experimental dose, root growth was encoded for each plant as a binary variable and interpreted as individual survival or mortality. Survival proportions S_{iX} of individuals from population i that did not reach their EC100 were estimated at each dose X for which an end of growth was observed in that population. Note that the number n_i of survival proportions S_{iX} was a secondary variable that could have differed between populations. The resulting survival curves [$S_i = f(X)$, where “ X ” is the concentration of Zn in the test solution] were fitted to sigmoidal dose-response curves with a variable slope by non-linear regression using GraphPad Prism version 4.00 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com. The equation used for regression was a two-parameter logistic equation established as following: $S_i = 100 / (1 + 10^{((\log(T_{50i}) - \log(X)) * B_i)})$ where the variable “ S_i ” represents the survival proportions of population i , expressed as a percentage, “ X ” is the concentration of Zn in the test solution, expressed in μM , the parameter

" T_{50i} " is the concentration for which 50% of individuals from population i had reached their EC100 (median EC100, $S_{iT_{50i}} = 50$), and the parameter " B_i " is the slope factor describing the steepness of the curve. Estimated T_{50} values were considered as overall estimations of the tolerance of populations whereas the slope factors B estimated the range of within-population polymorphism for tolerance. For both parameters and for all populations, the standard error (SE) and 95% confidence intervals of best-fit values were estimated by GraphPad prism during curve fitting.

In order to test for differences in average levels of tolerance abilities or within-population polymorphism between groups of *A. h. halleri* populations either tested separately in time (2000 vs 2004) or corresponding to distinct ecological categories (NM vs NMp vs M), box-plots were drawn for both parameters and for each category and Mann-Whitney tests were performed using Minitab 13.20 (Minitab Inc. State College, Pennsylvania, USA).

In order to estimate precise confidence intervals for T_{50} and to be able to perform pairwise comparisons of tolerance abilities of populations, the equation used for regression was therefore rewritten so as to directly estimate best-fit values and SE for $\log(T_{50})$ rather than T_{50} (Motulsky 1999) [$S=100/(1+10^{((\text{LOG}T_{50}-\log(X))*B})}$, where $\text{LOG}T_{50}$ and B are the estimated parameters]. $\text{LOG}T_{50}$ values were then compared through unplanned comparisons between pairs of populations using the GT2-method for unequal sample size (Sokal and Rohlf 1981). The method computes comparison intervals for $\text{LOG}T_{50}$ for each i population using the two-tailed studentized maximum modulus $m_{\alpha[k^*,v]}$ distribution as a critical value (Sokal and Rohlf 1981). Best-fit values from distinct populations were declared significantly different if their intervals did not overlap. For a 5% experiment-wise error rate, $k^* = (35*(35-1))/2 = 595$ and $v = 124$, we used the tabled value $m_{.05[561,120]} = 4.038$ as a conservative approximation to the desired value $m_{.05[595,153]}$.

Table 2 Best-fit values and standard errors of T_{50} , $\text{LOG}T_{50}$ and B . n_i : number of doses for which an end-of-growth event was observed in the population i , corresponds to the number of survival proportions available for survival curves fitting; d.f.: degree of freedom of parameters (note that df are equal to n_i-2 rather than n_i-1 because both parameters were estimated at the same time, Motulsky, 1999); SE: Standard Error. NM, NMP and M populations are respectively indicated in light grey, dark grey and black

Population	n_i	T_{50} (\pm SE)	$\text{LOG}T_{50}$ (\pm SE)	B (\pm SE)	d.f.
A05	6	901.9 (\pm 17.71)	2.955 (\pm 8.529 10^{-3})	-6.818 (\pm 0.765)	4
A08	6	1204 (\pm 19.72)	3.081 (\pm 7.113 10^{-3})	-7.218 (\pm 0.805)	4
B01	6	1011 (\pm 52.69)	3.005 (\pm 2.262 10^{-2})	-4.237 (\pm 0.681)	4
CH01	7	1129 (\pm 24.75)	3.053 (\pm 9.523 10^{-3})	-5.868 (\pm 0.724)	5
CH03	5	1099 (\pm 33.02)	3.041 (\pm 1.305 10^{-2})	-7.306 (\pm 1.468)	3
D01	6	1233 (\pm 31.68)	3.091 (\pm 1.115 10^{-2})	-7.533 (\pm 1.347)	4
D08	4	1437 (\pm 21.03)	3.157 (\pm 6.355 10^{-3})	-5.128 (\pm 0.615)	2
D09	5	1335 (\pm 44.04)	3.125 (\pm 1.433 10^{-2})	-7.31 (\pm 1.365)	3
D11	6	1247 (\pm 34.96)	3.096 (\pm 1.217 10^{-2})	-5.727 (\pm 0.865)	4
D12	6	1432 (\pm 35.30)	3.156 (\pm 1.07 10^{-2})	-7.910 (\pm 1.391)	4
D13	5	1384 (\pm 23.06)	3.141 (\pm 7.239 10^{-3})	-8.029 (\pm 1.047)	3
F01-2000	5	1503 (\pm 18.02)	3.177 (\pm 5.205 10^{-3})	-9.909 (\pm 1.083)	3
F01-2004	6	1569 (\pm 39.06)	3.196 (\pm 1.081 10^{-2})	-9.13 (\pm 1.948)	4
I01	4	1092 (\pm 31.44)	3.038 (\pm 1.25 10^{-2})	-4.921 (\pm 0.697)	2
I02	6	993.6 (\pm 12.97)	2.997 (\pm 5.669 10^{-3})	-8.986 (\pm 1.036)	4
I07	6	1348 (\pm 13.99)	3.13 (\pm 4.508 10^{-3})	-8.876 (\pm 0.720)	4
I08	5	1144 (\pm 27.37)	3.058 (\pm 1.039 10^{-2})	-7.451 (\pm 1.276)	3
I09	7	1007 (\pm 23.59)	3.003 (\pm 1.017 10^{-2})	-5.516 (\pm 0.629)	5
PL01	6	1407 (\pm 16.84)	3.148 (\pm 5.196 10^{-3})	-9.078 (\pm 0.912)	4
PL04	5	1509 (\pm 24.02)	3.174 (\pm 1.299 10^{-2})	-5.827 (\pm 0.742)	3
PL06	4	1648 (\pm 15.13)	3.217 (\pm 03.99 10^{-3})	-15.58 (\pm 1.778)	2
PL07	6	1495 (\pm 38.91)	3.175 (\pm 1.13 10^{-2})	-8.115 (\pm 1.574)	4
PL08	7	1390 (\pm 22.05)	3.143 (\pm 06.89 10^{-3})	-7.324 (\pm 0.768)	5
RO05	5	1262 (\pm 66.1)	3.101 (\pm 2.275 10^{-2})	-5.886 (\pm 1.607)	3
RO06	8	1208 (\pm 42.52)	3.082 (\pm 1.528 10^{-2})	-6.06 (\pm 0.451)	6
RO08	6	1199 (\pm 13.01)	3.079 (\pm 4.714 10^{-3})	-4.375 (\pm 0.923)	4
RO09	6	1377 (\pm 49.84)	3.139 (\pm 1.572 10^{-2})	-6.796 (\pm 2.316)	4
RO14	6	1368 (\pm 65.99)	3.136 (\pm 2.096 10^{-2})	-5.942 (\pm 1.105)	4
RO15	6	1162 (\pm 15.05)	3.065 (\pm 5.627 10^{-3})	-7.092 (\pm 0.495)	4
RO18	7	1170 (\pm 59.53)	3.068 (\pm 2.21 10^{-2})	-6.045 (\pm 1.379)	5
SK02	5	1022 (\pm 51.61)	3.01 (\pm 2.192 10^{-2})	-5.136 (\pm 0.897)	3
SK05-2000	5	1408 (\pm 35.23)	3.149 (\pm 1.087 10^{-2})	-11.16 (\pm 2.746)	3
SK05-2004	5	1208 (\pm 28.06)	3.082 (\pm 1.009 10^{-2})	-7.165 (\pm 1.18)	3
<i>A. thaliana</i>	3	245.8 (\pm 0.133)	2.389 (\pm 1.492 10^{-4})	-33.24 (\pm 1.072)	1
<i>A. lpetrea</i>	3	232.4 (\pm 1.175)	2.36 (\pm 2.229 10^{-3})	-30.13 (\pm 1.272)	1

Results

The distribution of EC100s is presented in Figure 2. As expected, *A. thaliana* and *A. l. petrea* non tolerant individuals reached their EC100 at first doses of exposure (*i.e.* 250 and 500 μM) and all *A. halleri* individuals showed clear tolerance abilities. Except for a few individuals that also reached their EC100 before 750 μM (2% and 1% of NM and M tested individuals, respectively, belonging to populations B1, D1, CH1 and F1), most of the *A. halleri* individuals ranged between 750 and 2000 μM of Zn in solution culture. Interestingly, mortality events were continuously distributed over all doses of exposure, in almost each type of population (*i.e.* NM, M and ovirensis, the only exception was for NMP populations in which no mortality event was observed at 1000 μM).

All survival data were well fitted by dose-response curves (all $R^2 > 0.93$). Estimated values for both T_{50} and B were precise (relatively low standard errors) and consistent (no aberrant results, Table 2). The number n_i of distinct survival proportion available for curve fitting varied between populations (from $n_i = 3$ to $n_i = 8$, Table 2) and was particularly low for the control species *A. thaliana* and *A. l. petrea* ($n_i = 3$) (Figure 2 and Figure 3). The n_i values of *A. halleri* populations were weakly related to sample size ($R^2 = 0.282$, $P = 0,001$). However, differences in n_i values did not explain variation in T_{50} and T_{50} standard error ($R^2 = 0.068$, $P = 0.144$ and $R^2 = 0.005$, $P = 0.692$, respectively). A weak but significant correlation was detected between n_i and B ($R^2 = 0.268$, $P = 0.002$), and, to some extent, between n_i and the standard error of B ($R^2 = 0.208$, $P = 0.007$). Since a low n_i always meant that the survival proportions dropped from 100% to 0% in a few successive concentrations of exposure, such a correlation was expected. Moreover, differences in B and the standard error of B were not related to differences in sample size ($R^2 = 0.012$, $P = 0.545$ and $R^2 = 0.006$, $P = 0.666$, respectively).

The T_{50} and B values obtained in 2000 and 2004 were very similar for the F01 population and slightly higher for SK05 in 2000 than in 2004 (Table 2). The GT2-test further attested that the $\text{LOG}T_{50}$ of both populations did not differ between 2000 and 2004 (see below). Mann-Whitney tests showed that T_{50} values of NM and M *A. halleri* populations obtained from distinct data sets did not differ from each other ($W = 41$, $P = 0.790$ and $W = 18$, $P = 0.561$, respectively). Finally, comparisons of B values and n_i revealed that they did not

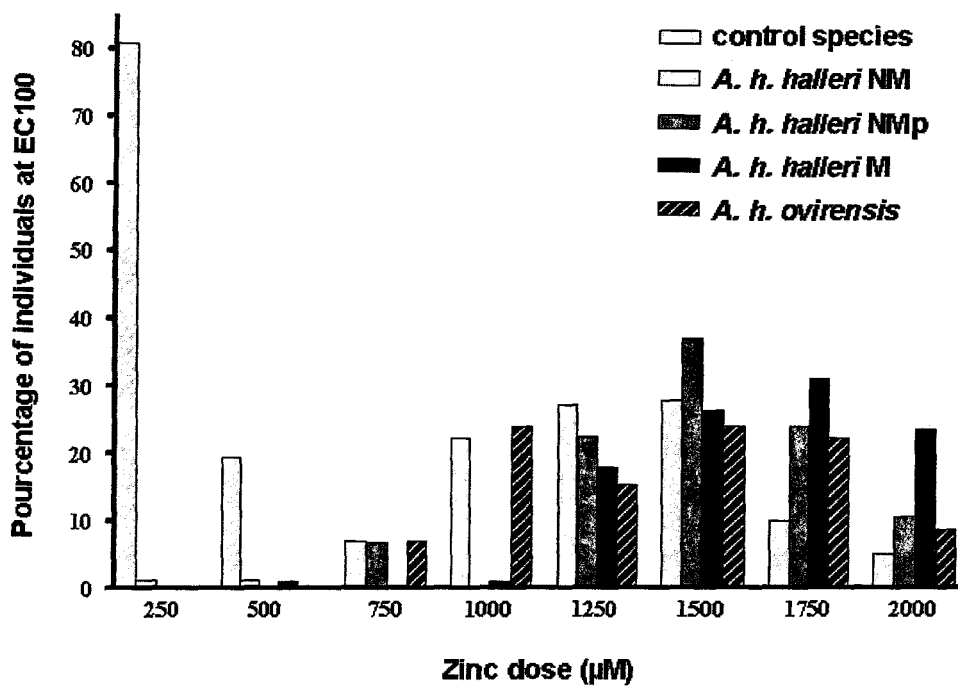


Fig. 2: Histograms showing frequency distribution of EC100s for each category of population. *A. h.*: *Arabidopsis halleri*; M: metallicolous; NM: non-metallicolous; NMP: non-metallicolous in polluted area.

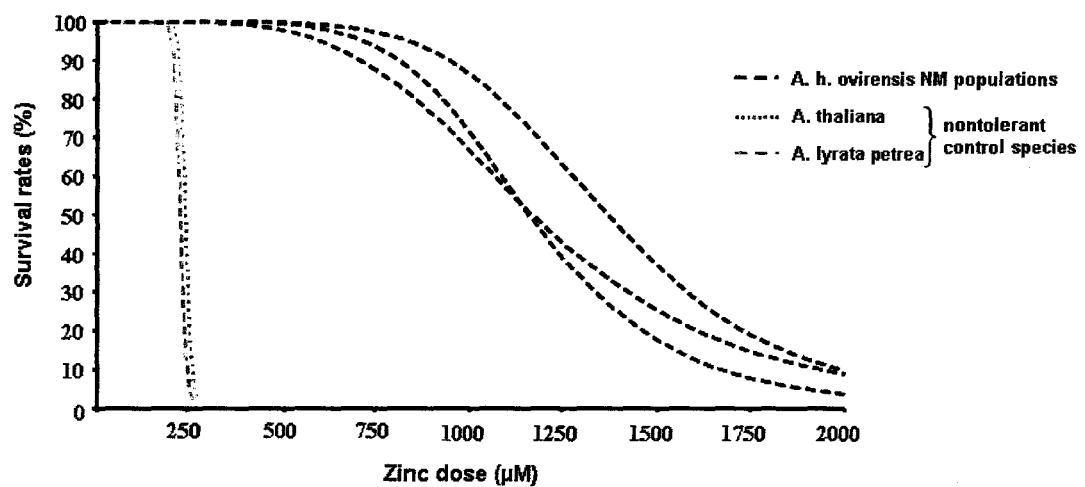
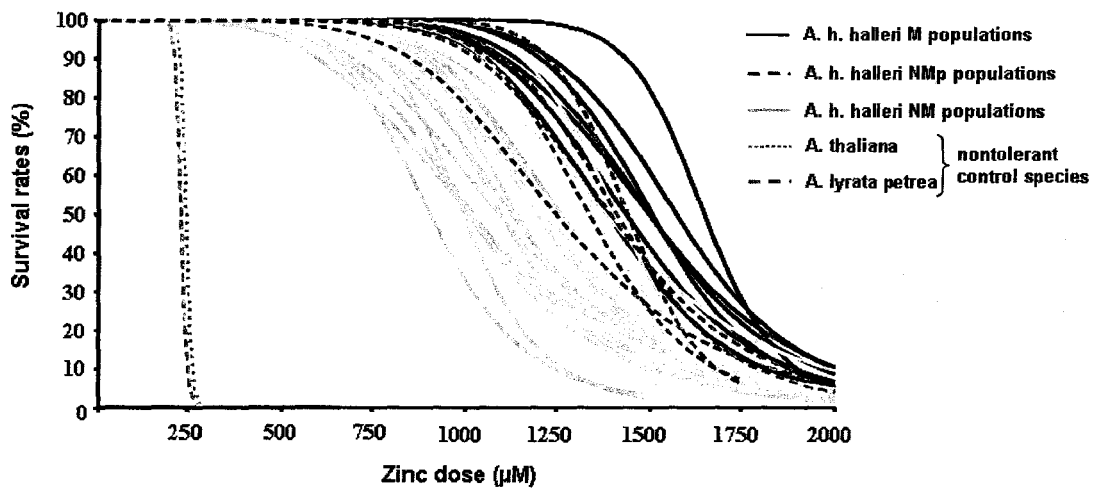


Fig. 3: Population survival curves obtained from fitting to a sigmoidal dose-response model. *A. h.*: *Arabidopsis halleri*.

differ between 2000 and 2004 ($W = 246, P = 0.066$ and $W = 307.5, P = 0.069$, respectively), in particular when M populations (over-represented in 2000 in comparison to 2004) were removed from the data sets ($W = 112, P = 0.317$ and $W = 196, P = 0.317$, respectively). These statistical similarities between experiments allowed us to analyse both data sets simultaneously. As both replicates from SK05 and F01 were then considered separately, this led to 33 *A. halleri* populations being tested.

Survival rates of both control species (*A. thaliana* and *A. l. petrea*) severely dropped in the lowest exposure doses (Figure 3). The respective T_{50} values were low and similar (Table 2). In our experiment, the T_{50} values obtained for control species should be considered to be indicative of the absence of Zn tolerance. In comparison, survival rates remained high to much higher Zn concentrations for all *A. h. halleri* populations (Figure 3) and T_{50} values were from 4 to 7 times higher than for control species (Table 2). This showed that all *A. h. halleri* populations had a high average level of Zn tolerance, irrespective of their edaphic origin. However, overall differences were observed between NM, NMp and M populations (Figure 4a). Comparisons of T_{50} revealed that tolerance abilities differed significantly between types (Table 3a). Unilateral tests further showed that T_{50} values were significantly increased from NM populations to M ones, with an intermediate position for NMp populations (results not shown, but see Figure 4a). Retrospectively, we ensured that such differences remained true in each data set (2000 and 2004) considered separately, at least when NM populations were compared to M ones ($W = 11, P = 0.037$ and $W = 112, P = 0.0321$, respectively). It appeared that the range of T_{50} values was higher for NM populations (mean (m) = 1158,1; standard deviation (S.D.) = 140,6; coefficient of variation (C.V.) = 0,12) than for M and NMp ones ($m = 1503,4$; S.D. = 86,4; C.V. = 0,06 and $m = 1363,2$; S.D. = 74,8; C.V. = 0,06, respectively). This was confirmed by a Levene test of equality of variances performed using Minitab 13.20 (data not shown).

Comparisons of $\text{LOG}T_{50}$ values confirmed that the control species did not differ from each other and that the Zn tolerance of all *A. h. halleri* populations tested was considerably higher (Figure 5). The trend towards enhanced tolerance in M populations was also confirmed, with most M populations statistically differing from 7 NM populations (out of 18). In particular, the tolerance of the PL06 population was statistically higher than the tolerance of most NM populations (16 out of 18). However, twenty-one *A. h. halleri* populations (M, NM or NMp) could not be distinguished from each other. This demonstrated that, rather than

Table 3 Results of the two-tailed Mann-Whitney tests for comparisons between median values of (a) T_{50} and (b) slope factor B of the different *A. h. halleri* population types (NM, NMp, M) and the *A. h. ovirensis* populations. NM: non-metallicolous; NMp: non-metallicolous in polluted area; M: metallicolous; ovirensis: *A. halleri* subsp. *ovirensis*; W: Mann-Whitney statistic.

(a) Populations	median value of T_{50} (μM)	NM	NMp	M
NM	1171.5	-	-	-
NMp	1390	W = 180, P = 0.008	-	-
M	1495	W = 153, P < 10^{-3}	W = 59, P = 0.035	-
ovirensis	1170	W = 40, P = 0.514	W = 8, P = 0.136	W = 6, P = 0.023

(b) Populations	median value of B	NM	NMp	M
NM	- 6.955	-	-	-
NMp	- 9.078	W = 84, P = 0.080	-	-
M	- 8.115	W = 187, P = 0.005	W = 48, P = 0.745	-
ovirensis	- 5.887	W = 19, P = 0.175	W = 8, P = 0.136	W = 6, P = 0.023

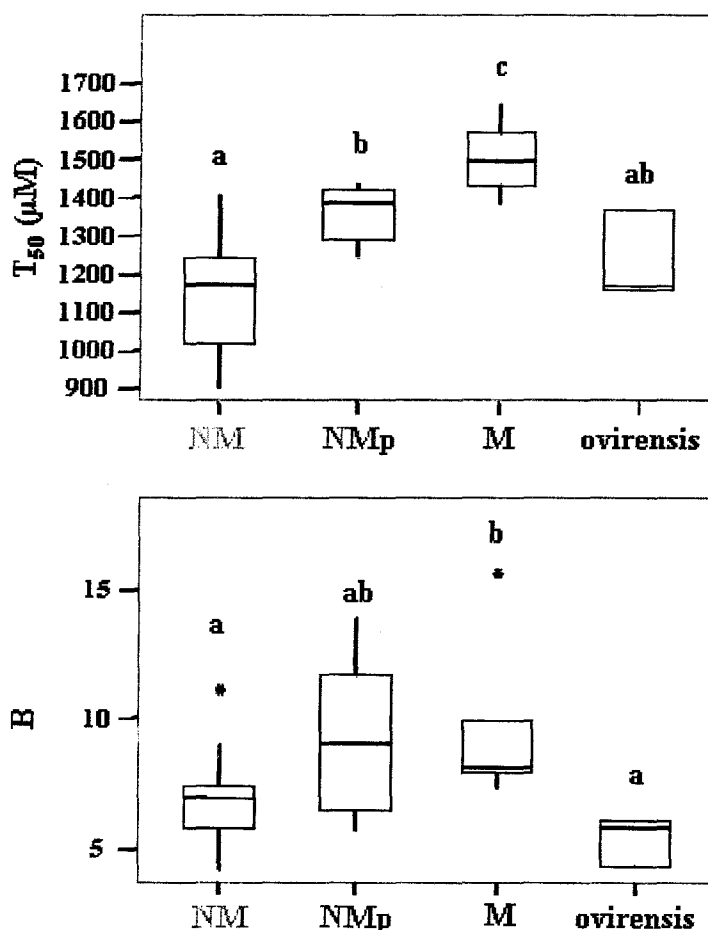


Fig. 4: Box plots of T_{50} and B obtained with Minitab 13.20 on the entire data set. Different letters above the box plots indicate a significant difference at the 5% level. M: metallicolous populations; NM: non-metallicolous populations; NMp: non-metallicolous populations in polluted area; ovirensis: *A. h. ovirensis* populations; *: outlier (i.e. any data that is distant from the upper or lower quartile by more than 1.5 times the standard error).

discrete statistical categories, a continuum existed from the least tolerant population (A05, NM) to the most tolerant one (PL06, M). Finally, the test showed that the more tolerant NM populations (I07 and RO09) differed significantly from the less tolerant ones (from A05 to CH03), thus confirming a higher heterogeneity of tolerance abilities in this group.

Comparisons of slope factors in *A. h. halleri* revealed that M populations had significantly higher slope factors than NM populations (Table 3b). Due to the wide range of values (Figure 4b), NMp populations did not differ either from M or NM populations (although a unilateral test showed that NMp populations had higher B values than NM ones, $W = 84$, $P = 0.04$). This suggested an overall tendency towards a reduction of polymorphism around the median EC100 in populations occurring in polluted areas.

In comparison to *A. h. halleri*, populations of the *A. h. ovirensis* subspecies showed moderate to average tolerance abilities (Table 2). Tolerance in these populations was intermediate between NM and NMp *A. h. halleri* populations (Figure 4a) from which they did not differ statistically (Table 3a). This was confirmed by the post-hoc comparisons of LOGT₅₀ that revealed that *A. h. ovirensis* populations did not differ from most of the *A. h. halleri* populations but did from the two extremes (A05 and PL06, see Figure 5). The levels of polymorphism in the *A. h. ovirensis* populations tested were more similar to NM populations of *A. h. halleri* (Figure 4b) and differed significantly from M populations (Table 3b).

Discussion

Constitutive zinc tolerance in the A. halleri population in Europe

Our survey confirmed that *A. halleri* is clearly pseudometallophyte in Europe. Most of its populations occurred in mountain areas, at medium to high altitudes, on soils with low levels of metal content. In contrast, in our sampling, both M and NMp *A. h. halleri* populations are not distributed throughout the species range but at the northern margin of the species range. However, most if not all individual we tested showed tolerance abilities and each population of *A. halleri* we tested showed high average levels of Zn tolerance, independently of the edaphic or geographic origin. We thus largely extended the results obtained from two M populations from northern France and three NM populations from Czech Republic by Bert *et al.* (2000) and strongly established that Zn tolerance is actually

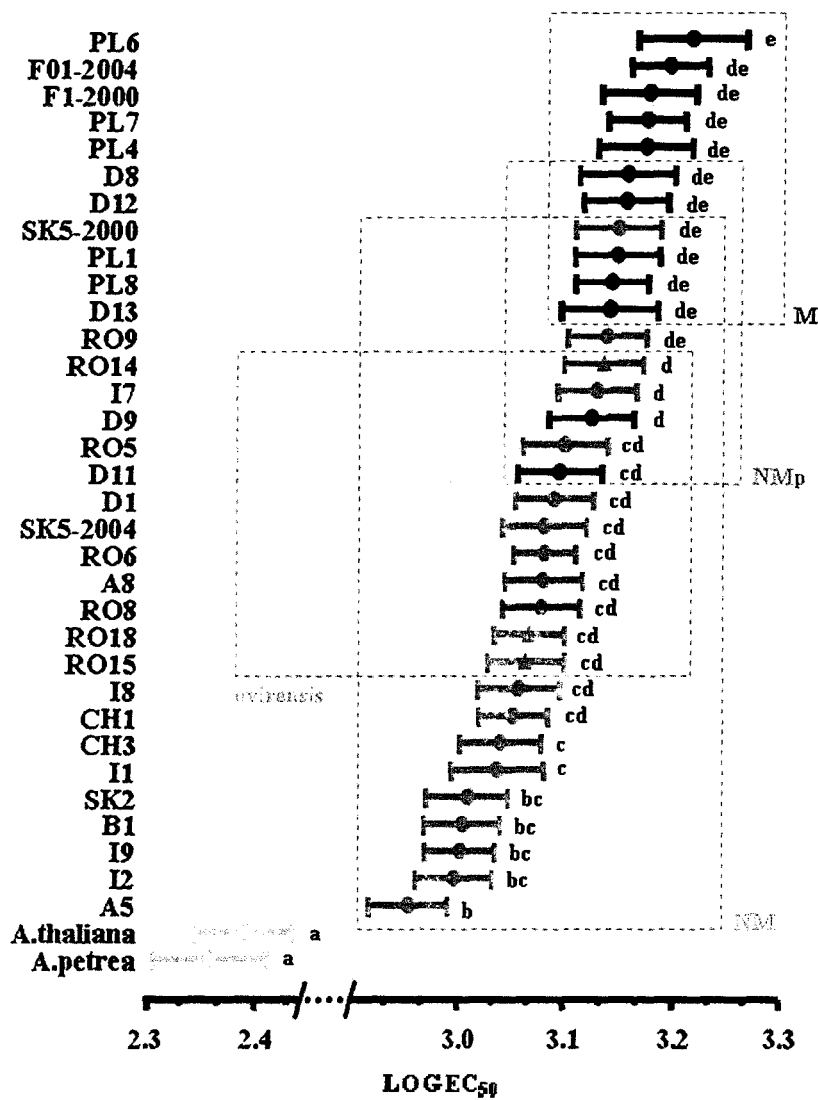


Fig. 5: Comparison intervals for $LOGT_{50}$ given by the GT2-method of multiple comparisons. Different letters indicate a significant difference at the 5% level. Best fit values of control species, *A. h. halleri* and *A. h. ovirensis* are respectively represented by open circles, closed circles and closed triangles. NM, NMp and M populations are respectively

constitutive in *A. h.* subsp. *halleri*. Nevertheless, in a few populations, in particular in NM populations, survival rates slightly fell in the first exposure doses, suggesting that rare genotypes with low tolerance abilities may occur. Such early stop in rooting could be also an artefact of the sequential test and have to be confirmed by measurement of the heritability of the trait. An appropriate screening of within-population polymorphism using a constant Zn concentration (e.g. 250 μ M) could be helpful in identifying putative non-tolerant or low tolerant genotypes that would prove particularly useful in genetic studies of metal tolerance mechanisms (Macnair *et al.* 1999, Bert *et al.* 2003, Willems *et al.*, submitted).

Zn tolerance was not only shown to be constitutive in *A. h.* subsp. *halleri* but also extended to *A. h.* subsp. *ovirensis*. Although no current local heavy metal exposure was detected, we showed Zn tolerance and within-population polymorphism in *A. h. ovirensis* populations. Average tolerance of *A. h. ovirensis* populations was quite similar to those of *A. h. halleri* NM populations, i.e. populations that had surely not been exposed to high levels of heavy metals in the recent times. Further studies should verify if Zn tolerance is not only constitutive in *A. h.* subsp. *halleri* but also species-wide in Europe, i.e. constitutive in both subspecies.

Origin of constitutive zinc tolerance in A. halleri

Constitutive metal tolerance is a rare phenomenon in pseudometallophytes (for a review, see Pollard *et al.* 2002). The most parsimonious hypothesis for the origin of constitutive tolerance would assume that it has been acquired only once. Given the very likely recent history of M populations, their marginal position and the extreme conditions they encountered on polluted soils, it is highly unlikely that metal tolerance spread from an initial M population to the entire species range (Levin 2000). Conversely, tolerance is likely to have evolved once, at an early stage of the species history, i.e. much earlier than the recent colonization of man-made habitats (Westerbergh 1994). Two major hypotheses remain equally possible. Firstly, metal tolerance in *A. halleri* could have evolved in response to an early exposure of the species to metal, as has been suggested for metalloendemics (Kruckeberg and Kruckeberg 1990). In these species, the adaptation to metal-enriched soils (including metal tolerance acquisition) is supposed to have led to reproductive isolation (Kruckeberg 1986, Macnair and Gardner 1998) and metal tolerance is consequently constitutive. Such a hypothesis would be difficult to confirm since *A. halleri* populations have

never been recorded as occurring on naturally metal-enriched soils. However, comparative phylogeography including both *A. halleri* and its close relatives (mainly *A. lyrata petrea* and *A. arenosa*), should be able to verify the likelihood of such a scenario. Secondly, Zn tolerance could have evolved in response to the hyperaccumulation trait. Metal hyperaccumulation (accumulation of metal in tissues to very high concentrations) is sometimes assumed to be a defence strategy against herbivores that could have evolved in the absence of metal exposure (Boyd 1998). Evolution of hyperaccumulation may have provoked a co-evolution of sequestration capacity which, in return could have improved the tolerance abilities of the hyperaccumulator species. However, the relationships between both Zn and Cd tolerance and hyperaccumulation (Macnair *et al.* 1999, Bert *et al.* 2003) and the defensive role of Zn/Cd hyperaccumulation are still highly debated both in *A. halleri* and in *T. caerulescens* (Assunção *et al.* 2003a, Frérot *et al.* 2005, Noret *et al.* 2005). Clearly, the support of this hypothesis will require a better understanding of the genetic make-up of both characters, involving the identification of candidate genes (Willems *et al.* submitted).

Evolution of Zn tolerance in Europe

Apart from constitutive tolerance, we revealed significant differences in Zn tolerance abilities of individuals and populations. We made it clear that an actual quantitative polymorphism exists between individuals and that a continuum exists in *A. h. halleri* from the least tolerant population to the most tolerant one. Up to now, Zn tolerance genetic architecture in *A. h. halleri* has been studied through interspecific crosses analyses with *A. lyrata petraea*: one genetic analysis assumed a control by only one major gene and modifiers (Macnair *et al.* 1999) and a recent QTL analysis detected three genomic regions of similar major effect (Willems *et al.* submitted). Therefore, the quantitative polymorphism we observed at the within species level could be explained by allelic variation at genes present in the identified QTL regions or by additional genes with minor effects not detected in interspecific crosses. Hypostatic modifier genes have commonly been evocated to explain quantitative polymorphism in tolerance in several pseudometallophytes (Macnair 1993), e.g. copper tolerance in *Mimulus guttatus* (Macnair 1983, Smith and Macnair 1998) and *Silene vulgaris* (Schat and Ten Bookum 1992, Schat *et al.* 1993, van Hoof *et al.* 2001).

The observed quantitative polymorphism is a prerequisite for allowing diversifying selection to adapt populations to local conditions (Latta 1998). Our results strongly attested

that there are significant differences in average Zn tolerance among populations. Although no simple correlation could be established between T_{50} measured in controlled conditions and Zn concentration in soils of sites in which seeds were collected (data not shown), heavy metal exposure of populations clearly distinguished groups of populations. M populations were shown to be significantly more tolerant. They were also shown to have a reduced level of within-population polymorphism for Zn tolerance. It has to be noticed, however, that the sequential test we used could have generated a bias towards an underestimation of variation in the more tolerant populations (Schat and Ten Bookum 1992).

The analysis of cpDNA population structure in the northern part of the species range has shown that the populations we sampled in man-made polluted areas have been founded from NM ones located in geographically proximate mountains area (Pauwels *et al.* 2005). In the absence of strong founder effects (Pauwels *et al.* 2005), the combination of a higher tolerance level and reduced polymorphism in M populations indicates that natural selection could have acted during the colonization of polluted areas (Meerts and van Isacker 1997). The location of the most tolerant populations (PL06, F01 and PL04) on the most polluted soils further suggests that the strength of selection was related to the local degree of heavy metal contamination of the soil. Thus, although it is constitutive, Zn tolerance appears to evolve secondary towards enhanced abilities in M populations. According to Pauwels *et al.* (2005) who concluded that the M population in northern and western Europe had an independent origin, genetic mechanisms that have been locally selected could differ in separately founded M populations (Schat *et al.* 1996). Testing of such a hypothesis in *A. halleri* will require mapping QTL for tolerance in recombinant populations from multiple intra-specific crosses between an NM population and independently founded M populations.

Post-hoc comparison of LOGT_{50} values demonstrated that many populations could not be significantly distinguished from each other. Comparison made it clear that, although the distinct types of populations differed according to tolerance abilities, they did not actually constitute discrete groups. In particular, an important quantitative polymorphism was observed between and within NM populations, with significant differences between the less and the more tolerant populations, and some NM populations were not distinguished from M ones. The occurrence of quantitative polymorphism among NM populations of *A. h. halleri*, *i.e.* in the absence of strong metal pressure, is more difficult to interpret in an evolutionary framework. The comparatively high level of tolerance detected in some NM populations (*e.g.*

SK05, RO09, I07 and RO05) suggests that genetic mechanisms conferring enhanced tolerance could be present at relatively high frequency in populations that have not experienced metal exposure in recent times. Overall, the polymorphism observed among NM populations was not related either to the Zn content in soils ($R^2 = 0.002$, $P = 0.867$) or to any available geographic variables. No correlation was found between T_{50} and either latitude ($R^2 = 2.10^{-4}$, $P = 0.953$) or altitude ($R^2 = 0.028$, $P = 0.505$); a weak but significant correlation was found between T_{50} and longitude ($R^2 = 0.241$, $P = 0.038$), although this result strongly suffered from the discontinuous distribution of NM populations according to longitude. Retrospectively, the absence of correlation between T_{50} and altitude in NM populations, which ranges from very low (160 m) to high (1530 m) elevations, attested that altitude could not explain the differences in tolerance abilities between NM and M populations whereas all M populations were at low elevation. Finally, our result also suggests that no selection acts against genetic mechanisms conferring enhanced tolerance in slightly or non-polluted habitats (Harper *et al.* 1997b). Obviously, such a hypothesis will require more appropriate experiment (*e.g.* reciprocal transplantation) to be more rigorously supported.

The features of NMp populations in polluted areas should be understood with respect to the narrow genealogical relationships they share with surrounding M ones (Pauwels *et al.* 2005). A first assumption is that NMp populations have been founded from M ones and have simply inherited enhanced tolerance. The low metal concentrations in soils and the absence of the directional selection that acts on polluted sites could have secondarily resulted in lower average tolerance levels of NMp population as compared to M ones (Harper *et al.* 1997a, Harper *et al.* 1998). Conversely, NMp could have acquired enhanced tolerance from M population through gene flow moving metal tolerance genes. The discussion of both hypotheses will require a population genetic structure analysis at a local scale, ideally replicated in both polluted areas where NMp populations have been mentioned. A combination of both hypotheses is not impossible and would explain the large range of within-population polymorphism observed.

Conclusion

The occurrence of constitutive metal tolerance in a pseudometallophyte species such as *A. halleri* is not in accordance with the classical model accounting for tolerance evolution in pseudometallophytes (Brooks 1987). In this model, mainly elaborated for the so-called

"mine taxa" whose M populations mostly occur in anthropogenic metal polluted areas and which include most pseudometallophyte species, metal tolerance is expected to have evolved as an adaptive response to recent metal exposure, *i.e.* in M populations only. In *A. halleri*, Zn metal tolerance could rather result from longer evolutionary history. As a result, the genetics of metal tolerance in these species could strongly differ from mine taxa. Thus, constitutive tolerance could distinguish *A. h. halleri* (as well as *T. caerulescens*, Ingrouille and Smirnoff 1986, Meerts and van Isacker 1997) from most pseudometallophytes (Pollard *et al.* 2002). The status of model species of these species is thus not straightforward and the extrapolation of results obtained from them will have to be done with caution.

Metal tolerance is usually defined as the ability to survive on metaliferous soils (Antonovics *et al.* 1971, Macnair and Baker 1994). Although tolerance tests involving root growth measurement are usually used to infer tolerance abilities of pseudometallophytes (Macnair 1993), it has to be noticed, however, that root growth could only be part of the adaptation to metalliferous soils. Thus, tolerance abilities revealed in NM populations in the present paper do not indicate that all these populations would be able either to lastingly develop on polluted soils or to found M populations. Reciprocal translocations on polluted and non polluted sites from a single geographic region using different M and NM accessions characterised in this study are currently in progress in southern Poland. The measurement of life history traits involved in the overall fitness of individuals and including both vegetative growth and reproduction should provide a more integrated view of the capacities of NM and M genotypes to respectively colonise or escape polluted sites.

Acknowledgement

The authors thank Drs. Vlastimil Mikolas, Mihai Puscas, Slavomir Sokol, Maciej Szczepka, Thomes Wilhelm, Patrick de Laguerie, Laurent Amsellem and Profs. Andreas Bresinsky, and Konrad Pagitz for help in finding populations. They are very grateful to Josep Escarré, Henk Schat, Maria Clauss and two anonymous referees for helpful comments on the manuscript. This work was supported by funding from the Contrat de Plan Etat/Région Nord-Pas de Calais (PRC) for the “Arabidopsis” project, from the European FEDER (contract No. 79/1769), from the Institut Français de Biodiversité (contract N°. 00N55/0083), and from the INSU-CNRS program ACI ECCO (contract No. 04 2 9 FNS) . M. Pauwels was funded by the French Ministry of Research and Technology.

References

- Al-Hiyaly, S. A., McNeilly, T. and Bradshaw, A. D. 1988. The effect of zinc contamination from electricity pylons - evolution in a replicated situation. *New Phytol.* **110**, 571-580.
- Al-Shehbaz, I. A. and O'Kane, S. L. J. 2002. Taxonomy and Phylogeny of *Arabidopsis* (Brassicaceae). In: *The Arabidopsis Book* (eds. C. R. Somerville and E. M. Meyerowitz). American Society of Plant Biologist, Rockville.
- Antonovics, J., Bradshaw, A. D. and Turner, R. G. 1971. Heavy metal tolerance in plants. *Advances in Ecological Research* **7**, 1-85.
- Assunção, A. G. L., Ten Bookum, W. M., Nelissen, H. J. M., Vooijs, R., Schat, H. and Ernst, W. H. O. 2003a. A cosegregation analysis of zinc (Zn) accumulation and Zn tolerance in the Zn hyperaccumulator *Thlaspi caerulescens*. *New Phytol.* **159**, 383-390.
- Assunção, A. G. L., Schat, H. and Aarts, M. G. M. 2003b. *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. *New Phytol.* **159**, 351-360.
- Becher, M., Talke, I. N., Krall, L. and Kramer, U. 2004. Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. *Plant J.* **37**, 251-268.
- Bert, V., Macnair, M. R., de Laguerie, P., Saumitou-Laprade, P. and Petit, D. 2000. Zinc tolerance and accumulation in metallicolous and non-metallicolous populations of *Arabidopsis halleri* (Brassicaceae). *New Phytol.* **146**, 225-233.
- Bert, V., Bonnin, I., Saumitou-Laprade, P., de Laguerie, P. and Petit, D. 2002. Do *Arabidopsis halleri* from nonmetallicolous populations accumulate zinc and cadmium more effectively than those from metallicolous populations? *New Phytol.* **155**, 47-57.

- Bert, V., Meerts, P., Saumitou-Laprade, P., Salis, P., Gruber, W. and Verbruggen, N. 2003. Genetic basis of Cd tolerance and hyperaccumulation in *Arabidopsis halleri*. *Plant and Soil* **249**, 9-18.
- Berton, A. 1946. Presentation de plantes *Arabis Halleri*, *Armeria elongata*, *Oenanthe fluviatilis*, *Galinsoga parviflora discoidea*. *Bulletin de la Société de Botanique de France* **93**, 139-145.
- Boyd, R. S. 1998. Hyperaccumulation as a plant defense strategy. In: *Plants that hyperaccumulate heavy metals* (ed. R. R. Brooks), pp. 181-201. CAB international, Oxford.
- Brooks, R. R. 1987. *Serpentine and its Vegetation: A Multidisciplinary Approach* Dioscorides Press, Portland, Oregon.
- Dahmani-Muller, H., van Oort, F., Gelie, B. and Balabane, M. 2000. Strategies of heavy metal uptake by three plant species growing near a metal smelter. *Env. Pollution* **109**, 231-238.
- Dräger, D. B., Desbrosses-Fonrouge, A.-G., Krach, C., Chardonens, A. N., Meyer, R. C., Saumitou-Laprade, P. and Krämer, U. 2004. Two genes encoding *Arabidopsis halleri* MTP1 metal transport proteins co-segregate with zinc tolerance and account for high MTP1 transcript levels. *Plant J.* **39**, 425-439.
- Ernst, W. H. O. 1990. Mine vegetation in Europe. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. A. J. Shaw), pp. 21-37. CRC Press, Boca Raton, Florida.
- Fabiszewski, J. 1986. Heavy metal tolerance of *Cardaminopsis halleri* (L.) Hayek populations in the Polish Tatra Mts. *Acta Societatis Botanicorum Poloniae* **55**, 421-428.
- Feder, M. E. and Mitchell-Olds, T. 2003. Evolutionary and ecological functional genomics. *Nat. Rev. Genet.* **4**, 649-655.
- Frérot, H., Lefèbvre, C., Petit, C., Collin, C., Dos Santos, A. and Escarre, J. 2005. Zinc tolerance and hyperaccumulation in F1 and F2 offspring from intra and interecotype crosses of *Thlaspi caerulescens*. *New Phytol.* **165**, 111-119.
- Harper, F. A., Smith, S. E. and Macnair, M. R. 1997a. Can an increased copper requirement in copper-tolerant *Mimulus guttatus* explain the cost of tolerance ? I. Vegetative growth. *New Phytol.* **136**, 455-467.
- Harper, F. A., Smith, S. E. and Macnair, M. R. 1997b. Where is the cost in copper tolerance in *Mimulus guttatus*? Testing the trade-off hypothesis. *Functional Ecology* **11**, 764-774.
- Harper, F. A., Smith, S. E. and Macnair, M. R. 1998. Can an increased copper requirement in copper-tolerant *Mimulus guttatus* explain the cost of tolerance? II. Reproductive phase. *New Phytol.* **140**, 637-654.
- Ingrouille, M. J. and Smirnoff, N. 1986. *Thlaspi caerulescens* J. & C. presl. (T. alpestre) in Britain. *New Phytol.* **102**, 219-233.
- Jackson, R. B., Linder, C. R., Lynch, M., Purugganan, M., Somerville, S. and Thayer, S. S. 2002. Linking molecular insight and ecological research. *TREE* **17**, 409-414.
- Koch, M., Mummenhoff, K. and Hurka, H. 1998. Systematics and evolutionary history of heavy metal tolerant *Thlaspi caerulescens* in Western Europe: evidence from genetic studies based on isozyme analysis. *Biochemical systematics and ecology* **26**, 823-838.

- Koch, M., Haubold, B. and Mitchell-Olds, T. 2001. Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. *Am. J. Bot.* **88**, 534-544.
- Kruckeberg, A. R. 1986. An essay: the stimulus of unusual geologies for plant speciation. *Syst. Bot.* **11**, 455-463.
- Kruckeberg, A. R. and Kruckeberg, A. L. 1990. Endemic metallophytes: their taxonomic, genetic, and evolutionary attributes. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. A. J. Shaw), pp. 301-312. CRC Press, Boca Raton, Florida.
- Lambinon, J. and Auquier, P. 1964. La flore et la végétation des terrains calaminaires de la Wallonie septentrionale et de la Rhénanie aixoise. Types chorologiques et groupes écologiques. *Natura mosana* **16**, 113-130.
- Latta, R. G. 1998. Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. *Am. Nat.* **153**, 283-292.
- Levin, D. A. 2000. *The Origin, Expansion and Demise of Plant Species* Oxford University Press, Oxford.
- Lexer, C. and Fay, M. F. 2005. Adaptation to environmental stress: a rare or frequent driver of speciation? *J. Evolution. Biol.* **18**, 893-900.
- Macnair, M. and Gardner, M. 1998. The Evolution of Edaphic Endemics. In: *Endless Forms, Species and Speciation* (eds. D. J. Howard and S. H. Berlocher), pp. 157-171. Oxford University Press, New York.
- Macnair, M. R. 1983. The genetic control of copper tolerance in the yellow monkey flower, *Mimulus guttatus*. *Heredity* **50**, 283-293.
- Macnair, M. R. 1993. Tansley review N°49: the genetic of metal tolerance in vascular plants. *New Phytol.*, **124**, 541-559.
- Macnair, M. R. and Baker, J. M. 1994. Metal-tolerant plant: An evolutionary perspective. In: *Plants and the chemical elements, biochemistry, uptake, Tolerance and toxicity* (ed. M. E. Farago), pp. 68-83. VCH, NY.
- Macnair, M. R., Tilstone, G. H. and Smith, S. E. 2000. The Genetics of Metal Tolerance and Accumulation in Higher Plants. In: *Phytoremediation of Contaminated Soil and Water* (eds. N. Terry, G. Banuelos and J. Vangronsveld), pp. 235-250. CRC Press, Boca Raton.
- Macnair, M. R., Bert, V., Huitson, S. B., Saumitou-Laprade, P. and Petit, D. 1999. Zinc tolerance and hyperaccumulation are genetically independent characters. *Proc. R. Soc. Lond. B* **266**, 2175-2179.
- Meerts, P. and van Isacker, N. 1997. Heavy metal tolerance and accumulation in metallicolous and nonmetallicolous populations of *Thlaspi caerulescens* from continental Europe. *Plant Ecology* **133**, 221-231.
- Mengoni, A., Barabesi, C., Gonnelli, C., Galardi, F., Gabbrielli, R. and Bazzicalupo, M. 2001. Genetic diversity of heavy metal-tolerant populations in *Silene paradoxa* L. (Caryophyllaceae): a chloroplast microsatellite analysis. *Mol. Ecol.* **10**, 1909-16.
- Mitchell-Olds, T. 2001. *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *TREE* **16**, 693-700.

- Motulsky, H. J. 1999. *Analyzing Data with GraphPad Prism* GraphPad Software Inc., San Diego CA.
- Noret, N., Meerts, P., Tolra, R., Poschenrieder, C., Barcelo, J. and Escarre, J. 2005. Palatability of *Thlaspi caerulescens* for snails: influence of zinc and glucosinolates. *New Phytol.* **165**, 763-772.
- Pauwels, M., Saumitou-Laprade, P., Holl, A.-C., Petit, D. and Bonnin, I. 2005. Multiple origin of metallicolous populations of the pseudometallophyte *Arabidopsis halleri* (*Brassicaceae*) in Central Europe: the cpDNA testimony. *Mol. Ecol.* **14**, 4403-4414.
- Pollard, A. J., Powell, K. D., Harper, F. A. and Smith, J. A. C. 2002. The genetic basis of metal hyperaccumulation in plants. *Cri. Rev. Plant Sci.* **21**, 539-566.
- Prat, S. 1934. Die Erbllichkeit der Resistenz gegen Kupfer. *Berichte der Deutschen Botanischen Gesellschaft* **52**, 65-78.
- Schat, H. and Ten Bookum, W. M. 1992. Genetic control of copper tolerance in *Silene vulgaris*. *Heredity* **68**, 219 - 229.
- Schat, H., Vooijs, R. and Kuiper, E. 1996. Identical major gene loci for heavy metal tolerances that have independently evolved in different local populations and subspecies of *Silene vulgaris*. *Evolution* **50**, 1888-1895.
- Schat, H., Kuiper, E., Ten Bookum, W. M. and Vooijs, R. 1993. A general model for the genetic control of copper tolerance in *Silene vulgaris*: evidence from crosses between plants from different tolerant populations. *Heredity* **70**, 142-147.
- Smith, S. E. and Macnair, M. 1998. Hypostatic modifiers cause variation in degree of copper tolerance in *Mimulus guttatus*. *Heredity* **80**, 760-768.
- Sokal, R. R. and Rohlf, F. J. 1981. *Biometry*, Second Edition edn. W. H. Freeman and Company, New York.
- van Hoof, N. A. L. M., Hassinen, V. H., Hakvoort, H. W. J., Ballintijn, K. F., Schat, H., Verkleij, J. A. C., Ernst, W. H. O., Karenlampi, S. O. and Tervahauta, A. I. 2001. Enhanced Copper Tolerance in *Silene vulgaris* (Moench) Garcke Populations from Copper Mines Is Associated with Increased Transcript Levels of a 2b-Type Metallothionein Gene. *Plant Physiol.* **126**, 1519-1526.
- van Rossum, F., Bonnin, I., Fenart, S., Pauwels, M., Petit, D. and Saumitou-Laprade, P. 2004. Spatial genetic structure within a metallicolous population of *Arabidopsis halleri*, a clonal, self-incompatible and heavy-metal-tolerant species. *Mol. Ecol.* **13**, 2959-2967.
- Vekemans, X. and Lefèbvre, C. 1997. On the evolution of heavy-metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. *J. Evolution. Biol.* **10**, 175-191.
- Weber, M., Harada, E., Vess, C., Roepenack-Lahaye, E. and Clemens, S. 2004. Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identifies nicotianamine synthase, a ZIP transporter and other genes as potential metal hyperaccumulation factors. *Plant J.* **37**, 269-281.
- Westerbergh, A. 1994. Serpentine and non-serpentine *Silene dioica* plants do not differ in nickel tolerance. **167**, 297-303.
- Westerbergh, A. and Saura, A. 1992. The effect of serpentine on the population structure of *Silene dioica* (Caryophyllaceae). *Evolution* **46**, 1537-1548.

- Willems, G., Godé, C., Dräger, D. B., Courbot, M., Verbruggen, N. and Saumitou-Laprade, P. submitted. Quantitative Trait Loci mapping of zinc tolerance in the metallophyte *Arabidopsis halleri* ssp. *halleri*. *Genetics*
- Wu, L., Bradshaw, A. D. and Thurman, D. A. 1975. The potential for evolution of heavy metal tolerance in plants. III. The rapid evolution of copper tolerance in *Agrostis stolonifera*. *Heredity* **34**, 165-187.

Chapitre 2: Multiple origin of metallicolous populations of the pseudometallophyte *Arabidopsis halleri* (Brassicaceae) in central Europe: the cpDNA testimony

Maxime Pauwels, Pierre Saumitou-Laprade, Anne Catherine Holl, Daniel Petit, Isabelle Bonnin

Abstract The population structure of the pseudo-metallophyte herb, *Arabidopsis halleri*, was studied using Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP) on chloroplast DNA (cpDNA). The history of metallicolous (M) populations showing increased zinc tolerance was investigated. Eight primer-enzyme combinations out of 72 tested were applied to a total of 625 individuals from 28 widespread populations, 14 of them being M. Eleven distinct chlorotypes were found: five were common to nonmetallicolous (NM) and M populations, whereas six were only observed in one edaphic type (five in NM and one in M). No difference in chlorotype diversity between edaphic types was detected. Computed on the basis of chlorotype frequencies, the level of population differentiation was high but remained the same when taking into account levels of molecular divergence between chlorotypes. Isolation by distance was largely responsible for population differentiation. Geographically isolated groups of M populations were more genetically related to their closest NM populations than to each other. Our results suggest that M populations have been founded separately from distinct NM populations without suffering founding events and that the evolution towards increased tolerance observed in the distinct M population groups occurred independently.

Keywords *Arabidopsis halleri*, cpDNA, heavy metal pollution, local adaptation, PCR-RFLP, pseudometallophyte

Multiple origin of metallicolous populations of the pseudometallophyte *Arabidopsis halleri* (Brassicaceae) in central Europe: the cpDNA testimony

MAXIME PAUWELS,*PIERRE SAUMITOU-LAPRADE,*ANNE CATHERINE HOLL,*DANIEL PETIT* and ISABELLE BONNIN†

*Laboratoire de Génétique et Evolution des Populations Végétales, UMR CNRS 8016, FR CNRS 1818, Université des Sciences et Technologies de Lille, Bâtiment SN2, F-59655 Villeneuve d'Ascq cedex, France, †UMR de Génétique Végétale (CNRS-INRA-UPS-INAPG), Ferme du Moulon, F-91190 Gif sur Yvette, France

Abstract

The population structure of the pseudo-metallophyte herb, *Arabidopsis halleri*, was studied using Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP) on chloroplast DNA (cpDNA). The history of metallicolous (M) populations showing increased zinc tolerance was investigated. Eight primer-enzyme combinations out of 72 tested were applied to a total of 625 individuals from 28 widespread populations, 14 of them being M. Eleven distinct chlorotypes were found: five were common to nonmetallicolous (NM) and M populations, whereas six were only observed in one edaphic type (five in NM and one in M). No difference in chlorotype diversity between edaphic types was detected. Computed on the basis of chlorotype frequencies, the level of population differentiation was high but remained the same when taking into account levels of molecular divergence between chlorotypes. Isolation by distance was largely responsible for population differentiation. Geographically isolated groups of M populations were more genetically related to their closest NM populations than to each other. Our results suggest that M populations have been founded separately from distinct NM populations without suffering founding events and that the evolution towards increased tolerance observed in the distinct M population groups occurred independently.

Keywords: *Arabidopsis halleri*, cpDNA, heavy metal pollution, local adaptation, PCR-RFLP, pseudometallophyte

Received 13 May 2005; revision accepted 10 August 2005

Introduction

Human activities can have acute effects on the dynamics of adaptation of natural plant populations. Industrial and mining activities, in particular, have left soils with very high levels of heavy metals which prove toxic to most living things (Antonovics *et al.* 1971). Such 'metalliferous' soils strongly differ from unpolluted ones with respect not only to heavy metal concentrations but also to numerous additional physical and biological factors (Macnair 1997). For most plant taxa, they constitute adaptive constraints that are too strong to be overcome. A few plant species (coined 'metallophytes') have nonetheless acquired the capacity to grow on them. Several of these species are

'pseudometallophytes' (also called facultative metallophytes), i.e. species having both metallicolous (M, that develop on metalliferous soils) and nonmetallicolous (NM, that do not) populations, sometimes only separated by a very sharp transition over a few metres (Jain & Bradshaw 1966). In pseudometallophytes, two edaphic types thus exist, sometimes referred to as ecotypes, which are exposed to very distinct ecological conditions. Pseudometallophytes therefore constitute highly relevant models for studying local adaptation in plants (Linhart & Grant 1996).

Pseudometallophytes have been the focus of many detailed studies for almost half a century (Baker 1987; Macnair 1987; Pollard *et al.* 2002) and the pseudometallophyte evolutionary framework has been largely discussed (Bradshaw & Mcneilly 1991; Macnair 1997). Overall, two hypotheses have been proposed to account for the origin and history of M populations. First, because the frequency

Correspondence: Pierre Saumitou-Laprade, Fax: +33 3 20436979; E-mail: pierre.saumitou-laprade@univ-lille1.fr

4404 M. PAUWELS ET AL.

of genotypes able to tolerate elevated heavy metal exposure is typically low in NM populations, substantial founder effects are expected to have occurred during the colonization of contaminated area (Lefèbvre & Vernet 1990; Wu 1990). As a result, the colonization of metalliferous soils from those populations may have caused strong genetic bottlenecks (Bradshaw 1984), thus leading to a reduction of gene diversity in M populations as compared to NM populations (Luikart *et al.* 1998). The second assumption concerns the origin of M populations. Considering that the distribution of metalliferous sites in Europe is highly disjunct (Buse *et al.* 2003), with site-to-site distances often too long to be easily overcome by natural dispersal and that the evolution of tolerance in NM populations suddenly exposed to heavy metal stress can occur rapidly and repeatedly (Al-Hiyaly *et al.* 1988), Schat *et al.* (1996) have argued that geographically distant conspecific M populations could have evolved independently and therefore constitute a polyphyletic group. In such a scenario, M and NM populations should not constitute distinct phylogenetic lineages and M populations should be more genetically related to geographically close NM populations than to more geographically distant M populations.

Antonovics *et al.* (1971) suggested that heavy metal tolerance is always populational or ecotypic in pseudometallophytes, i.e. it occurs with no within-population polymorphism in M populations but only at low frequencies (when present) in NM populations. As a result, major assumptions about the evolution of M populations in pseudometallophytes have focused on species exhibiting populational tolerance and have therefore failed to distinguish the colonization of metalliferous soil from the acquisition of heavy metal tolerance (see above). More recently, this initial view was challenged by the demonstration that several facultative metallophytes actually possess constitutive (species-wide) tolerance to heavy metals. In these species, individuals from any population have the capacity to grow on polluted soils, regardless of whether they currently grow on polluted or unpolluted soils (for a review, see Pollard *et al.* 2002). Intriguingly, however, recent studies have reported quantitative genetic variation for tolerance among edaphic types. For example, in the pseudometallophytes *Thlaspi caerulescens* (Meerts & van Isacker 1997) and *Arabidopsis halleri* (Bert *et al.* 2000), M populations were found to be more zinc-tolerant than NM populations in controlled conditions. This suggests that improved adaptation associated with metal exposure has influenced population genetics despite constitutive tolerance. Up to now, little is known about relationships between metal exposure and M population history in pseudometallophytes with constitutive tolerance.

Information from nonrecombinant and maternally inherited cpDNA markers has proven quite useful in population level studies in recent decades (McCauley 1995; Ennos *et al.* 1999). Several studies have investigated

cpDNA polymorphism to define intraspecific lineages and infer population history (for reviews, see Petit *et al.* 2003; Petit *et al.* 2005). Such an approach, which takes into account phylogenetic data from the organelle genome has recently been recommended to successfully detect the impact of environmental contamination on the genetic structure of a species (Staton *et al.* 2001). However, no study has yet attempted to determine M population history in pseudometallophytes using gene phylogeny. In this study, we used cpDNA to study the population genetic structure in the constitutively tolerant pseudometallophyte relative of *Arabidopsis thaliana*, *A. halleri*. We used the polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP) technique to investigate cpDNA polymorphism and to define genealogical lineages. In order to be representative of the intraspecific polymorphism over the whole species range, we sampled populations over a broad geographical range covering the main known polluted areas the species has colonized. We used our results to answer two questions that concern major assumptions about M population history in pseudometallophytes: did M populations suffer from genetic bottlenecks during the colonization of polluted areas? Were these populations founded from a single or from several independent events?

Materials and methods

Arabidopsis halleri (syn. *Cardaminopsis halleri*, Brassicaceae) is a clonal, self-incompatible and highly outcrossing perennial weed. It is a close relative of the model organism *A. thaliana* which is supposed to have diverged from the other *Arabidopsis* species from about 5 million years ago (Ma) (Koch *et al.* 2000; Koch *et al.* 2001; Mitchell-Olds 2001). According to Al-Shehbaz & O'Kane (2002), three *A. halleri* subspecies can be distinguished: the montane subspecies *Arabidopsis halleri halleri* evolved from the subspecies *Arabidopsis halleri ovirensis*, an alpine plant of the southeastern Alps, the carpathians and the northern Balkan Peninsula and is ancestral to *Arabidopsis halleri gemmifera* distributed in eastern Asia. *Arabidopsis halleri* ssp. *halleri* (henceforth called *A. halleri* in this study) is a pseudometallophyte that occurs mainly in central Europe on both metalliferous soils (extractable [Zn] up to more than 30 000 µg/g, Bert *et al.* 2002; Van Rossum *et al.* 2004) and nonmetalliferous soils (Bert *et al.* 2002). To be more precise, almost all habitats suitable for *A. halleri* in western Europe and at the northern border of the species range are of anthropogenic origin and polluted. This suggests that the colonization of these regions has been greatly affected by humans and must be recent.

Sampling and DNA extraction

In 1999, 28 scattered populations of *A. halleri* were sampled in the northern half part of the species range (Table 1). The

Table 1 Locations of population samples of *Arabidopsis halleri* used for the cpDNA PCR-RFLP study

Population	References in previous studies	Edaphic type	Origin	Ecological background	GPS coordinates			
					E	N	n_i	N_i est.
D1	G1*	NM	Bohemian forest (Germany)	Nitrogenous Regen river's bank	12°09'8.8	49°10'64	7	< 50
D2	G2*	NM	Bohemian forest (Germany)	Regen river's bank, in front of G1	12°09'52	49°11'31	8	< 50
D3	G3*	NM	Bohemian forest (Germany)	Shady lawn near the Regen river bank	12°39'66	49°13'10	9	< 50
D4	G4*	NM	Bohemian forest (Germany)	Regen river bank	12°47'75	49°09'85	11	< 50
D6	G6*	NM	Bohemian forest (Germany)	Cham river bank	12°49'78	49°17'42	10	< 50
D8	G8*	M	Harz (Germany)	Old mine (19th century)	10°29'04	51°53'79	13	< 50
D9	G9*	M	Harz (Germany)	Roadside	10°25'16	51°53'46	18	< 50
D11	G11*	M	Harz (Germany)	Underwood	10°21'95	51°51'27	18	< 50
D12	G12*	M	Harz (Germany)	Mine rubble	10°17'90	51°51'91	18	> 500
D13	G13*	M	Harz (Germany)	Roadside and lawn	10°18'50	51°55'22	20	50–500
F1	Auby†	M	North (France)	Wood near a smelter plant	03°03'	50°25'	23	> 1000
F2	—	M	North (France)	Lawn near a smelter plant	03°01'	50°25'	43	> 1000
F3	—	M	North (France)	Metallicolous lawn	03°25'	50°30'	37	> 1000
B1	—	NM	Hautes Fagnes (Belgium)	Underwood	06°40'	50°29'63	29	500–1000
PL1	P1*	M	Silesia (Poland)	Wood in Katowice suburbs	18°57'04	50°14'80	63	> 500
PL2	P2*	M	Silesia (Poland)	Metallurgical factory	18°56'67	50°29'68	21	50–500
PL3	P3*	M	Silesia (Poland)	Metallurgical factory	18°57'58	50°29'50	15	50–500
PL4	P4*	M	Silesia (Poland)	Metallurgical factory	18°55'79	50°29'98	21	> 1000
PL6	P5*	M	Silesia (Poland)	Old mine	19°01'52	50°16'95	10	> 1000
SK2	SI2*	NM	Tatras (Slovakia)	Shady meadow	21°07'81	48°46'17	22	> 500
SK5	SI4*	NM	High Tatras (Slovakia)	Tatranska javorina (Nature reserve)	20°09'24	49°16'98	46	> 500
CZ4	Cz1*	NM	Sumava (Czech Republic)	Forest above a railway	13°45'841	49°02'87	15	50–500
CZ5	Cz2*	NM	Sumava (Czech Republic)	Meadow along a railway	13°46'39	48°59'25	24	50–500
CZ6	Cz3*	NM	Sumava (Czech Republic)	Beside road next to a hay meadow	13°46'	48°59'	12	> 500
CZ8	Cz7*	NM	Sumava (Czech Republic)	River bank next to a meadow	13°48'	48°57'	26	> 1000
CZ14	Cz8*	M	Sumava (Czech Republic)	Slope in a cool, shady wood	12°42'92	49°28'37	20	> 500
CZ16	Cz9*	NM	Sumava (Czech Republic)	Wet meadow	13°33'31	49°03'35	57	> 5000
CZ18	Cz11*	NM	Sumava (Czech Republic)	Woody footpath	13°32'18	49°05'70	9	50–500

M, metallicolous; NM, non metallicolous; n_i , number of genotyped individuals; N_i est., approximative population size; * in Bert *et al.* (2002); † in Bert *et al.* (2000).

4406 M. PAUWELS ET AL.

scale of sampling ranged from a few kilometres between populations (e.g. D8 and D9) to more than 1300 kilometres for the most distant populations (e.g. F1 and SK2). Sampled M populations of *A. halleri* were discontinuously distributed in three geographically isolated groups (Poland, northern Germany and northern France) plus one isolated population in the western Czech Republic, i.e. at the border of the species range. Except for this isolated population where high metal concentration has been detected far from industrial activities, all M populations occurred in close proximity to zinc smelters or zinc mining sites (abandoned or still in activity). NM populations were sampled as close as possible to the M ones. The existence of closest NM populations can not be excluded but was not known at the moment of sampling. Finally, we sampled 14 M and 14 NM populations.

In each of these populations, leaves were collected from 7 to 63 individuals separated by at least 3 m to avoid clone sampling (Van Rossum *et al.* 2004). Overall, 623 genotypes were collected. Leaves were immediately dried in silica gel prior to molecular analysis. DNA from each genotype was extracted from 15 to 20 mg dry material using QIAGEN DNeasy® Kit and PCR amplification was performed on 1/100 dilutions.

PCR-RFLP analysis of chloroplast intergenic regions

Choice of intergenic regions. To assess cpDNA polymorphism in *A. halleri*, we first screen a subset of 12 individuals using PCR-RFLP. Individuals were randomly chosen in 12 scattered populations so as to be representative of the sample range. Sampled populations included populations D2, D5, D9, D12, F2, PL4, PL6, SK2, SK5, CZ5, CZ14 and CZ16 (Table 1). From these individuals, nine noncoding cpDNA regions were amplified and restricted using eight 4 bp cutter restriction endonucleases. Amplified regions included eight intergenic regions (AS, CD, CS, DT, K2Q, SfM, TC, VL) and the *trnK* region containing the *matK* intron (K1K2). The primer pairs used were those defined by Demesure *et al.* (1995) and Dumolin-Lapegue *et al.* (1997). Only the CD primer pair was modified to match the *A. thaliana* cpDNA sequence and improve PCR yield (*TmC7F*: CCAGTTCAAATCCGGGTGCC and *trnD7R*: GGGATTGTAGTTCAATTGGT). Amplification products were digested with the 4-bp cutter restriction endonucleases: *AluI*, *HinfI*, *Trt9I*, *HpaII*, *DpnII*, *Sau96I*, *AseI*, *AclI*. Overall, 72 different primer-enzyme combinations (PEC) were tested. Based on this survey, PECs that gave good PCR yield, nonambiguous and nonredundant polymorphisms were selected to be used over the whole sample. We qualified the polymorphisms as 'redundant', when the variation detected by distinct enzymes in the same amplified region clustered the individuals in the same way.

PCR and restriction conditions. The following PCR conditions were utilized: a total volume of 15 µL consisting of 5 µL (~5 ng) of template DNA, 3.5 mM MgCl₂, 200 µg/mL BSA, 200 µM of each of the four dNTPs, 0.2 µM of each primer, 0.625 U of AmpliTaq Standard (Perkin Elmer®) and 1× AmpliTaq Standard buffer (Perkin Elmer). PCRs were carried out in Perkin Elmer thermocyclers (model 2400 or 9700), using one cycle of 5 min at 95 °C, 36 cycles of 45 s at 92 °C, 45 s at 52.5–62 °C (depending on the primers sequences) and 2–4 min at 72 °C (depending on the size of the amplified fragment, precise protocols available upon request), followed by one cycle of 10 min at 72 °C. Restriction enzyme reactions were performed on a total volume of 10 µL consisting of 5 µL of PCR product, 200 µM of spermidine, 1× of restriction buffer (provided by supplier) and 1–1.5 U of restriction enzyme. Reagent mixtures were incubated at 37 °C (except for *AclI*, 50 °C) for 2–12 h.

RFLP analysis. Restriction fragments were separated by electrophoresis on either 1.5–2% agarose gels (to detect variation in fragments larger than 300 bp) or 8–9% polyacrylamide gels (prepared with a 19:1 acrylamide-bisacrylamide solution; for fragments sized between 75 and 300 bp). Agarose gels were stained with ethidium bromide and photographed with the Biolumage system (Bioprobe) under UV light, whereas polyacrylamide gels were silver stained. Restriction fragments were labelled by decreasing order of molecular weight as described by Demesure *et al.* (1996). Length polymorphisms (indel) and restriction polymorphism were encoded as in Dumolin-Lapegue *et al.* (1997). Fragment size polymorphisms were interpreted as insertion/deletions and were encoded as '1' or '2', according to decreasing size of the variable fragment. Polymorphisms corresponding to replacement of a large fragment by two smaller fragments of complementary size were interpreted as the absence or presence of a restriction site and encoded by '1' or '9', respectively.

Statistical analysis

Due to the nonrecombining nature of the chloroplast genome, alleles observed at all nine loci were combined into chlorotypes (cpDNA haplotypes).

Relationship among chlorotypes. A minimum-spanning tree (MST) of the chlorotypes was computed from a distance matrix containing the number of differences between each pair of chlorotypes using MINSNET, a program implemented within the software ARLEQUIN (Schneider *et al.* 2000). The same mutational weight was given to substitutions and to insertions/deletions. MST construction assumes that each chlorotype is linked by a single or a series of mutational events to all other chlorotypes through

a unique pathway and the construction method minimizes the number of such events (Excoffier & Smouse 1994).

Chlorotype diversity. Allelic richness (El Mousadik & Petit 1996; Kalinowski 2004) was estimated for each population i using the rarefaction method in *FSTAT* version 2.9.3 (Goudet 2001). Given that the smallest sample concerned population D1 with $n = 7$, estimates of a_i^r were standardized to a standard sample size $g = 7$. Chlorotypic diversity and its sampling variance (Nei 1987) were estimated within each population (H_{Si}), and over the whole sample (H_T) using *ARLEQUIN* (Schneider *et al.* 2000). To test for differences among edaphic types in allelic richness (a_i^{NM} vs. a_i^M) and chlorotype diversity (H_{SiNM} vs. H_{SiM}) permutation procedures (1000 permutations of populations between edaphic types) were carried out using *FSTAT* version 2.9.3 (Goudet 2001).

Genetic structure. Genetic differentiation was estimated using Weir and Cockerham's estimators in *ARLEQUIN* (Schneider *et al.* 2000): (i) between all populations (F_{ST}); (ii) by hierarchical analysis of variance among populations within edaphic type (F_{SC}) and between edaphic types (F_{CT}); (iii) between populations within each edaphic type separately (F_{STM} and F_{STNM}). Estimations were conducted using either allele frequencies only (F -statistics) or both allele frequencies and molecular information as given by the distance matrix among chlorotypes in an AMOVA framework (Φ -statistics). The difference between the two estimators with greatest value of Φ_{ST} compared to F_{ST} is an indication that the species presents some degree of phylogenetic subdivision (Excoffier *et al.* 1992). The statistical significance of estimators was computed by nonparametric permutational procedures of chlorotypes between populations within edaphic type, and between populations among edaphic type. To test for differences in population differentiation between edaphic types (F_{STM} vs. F_{STNM}) 1000 permutations of populations between edaphic types were carried out using *FSTAT* version 2.9.3. A neighbour-joining (NJ) population tree was drawn from Cavalli-Sforza and Edwards distances based on allele frequencies using the *POPULATIONS* software (O. Langella, UMR de Génétique Végétale, Ferme du Moulon, Gif/Yvette, France). To test for node robustness, bootstrapping was performed on individuals using 10 000 resamplings. Isolation-by-distance patterns between populations were tested considering all populations and then considering metallicolous and nonmetallicolous populations separately. The Mantel correlation coefficient between geographic and genetic distances was compared with its random distribution obtained from Mantel-like permutations of the genetic [$F_{ST}/(1 - F_{ST})$] and geographic [$\ln(\text{geographical distance})$] matrices as described in Rousset (1997), and implemented in the *GENEPOP* 3.4 software (Raymond & Rousset 1995).

Results

cpDNA polymorphism

Screening cpDNA of the subset of individuals using 72 different primer-enzyme combinations (PEC) revealed low polymorphism. Most intergenic regions were either monomorphic (Sfm) or presented polymorphism for a single restriction reaction only (AS-*AclI*; DT-*AclI*; K2Q-*HinfI*; TC-*AluI*; VL-*HinfI*). In these cases, the revealed polymorphism was always unique to a particular population (singleton). Three amplification products gave at least two nonredundant polymorphisms. They concerned the CD, CS and *trnK* (K1K2) fragments differentially restricted by *AluI*, *HinfI*, *Tru9I*, *HpaII* or *AclI*. Overall, eight non-redundant PECs involving these fragments were retained for use over the whole sample. These PECs generated 13 independent polymorphic sites, nine of which consisted in gain/loss of a restriction site whereas four were indels from 5 pb to 34 pb (Table 2). By sequencing three individuals representative of the variations detected, we confirmed the substitutive nature of the four restriction-site polymorphisms identified in the K1K2 region (GenBank Accession nos: DQ149105, DQ149106 and DQ149107).

Chlorotype relationship and diversity

Overall, 11 distinct chlorotypes labelled from A to K were defined (Tables 2 and 3). The MST revealed that they were linked by 13 nonhomoplasic mutation events (Fig. 1). In three cases (chlorotypes E-G, G-J and G-H), chlorotypes were indeed separated by more than a single mutation. Although they were not represented in our sampling, we considered that intermediate chlorotypes exist but are missing because of very low frequencies in the sampled regions. Anyway, these infrequent chlorotypes should have no influence on the estimation of the genetic structure. Most chlorotypes that were shared between edaphic types (four out of five) were interior chlorotypes. In contrast, chlorotypes that were specific to a given edaphic type were spread at the tips of the MST and did not cluster together.

Chlorotypes were not equally frequent overall (Table 3). Whereas chlorotypes A to F, that clustered together in the MST, were widely represented (78.7% of individuals), chlorotypes G to K were much less widespread (21.3% of individuals). Twenty-two out of the 28 populations were polymorphic and allelic richness varied from 1 to 4.983 across populations. Within-populations genetic diversity H_{Si} varied from 0 to 0.893 (± 0.086) with a high level of overall chlorotype diversity ($H_T = 0.853$, see Table 3). Since either monomorphy or a high level of diversity could be found for both large (F3 and SK5, respectively) and small sample sizes (Cz18 and D2, respectively), allelic richness and genetic diversity did not appear to be correlated with

4408 M. PAUWELS ET AL.

 Table 2 Description of the 11 PCR-RFLP chlorotypes identified in 28 populations of *Arabidopsis halleri*. The length variants are noted from 1 to 2; 9 correspond to restriction mutation site (Demesure *et al.* 1996)

Mutation number	1	2	3	4	5	6	7	8	9	10	11	12	13
PCR fragment	K1K2	K1K2	K1K2	CS	CS	CD	CD	CD	K1K2	K1K2	K1K2	CD	CD
Endonuclease	<i>HpaII</i>	<i>AcsI</i>	<i>Tru9I</i>	<i>AluI</i>	<i>HinfI</i>	<i>AcsI</i>	<i>AcsI</i>	<i>AcsI</i>	<i>HinfI</i>	<i>HinfI</i>	<i>Tru9I</i>	<i>HinfI</i>	<i>HinfI</i>
Polymorphic fragment	3	3	5	1	4	1	1	2	6	5	8	7	8
Haplotype													
A	9	9	9	1	1	1	1	9	1	1	1	1	2
B	9	9	9	9	1	1	1	9	1	1	1	1	2
C	9	9	9	1	1	1	9	9	1	1	1	1	2
D	9	1	9	1	1	1	1	9	1	1	1	1	2
E	1	9	9	1	1	1	1	9	1	1	1	1	2
F	1	9	9	1	1	1	1	9	1	1	2	1	2
G	1	9	9	1	1	1	1	1	1	2	1	1	2
H	1	9	1	1	9	1	1	1	1	2	1	1	2
I	1	9	9	1	1	1	1	1	9	2	1	1	2
J	1	9	9	1	1	9	1	1	1	2	1	2	2
K	1	9	9	1	1	9	1	1	1	2	1	2	1

 Table 3 Distribution of chlorotypes among populations of *Arabidopsis halleri*

Pop.	Edaphic type	n_i	Chlorotype											a_i^j	$H_{SI} (\pm SE)$
			A	B	C	D	E	F	G	H	I	J	K		
D1	NM	7	.	.	2	.	4	1	.	3	0.667 (± 0.160)
D2	NM	8	2	.	2	.	1	.	.	1	.	2	.	4.983	0.893 (± 0.086)
D3	NM	9	2	.	.	.	6	.	.	1	.	.	.	2.960	0.556 (± 0.165)
D4	NM	11	5	.	.	.	2	.	.	4	.	.	.	2.990	0.691 (± 0.086)
D6	NM	10	10	1	0
D8	M	13	1	.	.	10	1	2.674	0.318 (± 0.164)
D9	M	18	3	.	.	15	1.962	0.294 (± 0.119)
D11	M	18	5	.	.	9	4	2.987	0.660 (± 0.069)
D12	M	18	4	.	.	6	8	2.989	0.680 (± 0.056)
D13	M	20	.	.	.	17	3	1.940	0.268 (± 0.113)
B1	NM	29	2	27	1.680	0.133 (± 0.081)
F1	M	23	.	.	.	21	2	1.780	0.166 (± 0.097)
F2	M	43	.	.	.	27	16	1.999	0.478 (± 0.041)
F3	M	37	.	.	.	37	1	0
P11	M	63	40	1	.	.	22	.	2.209	0.482 (± 0.040)
P12	M	21	14	.	.	.	7	.	1.999	0.467 (± 0.075)
P13	M	15	13	.	.	.	2	.	1.934	0.248 (± 0.131)
P14	M	21	7	1	.	.	13	.	2.560	0.529 (± 0.079)
P16	M	10	10	1	0
SK2	NM	22	4	18	1.967	0.312 (± 0.106)
SK5	NM	46	19	1	.	25	1	.	2.565	0.545 (± 0.037)
CZ4	NM	15	.	13	1	.	.	1.759	0.143 (± 0.119)
CZ5	NM	24	.	22	.	.	.	2	1.762	0.159 (± 0.094)
CZ6	NM	12	.	12	1	0
CZ8	NM	26	.	26	1	0
CZ14	NM	20	20	.	.	1	0
CZ16	NM	57	56	.	.	.	1	.	.	1.231	0.035 (± 0.034)
CZ18	NM	9	9	1	0
Total		623	32	73	4	142	116	123	3	6	26	71	27	6.237	

NM, non metallicolous; M, metallicolous; n_i , number of genotyped individuals per population; a_i^j , standardized allelic richness; $H_{SI} (\pm SE)$, chlorotype diversity ($\pm SE$).

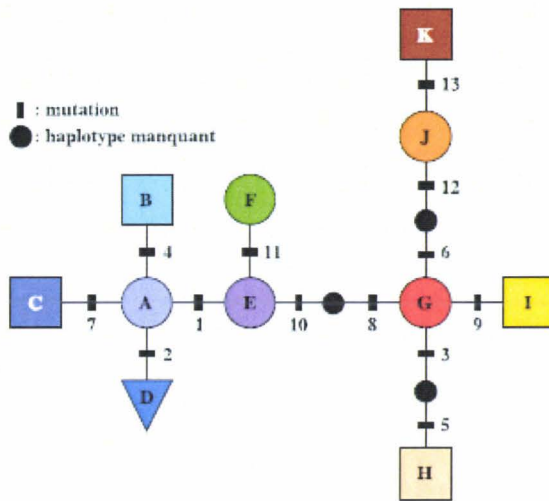


Fig. 1 Minimum-spanning tree representing the relationship between haplotypes inferred by polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP) analysis of *Arabidopsis halleri* cpDNA fragments. Triangles and squares represent haplotypes private to M and NM populations, respectively. Haplotypes occurring in both M and NM populations are indicated by circles. Numbers refer to mutation numbers as presented in Table 2.

sample size (respectively, Pearson's $r = -0.195$, $P = 0.320$; $r = -0.078$, $P = 0.693$). Interestingly, 6 chlorotypes out of 11 (B, C, D, H, I, K) were not shared between NM and M populations but were specific to one edaphic type (Fig. 2). Most of them (B, C, H, I, K) were NM specific, thus providing partial support to the hypothesis of a founder effect in M populations. However, neither allelic richness nor chlorotype diversity differed significantly between NM and M populations ($a_7^{NM} = 2.064$, $a_7^M = 2.002$, $P = 0.42$; $H_{SNM} = 0.237$, $H_{SM} = 0.349$, $P = 0.31$). NM-specific chlorotypes were among the least represented chlorotypes and

probably did not contribute much in terms of gene diversity estimates.

Genetic structure

Considering the whole sample, estimates of population differentiation revealed a high level of population structure ($F_{ST} = 0.678$, $P < 10^{-5}$). Taking into account molecular distances only slightly modified the inferred population genetic structure ($\Phi_{ST} = 0.677$, $P < 10^{-5}$). A Mantel test between the matrices of pairwise F_{ST} and Φ_{ST} values showed that weighting the estimates of genetic differentiation by molecular distances among chlorotypes did not affect the pattern of population differentiation ($P < 10^{-5}$). Population differentiation within each edaphic type was also high but did not differ significantly between NM and M populations ($F_{STM} = 0.538$, $F_{STNM} = 0.731$, $P = 0.123$).

Differentiation between edaphic types was found to be significant in terms of chlorotype frequencies only ($F_{CT} = 0.124$, $P = 4.9 \times 10^{-3}$, Table 4). Nevertheless, the level of differentiation due to variation between edaphic types dropped strongly when molecular distances between chlorotypes were taken into account ($\Phi_{CT} = 0.036^{ns}$, Table 4). This result suggests that, although edaphic types differed in chlorotype frequencies, chlorotypes in M populations were not more closely related to each other than randomly chosen chlorotypes. The distribution of chlorotypes in populations (Fig. 2) showed that chlorotypes present in eastern and western M populations (A, D, E and F, G, J, respectively) were indeed among the most distant in the MST (Fig. 1).

The NJ tree further confirmed that M and NM populations did not constitute distinct genetic groups (Fig. 3). Despite low bootstrap values, which might be explained by monolocus data combined with a relatively low sample size, M populations clustered in three distinct groups dispersed among NM populations. Each of these three groups

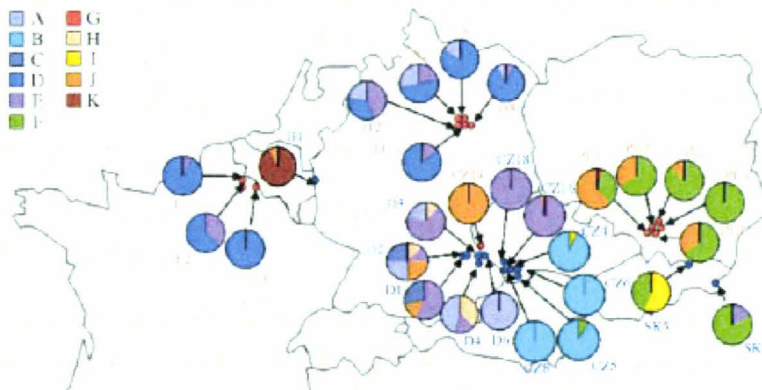


Fig. 2 Geographic distribution of *Arabidopsis halleri* chlorotypes. Metallicolous populations are noted in red, nonmetallicolous ones in blue.

4410 M. PAUWELS ET AL.

Table 4 Analysis of molecular variance (AMOVA) for 28 populations of *Arabidopsis halleri* using (a) haplotype frequencies only (*F*-statistics) or (b) both haplotype frequencies and molecular information (Φ -statistics). Statistics include degrees of freedom (d.f.); sums of squares deviations (SSD), variance component (VC), and percentages of total variance contributed by each component (%TV)

(a) Source of variation	d.f.	SSD	VC	%TV	<i>P</i>
Between edaphic types	1	25.890	0.057	12.37	**
Between populations within edaphic type	26	147.523	0.256	55.43	***
Within populations	595	88.641	0.149	32.20	***
Total	622	262.055	0.463		
(b) Source of variation	d.f.	SSD	VC	%TV	<i>P</i>
Between edaphic types	1	46.267	0.05306	3.63	NS
Between populations within edaphic type	26	536.717	0.93632	64.08	***
Within populations	595	280.673	0.47172	32.29	***
Total	622	863.657	1.46111		

NS, nonsignificant; ** *P* < 0.01; *** *P* < 0.001.

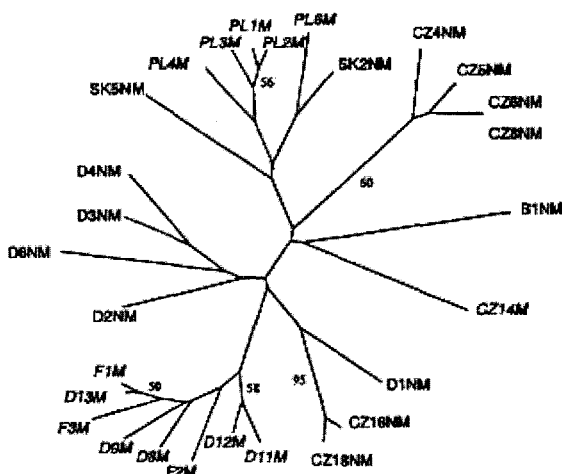


Fig. 3 Neighbour-joining population tree using Cavalli-Sforza and Edwards distances. Only bootstrap values greater than 50% are indicated. NM, nonmetallicolous; M, metallicolous.

clustered together geographically close M populations and was more genetically distant from the other ones than from geographically close NM populations. Thus, French and German M populations clustered with German NM ones (with which they shared two out of three chlorotypes: chlorotypes A and E, chlorotype D being M-specific), Polish M populations clustered with Slovak NM ones (mainly due to the presence of chlorotype F) and the CZ14 population clustered with the Belgian one. Inspection of Fig. 3 indicated that Cavalli-Sforza genetic distances were generally related to geographic distance.

This pattern was confirmed by a significant isolation-by-distance (IBD) pattern for either all populations

(*P* = 0.003) or each edaphic type separately (*P* < 10⁻⁵ for M sites and *P* = 0.022 for NM sites). The chlorotype distribution confirmed that chlorotype D was dominant in northern Germany and northern France (74.7% of individuals), closely related chlorotypes A and E were dominant in southern Germany and western Czech (74% of individuals), eastern Czech populations were almost exclusively composed of chlorotype B (94.8% of individuals), whereas Polish and Slovak populations were composed of chlorotype F (61.7% of individuals).

Discussion

Level of cpDNA polymorphism and population structure in *Arabidopsis halleri*

Few studies have explored population genetic structure in metallicolous plant populations. To our knowledge, this is the first study in pseudometallophytes using PCR-RFLP on the maternally inherited genome over a broad geographical range. Despite our exhaustive initial screening effort, which involved scattered populations and numerous PECs, we found low overall polymorphism in *Arabidopsis halleri*. Nevertheless, as usually observed (King & Ferris 1998; Raspé *et al.* 2000; Rendell & Ennos 2002; Magni *et al.* 2005), fragments K1K2 and CD were among the most variable cpDNA regions. Contrary to more general observations in the chloroplast genome (Clegg *et al.* 1994) and to results obtained in other species with the same approach (Raspé *et al.* 2000; Grivet & Petit 2002; Palme *et al.* 2003a, b), most polymorphisms revealed in *A. halleri* were due to point mutations rather than to size polymorphism. Although our data are not appropriate for rigorously discussing the molecular evolution of cpDNA (sequencing analysis should be more relevant), these results may

suggest divergence of molecular evolutionary processes in the chloroplast genome between taxa.

Most studies that have analysed population structure and history at a large geographic scale using PCR-RFLP on cpDNA have been performed on woody species (Petit *et al.* 2005), and similar reports on herb species are scarce (Stehlik 2002; Stehlik *et al.* 2002; McCauley *et al.* 2003). In comparison to the detected number of cpDNA chlorotypes and level of population structure typically reported in woody species [respectively mean = 16.9 (min: 4 – max: 50) and 0.54 (min: 0.09 – max: 1) for 22 European trees and shrubs studied over Europe, in Petit *et al.* 2003], values in *A. halleri* (11 chlorotypes, $F_{ST} = 0.678$) can be considered typically average. The level of population differentiation we observed could be due to the relatively recent colonization of industrially polluted sites or to metapopulation dynamics of the species in the sampled area (McCauley 1995; Ouborg *et al.* 1999). However, since we did not find any significant difference in the level of population differentiation between edaphic types, high values of the estimators cannot be linked to a particular feature of pseudometallophyte species.

For most studies which have compared estimators of population differentiation either taking into account or not the molecular distance between haplotypes in European woody species (for a review see Petit *et al.* 2003), the existence of a 'phylogeographic' structure (Schaal *et al.* 1998) has been revealed by greater value of the estimator when the molecular distance is taken into account. Overall patterns of postglacial colonization of Europe have thus been defined which have confirmed that the southern peninsulas of Europe acted as major ice age refugia where species were restricted during the last glacial maximum (LGM; Hewitt 1999). In our study, sampling was adapted to the study of M population history and thus did not cover southern parts of the species range. However, the striking similarity of F_{ST} and Φ_{ST} strongly suggests that *A. halleri* shows a different pattern than would be expected in a species that was confined to southern refugia during the LGM. Further sampling, including populations from southern Europe, will be needed to test the existence of refugia at higher latitudes in *A. halleri*.

Colonization of polluted areas and genetic bottleneck

The founder-effect hypothesis in M populations initially relied on the assumptions that tolerant genotypes have low frequency in founder populations and that evolution towards heavy metal tolerance causes a strong genetic bottleneck during colonization of polluted areas (Bradshaw 1984). Many studies of pseudometallophytes exhibiting population tolerance have therefore attempted to detect reduced genetic variation in M populations as compared to NM ones. Surprisingly, most of them failed to detect strong founder effect using allozymes (for a review, see Vekemans &

Lefèbvre 1997; Mengoni *et al.* 2001). Recently, however, Mengoni *et al.* (2001) compared studies of eight geographically closed populations of the pseudometallophyte *Silene paradoxa* using either random amplified polymorphic DNA (RAPD) or chloroplast microsatellite loci (cpSSR) and highlighted the greater relevance of cpDNA markers for successfully detecting reduction in genetic diversity in M populations. Using cpSSR data, the authors detected a founder effect in tolerant populations of *Silene vulgaris* from copper-contaminated soils in comparison to copper-intolerant populations from serpentine soils or uncontaminated sites that were not detected from RAPD data. The founder effect at nuclear loci may indeed be eroded by subsequent pollen flow from populations neighbouring contaminated sites which could increase genetic variation in M populations (Vekemans & Lefèbvre 1997; Mengoni *et al.* 2001).

The study of Mengoni *et al.* (2001) provided some evidence that, as expected from the founder-effect hypothesis, adaptation to metal-contaminated environments is associated with a genetic bottleneck in species with population tolerance. Conversely, our large-scale survey suggests that it is not the case for species with constitutive tolerance. Although the number of PCR-RFLP chlorotypes is overall lower in *A. halleri* M populations than in NM ones, we did not observe any significant difference between edaphic groups in allelic richness estimated by the rarefaction method or in gene diversity. Chlorotypes that are specific to NM populations are tip-chlorotypes in the MST and are among the less geographically widespread, as expected (Crandall & Templeton 1996). Although they were probably not represented in the pool of migrants that colonized polluted sites, their absence in M populations does not appear to be a consequence of strong gene diversity reduction. Moreover, only one chlorotype is specific to the M populations (chlorotype D, private to French and German M populations), suggesting that the diversity observed in M populations cannot be explained only by new chlorotypes produced by mutation events after colonization. Our results differ from a previous survey using 42 nuclear microsatellites loci to study population genetic structure in *Arabidopsis* species (Clauss *et al.* 2002). The survey revealed an excess of heterozygotes in one M population of *A. halleri* from Thuringia, Germany, that could coincide with a hypothesized founder event during the expansion of mining in Thuringia. However, as highlighted by the authors, analysis of variation in additional populations and comparisons with further M or NM populations were lacking. As a result, the possibility that it revealed a general feature of *A. halleri* populations, or, conversely, a particular feature of the sampled population independently of its edaphic type cannot be excluded from this study.

Recently, Macnair (1997) demonstrated that polluted areas differ from unpolluted ones in many components

4412 M. PAUWELS ET AL.

apart from higher metal contamination (soil structure, organic and mineral content, dryness, etc.). He suggested that the evolution of metal tolerance could be only a part of the evolution of M populations. In this context, the founder-effect hypothesis could be extended to other adaptive characters required for the colonization of polluted areas and apply to species with constitutive tolerance. Nevertheless, our study suggests that when the capacity to grow on polluted soils exists in founder populations, it could be sufficient to prevent a founder effect in M populations. In consequence, founder effects in pseudometallophyte M populations may only be due to adaptation for heavy metal tolerance and may not be detected in species exhibiting constitutive tolerance despite the numerous ecological constraints encountered in metal-polluted areas.

The independent origin of metallicolous populations

The hypothesis of multiple independent origins of M populations within pseudometallophytes has already been validated for species showing tolerance in M populations only (Westerbergh & Saura 1992; Vekemans & Lefebvre 1997; Koch *et al.* 1998; Mengoni *et al.* 2001), or in both M and NM populations as for *Thlaspi caerulescens* (Koch *et al.* 1998). Our results strongly suggest that in *A. halleri* the foundation of M populations also occurred several times and that the increased tolerance observed in M populations has a polyphyletic origin. Hierarchical AMOVA clearly showed that M populations are not distinct from NM ones when phylogenetic information is taken into account and no correspondence between edaphic type and chlorotype lineage was observed. Population structure was instead due to isolation by distance rather than to zinc exposure; population haplotypic composition rather depended on geographical location than on edaphic origin. As a result, M populations did not cluster together in the NJ tree but in three separated groups dispersed among NM populations and corresponding to the different polluted areas we sampled. Except for French M populations, genetic relationships between groups of M populations and their spatially closest NM populations allow us to draw assumptions about foundation events: German M populations may have been founded by German NM populations (haplotypes A), Polish M populations may come from the Slovakian NM population (haplotype F), and sharing of haplotype J by CZ14 and the geographically closest D1 and D2 strongly suggests the foundation of CZ14 from D1 or D2. In such a scenario, almost all M populations have been founded either from distinct NM population groups (e.g. D8-13 and PL1-6) or from the same group but separately in time (e.g. D8-13 and CZ14). Accordingly, both historical (Berton 1946) and genealogical data (Fig. 2) suggest that *A. halleri* was introduced into northern France during the first part of

the 20th century from metallicolous sites located in the Harz (Germany), where its distribution was associated with mining activities since the Middle Ages (Liessmann 1997; Brooks 1998).

The fact that different independently founded M populations could present different levels of tolerance has led several authors to consider that locally selected tolerance mechanisms could differ between M populations (for a review, see Schat *et al.* 1996). In *S. vulgaris*, Schat *et al.* (1996) found identical major zinc-tolerance genes in tolerant populations which were supposed to have evolved independently from intolerant ones and suggested that variation in their tolerance abilities should rather be due to allelic variation between major loci or to the presence of 'modifiers' (Smith & Macnair 1998). In *A. halleri*, tolerance is species-wide but quantitative variation exists among and within edaphic type (Bert *et al.* 2000; Pauwels *et al.* in prep.). Parsimony suggests that fixed tolerance has occurred only once, probably early in the species' history, and that major gene(s) that produce constitutive tolerance are identical for all individuals. However, secondary mechanisms (allelic variation or presence of modifiers) may exist that confer enhanced tolerance to M populations. According to Schat *et al.* (1996), the genetic make-up of such mechanisms may differ in separately founded M populations. Mapping QTL for tolerance in recombinant populations from multiple crosses between a NM population and independently founded M populations should provide further support for this hypothesis (Symonds *et al.* 2005).

The findings of this study clearly demonstrate that integrating the phylogenetic component of the spatial distribution of gene lineages and relocating the colonization of recently polluted sites in the overall background of the species' history makes it easier to identify the impact of contaminant exposure on the genetic structure of a species (Staton *et al.* 2001). Such an approach should therefore be successful in investigating genealogical relationships between M populations in pseudometallophytes from both ecological categories (i.e. presenting either constitutive or populational metal tolerance). It has to be noticed that some of the results from intraspecific studies could partially be due to the sharing of alleles with other subspecies or close relatives. In our study, for example, region-specific results could result from a 'swamping out' of adaptive lineage divergence. In this case, a comparative approach including other subspecies or close relatives should be successful in determining how other underlying divergence processes may have accelerated or possibly impeded divergence of chlorotypes. Moreover, in all pseudometallophytes such as *A. halleri* with species-wide tolerance, the evolutionary origin of the tolerance character cannot be determined from intraspecific sampling and a comparative approach covering the closest nontolerant relatives is necessary.

Acknowledgements

The authors thank Cécile Godé for technical advice and support in sequencing. They are very grateful to Dr Vlastimil Mikolas, Slavomir Sokol, Maciej Szczepka, Patrick de Laguerie and Prof Andreas Bresinsky for help in finding populations. They are also very grateful to Marcus Koch, Vincent Castric, Xavier Vekemans, Jean-François Arnaud, and one anonymous referee for helpful comments on the manuscript. This work was supported by funding from the Contrat de Plan État/Région Nord-Pas de Calais (PRC), from the European FEDER (contract no. 79/1769), from the 'Programme National/ACI ECCO' (contract no. 04 2 9 FNS) from the CNRS, and from the Institut Français de la Biodiversité (contract no. 00N55/0083).

References

- Al-Hiyaly SA, McNeilly T, Bradshaw AD (1988) The effect of zinc contamination from electricity pylons — evolution in a replicated situation. *New Phytologist*, **110**, 571–580.
- Al-Shehbaz IA, O'Kane SLJ (2002) Taxonomy and Phylogeny of *Arabidopsis* (Brassicaceae). In: *The Arabidopsis Book* (eds Somerville CR, Meyerowitz EM). American Society of Plant Biologists, Rockville, MD.
- Antonovics J, Bradshaw AD, Turner RG (1971) Heavy metal tolerance in plants. *Advances in Ecological Research*, **7**, 1–85.
- Baker AJM (1987) Metal tolerance. *New Phytologist*, **106** (Suppl.), 93–111.
- Bert V, Bonnin I, Saumitou-Laprade P, de Laguerie P, Petit D (2002) Do *Arabidopsis halleri* from nonmetallicolous populations accumulate zinc and cadmium more effectively than those from metallicolous populations? *New Phytologist*, **155**, 47–57.
- Bert V, Macnair MR, de Laguerie P, Saumitou-Laprade P, Petit D (2000) Zinc tolerance and accumulation in metallicolous and non-metallicolous populations of *Arabidopsis halleri* (Brassicaceae). *New Phytologist*, **146**, 225–233.
- Berton A (1946) Présentation de plantes *Arabis Halleri*, *Armeria elongata*, *Oenanthe fluviatilis*, *Galinsoga parviflora discoidea*. *Bulletin de la Société de Botanique de France*, **93**, 139–145.
- Bradshaw AD (1984) The importance of evolutionary ideas in ecology and vice versa. In: *Evolutionary Ecology* (ed. Shorrocks B), pp. 1–25. Blackwell Science Publications, Oxford, UK.
- Bradshaw AD, McNeilly T (1991) Stress tolerance in plants: the evolutionary framework. In: *Ecological Responses to Environmental Stresses* (eds Rozema J, Verkleij JAC), pp. 2–5. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Brooks R (1998) Phytoarcheology and hyperaccumulators. In: *Plants That Hyperaccumulate Heavy Metals* (ed. Brooks R), pp. 181–202. CAB International, Oxford, UK.
- Buse A, Norris D, Harmens H *et al.* (2003) *Heavy Metals in European Mosses: 2000/2001 Survey*. Centre for Ecology and Hydrology, University of Wales Bangor.
- Clauss MJ, Cobban H, Mitchell-Olds T (2002) Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaceae). *Molecular Ecology*, **11**, 591–601.
- Clegg MT, Gaut BS, Learn GH Jr, Morton BR (1994) Rates and patterns of chloroplast DNA evolution. *Proceedings of the National Academy of Sciences, USA*, **91**, 6795–6801.
- Crandall KA, Templeton AR (1996) Applications of intraspecific phylogenetics. In: *New Uses for New Phylogenies* (eds Harvey PH, Brown AJL, Maynard Smith J, Nee S). Oxford University Press, Oxford, UK.
- Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution*, **50**, 2515–2520.
- Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology*, **4**, 129–131.
- Dumolin-Lapegue S, Fomonge MH, Petit RJ (1997) An enlarged set of consensus primers for the study of organelle DNA in plants. *Molecular Ecology*, **6**, 393–397.
- El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, **92**, 832–839.
- Ennos RA, Sinclair WT, Hu X-S, Langdon A (1999) Using organelle markers to elucidate the history, ecology and evolution of plant populations. In: *Molecular Systematics and Plant Evolution* (eds Hollingsworth PM, Bateman RM, Gornall RJ), pp. 1–19. Taylor & Francis, London and New York.
- Excoffier L, Smouse PE (1994) Using allele frequencies and geographic subdivision to reconstruct gene tree within a species: molecular variance parsimony. *Genetics*, **136**, 343–359.
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Goudet J (2001) *FSTAT*. A program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from <http://www.unil.ch/izea/software/fstat.html>
- Grivet D, Petit RJ (2002) Phylogeography of the common ivy (*Hedera* sp.) in Europe: genetic differentiation through space and time. *Molecular Ecology*, **11**, 1351–1362.
- Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.
- Jain SK, Bradshaw AD (1966) Evolutionary divergence among adjacent plant populations. I. The evidence and its theoretical analysis. *Heredity*, **20**, 407–441.
- Kalinowski ST (2004) Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. *Conservation Genetics*, **5**, 539–543.
- King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151–1161.
- Koch M, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis* and related genera (Brassicaceae). *Molecular Biology and Evolution*, **17**, 1483–1498.
- Koch M, Haubold B, Mitchell-Olds T (2001) Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. *American Journal of Botany*, **88**, 534–544.
- Koch M, Mummenhoff K, Hurka H (1998) Systematics and evolutionary history of heavy metal tolerant *Thlaspi caerulescens* in Western Europe: evidence from genetic studies based on isozyme analysis. *Biochemical Systematics and Ecology*, **26**, 823–838.
- Lefebvre C, Vernet P (1990) Microevolutionary processes on contaminated deposits. In: *Heavy Metal Tolerance in Plants: Evolutionary Aspects* (ed. Shaw J), pp. 285–300. CRC Press, Boca Raton.
- Liessmann W (1997) *Historischer Bergbau Im Harz*. Springer-Verlag, Berlin, Heidelberg, New York.
- Linhart YB, Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*, **27**, 237–277.
- Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology*, **7**, 963–974.

4414 M. PAUWELS ET AL.

- Macnair MR (1987) Heavy metal tolerance in plants: a model evolutionary system. *Trends in Ecology & Evolution*, **2**, 354.
- Macnair MR (1997) The evolution of plants in metal-contaminated environments. In: *Environmental Stress, Adaptation and Evolution* (eds Bijlsma R, Loeschcke V), pp. 1–24. Birkhäuser Verlag, Basel, Switzerland.
- Magni CR, Ducouso A, Caron H, Petit RJ, Kremer A (2005) Chloroplast DNA variation of *Quercus rubra* L. in North America and comparison with other Fagaceae. *Molecular Ecology*, **14**, 513–524.
- McCauley DE (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology & Evolution*, **10**, 198–202.
- McCauley DE, Smith RA, Lisenby JD, Hsieh C (2003) The hierarchical spatial distribution of chloroplast DNA polymorphism across the introduced range of *Silene vulgaris*. *Molecular Ecology*, **12**, 3227–3235.
- Meerts P, van Isacker N (1997) Heavy metal tolerance and accumulation in metallicolous and nonmetallicolous populations of *Thlaspi caerulescens* from continental Europe. *Plant Ecology*, **133**, 221–231.
- Mengoni A, Barabesi C, Gonnelli C *et al.* (2001) Genetic diversity of heavy metal-tolerant populations in *Silene paradoxa* L. (Caryophyllaceae): a chloroplast microsatellite analysis. *Molecular Ecology*, **10**, 1909–1916.
- Mitchell-Olds T (2001) *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends in Ecology & Evolution*, **16**, 693–700.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Ouborg NJ, Piquot Y, Van Groenendael JM (1999) Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology*, **87**, 551–568.
- Palme AE, Semerikov V, Lascoux M (2003a) Absence of geographical structure of chloroplast DNA variation in sallow, *Salix caprea* L. *Heredity*, **91**, 465–474.
- Palme AE, Su Q, Rautenberg A, Manni F, Lascoux M (2003b) Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. *Molecular Ecology*, **12**, 201–212.
- Petit RJ, Aguinagalde I, de Beaulieu J-L *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, **300**, 1563–1565.
- Petit RJ, Duménil J, Fineschi S *et al.* (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, **14**, 689–701.
- Pollard AJ, Powell KD, Harper FA, Smith JAC (2002) The genetic basis of metal hyperaccumulation in plants. *Critical Reviews in Plant Sciences*, **21**, 539–566.
- Raspé O, Saumitou-Laprade P, Cuguen J, Jacquemart A-L (2000) Chloroplast DNA haplotype variation and population differentiation in *Sorbus aucuparia* L. (Rosaceae: Maloideae). *Molecular Ecology*, **9**, 1113–1122.
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
- Rendell S, Ennos RA (2002) Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe. *Molecular Ecology*, **11**, 69–78.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.
- Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, **7**, 465–474.
- Schat H, Vooijs R, Kuiper E (1996) Identical major gene loci for heavy metal tolerances that have independently evolved in different local populations and subspecies of *Silene vulgaris*. *Evolution*, **50**, 1888–1895.
- Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN: A software for population genetics data analysis. Version 2.000, Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Switzerland.
- Smith SE, Macnair M (1998) Hypostatic modifiers cause variation in degree of copper tolerance in *Mimulus guttatus*. *Heredity*, **80**, 760–768.
- Staton JL, Schizas NV, Chandler GT, Coull BC, Quattro JM (2001) Ecotoxicology and population genetics: the emergence of 'phylogeographic and evolutionary ecotoxicology'. *Ecotoxicology*, **10**, 217–222.
- Stehlik I (2002) Glacial history of the Alpine herb *Rumex nitidus* (Polygonaceae): a comparison of common phylogeographic methods with nested clade analysis. *American Journal of Botany*, **89**, 2007–2016.
- Stehlik I, Blattner FR, Holderegger R, Bachmann K (2002) Nunatak survival of the high Alpine plant *Ertrichium nanum* (L.) Gaudin in the central Alps during the ice ages. *Molecular Ecology*, **11**, 2027–2036.
- Symonds VV, Godoy AV, Alconada Tet *et al.* (2005) Mapping quantitative trait loci in multiple populations of *Arabidopsis thaliana* identifies natural allelic variation for trichome density. *Genetics*, **169**, 1649–1658.
- Van Rossum F, Bonnin I, Fenart S *et al.* (2004) Spatial genetic structure within a metallicolous population of *Arabidopsis halleri*, a clonal, self-incompatible and heavy-metal-tolerant species. *Molecular Ecology*, **13**, 2959–2967.
- Vekemans X, Lefévre C (1997) On the evolution of heavy-metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. *Journal of Evolutionary Biology*, **10**, 175–191.
- Westerbergh A, Saura A (1992) The effect of serpentine on the population structure of *Silene dioica* (Caryophyllaceae). *Evolution*, **46**, 1537–1548.
- Wu L (1990) Colonization and establishment of plants in contaminated sites. In: *Heavy Metal Tolerance in Plants: Evolutionary Aspects* (ed. Shaw AJ), pp. 269–284. CRC Press, Boca Raton, Florida.

The authors of the laboratory 'Génétique et Evolution des Populations Végétales' are part of the 'genetics and evolution of plant adaptation to sites polluted by heavy metal' team headed by Pierre Saumitou-Laprade. They are involved in population genetics of plant species and QTL mapping with a special attention given to the genetics and evolution of zinc tolerance and zinc hyperaccumulation in *Arabidopsis halleri*. The present manuscript will form part of Maxime Pauwels' thesis mainly interested in comparing geographical distribution of zinc tolerance abilities over the whole species range and population structure of both chloroplast and nuclear genomes. I. Bonnin is population geneticist particularly interested in evolution of selected characters and relationships between migration, diversity and adaptation in experimental populations of wheat and *Arabidopsis*.

Chapitre 3: Population genetic structure of the pseudometallophyte *A. halleri* in Europe: evidences from cpDNA and nuclear microsatellites (in prep).

Abstract. To unravel the fingerprints of postglacial history in the current population genetic structure of the pseudometallophyte *A. halleri*, the population genetic structure of sixty-five European populations was assessed using cpDNA and nuclear microsatellite loci. Both types of markers revealed a high level of population structure at the European scale ($F_{STc} = 0.727$; $F_{STn} = 0.31$). Genetic structure was clearly hierarchical. At the upper-most level, two latitudinally separated and longitudinally organized genetic groups were distinguished north and south of the Alps, among which a majority of cpDNA variation and a substantial level of nuclear microsatellite variation were partitioned. The Alps appeared as a strong barrier to gene flow causing genetic divergence among populations. The two majors groups identified did not differ in average levels of gene diversity and a significant phylogeographic signal was detected among groups, suggesting the maintenance of *A. halleri* north of the Alps during the Glacial Ages. Thirteen genetic subgroups were distinguished within groups. We used comparisons of genetic diversity indices among subgroups to discuss the location of glacial refuges and postglacial colonization routes. The impact of the recent colonization of metal-polluted sites on population structure was also debated.

Keywords: cpDNA, phylogeography, nuclear microsatellite, population genetic structure, glacial refugia, colonization, Alps

Introduction

Arabidopsis halleri (L.) (O'Kane & Al-Shehbaz) (syn. *Cardaminopsis halleri* (L.) Hayek) is a clonal, self incompatible and highly outcrossing perennial weed largely distributed in Europe. The species mainly occurs in occurring on acidic, fresh and oligotrophic soils (Clapham & Akeroyd 1993) but also on limestones in the Tatra Mountains (Saumitou-Laprade, pers. com.). Two European *A. halleri* subspecies are currently described. *A. h.* subsp. *halleri* is distributed throughout Europe, except in the western part, the far north and most of Italy (O' Kane & Al-Shehbaz 1997). It is encountered on montane slopes, forest margins and rocky crevices from 600 to 2200 m upper the sea level, generally below timberline (O' Kane & Al-Shehbaz 1997; Al-Shehbaz & O'Kane 2002). *A. h.* subsp. *ovirensis* is distributed on rocky slopes and crevices in eastern and southern Europe, between 1200 and 2400 m, above timberline (O' Kane & Al-Shehbaz 1997; Al-Shehbaz & O'Kane 2002). As a pseudometallophyte, *A. halleri* also belongs to the minority of plant species that occurs on both metal-polluted (metalliferous) and non polluted soils (Bert *et al.* 2000; Pauwels *et al.* submitted). *A. halleri* indeed evolved the amazing ability to tolerate high amount of metal in soils on which it develops (Bert *et al.* 2000; Dahmani-Muller *et al.* 2000; Bert *et al.* 2003; Pauwels *et al.* submitted).

Given its close phylogenetic relationships with the model system for plant molecular biology *A. thaliana* (Koch *et al.* 2001; Al-Shehbaz & O'Kane 2002), *A. halleri* has recently been proposed to be an appropriate model for the study of metal tolerance and hyperaccumulation in plants (Mitchell-Olds 2001; Lexer & Fay 2005). The easy shift of the huge amount of molecular tools available for *A. thaliana* is undeniably an incomparable asset for the application of the high-throughput technologies used in the identification of candidate genes involved in adaptive characters (Assunção *et al.* 2003). Consequently, *A. halleri* has already been used in several molecular studies benefiting from these technologies and which have focused on the characterization of the molecular physiological mechanisms underlying the adaptation to metal stress conditions (Becher *et al.* 2004; Dräger *et al.* 2004; Weber *et al.* 2004). So far, studies of Zn metal tolerance and hyperaccumulation of *A. halleri* relied on a very few M accessions of the species (when mentioned), mainly located in Langelsheim, Harz

Mountains, Germany (Becher *et al.* 2004; Dräger *et al.* 2004; Weber *et al.* 2004) or Auby, Northern France (Dräger *et al.* 2004; Willems *et al.* submitted). M populations indeed suffered metal exposure and are therefore supposed to have evolved physiological mechanisms in response to metal-induced stress.

Evolution of tolerance mechanisms in M populations highly depends on the locally available genetic variance for tolerance (Macnair 1997). Traditionally, population genetics models considered that, at least at the local geographic scale, genetic variation in populations was moulded principally by four contemporary factors: mutation, migration, genetic drift, and natural selection (Staton *et al.* 2001). Conversely, phylogeography has recently revealed that, when it is described over a wide distribution range, intraspecific genetic variation could have been greatly influenced by the historical processes in population demographics over extended periods of time (Avice 2000). It has thus been sustained that the understanding of the consequences of demographic history on the neutral genetic structure of a species might enable to interpret current variation at genes underlying ecologically important traits in populations (Feder & Mitchell-Olds 2003; Storz 2005; Wright & Gaut 2005) and was necessary to understand effects of a stressor, e.g. metal pollution, on the evolution of a particular population (Staton *et al.* 2001). Moreover, the knowledge of the genetic structure of neutral loci can serve as a background for the detection and analysis of loci of adaptive significance (Lascoux *et al.* 2004).

In the last decades, the phylogeographic framework, initiated by Avice (Avice *et al.* 1987), gave new means of understanding demographic history of a particular species from its current population genetic structure. More precisely, the phylogeographic approach proposed to seek the footprints of historical processes (e.g. range expansion, bottlenecks, vicariance, migration) in present-day patterns of large scale genetic structure (Schaal *et al.* 1998; Avice 2000) from the pattern of congruence or lack of congruence between the genealogical relationships among neutral genes and their current geographical distribution. Phylogeographic studies have for example been successful in confirming from the analysis of current spatial distribution of genetic lineages in European tree and shrub species that most of them, although common in northern Europe today, survived the Last Glacial Maximum (LGM, 115,000 to 15,000 ¹⁴C years ago) in a range of different refugia located in Southern European peninsula (e. g. King & Ferris 1998; Grivet & Petit 2002; Rendell & Ennos 2002). They also inferred migration routes implicated in the postglacial recolonization of Northern

Europe from the southern refugia (Taberlet *et al.* 1998; Hewitt 1999; Hewitt 2000). Northward expansion have been shown to imply recurrent founder events resulting in latitudinal gradients of genetic divergence between populations, allelic richness and intraspecific diversity (Hewitt 1996; Comps *et al.* 2001; Petit *et al.* 2003). Interestingly, steep latitudinal gradients have also been observed in life-history traits of forest trees, such as budburst or budset, although nothing is known about how it relates to the post-glacial history of the species (Lascoux *et al.* 2004).

The chloroplast genome (cpDNA) has been the sole source of genetic variation of most published plant phylogeography studies, especially those carried at the European scale (Petit *et al.* 2005). Several factors account for this observation. In flowering plants, the mutation rate of cpDNA is sufficient to detect intraspecific polymorphism (Soltis *et al.* 1992) and low enough (Clegg *et al.* 1994) to allow cpDNA to retain historical imprints over long time periods. The cpDNA is haploid, non-recombining, and usually maternally inheritance (Birky 1995). This makes the construction of genealogical relationships between genes easier than for nuclear loci, because new mutations are consistently transmitted along maternal lineages and cannot be exchange among lineages (Cruzan & Templeton 2000). Effective population size of cpDNA is reduced compared to a diploid locus and population structure is usually expected to be stronger than for nuclear markers (Birky *et al.* 1989; Petit *et al.* 1993; Ennos 1994). Finally, limited dispersal of cpDNA preserves the patterns of genetic variation that were created during the establishment of current distribution (Crandall & Templeton 1996; Templeton 1998).

CpDNA represents only a single gene genealogy and conclusions that can be drawn from a single non-recombining genome are in general limited (Nordborg & Innan 2002). Contradictions between genetic structure based on cpDNA or nuclear markers have for example been reported (Avice 1994), in particular with respect to the genetic divergence of populations (Petit *et al.* 1998). It has been argued that studies rigorously designed to investigate population demography should require a large number of genome-wide markers allowing for recombination and integrating several genealogical processes (Nordborg & Innan 2002). Microsatellites are abundant and widely distributed markers throughout the eukaryote genome (Ruzzante 1998). They often exhibit high levels of allelic polymorphism and therefore ensure low standard errors in estimates of population genetic statistics (Goudet *et al.* 1996). Thus, microsatellite markers have been shown to be among the best markers available

to detect bottleneck (Cornuet & Luikart 1996). They can also provide a good description of the boundaries between populations and an accurate representation of their historical relationships (Goldstein *et al.* 1999). Microsatellites are thus among the most suitable markers to detect genetic or demographic structure of populations (Ruzzante 1998; Goldstein *et al.* 1999; Evanno *et al.* 2005). They have been recently largely used in study of population structure, mainly those involving Bayesian approaches (Natoli *et al.* 0; Heuertz *et al.* 2004; Luo *et al.* 2004; Song *et al.* 2006; Clauss & Mitchell-Olds in press).

Despite the current status of *A. halleri* as a model for the study of plant metal tolerance, the available knowledge about the species ecology and evolutionary history remains sparse. In this study, we attempted to improve the understanding of the postglacial history of the species investigating the neutral genetic structure of *A. halleri* at the European scale. First, we extended the cpDNA phylogeographic study initiated by Pauwels *et al.* (2005) to the entire European species range but focusing mainly on *A. h. subsp. halleri*, the only subspecies that is currently mentioned on metalliferous sites in Europe. For reasons described above, we also studied population genetic structure using nuclear microsatellite loci on the same sampling. Demographic history of *A. halleri* during and from the L.G.M. as well as the impact of the recent colonization of polluted area in the overall patterns of genetic structure were discussed.



Material and methods

Sampling and DNA extraction

Sixty-one scattered populations of *A. h.* subsp. *halleri* were sampled so as to cover the subspecies range (Annexe A). The scale of sampling ranged from less than one kilometre between populations (e.g. I06 and I07) to more than 1500 kilometres for the most distant populations (e.g. F01 and RO06). Mean geographic distance was 573.76 ± 339.96 km. In each of samples, leaves were collected from 7 to 70 individuals separated by at least three meters to avoid sampling of clones (van Rossum *et al.* 2004). Variations in sample size more or less reflected variation in population size. In particular, smallest sample sizes were almost exhaustive but corresponded to smallest populations.

Most populations (45) were scattered in the European mountain ranges (Figure 1), at middle to high altitudes (mean = 702 (\pm 360) m), far from any known metal pollution source, in non-contaminated environments (NM populations, Annexe A, Fig. 1). Sixteen populations (D08-09, D11-14, F01-03, PL01-04, PL06-08, henceforth referred as metallicolous (M) populations) were sampled outside the mountain ranges, at low to moderate altitudes (mean = 255 (\pm 189) m), in the three known polluted areas the species has colonized (Annexe A, Fig. 1). The extension of the species range in these disturbed habitats is assumed to be recent and has to be related to anthropogenic metal pollution via industrial or mining activities (Berton 1946; Fabiszewski 1986; Ernst 1990). Moreover, given the high selection pressure usually encountered in metal contaminated area, substantial founder effects are expected to have occurred during their colonization (Lefèbvre & Vernet 1990; Wu 1990).

In addition to *A. h. halleri* samples, leaves were similarly collected in 3 NM populations of *A. h.* subsp. *ovirensis* located in the Fagaras Mountains (Southern Carpathian, Romania).

Leaves were immediately dried in silica gel prior to molecular analysis. DNA from each genotype was extracted from 15-20 mg dry material using Qiagen® Dneasy® kit and PCR amplification was performed on 1/100 dilutions.

cpDNA analysis

Genotyping. Part of the cpDNA analysis of population samples was done in (Pauwels *et al.* 2005). Genotyping of remaining samples (except CZ07 and D14) were performed after modification of the molecular technique used. Polymorphic sites identified through PCR-RFLP analysis in the K1K2 and CS cpDNA fragments in Pauwels *et al.* (2005) were precisely characterized by sequencing analysis and new primers were defined so as to convert polymorphic sites into SNP markers, more adapted to large sample sizes (Annexe 1). In order to homogenise the global dataset, results of SNP genotyping were subsequently encoded to correspond to PCR-RFLP data (see Pauwels *et al.* 2005). Due to the non-recombining nature of the chloroplast genome, alleles observed at all loci were combined into chlorotypes (cpDNA haplotypes).

Relationship among chlorotypes. A minimum spanning tree (MST) of the chlorotypes was computed from a distance matrix containing the number of differences between each pair of chlorotypes using ARLEQUIN 2.0 (Schneider *et al.* 2000). MST construction assumes that each chlorotype is linked by a single or a series of mutational events to all other chlorotypes through a unique pathway and the construction method minimizes the number of such events (Excoffier & Smouse 1994).

Genetic diversity and overall population structure analysis. Allelic / chlorotypic richness (A_{Sc}) was estimated for each population and over the whole sample using FSTAT 2.9.3. (Goudet 2001). Given that the smallest sample concerned population SLO8 with $n = 6$, estimates of A_{Sc} were standardised to a standard sample size equal to 6. Chlorotypic diversity (H_S) and its sampling variance (Nei 1987) were estimated within each population, and over the whole sample (H_T) using ARLEQUIN.

Global F_{STc} (for chloroplast F_{ST}) and N_{ST} was computed using the software SPaGeDi 1.2b (Hardy & Vekemans 2002). To test if the phylogenetic distance between chlorotypes is informative with respect to population differentiation (phylogeographic structure), 10 000 permutations of rows and columns of distance matrices between chlorotypes were performed and the probability that observed values exceed mean G_{ST} permutation values was computed.

Microsatellite analysis

Genotyping. Six primer pairs of microsatellite loci which had previously been shown to display easy to read band patterns and a high degree of polymorphism (van Rossum *et al.* 2004) were used for polymerase chain reaction (PCR). Five microsatellites (ATH, GC16, LYR132, LYR133, LYR417) have been previously used at the within-population level in a previous study (van Rossum *et al.* 2004, Annexe 2). The remaining one (LYR104) was kindly provided by Thomas Mitchell-Olds. For all markers, PCR conditions and amplification product analysis were as previously described in van Rossum *et al.* (2004).

Genetic diversity and overall population structure analysis. The level of polymorphism was estimated for each microsatellite loci and overall as the number of alleles (N_a), the range of allele size and the total gene diversity H_T using SPaGeDi. We also calculated Wright's inbreeding coefficient F_{IS} , providing information on the cumulative action of inbreeding, population substructure and potential null alleles. Overall differentiation among populations was assessed based on allele identity with the statistic F_{ST} or on allele size with R_{ST} using SPaGeDi.

For each population, we calculated observed heterozygosity (H_O) and expected heterozygosity (H_E) using GENETIX 4.01 (Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France), F_{IS} and mean allelic richness (A_{Sn}) were calculated using FSTAT. Deviation of F_{IS} from Hardy-Weinberg equilibrium expectations was assessed performing 7800 randomization. Allelic richness is a measure of the number of alleles independent of sample size, hence allowing comparing this quantity between different sample sizes. Mean allelic richness were estimated per samples as the (non-weighted) average over loci. Deviations in allelic diversity and expected heterozygosity from mutation-drift equilibrium under the Infinite Allele Model (IAM), the Stepwise Mutation Model (SMM), and an intermediate Two-Phase Model (TPM, recommended for microsatellites and consisting of mostly (90%) one-step mutations and 10% of multi-step changes) were tested for each population computing the statistic T_2 with the program BOTTLENECK 1.2.02 (Cornuet & Luikart 1996). T_2 contrasts an average over loci of the deviation of the current gene diversity H_E from the heterozygosity predicted on the basis of the observed number of alleles under the assumption of mutation-drift equilibrium (Cornuet & Luikart 1996). Positive values of T_2 reflect a gene diversity excess possibly caused by recent

founder events, whereas negative values are consistent with recent population expansion without immigration. Significant deviation from equilibrium gene diversity was determined with the Wilcoxon signed rank test, which is the most appropriate test when only few polymorphic loci are analyzed (Piry et al. 1999).

Overall differentiation among populations was assessed based on allele identity with the statistic F_{STn} (for nuclear F_{ST}) or on allele size with R_{ST} using SPaGeDi. The significance of the difference of F_{STn} and R_{ST} values (namely, $R_{ST} > F_{STn}$) was investigated by testing whether the observed R_{ST} was larger than its value obtained after permuting allele sizes among alleles within locus.

Spatial analysis

Isolation by distance. For both datasets, the Mantel correlation coefficients between geographic and genetic distances were compared with random distributions obtained from 10000 Mantel-like permutations of the genetic [$F_{ST}/(1 - F_{ST})$] and geographic [$\ln(\text{geographical distance})$] matrices to test for an overall pattern of isolation by distance using FSTAT.

Identification of genetic boundaries. To test the possibility of reduced gene flow between some populations, we used the computational geometry approach implemented in BARRIER 2.2, which identify the zones where genetic differences between pairs of populations are largest ("boundaries" or "barriers" Manni & Guérard 2004). From the geographic coordinates of samples, a map and a Voronoï tessellation are first constructed. The Voronoï tessellation represents a polygonal neighbourhood for each sample that is constituted of the points, on the plane, that are closer to the location of the sample than to any other sample. Two samples are neighbours if their respective Voronoï polygons have a common edge. Each edge of the tessellation is thus associated with an estimated value of genetic distance between the two populations that it separates. Monmonier's maximum-difference algorithm is then applied to identify barriers. Starting from the edge for which the distance value is highest and proceeding across adjacent edge whose genetic distance is highest. The procedure is continued until the forming barrier has reached either the limits of the map or another pre-existing barrier.

BARRIER allows the computation of barriers with multiple matrices. The thickness of the segments representing boundaries are thus plotted proportionally to the numbers of times each one of them is included in one of the boundaries computed from distinct matrices. We used pairwise F_{ST} matrices computed from both cpDNA and microsatellite data with SPaGeDi to identify boundaries. BARRIER allows computing as many barriers as there are populations. In this case, however, each population will finally be surrounded by barriers. Alternatively, we chose to compute the first barrier from each matrix only, i.e. the most important one (Manni & Guérard 2004).

Cluster analysis.

The application of Bayesian methods to population genetics (Shoemaker *et al.* 1999; Beaumont & Rannala 2004) provided new opportunities to test the effect of the recolonization process on patterns of genetic structure. Combining information from large multilocus dataset into a single probability model, Bayesian approaches enabled to detect hidden substructure through pooling of either combinations of the sampled populations (Corander *et al.* 2003) or sets of individuals independently of the current sample structure (Pritchard *et al.* 2000). We used Bayesian methods implemented in the program STRUCTURE 2.1 (Pritchard *et al.*, 2000) to estimate the most probable number of putative cluster (K ; syn.: "gene pool") that best explains the pattern of genetic variability. STRUCTURE assumes that the marker loci are unlinked and at linkage equilibrium within populations. Linkage disequilibrium was assessed for all pairs of loci in each sampled population with exact test using GENEPOP 3.4 (Raymond & Rousset 1995). We assumed the admixture model in which the fraction of ancestry from each cluster is estimated for each individual. The fraction of ancestry was average over individuals within each sample and plotted onto the map of Europe. The parameter of individual admixture α was chosen to be the same for all clusters and it was given a uniform prior. The user-defined number K of clusters or gene pools was set from 2 to 35. Each K was replicated 10 times for 10^5 iterations after a burn-in period of $5 \cdot 10^4$ without any prior information on the population of origin. According to Evanno *et al.* (Evanno *et al.* 2005), we used ΔK , the second order rate of change of the log likelihood function with respect to K , to choose the most likely number of cluster. The modal value of the distribution of ΔK is expected to be located at the real K and the height of this modal value is used as an indicator of the strength of the signal detected.

The STRUCTURE analyses gave a bimodal-like distribution of ΔK (see above), which should be interpreted as a hierarchical structure of genetic variation (Evanno *et al.* 2005). According to chlorotype composition, geographic proximity and STRUCTURE results, we gathered populations in distinct nested groups and subgroups corresponding to a two-level hierarchical structure (Table2, AnnexeB).

At the upper-most level of structure, the role of phylogenetic distances between chlorotypes in group differentiation was assessed computing the probability that mean pairwise N_{ST} values within and among group exceed mean G_{ST} values using SPaGeDi. Mantel tests between the matrix of pairwise $F_{ST} / (1 - F_{ST})$ ratios (computed in SPaGeDi) and that of the logarithm of geographical distances were performed to test for isolation by distance within and among groups. To allow comparison, mantel test were performed only on the largest distance interval defined by pairs of populations from distinct groups.

At both levels of structure, we determine the proportion of variation among groups and subgroups in an AMOVA based on either molecular distance among chlorotypes or on pairwise F_{ST} (microsatellite data) using ARLEQUIN. For each group, we computed microsatellites F_{ST} and average estimates of A_{Sc} , A_{Sn} , H_O , H_E , F_{IS} and performed comparisons among group through permutation tests using FSTAT. Differences for T_2 TPM were analyzed with Kruskal-Wallis tests in Minitab 13.20 (Minitab Inc. State College, Pennsylvania, USA).

Results

Overall diversity and population structure

CpDNA data. In comparison to Pauwels *et al.* (2005), three additional chlorotypes were identified, thus leading to 14 the total number of chlorotypes, labelled from A to O (Fig. 1). Two of them (chlorotypes M and O) corresponded to expected but missing chlorotypes in Pauwels *et al.* (2005) and were confirmed to be in very low frequencies in the overall sampling (0,32% and 0,08%, respectively). The third one, chlorotype L, derived from chlorotype G (considered to be ancestral, Pauwels *et al.* 2005) by the same mutation that linked chlorotypes A and D. The homoplastic nature of the mutation was confirmed by sequencing (unpublished data). Thus, no new mutation was identified.

Table 1 Allelic diversity of the nuclear microsatellite loci scored. Allele size: in base pairs (bp), number in brackets indicates an odd-sided allele outside the main range of allele size; Na: number of alleles; HT: total gene diversity; FIS: Wright's inbreeding coefficient; FST: population differentiation based on allele frequencies; RST: population differentiation based on allele size.

Locus	Allele sizes	Na	H _T	F _{IS}	F _{ST}	R _{ST}
ATH	157-191(204)	17	0.612	0.139	0.26	0.365
GC16	153-201	16	0.767	0.004	0.321	0.303
Lyr104	127-160(235)	10	0.758	0.062	0.322	0.199
Lyr132	227-238	7	0.57	0.1	0.286	0.158
Lyr133	151-183	19	0.783	0.158	0.344	0.255
Lyr417	173-317	49	0.927	0.123	0.306	0.353
Overall		118	0.736	0.097	0.309	0.370

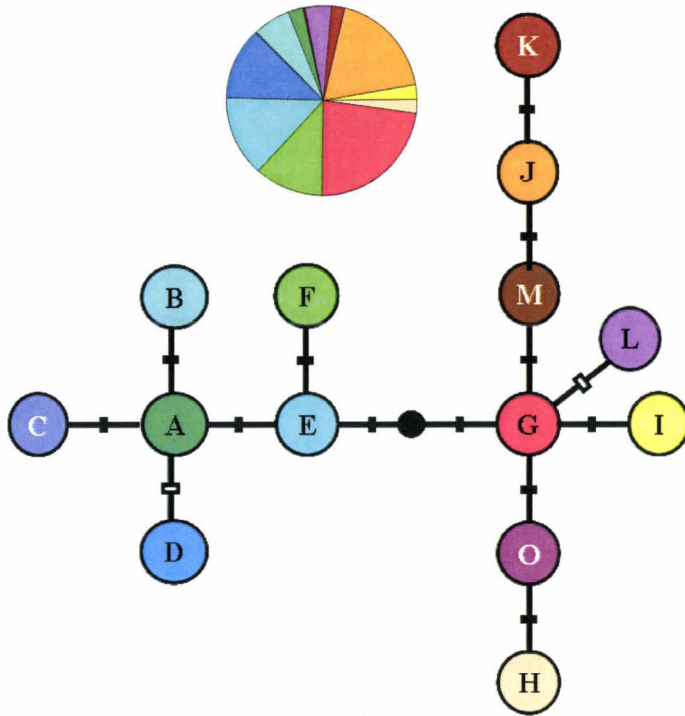


Figure 1 Minimum-spanning tree representing the relationship between chlorotypes and relative frequencies of chlorotypes in the overall sampling. Black circle: missing chlorotype; black rectangle: mutation event; empty rectangle: homoplasy.

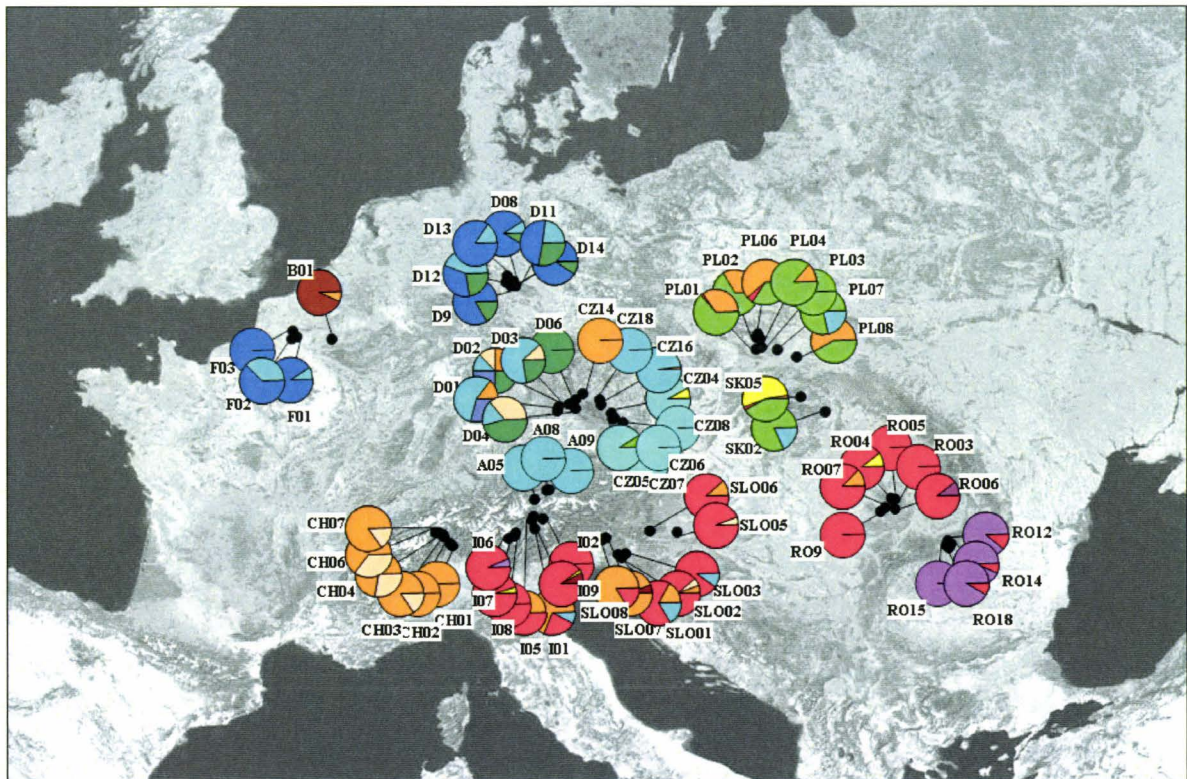


Figure 2 Geographic distribution of *Arabidopsis halleri* chlorotypes

Chlorotypes were not equally frequent overall (Fig. 1 and Fig. 2). As usually expected (Crandall & Templeton 1996), the most widely represented chlorotypes were internal ones (e.g. in order of decreasing importance: chlorotypes G, J and E with 21.66%, 18.47% and 16.24% of individuals, respectively). Chlorotypic richness A_{Sc} varied from 1 to 4.898 across populations and within-population gene diversity H_S varied from 0 to 0.893 (± 0.086) with a high level of overall chlorotype diversity ($H_T = 0.856$, see Annexe C). No correlation was found between sample size and either A_{Sc} or H_S ($R^2 = 9.4 \cdot 10^{-3}$, $P = 0.442$ and $R^2 = 3.69 \cdot 10^{-4}$, $P = 0.879$, respectively).

Estimates of population differentiation revealed a high level of population structure, considering either all samples ($F_{STc} = 0.727^{***}$) or *A. h. halleri* samples only ($F_{STc} = 0.718^{***}$). Although taking into account molecular distances slightly increased the inferred population genetic structure ($N_{ST} = 0.766$ and $N_{ST} = 0.753$, respectively), no significant phylogeographic signal was detected ($P = 0.074$ and $P = 0.101$).

Microsatellite data. All six microsatellite loci scored in this study were highly polymorphic, displaying a high number of alleles observed in the overall sample range (from 7 to 49 alleles per locus, Table 1). When shared, microsatellite markers displayed from 2 to 8 times more polymorphism at the European scale than previously observed at the within-population level (Table 1 in van Rossum *et al.* 2004). Except for ATH and Lyr104, which displayed one allele of particularly large size, the variation of allele sizes seemed almost continuous. Moderate variation was observed over loci in H_T (C.V. = 17.83%), F_{ST} (C.V. = 9.91%) and R_{ST} (C.V. = 30.69%) and a stronger variation in F_{IS} (C.V. = 59.82%).

Among a total of 915 tests for linkage disequilibrium between pairs of loci, 11 were significant at $P \leq 0.01$ and 49 at $P \leq 0.05$. The global tests (Fisher's method) for each pair of loci across samples were all not significant. We concluded that the analyzed loci were sufficiently independent for the application of Bayesian methods for the analysis of population structure.

Mean allelic richness (A_{Sn}), observed (H_O) and expected (H_E) heterozygosity and F_{IS} were calculated for each population (Annexe C). A_S varied from 1.95 to 4.2, H_O from 0.27 to 0.63 and H_E from 0.2 to 0.76. Coefficients of variation of estimators were singularly similar (C.V. \sim 20%). Surprisingly, F_{IS} varied from 0 to 0.33, *i.e.* strongly differed among populations (C.V. = 83.06%), and some populations were significantly inbred. A significant gene diversity excess was detected in most samples (87.5% of samples under IAM, 70.31% under TPM, 62.5% under SMM). However, important variation among samples was observed (C.V. = 19.44%, C.V. = 37.97%, C.V. = 53.04%, respectively). No correlation was found between genetic diversity, F_{IS} or T_2 and sample size (except for T_2 IAM, data not shown).

As expected (Petit *et al.* 2005), population differentiation was lower for nuclear loci than for cpDNA. Differentiation among population taking into account the allele sizes R_{ST} did not differ from differentiation based on allele identities, neither considering all samples ($R_{ST} = 0.37$; $F_{STn} = 0.31$; $P = 0.139$) or *A. h. halleri* samples only ($R_{ST} = 0.36$; $F_{STn} = 0.31$; $P = 0.219$), indicating that stepwiselike mutations did not contributed to overall differentiation among populations. We therefore focused on F_{ST} in remaining analysis.

Spatial analysis.

A significant pattern of isolation by distance was detected overall for both cpDNA ($P = 10^{-5}$) and microsatellite ($P = 10^{-5}$) data.

Several edges of the Voronoï tessellation were included in the first barrier independently computed from both cpDNA and microsatellite pairwise F_{ST} matrices, thus resulting in a so-called "major barrier" in comparison to edges that belong to either barrier. This barrier mainly isolates Austrian populations from most of the surrounding ones (Fig. 3).

Cluster analysis.

The distribution of $\Delta(K)$ as a function of K showed 3 modes at $K = 2$, 15 and 20 (Fig. 4), suggesting a hierarchical structure of genetic variation. Interestingly, $K = 2$ separated neither the two subspecies nor populations from distinct edaphic types but separated *A. halleri* samples according to their latitude in two clusters labelled 1 (51% of individuals) and 2 (49% of individuals, Fig. 5). Samples located south of the latitude 47°N ("southern samples", *i.e.*

Table 2 Genetic variation within and among groups and subgroups of *A. halleri* populations. A_{Sc} : chlorotypic richness; H_S : chlorotype diversity; A_{Sn} : nuclear allelic richness; H_O : observed heterozygosity; H_E : expected heterozygosity; F_{IS} : Wright inbreeding coefficient; T_2 TPM: bottleneck statistic, computed on the basis of an intermediate Two-Phase Model, recommended for microsatellites; F_{STn} : nuclear population differentiation; NWE: North Western Europe; Hz: Harz; BF: Bohemian Forest; Su: Sumava; NLA: Northern Limestone Alps; TaM: Tatra Moutain; Si: Silesia; LA: Lepontine Alps; Do: Dolomites; SAF: Southeastern Alpine Forelands; ApM: Apuseni Moutain; FaM: Fagaras Moutain; N: Northern group; S: Southern group; n.s.: nonsignificant; *: $P \leq 0.05$; **: $P \leq 0.01$; ***: $P \leq 0.001$.

Clusters	Populations	A_{Sc} (\pm SE)	H_S (\pm SE)	F_{STc}	A_{Sn} (\pm SE)	H_O (\pm SE)	H_E (\pm SE)	F_{IS} (\pm SE)	T_2 TPM (\pm SE)	F_{STn}
NWE	F1-3	1,571 (\pm 0,514)	0,215 (\pm 0,243)	0.258	2,814 (\pm 0,035)	0,502 (\pm 0,011)	0,531 (\pm 0,012)	0,055 (\pm 0,01)	2,085 (\pm 0,723)	0.019
Hz	D8-9; D11-14	2,35 (\pm 0,543)	0,4 (\pm 0,214)	0.146	3,254 (\pm 0,311)	0,451 (\pm 0,057)	0,567 (\pm 0,052)	0,208 (\pm 0,101)	1,527 (\pm 0,678)	0.077
BF	D1-4; D6; CZ14	2,63 (\pm 1,469)	0,468 (\pm 0,337)	0.554	3,35 (\pm 0,54)	0,538 (\pm 0,062)	0,615 (\pm 0,075)	0,132 (\pm 0,082)	1,449 (\pm 0,775)	0.146
Su1	CZ4-8;	1,276 (\pm 0,378)	0,06 (\pm 0,083)	0.008	3,423 (\pm 0,185)	0,514 (\pm 0,035)	0,61 (\pm 0,015)	0,161 (\pm 0,061)	1,298 (\pm 0,787)	0.011
Su2	CZ16; CZ18	1,1 (\pm 0,141)	0,018 (\pm 0,025)	-0.057	2,355 (\pm 0,029)	0,409 (\pm 0,011)	0,434 (\pm 0,001)	0,06 (\pm 0,026)	2,515 (\pm 0,593)	0.122
NLA	A5; A8-9	1 (\pm 0,395)	0 (\pm 0,165)	0.000	2,024 (\pm 0,403)	0,31 (\pm 0,049)	0,336 (\pm 0,042)	0,083 (\pm 0,022)	2,34 (\pm 1,551)	0.255
TaM	SK2;SK5	2,21 (\pm 0,291)	0,429 (\pm 0,208)	0.334	3,228 (\pm 0,655)	0,502 (\pm 0,108)	0,575 (\pm 0,115)	0,131 (\pm 0,031)	0,617 (\pm 0,277)	0.096
Si	PL1-4; PL6-8	1,932 (\pm 0,457)	0,372 (\pm 0,194)	0.152	3,123 (\pm 0,508)	0,495 (\pm 0,078)	0,566 (\pm 0,09)	0,127 (\pm 0,077)	2,487 (\pm 0,323)	0.092
N		1,878 (\pm 0,865)	0,28 (\pm 0,265)	0.686	3,045 (\pm 0,556)	0,474 (\pm 0,087)	0,546 (\pm 0,102)	0,132 (\pm 0,081)	1,813 (\pm 0,824)	0.272
	Differences among N subgroups, with/without NLA	*n.s.	*n.s.	n.s. / n.s.	** / n.s.	n.s. / n.s.	** / n.s.	n.s. / n.s.	* / *	n.s. / n.s.
LA	CH1-7	1,641 (\pm 0,497)	0,233 (\pm 0,195)	0.107	2,362 (\pm 0,165)	0,37 (\pm 0,045)	0,418 (\pm 0,044)	0,114 (\pm 0,114)	2,546 (\pm 0,402)	0.166
Do	I1-2; I5-9	1,647 (\pm 0,584)	0,182 (\pm 0,203)	0.517	2,52 (\pm 0,551)	0,405 (\pm 0,1)	0,416 (\pm 0,09)	0,033 (\pm 0,053)	2,787 (\pm 0,295)	0.219
SAF	SLO1-3; SLO5-8	1,964 (\pm 0,501)	0,25 (\pm 0,173)	0.615	3,403 (\pm 0,443)	0,539 (\pm 0,08)	0,572 (\pm 0,076)	0,06 (\pm 0,051)	1,9 (\pm 0,523)	0.132
ApM	RO3-9	1,443 (\pm 0,486)	0,108 (\pm 0,12)	-0.023	3,287 (\pm 0,452)	0,462 (\pm 0,086)	0,54 (\pm 0,095)	0,149 (\pm 0,098)	1,721 (\pm 0,626)	0.097
FaM	RO12; RO14-15; RO18	1,578 (\pm 0,399)	0,131 (\pm 0,094)	-0.036	3,341 (\pm 0,112)	0,555 (\pm 0,04)	0,557 (\pm 0,03)	0,005 (\pm 0,051)	2,433 (\pm 0,372)	0.013
S		1,67 (\pm 0,505)	0,186 (\pm 0,166)	0.690	2,957 (\pm 0,596)	0,461 (\pm 0,102)	0,496 (\pm 0,1)	0,075 (\pm 0,089)	2,284 (\pm 0,601)	0.273
	Differences among S subgroups	n.s.	n.s.	**	***	**	**	*	**	n.s.
Europe	all	1,782 (\pm 0,724)	0,237 (\pm 0,228)	0.727	3,004 (\pm 0,572)	0,468 (\pm 0,094)	0,523 (\pm 0,103)	0,106 (\pm 0,089)	2,027 (\pm 0,764)	0.31
	Differences among groups	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	**	*	n.s.

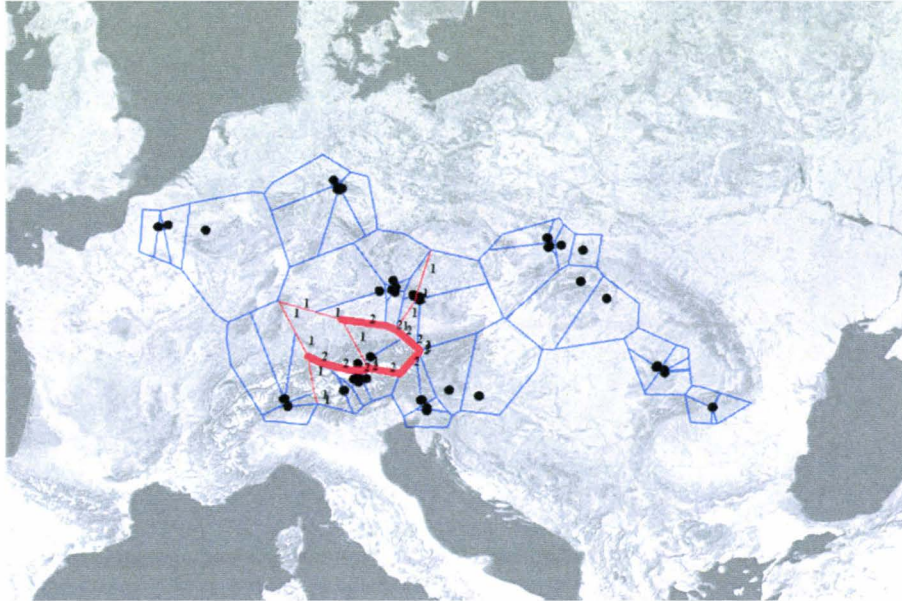


Figure 3 Genetic barriers (red line) obtained with Monmonier's maximum difference algorithm. A score is associated to all the different edges of the Voronoï tessellation (blue lines) that constitute barriers, thus indicating in how many barriers computed from both matrices each one of them is included. The scores are visualized by representing the thickness of each edge proportionally to its score. Black circles represent populations.

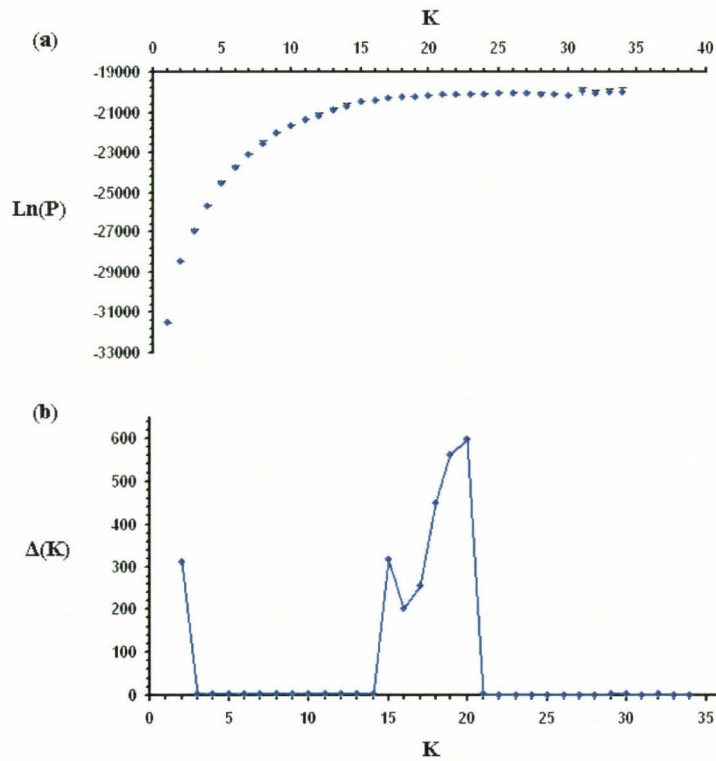


Figure 4 (a): Mean values and standard errors of log likelihood of the multilocus genotypic data, $\ln(P)$, as a function of the number of cluster, K , obtained with STRUCTURE. Mean values and standard errors have been obtained after 10 runs for each K ; (b): corresponding distribution of $\Delta(K)$.

CH1-7; I1-2; I5-9; SLO1-3; SLO5-8; RO3-9; RO12; RO14-15; RO18) mostly belonged to gene pool 1 (87% of individuals). Conversely, samples located north of this latitude ("northern samples", *i.e.* samples B1;F1-3; D8-9; D11-14; D1-4; D6; CZ4-6; CZ7-8; CZ16; CZ18; A5; A8-9; SK2;SK5; PL1-4; PL6-8) mostly belonged to the gene pool 2 (78% of individuals). Beside this overall pattern, an elevated level of admixture was nevertheless observed in Slovenian and Polish samples (respectively 54% and 57% only of individuals in gene pool 1).

Populations that shared a common nuclear gene pool also shared specific lineages of chlorotypes (Fig. 1 and Fig. 2). Southern samples were essentially composed of the lineage including chlorotypes G to O (99% of individuals) whereas northern ones were essentially composed of the lineage including chlorotypes A to F (81% of individuals). Samples were therefore gathered in two groups labelled S (South) and N (North, Table 2) to performed among-group comparisons. AMOVAs revealed that 42% of cpDNA variation and 10% of microsatellite variation was partitioned among groups (d.f. = 1; S.S. = 463.99; V.C. = 0.741; $P < 10^{-5}$ and d.f. = 1; S.S. = 333.562; V.C. = 0.221; $P < 10^{-5}$, respectively). Percentages of variation among populations within a group and among individuals within population were respectively 39% and 19% for cpDNA and 25% and 65% for microsatellites. The comparison of pairwise N_{ST} revealed that they were significantly higher than mean permutation value (G_{ST}) in comparison among groups ($N_{ST} = 0.793$; $G_{ST} = 0.721$; $P = 0.014$) but that they did not differ within N and S groups ($N_{ST} = 0.605$; $G_{ST} = 0.596$; $P = 0.382$ and $N_{ST} = 0.474$; $G_{ST} = 0.483$; $P = 0.66$, respectively), suggesting phylogeographic relevance of the clustering. Mantel tests revealed a significant increase of nuclear $F_{STn} / (1 - F_{STn})$ with the log of geographical distance in the N and S group ($P = 10^{-4}$), but a significant decrease of nuclear $F_{STn} / (1 - F_{STn})$ in among group pairwise comparisons ($P = 1.2 \cdot 10^{-3}$, Fig. 6), suggesting a reduction of gene flow between groups. Finally, groups did not differ from each other in gene diversity indices and level of population differentiation (Table 2). Conversely, they significantly differed in Wright's inbreeding coefficient ($P = 0.005$), and T_2 TPM was significantly higher in group S ($P = 0.016$).

STRUCTURE analysis identified a second level of organization at either $K = 15$ and $K = 20$ (Fig. 4). Gene pools were labelled from 1 to 20. Although ΔK was higher at $K = 20$

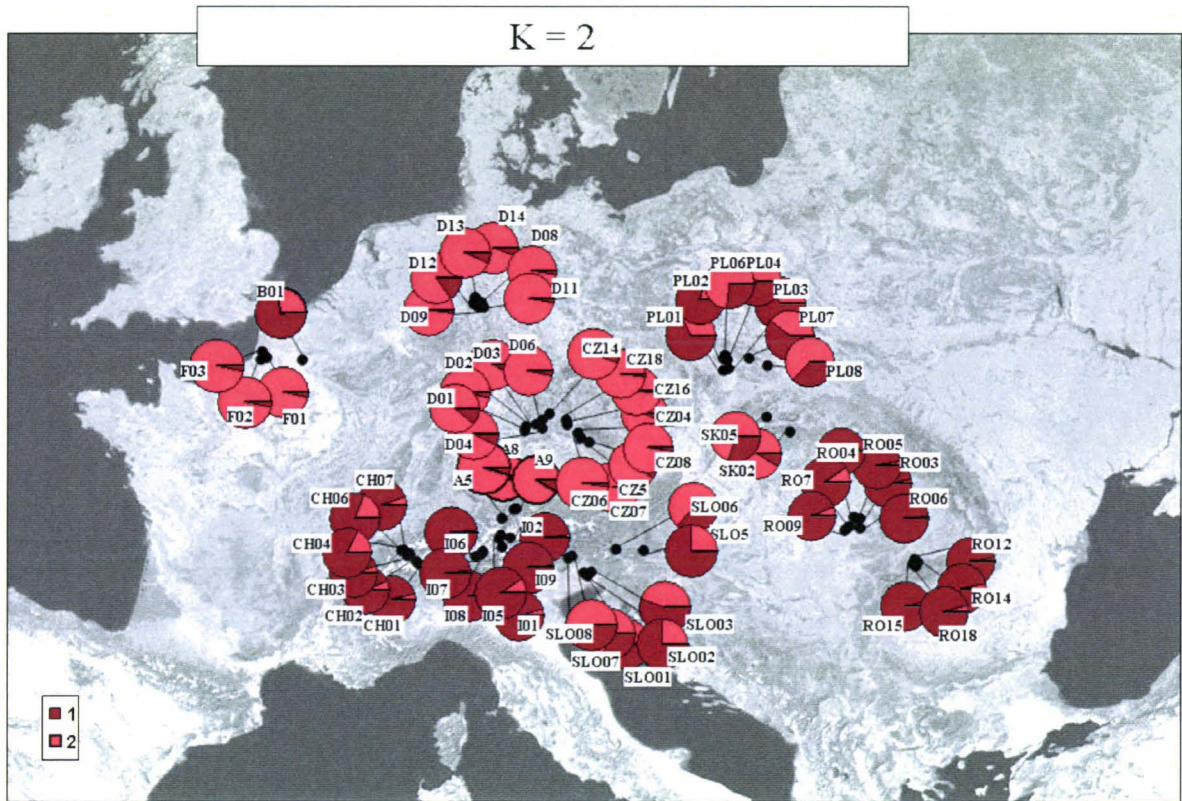


Figure 5 Proportions of ancestry of each population in each of the cluster identified at K=2 in STRUCTURE.

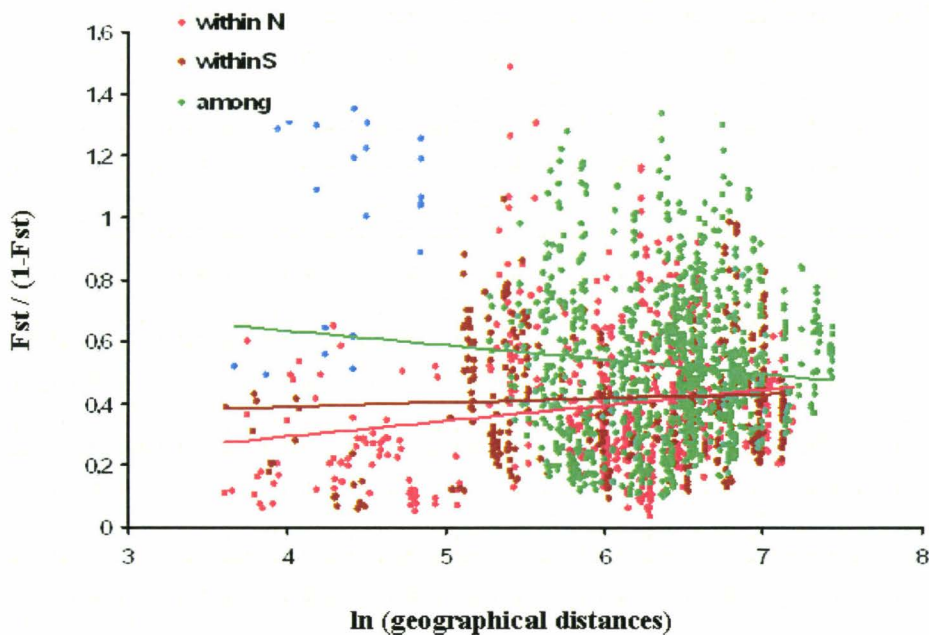


Figure 6 Patterns of isolation by distance within and among groups defined at the upper-most level of population structure. N-N pairs of population are represented in red, S-S pairs in brown. N-S pairs are represented in green, except for pairs involving populations from Austria and Italy represented in blue to emphasize the high level of differentiation compared to within-group comparisons at the same spatial scale.

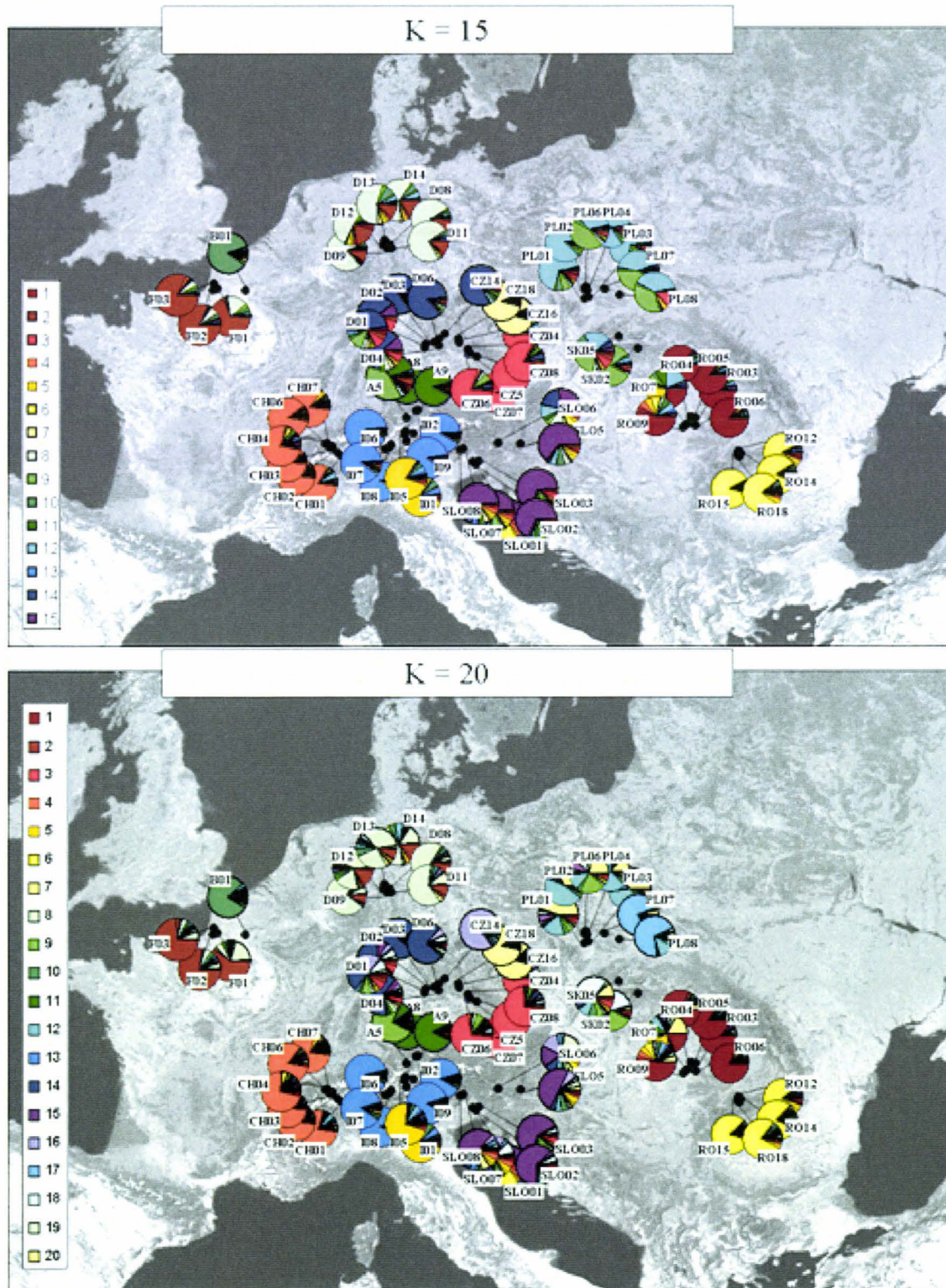


Figure 7 Proportions of ancestry of each population in each of the cluster identified at K = 15 and K = 20 in STRUCTURE.

(Fig. 4), the passage from $K = 15$ to $K = 20$ occurred with only minor changes in the overall population structure (Fig. 7a and 7b). Except a high level of admixture observed in some Polish samples (only 43.4%, 42.2% and 35.6% of individuals from samples PL6, PL7 and PL8 in gene pool 9), most samples belonged to one major gene pool only.

According to the pattern of isolation by distance observed within each group, geographically proximate samples were composed by the same gene pools. Considering these results, but also chlorotype composition and geographic location, we defined 12 distinct subgroups of geographically samples comprising enough samples to allow among-subgroups comparisons (Table 2, Fig. 8). Eight subgroups belonged to the group N ("northern subgroups"): North-Western Europe (NWE), Harz (Hz), Bavarian Forest (BF), Sumava 1 (Su1), Sumava 1 (Su2), Northern Limestone Alps (NLA), Tatras Mountains (TaM) and Silesia (Si); five subgroups belonged to the group S ("southern subgroups"): Lepontine Alps (LA), Dolomites (Do), Southeastern Alpine Forelands (SAF), Apuseni Mountains (ApM), Făgăraș Mountains (FaM).

In the group S, 43% of cpDNA and 17% of microsatellite variation was partitioned among subgroups (AMOVAs; d.f. = 4; S.S. = 125.221; V.C. = 0.299; $P < 10^{-5}$ and d.f. = 4; S.S. = 345.246; V.C. = 0.344; $P < 10^{-5}$, respectively). Significant differences between subgroups were detected in average estimates of genetic diversity at nuclear loci but not in chloroplast genetic diversity and nuclear population differentiation (Table 2). Significant positive correlations were identified between longitude and nuclear allelic richness A_{Sn} , observed heterozygosity H_O and expected heterozygosity H_E (Fig. 8). The evolution of nuclear allelic richness and gene diversity with longitude was confirmed by a significant negative correlation of T_2 TPM with longitude, suggesting recurrent founder events from East to West (Fig. 9).

In the group N, 34% of cpDNA and 16% of microsatellite variation was partitioned among subgroups (AMOVAs; d.f. = 6; S.S. = 323.938; V.C. = 0.421; $P < 10^{-5}$ and d.f. = 6; S.S. = 583.410; V.C. = 0.357; $P < 10^{-5}$, respectively). Subgroups significantly differed from each other in all average genetic diversity indices (Table 2) but not in F_{IS} ($P = 0.318$) nor in F_{ST} ($P = 0.289$). However, observed differences were mostly due to the very low values observed in Austrian populations and statistical tests turned insignificant when these populations were removed from the dataset. Given the presence of M population in this group,

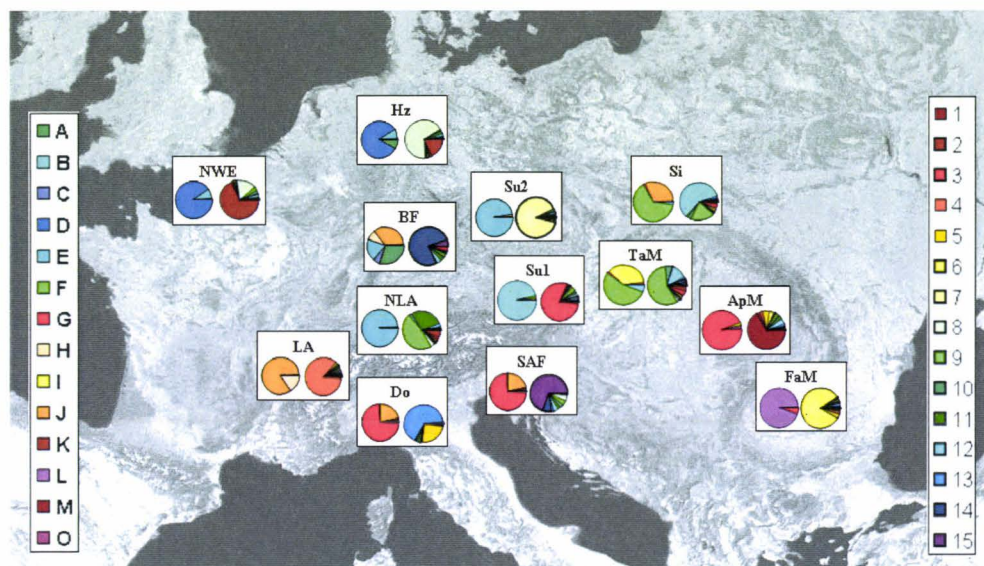
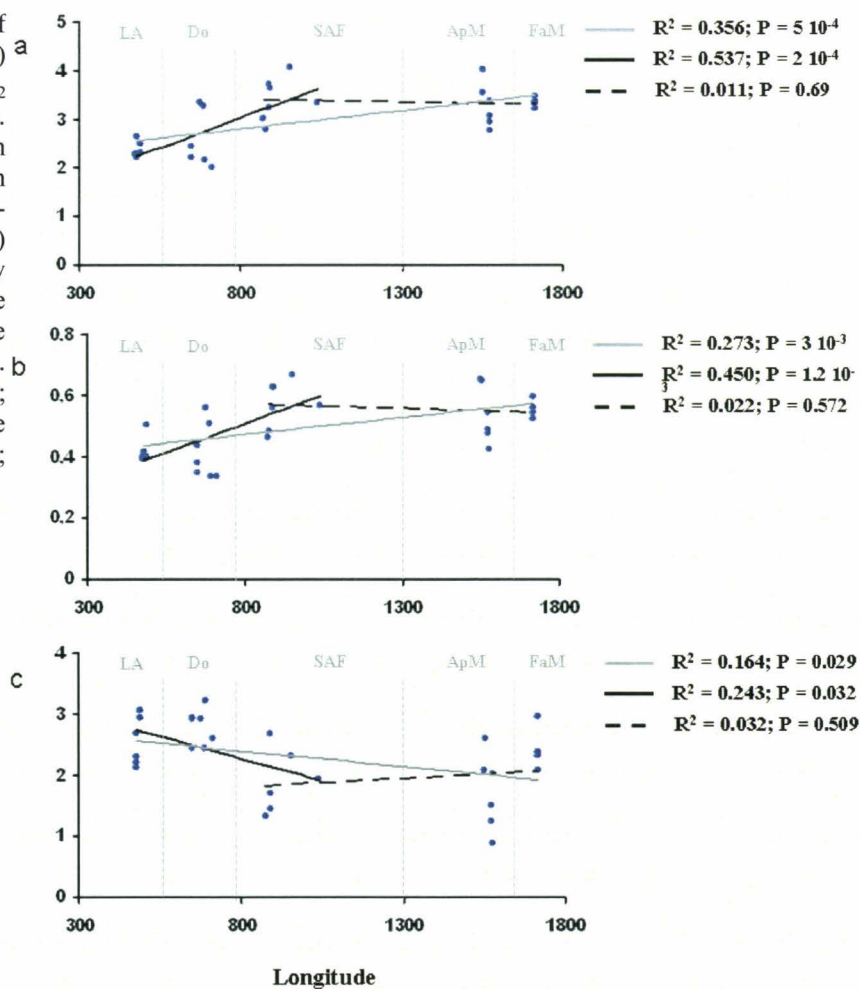


Figure 8 Geographic location of defined subgroups of populations, distribution of chlorotypes A-O in subgroups (left pie charts) and proportions of ancestry of each subgroup in each of the fifteen gene pools (1-15) identified at $K = 15$ (right pie charts). Colours corresponding to distinct chlorotypes and gene pools are indicated. NWE: North Western Europe; Hz: Harz; BF: Bohemian Forest; Su: Sumava; NLA: Northern Limestone Alps; TaM: Tatras Moutain; Si: Silesia; LA: Lepontine Alps; Do: Dolomites; SAF: Southeastern Alpine Forelands; ApM: Apuseni Moutain; FaM: Fagaras Moutain.

Figure 9 Longitudinal gradients of (a) nuclear allelic richness A_{Sn} , (b) expected heterozygosity H_E and (c) T_2 TPM in southern populations. Regression analyses have been performed considering all Southern populations (grey line), South-Western populations only (black line) or South-Eastern populations only (broken black line). Longitudes are expressed as distance from the Greenwich meridian, in kilometres. LA: Lepontine Alps; Do: Dolomites; SAF: South-Eastern Alpine Forelands; ApM: Apuseni Mountains; FaM: Făgăraş Mountains.



in which founder effect could be expected (Pauwels et al. 2005), we also tested for difference in gene diversity estimators and Wright's F statistics between M and NM populations. All tests were not significant. Conversely, there was a tendency towards higher T_2 TPM values in M population ($P = 0.105$) that turned significant when Austrian population were removed from the dataset ($P = 0.036$).

Discussion

Both types of markers revealed a high level of population structure at the European scale ($F_{STc} = 0.727$; $F_{STn} = 0.31$). Estimations were higher than mean values calculated from a meta-analysis of available data for angiosperm species ($0.637 (\pm 0.002)$ for cpDNA markers, $0.184 (\pm 0.002)$ for nuclear markers, Petit *et al.* 2005). However, means values might reflect the bias towards tree and shrub species in plant phylogeographic studies (Comes & Kadereit 1998). Our result was rather typical to short-lived perennial species (mean value for 12 distinct species $F_{STn} = 0.31$, Nybom 2004). They also conformed to expectations for a small herb that produces seeds without a specialized dispersal mechanism. Species with seeds that lack mechanisms for long distance dispersal are indeed generally expected to have higher levels of population differentiation than species that use particular agents for seed dispersal, *i.e.* wind, water or ingestion by animals (Nybom 2004). The low ability of seed dispersal in *A. halleri* has already been inferred from an empirical study of fine-scale spatial genetic structure within a population (van Rossum *et al.* 2004). Finally, among the few available studies using PCR-RFLP cpDNA polymorphism to assess large-scale population structure in herb species, results obtained for the arctic-montane herb *Saxifraga hirculus* were quite similar to ours ($G_{STc} = 0.731$ Oliver *et al.* 2006).

Factors influencing the overall genetic structure of Arabidopsis halleri in Europe

Both the distribution of chlorotypes among populations and the Bayesian analysis of nuclear gene pools supported the existence of a hierarchical pattern of population genetic structure. Although the lower sampling of *A. h. ovirensis* and *A. h. halleri* M populations in comparison to *A. h. halleri* NM populations could affect splitting order (Rosenberg *et al.* 2002), the genetic structure observed was not explained by the partitioning of population among subspecies or edaphic types.

At the upper-most level of genetic structure, two latitudinally separated and longitudinally organized genetic groups were distinguished among which a majority of cpDNA variation and a substantial level of nuclear microsatellite variation were partitioned. Despite the pattern of isolation by distance observed overall for both cpDNA and microsatellite variation, both groups clustered populations that could be geographically more distant than populations belonging to the two major groups (*e.g.* populations CH_i and RO_i are clustered in gene pool 1 although they are more distant to each other than to, for example, populations A_i or CZ_i, Fig. 5). This was confirmed by the absence of spatial pattern of genetic variation in comparisons of population from distinct groups, suggesting the occurrence of reproductive barriers between groups (Hardy & Vekemans 2001). This assumption is supported by the results of the BARRIER analysis which confirmed the occurrence of reduced gene flow between populations from Austria (NLA) and Northern Italy (Do) that was not explained by isolation by distance (Fig. 3). Populations were indeed separated by 40 to 120 km only and genetic differentiation between populations between 5 to more than 10 times higher than pairwise difference involving within-group comparisons at the same spatial scale (Fig. 6). Conversely, the fact that the identified barrier also distinguished Austrian populations from more northern ones fitted the pattern of isolation by distance identified within the N group of populations.

Populations from Austria (NLA) and Northern Italy (Do) were mainly separated by the Alps. We assumed that the Alps Mountains constitute a physical barrier to gene flow for *A. halleri* in Europe and that they might have played and still play a main role in the current observed genetic structure of *A. halleri*, at least in the western part of the European species range. Alps have been mentioned to reduce gene dispersal in many phylogeographic studies (Hewitt 1999). Alps could have opposed an initial barrier to the post-glacial northward expansion of temperate tree and shrub populations located in the Italian peninsula during the glacial ages, and the recolonization of Northern Europe by these species then implicated population expanding from the Iberic and Balkan refugia (Taberlet *et al.* 1998; Hewitt 1999; Hewitt 2000). This pattern established for tree and shrub species implies an east/west division of genetic variation that does not fit our data. It also involves a north-south gradient of genetic divergence between population, allelic richness and intraspecific diversity (Hewitt 1996; Comps *et al.* 2001; Petit *et al.* 2003). Conversely, we detected differences neither in level of

population differentiation nor in average cpDNA or nuclear genetic diversity indices between groups of population north and south of continental Europe.

Considering the distribution of genetic variation, we assumed that the overall genetic structure observed across Europe did not result of a postglacial recolonization from southern refugia but reveal the maintenance of *A. halleri* north of the Alps during the LGM. The distribution of cpDNA chlorotypes further supports the hypothesis of glacial survival of *A. halleri* in central Europe. The strong amount of variation among groups and the presence of distinct cytoplasmic evolutionary lineages within groups indeed generally support the hypothesis of vicariant population differentiation rather than dispersal events (Avice 2000). Thus, the advance of glaciers during the cold periods of the Quaternary might have isolated populations north and south of the Alps, resulting in a genetic divergence of populations. In this context, the high level of admixture observed in Si and SAF populations, in accordance with the presence of haplotypes from distinct lineages within these populations (G and E, J and F, respectively), should be interpreted as the result of secondary contact between differentiated gene pool.

Postglacial history of A. halleri south of the Alps

Populations located south of the Alps showed a clear East-West decreasing gradient of genetic diversity. A decrease of allelic richness, and to some extent, gene diversity is usually interpreted as the consequence of a postglacial recolonization process in phylogeographic studies (Hewitt 1996; Comps *et al.* 2001; Petit *et al.* 2003). Glacial refugia are expected to harbour higher levels of genetic diversity than do areas that have been colonized after the retreat of glaciers (Widmer & Lexer 2001). We assumed that the observed gradients indicated a postglacial recolonization of Southern Alps from an eastern refugium. This is well supported by the increasing of T2 values, i.e. more intense bottleneck signals in the West. South-eastern Alps were not glaciated during the LGM and have already been hypothesized to have acted as potential refugia for several silicicolous subalpine to lower alpine plant species (Tribusch & Schonswetter 2003; Schonswetter *et al.* 2005). Our results did not allow to distinguished populations from North western Balkans (SAF) and Carpathians (ApM and FaM) as putative sources for the colonization process. However, the complete absence of *A. halleri* in Hungary strongly suggested that postglacial recolonization of Southern Alps should have been initiated from SAF populations.

Postglacial history of A. halleri north of the Alps

Although differences between subgroups were conditioned by the low values of NLA populations, it appeared that population from Southern Germany (BF) arrayed a high level of chloroplast and nuclear genetic diversity (Table 2). Given the low sizes of those populations, this was unexpected. Northern-Alpine peripheral refugia have already been inferred to be located in Southern Germany for mountain species from molecular data (Schonswetter et al. 2005). The possibility of persistence between Central and Southern Germany has also been inferred from high level of gene diversity at nuclear microsatellite loci for the closest relative species of *A. halleri*: *A. lyrata* subsp. *petrea* (Clauss & Mitchell-Olds in press). The authors have sustained that the steppe-tundra habitat of central Europe during the Last Glacial Maximum was likely to have been suitable for cold tolerant species. We assume that Southern Germany could also have constituted glacial refugia for *A. halleri*.

The colonization of metal polluted areas in the north of the Alps is an important feature of the recent history of *A. halleri* in northern Europe and could have greatly influence the genetic structure. In particular, foundation of M populations in the last centuries could have caused strong genetic bottlenecks (Bradshaw 1984), thus leading to a reduction of gene diversity in M populations as compared to NM populations (Luikart et al. 1998). However, this was only slightly supported by our analysis. Significant differences in genetic diversity were detected in neither the chloroplast nor the nuclear genome, suggesting that the foundation of M populations was not accompanied by strong founder events. Moreover, although a tendency towards selfing is generally expected in M populations as compared to NM ones (Antonovics 1968; McNeilly & Antonovics 1968; Lefebvre 1970; Cuguen et al. 1989), we detect no difference in Wright's inbreeding coefficient between edaphic types. This generalized the results of van Rossum et al. (van Rossum et al. 2004) who concluded from a fine spatial scale genetic study conducted in a metallicolous population of *A. halleri* that metal exposure did not influence breeding system. Conversely, the BOTTLENECK statistic T_2 was significantly higher in M populations than in NM ones. This might be interpreted as a consequence of the recent origin of M populations in polluted areas and attested that the presence of *A. halleri* is linked to the extension of human activities in these regions. Interestingly, the mean T_2 values were higher in the industrial sites of Silesia and North of France (mean $T_2 = 2.085$ and 2.487 , respectively), where industrial activities only began in the 19th and 20th centuries (Berton 1946; Grodzinska *et al.* 2000), than in the Harz Mountains

($T_2 = 1.527$, $P = 0.015$) where mining is reported since the Middle Ages (Liessmann 1997; Ernst *et al.* 2004).

The sharing of gene pools 9 and 12 at $K = 15$ by Si populations (M) and TaM populations (NM) strongly suggested that, although they occurred on different edaphic types and differed in their level of metal tolerance (Pauwels *et al.* submitted), the level of differentiation at neutral loci between these population is low. It is very likely from our study that geographic distance remains the major factor of population differentiation at the within-group level. However, it would be interesting to perform a STRUCTURE analysis taking into account populations north of the Alps only. Although it is not expected from our results, this could help to identify a clustering of nuclear polymorphism corresponding to the distinct edaphic types that was not detectable in the analysis including all European samples.

Conclusion

Our results from both chloroplast and nuclear markers attested that geographical distances were an important factor of the genetic structure observed in *A. halleri*. However, Alps proved to play a main role, separating two latitudinally separated and longitudinally organized genetic assemblages. These groups might have been separated during long periods of time, likely during the glacial ages, and have diverged through vicariance.

The knowledge of the genetic structure of neutral loci can serve as a background for the detection and analysis of loci of adaptive significance. Our results gave new insights in the choice of accessions that could be compared in the context of evolutionary study of metal tolerance. It appeared clearly that the genome of NM accessions located south of the Alps might differ from northern M accessions for many reasons apart from metal exposure. Conversely, we observed that NM accessions did not differ from M ones, from a neutral point of view, when they are geographically proximate (*e.g.* populations from the Tatras Mountain and from Silesia). Finally, M populations of *A. halleri* have recently been mentioned in Italia (W. H. O. Ernst, pers. com.). It will be interesting to compare genetic mechanisms that have locally evolved in populations that are highly divergent at neutral loci.

References

- Al-Shehbaz IA, O'Kane SLJ (2002) Taxonomy and Phylogeny of *Arabidopsis* (Brassicaceae). In: *The Arabidopsis Book* (eds. Somerville CR, Meyerowitz EM). American Society of Plant Biologist, Rockville.
- Antonovics J (1968) Evolution in closely adjacent plant populations. V. The evolution of self-fertility. *Heredity*, **23**, 219-238.
- Assunção AGL, Schat H, Aarts MGM (2003) *Thlaspi caerulescens*, an attractive model species to study heavy metal hyperaccumulation in plants. *New Phytologist*, **159**, 351-360.
- Avise JC (1994) *Molecular markers, natural history and evolution* Chapman and Hall, New York.
- Avise JC (2000) *Phylogeography: the history and formation of species* Harvard University Press, Cambridge, Massachusetts and London, England.
- Avise JC, Arnold J, Ball RMJ, *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **7**, 489-522.
- Beaumont MA, Rannala B (2004) The Bayesian revolution in genetics. *Nature Reviews Genetics*, **5**, 251-261.
- Becher M, Talke IN, Krall L, Kramer U (2004) Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. *The Plant Journal*, **37**, 251-268.
- Bert V, Macnair MR, de Laguerie P, Saumitou-Laprade P, Petit D (2000) Zinc tolerance and accumulation in metallicolous and non-metallicolous populations of *Arabidopsis halleri* (Brassicaceae). *New Phytologist*, **146**, 225-233.
- Bert V, Meerts P, Saumitou-Laprade P, *et al.* (2003) Genetic basis of Cd tolerance and hyperaccumulation in *Arabidopsis halleri*. *Plant and Soil*, **249**, 9-18.
- Berton A (1946) Presentation de plantes *Arabis Halleri*, *Armeria elongata*, *Oenanthe fluviatilis*, *Galinsoga parviflora discoidea*. *Bulletin de la Société de Botanique de France*, **93**, 139-145.
- Birky C, Jr (1995) Uniparental Inheritance of Mitochondrial and Chloroplast Genes: Mechanisms and Evolution. *PNAS*, **92**, 11331-11338.
- Birky CW, Jr, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration, mutation and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, **121**, 613-627.
- Bradshaw AD (1984) The importance of evolutionary ideas in ecology and vice versa. In: *Evolutionary Ecology* (ed. Shorrocks B), pp. 1-25. Blackwell Sc. Publications, Oxford, U. K.
- Clapham AR, Akeroyd JR (1993) *Cardaminopsis*. In: *Flora Europea* (eds. Tutin TG, Heywood VH, Burges NA, *et al.*), p. 290. Cambridge University Press, Cambridge.
- Clauss MJ, Mitchell-Olds T (in press) Population genetic structure of *Arabidopsis lyrata* in Europe. *Molecular Ecology*.

- Clegg MT, Gaut BS, Learn GH, Jr., Morton BR (1994) Rates and patterns of chloroplast DNA evolution. *Proceeding of the National Academy of Science USA*, **91**, 6795-6801.
- Comes HP, Kadereit JW (1998) The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, **3**, 432-438.
- Comps B, Gomory D, Letouzey J, Thiebaut B, Petit RJ (2001) Diverging Trends Between Heterozygosity and Allelic Richness During Postglacial Colonization in the European Beech. *Genetics*, **157**, 389-397.
- Corander J, Waldmann P, Sillanpaa MJ (2003) Bayesian analysis of genetic differentiation between populations. *Genetics*, **163**, 367-374.
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001-2014.
- Crandall KA, Templeton AR (1996) Applications of intraspecific phylogenetics. In: *New Uses for New Phylogenies* (eds. Harvey PH, Brown AJL, Maynard Smith J, Nee S). Oxford University Press, Oxford, U. K.
- Cruzan MB, Templeton AR (2000) Paleocology and coalescence: phylogeographic analysis of hypotheses from the fossil record. *Trends in Ecology and Evolution*, **15**, 491-496.
- Cuguen J, Acheroy M, Loutfi AL, Vernet P (1989) Breeding system differentiation in *Arrhenatherum elatius* populations: evolution toward selfing? *Evolutionary Trends in Plants*, **2**, 17-24.
- Dahmani-Muller H, van Oort F, Gelie B, Balabane M (2000) Strategies of heavy metal uptake by three plant species growing near a metal smelter. *Environmental Pollution*, **109**, 231-238.
- Dräger DB, Desbrosses-Fonrouge A-G, Krach C, *et al.* (2004) Two genes encoding *Arabidopsis halleri* MTP1 metal transport proteins co-segregate with zinc tolerance and account for high *MTP1* transcript levels. *The Plant Journal*, **39**, 425-439.
- Ennos RA (1994) Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*, **72**, 250-259.
- Ernst WHO (1990) Mine vegetation in Europe. In: *HEAVY METAL TOLERANCE in PLANTS: Evolutionary Aspects* (ed. Shaw AJ), pp. 21-37. CRC Press, Boca Raton, Florida.
- Ernst WHO, Knolle F, Kratz S, Schnug E (2004) Aspects of ecotoxicology of heavy metals in the Harz region – a guided excursion. *Landbauforschung Völkenrode*, **54**, 53-71.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.
- Excoffier L, Smouse PE (1994) Using allele frequencies and geographic subdivision to reconstruct gene tree within a species: molecular variance parsimony. *Genetics*, **136**, 343-359.
- Fabiszewski J (1986) Heavy metal tolerance of *Cardaminopsis halleri* (L.) Hayek populations in the Polish Tatra Mts. *Acta Societatis Botanicorum Poloniae*, **55**, 421-428.
- Feder ME, Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Nature Reviews Genetics*, **4**, 649-655.

- Goldstein DB, Roemer GW, Smith DA, *et al.* (1999) The Use of Microsatellite Variation to Infer Population Structure and Demographic History in a Natural Model System. *Genetics*, **151**, 797-801.
- Goudet J (2001) *Fstat, a Program to Estimate and Test Gene Diversities and Fixation Indices*.
- Grivet D, Petit RJ (2002) Phylogeography of the common ivy (*Hedera* sp.) in Europe: genetic differentiation through space and time. *Molecular Ecology*, **11**, 1351-1362.
- Grodzinska K, Korzeniak U, Szarek-Lukaszewska G, Godzik B (2000) Colonization of zinc mine spoils in southern Poland - preliminary studies on vegetation, seed rain and seed bank. *Fragm. Flor. Geobot.*, **45**, 123-145.
- Hardy OJ, Vekemans X (2001) Patterns of allozyme variation in diploid and tetraploid *Centaurea jacea* at different spatial scales. *Evolution*, **55**, 943-954.
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology*, **2**, 618-620.
- Heuertz M, Hausman JF, Hardy OJ, *et al.* (2004) Nuclear microsatellites reveal contrasting patterns of genetic structure between western and southeastern European populations of the common ash (*Fraxinus excelsior* L.). *Evolution*, **58**, 976-988.
- Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247-276.
- Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87-112.
- Hewitt GM (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907-913.
- King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology*, **7**, 1151-1161.
- Koch M, Haubold B, Mitchell-Olds T (2001) Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. *American Journal of Botany*, **88**, 534-544.
- Lascoux M, Palme AE, Cheddadi R, Latta RG (2004) Impact of Ice Ages on the genetic structure of trees and shrubs. *Philosophical Transactions of the Royal Society of London Serie B*, **359**, 197-207.
- Lefebvre C (1970) Self-fertility in maritime and zinc mine populations of *Armeria maritima* (Mill.) Willd. *Evolution*, **24**, 571-577.
- Lexer C, Fay MF (2005) Adaptation to environmental stress: a rare or frequent driver of speciation? *Journal of Evolutionary Biology*, **18**, 893-900.
- Liessmann W (1997) *Historischer Bergbau Im Harz* Springer-Verlag, Berlin Heidelberg New York.
- Luikart G, Sherwin WB, Steele BM, Allendorf FW (1998) Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology*, **7**, 963-974.
- Luo S-J, Kim J-H, Johnson WE, *et al.* (2004) Phylogeography and Genetic Ancestry of Tigers (*Panthera tigris*). *PLoS Biology*, **2**, e442.

- Macnair MR (1997) The evolution of plants in metal-contaminated environments. In: *Environmental stress, Adaptation and Evolution* (eds. Bijlsma R, Loeschke V), pp. 1-24. Birkhäuser Verlag Basel.
- Manni F, Guérard E (2004) Barrier vs. 2.2. Manual of the user (ed. Population genetics team MoMMdIH). Publication distributed by the authors, Paris.
- McNeilly T, Antonovics J (1968) Evolution in closely adjacent plant population. IV. Barriers to gene flow. *Heredity*, **23**, 205-218.
- Mitchell-Olds T (2001) *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends in Ecology & Evolution*, **16**, 693-700.
- Natoli A, Canadas A, M. PV, *et al.* (0) Phylogeography and alpha taxonomy of the common dolphin (*Delphinus sp.*). *Journal of Evolutionary Biology*, **0**.
- Nei M (1987) *Molecular Evolutionary Genetics* Columbia University Press, New York.
- Nordborg M, Innan H (2002) Molecular population genetics. *Current Opinion in Plant Biology*, **5**, 69-73.
- Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology*, **13**, 1143-1155.
- O' Kane SL, Al-Shehbaz IA (1997) A synopsis of *Arabidopsis* (Brassicaceae). *Novon*, **7**, 323-327.
- Oliver C, Hollingsworth PM, Gornall RJ (2006) Chloroplast DNA phylogeography of the arctic-montane species *Saxifraga hirculus* (Saxifragaceae). *Heredity*, **96**, 222-231.
- Pauwels M, Frérot H, Bonnin I, Saumitou-Laprade P (submitted) A broad-scale analysis of population differentiation for Zn-tolerance in an emerging model species for tolerance study: *A. halleri* (Brassicaceae). *Journal of Evolutionary Biology*.
- Pauwels M, Saumitou-Laprade P, Holl A-C, Petit D, Bonnin I (2005) Multiple origin of metal-tolerant populations of the pseudometallophyte *Arabidopsis halleri* (Brassicaceae) in Central Europe: the cpDNA testimony. *Molecular Ecology*, **14**, 4403-4414.
- Petit RJ, Aguinagalde I, de Beaulieu J-L, *et al.* (2003) Glacial Refugia: Hotspots But Not Melting Pots of Genetic Diversity. *Science*, **300**, 1563-1565.
- Petit RJ, Duminil J, Fineschi S, *et al.* (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology*, **14**, 689-701.
- Petit RJ, El Mousadik A, Pons O (1998) Identifying Populations for Conservation on the Basis of Genetic Markers. *Conservation Biology*, **12**, 844-855.
- Petit RJ, Kremer A, Wagner DB (1993) Finite island model for organelle and nuclear genes in plant. *Heredity*, **71**, 630-641.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248-249.
- Rendell S, Ennos RA (2002) Chloroplast DNA diversity in *Calluna vulgaris* (heather) populations in Europe. *Molecular Ecology*, **11**, 69-78.

- Rosenberg NA, Pritchard JK, Weber JL, *et al.* (2002) Genetic structure of human populations. *Science*, **298**, 2381-2385.
- Ruzzante DE (1998) A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Canadian Journal of Fish and Aquatic Science*, **55**, 1-14.
- Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology*, **7**, 465-474.
- Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetics data analysis. Ver 2.000, Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.
- Schonswetter P, Stehlik I, Holderegger R, Tribsch A (2005) Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, **14**, 3547-3555.
- Shoemaker JS, Painter IS, Weir BS (1999) Bayesian statistics in genetics: a guide for the uninitiated. *Trends in Genetics*, **15**, 354-358.
- Soltis DE, Soltis PS, Milligan BG (1992) Intraspecific Chloroplast DNA Variation: Systematic and Phylogenetic Implications. In: *Molecular systematics of plants* (eds. Soltis PS, Soltis DE, Doyle JJ), pp. 117-150. Chapman and Hall, New York.
- Song B-H, Clauss MJ, Pepper A, Mitchell-Olds T (2006) Geographic patterns of microsatellite variation in *Boechera stricta*, a close relative of *Arabidopsis*. *Molecular Ecology*, **15**, 357-369.
- Staton JL, Schizas NV, Chandler GT, Coull BC, Quattro JM (2001) Ecotoxicology and population genetics: the emergence of "phylogeographic and evolutionary ecotoxicology". *Ecotoxicology*, **10**, 217-222.
- Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671-688.
- Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453-464.
- Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381-397.
- Tribsch A, Schonswetter P (2003) Patterns of endemism and comparative phylogeography confirm palaeoenvironmental evidence for Pleistocene refugia in the Eastern Alps. *Taxon*, **52**, 477-497.
- van Rossum F, Bonnin I, Fenart S, *et al.* (2004) Spatial genetic structure within a metallicolous population of *Arabidopsis halleri*, a clonal, self-incompatible and heavy-metal-tolerant species. *Molecular Ecology*, **13**, 2959-2967.
- Weber M, Harada E, Vess C, Roepenack-Lahaye E, Clemens S (2004) Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identifies nicotianamine synthase, a ZIP transporter and other genes as potential metal hyperaccumulation factors. *The Plant Journal*, **37**, 269-281.
- Widmer A, Lexer C (2001) Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology & Evolution*, **16**, 267-269.

Willems G, Godé C, Dräger DB, *et al.* (submitted) Quantitative Trait Loci mapping of zinc tolerance in the metallophyte *Arabidopsis halleri* ssp. *halleri*. *Genetics*.

Wright SI, Gaut BS (2005) Molecular Population Genetics and the Search for Adaptive Evolution in Plants. *Molecular Biology and Evolution*, **22**, 506-519.

Annexe A Locations of sampled populations of *A. halleri* and sample sizes for cpDNA (n_c) and nuclear microsatellite (n_n) analysis. *A. h. halleri*: *Arabidopsis halleri* subsp. *halleri*; *A. h. ovirensis*: *Arabidopsis halleri* subsp. *ovirensis*.

Population	Subspecies	Origin	GPS coordinates		Altitude (m)	n_c	n_n
			North	East			
A05	<i>A. h. halleri</i>	Northern Tyrol (Austria)	47°13'778	11°22'779	807	14	14
A08	<i>A. h. halleri</i>	Northern Tyrol (Austria)	47°25'177	11°51'957	522	45	55
A09	<i>A. h. halleri</i>	Northern Tyrol (Austria)	47°25'243	11°51'895	522	20	21
B01	<i>A. h. halleri</i>	Hautes Fâgnes (Belgium)	50°29'63"	4°60'00"	160	29	32
CH01	<i>A. h. halleri</i>	Tessin (Switzerland)	45°59'233	08°50'387	230	16	26
CH02	<i>A. h. halleri</i>	Tessin (Switzerland)	45°59'814	08°50'244	520	19	20
CH03	<i>A. h. halleri</i>	Tessin (Switzerland)	46°10'632	08°43'182	290	19	20
CH04	<i>A. h. halleri</i>	Tessin (Switzerland)	46°10'685	08°42'261	317	11	12
CH06	<i>A. h. halleri</i>	Tessin (Switzerland)	46°10'134	08°40'502	475	20	20
CH07	<i>A. h. halleri</i>	Tessin (Switzerland)	46°10'134	08°40'528	475	13	14
CZ04	<i>A. h. halleri</i>	Sumava (Czech Republic)	49°02'872	13°45'841	1067	14	17
CZ05	<i>A. h. halleri</i>	Sumava (Czech Republic)	48°59'258	13°46'390	352	24	25
CZ06	<i>A. h. halleri</i>	Sumava (Czech Republic)	48°59'	13°46'	362	12	13
CZ07	<i>A. h. halleri</i>	Sumava (Czech Republic)	48°57'	13°48'	430	8	9
CZ08	<i>A. h. halleri</i>	Sumava (Czech Republic)	48°57'	13°48'	430	26	27
CZ14	<i>A. h. halleri</i>	Sumava (Czech Republic)	49°28'37	12°42'92	897	20	22
CZ16	<i>A. h. halleri</i>	Sumava (Czech Republic)	49°03'35	13°33'31	1084	57	62
CZ18	<i>A. h. halleri</i>	Sumava (Czech Republic)	49°05'70	13°32'18	1156	9	12
D01	<i>A. h. halleri</i>	Bohemian forest (Germany)	49°10'64	12°09'88	340	7	10
D02	<i>A. h. halleri</i>	Bohemian forest (Germany)	49°11'31	12°09'52	348	8	10
D03	<i>A. h. halleri</i>	Bohemian forest (Germany)	49°13'10	12°39'66	360	9	10
D04	<i>A. h. halleri</i>	Bohemian forest (Germany)	49°09'85	12°47'75	590	11	12
D06	<i>A. h. halleri</i>	Bohemian forest (Germany)	49°17'42	12°49'78	380	10	10
D08	<i>A. h. halleri</i>	Harz moutains (Germany)	51°53'79	10°29'04	190	12	15
D09	<i>A. h. halleri</i>	Harz moutains (Germany)	51°53'46	10°25'11	190	18	20
D11	<i>A. h. halleri</i>	Harz moutains (Germany)	51°51'27	10°21'95	674	18	20
D12	<i>A. h. halleri</i>	Harz moutains (Germany)	51°51'90	10°17'90	673	18	20
D13	<i>A. h. halleri</i>	Harz moutains (Germany)	51°55'22	10°18'50	237	20	21
D14	<i>A. h. halleri</i>	Harz moutains (Germany)	52°06'21	10°06'93	67	11	11
F01	<i>A. h. halleri</i>	North (France)	50°24'215	3°05'345	52	23	23
F02	<i>A. h. halleri</i>	North (France)	50°25'	3°01'	66	43	48
F03	<i>A. h. halleri</i>	North (France)	50°29'647	3°27'387	36	37	40
I01	<i>A. h. halleri</i>	Trentin (Italy)	46°52'783	11°24'308	967	26	27
I02	<i>A. h. halleri</i>	Trentin (Italy)	46°49'920	11°43'823	840	12	15
I05	<i>A. h. halleri</i>	Trentin (Italy)	46°48'516	11°14'738	680	39	49
I06	<i>A. h. halleri</i>	Trentin (Italy)	46°29'500	10°52'984	1386	18	19
I07	<i>A. h. halleri</i>	Trentin (Italy)	46°29'623	10°53'349	1388	24	24

Annexe A fin.

Population	Subspecies	Origin	GPS coordinates		Altitude (m)	n _c	n _n
			North	East			
I08	<i>A. h. halleri</i>	Trentin (Italy)	46°29'731	10°53'739	1363	31	33
I09	<i>A. h. halleri</i>	Trentin (Italy)	46°43'882	11°25'958	1530	41	50
PL01	<i>A. h. halleri</i>	Silesia (Poland)	50°14'80	18°57'04	272	63	70
PL02	<i>A. h. halleri</i>	Silesia (Poland)	50°29'68	18°56'67	268	21	21
PL03	<i>A. h. halleri</i>	Silesia (Poland)	50°29'50	18°57'58	262	15	15
PL04	<i>A. h. halleri</i>	Silesia (Poland)	50°29'98	18°55'79	269	21	25
PL06	<i>A. h. halleri</i>	Silesia (Poland)	50°16'95	19°01'52	306	10	10
PL07	<i>A. h. halleri</i>	Silesia (Poland)	50°17'014	19°29'051	335	19	20
PL08	<i>A. h. halleri</i>	Silesia (Poland)	50°06'143	20°21'563	190	12	13
RO03	<i>A. h. halleri</i>	Carpathians (Romania)	46°43'000	23°02'621	1003	7	8
RO04	<i>A. h. halleri</i>	Carpathians (Romania)	46°43'891	23°02'291	1018	10	20
RO05	<i>A. h. halleri</i>	Carpathians (Romania)	46°43'305	23°02'637	990	8	9
RO06	<i>A. h. halleri</i>	Carpathians (Romania)	46°39'083	23°02'364	1154	10	10
RO07	<i>A. h. halleri</i>	Carpathians (Romania)	46°51'799	22°48'474	559	8	9
RO09	<i>A. h. halleri</i>	Carpathians (Romania)	46°49'475	22°45'635	666	12	13
SK02	<i>A. h. halleri</i>	Tatras (Slovakia)	48°46'17	21°07'81	690	22	23
SK05	<i>A. h. halleri</i>	High Tatras (Slovakia)	49°16'98	20°09'24	1027	46	53
SLO01	<i>A. h. halleri</i>	Slovenia	46°03'281	14°02'989	950	6	7
SLO02	<i>A. h. halleri</i>	Slovenia	46°00'022	14°02'994	960	19	20
SLO03	<i>A. h. halleri</i>	Slovenia	46°04'942	14°03'789	540	20	20
SLO05	<i>A. h. halleri</i>	Slovenia	46°22'366	15°59'756	210	20	20
SLO06	<i>A. h. halleri</i>	Slovenia	46°32'569	14°54'175	465	18	20
SLO07	<i>A. h. halleri</i>	Slovenia	46°16'980	13°53'366	538	18	20
SLO08	<i>A. h. halleri</i>	Slovenia	46°16'713	13°49'924	530	6	7
RO12	<i>A. h. ovirensis</i>	Fagaras Mountains (Romania)	45°36'246	24°37'162	2020	9	10
RO14	<i>A. h. ovirensis</i>	Fagaras Mountains (Romania)	45°36'139	24°37'064	2050	15	23
RO15	<i>A. h. ovirensis</i>	Fagaras Mountains (Romania)	45°36'139	24°37'064	2050	17	17
RO18	<i>A. h. ovirensis</i>	Fagaras Mountains (Romania)	45°36'163	24°37'336	2230	12	12

Annexe B Distribution of chlorotypes among populations, subgroups and groups of populations of *Arabidopsis halleri*

Chlorotype Population	A	B	C	D	E	F	G	H	I	J	K	L	M	O	total
A5					14										14
A8					45										45
A9					20										20
B1										2	27				29
CH1										16					16
CH2										19					19
CH3								3		16					19
CH4								3		8					11
CH6								7		13					20
CH7								2		11					13
CZ4		13							1						14
CZ5		22				2									24
CZ6		12													12
CZ7		8													8
CZ8		26													26
CZ14										20					20
CZ16					56					1					57
CZ18					9										9
D1			2		4					1					7
D2	2		2		1			1		2					8
D3	2				6			1							9
D4	5				2			4							11
D6	10														10
D8	1			10	1										12
D9	3			15											18
D11	5			9	4										18
D12	4			6	8										18
D13				17	3										20
D14	1			10											11
F1				21	2										23
F2				27	16										43
F3				37											37
I1					2		6			18					26
I2							12								12
I5							13			26					39
I6							17					1			18
I7							23		1						24
I8							31								31
I9							37		1				3		41
PL1						40	1			22					63
PL2						14				7					21
PL3						13				2					15
PL4						7	1			13					21
PL6						10									10
PL7					4	15									19
PL8						5				7					12
RO3							7								7
RO4							9		1						10

Chapitre 3: Population genetic structure of *A. halleri*

Annexe B fin

Chlorotype Population	A	B	C	D	E	F	G	H	I	J	K	L	M	O	total
RO5							8								8
RO6							9							1	10
RO7							7			1					8
RO9							12								12
RO12							1					8			9
RO14							1					14			15
RO15												17			17
RO18							1					11			12
SK2					4	18									22
SK5						19	1		25	1					46
SLO1					1		4			1					6
SLO2							17	1		1					19
SLO3					2		18								20
SLO5							19	1							20
SLO6							16			2					18
SLO7										17			1		18
SLO8							1			5					6
total	33	81	4	152	204	143	272	23	29	232	27	51	4	1	1256
% total	2.63	6.45	0.32	12.10	16.24	11.39	21.66	1.83	2.31	18.47	2.15	4.06	0.32	0.08	100

N	33	81	4	152	199	143	3	6	26	76					723
NWE				85	18										103
Hz	14			67	16										97
BF	19		4		13			6		23					65
Su1		81				2			1						84
Su2					65					1					66
NLA					79										79
TaM					4	37	1		25	1					68
Si					4	104	2			51					161

S					5		269	17	3	154		51	4	1	504
LA								15		83					98
Do					2		139		2	44		1	3		191
SAF					3		75	2		26			1		107
ApM							52		1	1				1	55
FaM							3					50			53

Annexe C Genetic diversity indices, F_{IS} and T_2 values for each sampled population.

Group	Subgroup	Population	A_{Sc}	$H_S (\pm SE)$	A_{Sn}	H_0	$H_R (\pm SE)$	F_{IS}	T_2 IAM	T_2 TPM	T_2 SMM
N	NLA	A05	1	0	1.956	0.274	0.319 (± 0.223)	0.148 n.s.	3.207*	2.51*	2.34*
N	NLA	A08	1	0	2.066	0.331	0.337 (± 0.287)	0.018 n.s.	3.598*	2.674*	2.378*
N	NLA	A09	1	0	2.050	0.325	0.353 (± 0.313)	0.082 n.s.	2.509*	1.837*	1.641*
N	Su1	CZ04*	1.683	0.143 (± 0.119)	3.614	0.486	0.610 (± 0.220)	0.209 ***	1.777ns	0.267ns	-0.358ns
N	Su1	CZ05*	1.697	0.159 (± 0.094)	3.636	0.507	0.632 (± 0.211)	0.203 ***	3.322*	2.432*	2.14*
N	Su1	CZ06*	1	0	3.287	0.498	0.607 (± 0.160)	0.186 *	2.194*	1.172ns	0.831ns
N	Su1	CZ07	1	0	3.294	0.574	0.608 (± 0.215)	0.061 n.s.	2.04*	1.069ns	0.765ns
N	Su1	CZ08*	1	0	3.285	0.507	0.591 (± 0.222)	0.146 **	2.748*	1.549ns	1.05ns
N	BF	CZ14*	1	0	2.688	0.371	0.473 (± 0.144)	0.221 ***	2.166ns	0.986ns	0.633ns
N	Su2	CZ16*	1.2	0.035 (± 0.034)	2.376	0.402	0.435 (± 0.182)	0.078 *	3.84*	2.934*	2.712*
N	Su2	CZ18*	1	0	2.335	0.417	0.433 (± 0.235)	0.041 n.s.	2.859*	2.095*	1.897*
N	BF	D01*	2.989	0.667 (± 0.16)	4.201	0.617	0.761 (± 0.075)	0.199 **	1.504ns	0.31ns	-0.179ns
N	BF	D02*	4.898	0.893 (± 0.086)	3.787	0.578	0.661 (± 0.245)	0.133 n.s.	2.846*	2.149*	1.964*
N	BF	D03*	2.897	0.556 (± 0.165)	2.967	0.465	0.562 (± 0.221)	0.181 *	2.734*	1.949*	1.689*
N	BF	D04*	2.994	0.691 (± 0.086)	3.514	0.617	0.624 (± 0.204)	0.014 n.s.	2.842*	2.055*	1.785*
N	BF	D06*	1	0	2.943	0.583	0.607 (± 0.194)	0.043 n.s.	2.233*	1.244ns	0.876ns
N	H _z	D08*	2.522	0.318 (± 0.164)	3.024	0.376	0.518 (± 0.216)	0.282 ***	2.722*	1.884*	1.58*
N	H _z	D09*	1.931	0.294 (± 0.119)	2.828	0.433	0.516 (± 0.232)	0.166 **	1.983ns	0.589ns	0.071ns
N	H _z	D11*	2.968	0.66 (± 0.07)	3.138	0.486	0.544 (± 0.36)	0.11 *	2.939*	2.085*	1.825*
N	H _z	D12*	2.973	0.68 (± 0.055)	3.564	0.543	0.593 (± 0.219)	0.087 n.s.	2.55*	1.444ns	1.056ns
N	H _z	D13*	1.902	0.268 (± 0.113)	3.355	0.430	0.579 (± 0.235)	0.263 ***	3.206*	2.277*	1.915ns
N	H _z	D14	1.805	0.182 (± 0.144)	3.616	0.439	0.652 (± 0.174)	0.338 ***	1.992ns	0.884ns	0.427ns
N	NWE	F01*	1.716	0.1660.097)	2.823	0.496	0.531 (± 0.276)	0.067 n.s.	3.486*	2.373*	1.958*
N	NWE	F02*	1.998	0.478 (± 0.041)	2.776	0.495	0.519 (± 0.274)	0.047 n.s.	2.633*	1.262ns	0.703ns
N	NWE	F03*	1	0	2.844	0.515	0.542 (± 0.261)	0.052 n.s.	3.444*	2.62*	2.367*
N	SI	PL01*	2.178	0.482 (± 0.039)	3.642	0.568	0.642 (± 0.127)	0.118 ***	3.906*	2.935*	2.554*

Annexe C suite

Group	Subgroup	Population	$A_{S\pm}$	$H_S (\pm SE)$	A_{S_1}	H_D	$H_E (\pm SE)$	F_{IS}	T_2 IAM	T_2 TPM	T_2 SMM
N	SI	PL02*	1.997	0.467 (± 0.075)	2.787	0.584	0.586 (± 0.102)	0.005 n.s.	3.027*	2.063*	1.68ns
N	SI	PL03*	1.888	0.247 (± 0.131)	2.963	0.489	0.585 (± 0.106)	0.17 *	3.261*	2.571*	2.355*
N	SI	PL04*	2.492	0.529 (± 0.079)	2.660	0.469	0.518 (± 0.149)	0.096 n.s.	3.631*	2.854*	2.599*
N	SI	PL06*	1	0	4.021	0.541	0.699 (± 0.128)	0.237 **	2.959*	2.194*	1.98*
N	SI	PL07	1.968	0.351 (± 0.111)	2.988	0.460	0.498 (± 0.305)	0.078 n.s.	3.167*	2.36*	2.142*
N	SI	PL08	2	0.53 (± 0.076)	2.804	0.355	0.432 (± 0.213)	0.185 *	3.172*	2.43*	2.16*
N	TaM	SK02*	1.93	0.312 (± 0.106)	2.943	0.467	0.545 (± 0.213)	0.146 *	2.564*	1.713ns	1.457ns
N	TaM	SK05*	2.489	0.545 (± 0.037)	3.513	0.536	0.605 (± 0.191)	0.115 ***	1.994ns	-0.481ns	-2.085ns
S	LA	CH01	1	0	2.479	0.423	0.506 (± 0.183)	0.167 **	3.751*	2.938*	2.715*
S	LA	CH02	1	0	2.304	0.405	0.399 (± 0.312)	-0.015 n.s.	3.729*	3.056*	2.872*
S	LA	CH03	1.917	0.281 (± 0.116)	2.218	0.394	0.393 (± 0.317)	-0.001 n.s.	3.036*	2.196*	1.96 ns
S	LA	CH04	1.997	0.356 (± 0.159)	2.643	0.305	0.416 (± 0.292)	0.277 **	2.849*	2.111*	1.897*
S	LA	CH06	1.998	0.48 (± 0.072)	2.277	0.333	0.402 (± 0.266)	0.179 *	3.119*	2.295*	2.07*
S	LA	CH07	1.933	0.282 (± 0.142)	2.253	0.362	0.391 (± 0.276)	0.079 n.s.	3.373*	2.682*	2.53*
S	Do	I01	2.635	0.48 (± 0.094)	3.270	0.529	0.507 (± 0.316)	-0.042 n.s.	3.292*	2.442*	2.175*
S	Do	I02	1	0	2.022	0.307	0.338 (± 0.214)	0.096 n.s.	3.244*	2.602*	2.439*
S	Do	I05	1.995	0.456 (± 0.053)	3.342	0.536	0.562 (± 0.306)	0.047 n.s.	3.852*	2.927*	2.647*
S	Do	I06	1.562	0.111 (± 0.096)	2.205	0.328	0.349 (± 0.231)	0.064 n.s.	3.185*	2.447ns	2.223ns
S	Do	I07	1.441	0.083 (± 0.075)	2.209	0.380	0.381 (± 0.270)	0.003 n.s.	3.553*	2.925*	2.785*
S	Do	I08	1	0	2.436	0.444	0.436 (± 0.197)	-0.018 n.s.	3.627*	2.945*	2.763*
S	Do	I09	1.898	0.142 (± 0.071)	2.156	0.308	0.336 (± 0.217)	0.084 *	3.909*	3.222*	3.095*
S	ApM	RO03	1	0	3.358	0.464	0.544 (± 0.262)	0.158 n.s.	1.742ns	0.885ns	0.643ns
S	ApM	RO04	1.853	0.2 (± 0.154)	3.072	0.418	0.489 (± 0.329)	0.149 *	2.453*	1.243ns	0.776ns
S	ApM	RO05	1	0	2.951	0.343	0.478 (± 0.206)	0.297 **	2.178*	1.499ns	1.33ns
S	ApM	RO06	1.853	0.2 (± 0.154)	2.775	0.428	0.425 (± 0.399)	-0.008 n.s.	2.603*	2.026*	1.888ns
S	ApM	RO07	1.95	0.25 (± 0.18)	3.554	0.574	0.650 (± 0.215)	0.124 n.s.	3.234*	2.594*	2.421*

Annexe C end

Group	Subgroup	Population	A_{S_c}	$H_S (\pm SE)$	A_{S_n}	H_D	$H_T (\pm SE)$	F_{IT}	T_2 IAM	T_2 TPM	T_2 SMM
S	ApM	RO09	1	0	4.011	0.545	0.654 (± 0.215)	0.173 **	2.851*	2.08*	1.872*
S	FaM	RO12	1.902	0.222 (± 0.166)	3.373	0.533	0.560 (± 0.225)	0.051 n.s.	2.87*	2.081*	1.828*
S	FaM	RO14	1.648	0.133 (± 0.112)	3.209	0.522	0.544 (± 0.200)	0.042 n.s.	3.259*	2.375*	2.081*
S	FaM	RO15	1	0	3.307	0.554	0.526 (± 0.183)	-0.055 n.s.	3.646*	2.958*	2.789*
S	FaM	RO18	1.761	0.167 (± 0.134)	3.475	0.611	0.596 (± 0.155)	-0.027 n.s.	3.064*	2.319*	2.063*
S	SAF	SLO01	3	0.6 (± 0.215)	3.230	0.571	0.560 (± 0.349)	-0.021 n.s.	3.221*	2.676*	2.557*
S	SAF	SLO02	2.075	0.205 (± 0.119)	3.717	0.629	0.627 (± 0.229)	-0.001 n.s.	2.595*	1.445ns	0.996ns
S	SAF	SLO03	1.776	0.189 (± 0.108)	3.654	0.579	0.629 (± 0.231)	0.083 n.s.	2.758*	1.703*	1.296ns
S	SAF	SLO05	1.515	0.1 (± 0.088)	3.336	0.523	0.568 (± 0.191)	0.082 n.s.	2.872*	1.94*	1.683*
S	SAF	SLO06	1.82	0.209 (± 0.116)	4.078	0.607	0.668 (± 0.204)	0.093 *	3.178*	2.319*	2.043*
S	SAF	SLO07	1.562	0.111	2.793	0.428	0.484 (± 0.272)	0.118 n.s.	2.257ns	1.317ns	0.979ns
S	SAF	SLO08	2	0.333 (± 0.215)	3.012	0.436	0.465 (± 0.241)	0.068 n.s.	-	-	-
-	-	B01*	1.615	0.133 (± 0.081)	2.287	0.352	0.385 (± 0.301)	0.088 n.s.	3.091*	2.18*	1.895*

Discussion générale et perspectives

L'intérêt des scientifiques, toutes disciplines confondues, pour *A. halleri* est seulement récent. Il est lié à deux résultats. Tout d'abord, l'identification de capacités élevées de tolérance et d'hyperaccumulation de Zn et de Cd chez l'espèce (Bert *et al.* 2000; Dahmani-Muller *et al.* 2000; Dahmani-Muller *et al.* 2001; Bert *et al.* 2002; Bert *et al.* 2003). Ensuite, et presque simultanément, le repositionnement taxonomique de *Cardaminopsis halleri* dans le genre *Arabidopsis* à l'aide à la fois de critères morphologiques (Al-Shehbaz & O'Kane 2002) et phylogénétiques (Koch *et al.* 2000; Koch *et al.* 2001; O'Kane & Al-Shehbaz 2003). L'intérêt de la possible application de l'ensemble des outils moléculaires disponibles sur *A. thaliana* sur une espèce possédant un caractère à la fois physiologiquement exceptionnel (Clemens 2001) et régulièrement présenté comme un exemple d'adaptation à un stress environnemental (Bradshaw 1970; Antonovics *et al.* 1971; Baker 1987) est vite devenu évident (Mitchell-Olds 2001; Lexer & Fay 2005).

L'espèce a donc aujourd'hui acquis le statut de modèle pour l'étude de la physiologie et l'évolution de la tolérance aux métaux chez les végétaux supérieurs. Cependant, à la différence d'*A. thaliana*, sur laquelle la somme de connaissances disponibles est symbolisée par *The Arabidopsis Book*, ou d'autres métallophytes telles que *Silene vulgaris* et *Mimulus guttatus*, utilisées dans d'autres études que celles intéressées par la tolérance aux métaux (e.g. McCauley *et al.* 2003), ou *Thlaspi caerulescens*, modèles d'études depuis plus de vingt ans (Allen & Sheppard 1971; Ingrouille & Smirnov 1986; Macnair, 1981 #746), les connaissances biologiques disponibles sur *A. halleri* sont maigres. Fondé sur ce constat, l'objectif de cette thèse était d'aborder *A. halleri* à l'échelle de son aire de répartition pour donner une image plus précise de la distribution des capacités de tolérance aux métaux de l'espèce et d'offrir un contexte historique et de génétique neutre aux études génétiques à venir.

I Amplitude écologique d'*A. halleri*

A. halleri a récemment été décrite comme typiquement trouvée sur des sites métallifères, bien qu'incluant des populations sur sols non métallifères (Fig. 9, espèce C, Pollard *et al.* 2002), ou encore comme métallophyte absolue (Dahmani-Muller *et al.* 2000). L'effort d'échantillonnage effectué au cours de la thèse a clairement établi qu'*A. halleri* se

comportait plutôt comme une pseudométallophyte classique, *i.e.* dont la majorité des populations se développent sur sites non pollués (c.f. espèce B, Fig. 9). La plupart des populations échantillonnées se développaient en effet dans des régions montagneuses, loin de toute pollution métallique.

II Origine de la tolérance au zinc d'*A. halleri*

Les résultats des tests effectués en hydroponie sur plus de trente populations réparties dans l'aire de répartition de l'espèce ont **clairement attesté que la tolérance au Zn d'*A. halleri* était constitutive**, *i.e.* présente chez tous les individus de l'espèce, quelle que soit leur position géographique, mais aussi, et surtout, quelle que soit leur régime édaphique. Avec *Thlaspi caerulescens* (Meerts & van Isacker 1997; Escarré *et al.* 2000), *A. halleri* appartient aux quelques rares espèces pseudométallophytes chez lesquelles la tolérance aux métaux n'est pas associée aux milieux métallifères. Chez la plupart des pseudométallophytes, la tolérance est en effet écotypique, *i.e.* rencontrée surtout dans les populations M et seulement à faible fréquence dans les populations NM.

La confirmation d'une tolérance constitutive a définitivement écarté de ce travail la perspective d'identifier l'origine du caractère. Elle permet néanmoins d'élaborer des hypothèses. Le scénario le plus parcimonieux est celui d'**une acquisition du caractère très tôt dans l'histoire de l'espèce**. La présence de la tolérance dans les deux ensembles phylogéographiques identifiés lors de l'analyse de l'organisation de la diversité génétique neutre suggère que son acquisition est antérieure à la séparation de ces groupes durant les épisodes glaciaires, dont le dernier a pris fin il y a 18 000 ans.

L'étude des autres espèces présentant une tolérance aux métaux constitutive, c'est-à-dire les métallophytes absolues endémiques des sols riches en métaux, suggère que l'adaptation aux sites pollués et l'acquisition du caractère de tolérance s'accompagne parfois d'un isolement reproducteur entre populations NM et M, d'abord conspécifiques, et résulte alors en un événement de spéciation écologique (Kruckeberg 1986; Macnair & Gardner 1998). La divergence d'*A. halleri* avec *A. lyrata*, espèce sœur non tolérante et non hyperaccumulatrice, est postérieure à la séparation d'avec *A. thaliana*, il y a environ 5 millions

d'années (Koch *et al.* 2000). La phylogéographie a révélé que les oscillations climatiques du Pléistocène avaient, de façon générale, entraîné des réductions des aires de répartition des espèces et l'isolement de populations dans des zones refuges aux conditions environnementales variables. Un rôle de cet isolement dans la formation de nouvelles espèces a été suggéré (Bennett 2004; Kadereit *et al.* 2004; Willis & Niklas 2004). L'étude phylogéographique comparée d'*A. halleri* et *A. lyrata* ainsi qu'une datation précise de la séparation des deux espèces devraient permettre de discuter **l'hypothèse de l'isolement de populations de l'espèce ancestrale dans des refuges glaciaires naturellement pollués par les métaux** et de son impact sur la séparation des deux espèces. La possibilité de tels refuges est soutenue par la présence d'*A. halleri* et *A. lyrata* dans des régions montagneuses, où l'on s'attend à trouver des affleurements de roches naturellement riches en métaux.

La différence majeure entre *A. halleri* et les espèces endémiques réside néanmoins dans sa large distribution sur sols non pollués, qu'il s'agirait d'expliquer. L'hypothèse rivale pour l'origine d'une tolérance aux métaux constitutive suggère que **la tolérance pourrait être apparue avec le développement des capacités d'hyperaccumulation** des métaux dans les tissus, sous l'effet de pressions évolutives différentes de la pression métallique (Boyd 1998; Boyd & Martens 1998; Boyd 2004). Aucune pression de sélection rendant compte de l'hyperaccumulation n'a cependant pour le moment été clairement mise en évidence (Assunção *et al.* 2003; Frérot *et al.* 2005; Noret *et al.* 2005) et l'hypothèse reste donc encore très spéculative.

Le test séquentiel de tolérance a révélé, chez quelques populations, des individus dont le niveau de tolérance était particulièrement faible. Le protocole du test n'avait pas été écrit avec la volonté d'obtenir une estimation fine du polymorphisme intra-population pour la tolérance ou d'identifier des individus mutants non tolérants (en comparaison, l'identification de géotypes tolérants dans des populations NM de pseudométallophytes a nécessité de suivre 20 000 semis pendant de semaines à dose d'exposition fixe, Tableau 5, d'après Ingram, 1988, cité dans Baker & Proctor 1990). Les résultats suggèrent tout de même que des géotypes (mutants?) pas ou peu tolérants pourraient exister. Les études de physiologie moléculaires menées actuellement sur *A. halleri* (Becher *et al.* 2004; Dräger *et al.* 2004; Weber *et al.* 2004) engagent la comparaison de profils d'expression de gènes avec *A. thaliana*, soit un génome

qui diverge de celui d'*A. halleri* depuis environ 5 millions d'années. De même, les approches QTL se fondent sur un croisement interspécifique avec *A. lyrata* (Willems 2006) alors que des barrières reproductives existent, suggérant un bruit potentiel dans les patrons de ségrégation observés. Dans ce contexte, **l'existence de génotypes *A. halleri* non tolérants permettrait des comparaisons et croisements intraspécifiques** et des études destinées à l'identification de tels génotypes sont souhaitables.

III Evolution de la tolérance au zinc chez *A. halleri*

Des différences significatives de capacités moyennes de tolérance au Zn entre populations ont été détectées, suggérant un important polymorphisme quantitatif pour la tolérance au sein de l'espèce. Ce polymorphisme existe aussi à l'intérieur des populations notamment des populations NM. Un tel polymorphisme est généralement associé à des mécanismes génétiques secondaires aux gènes principaux (gènes modificateurs, Macnair 1993; Macnair *et al.* 2000). Etant donné les résultats récents d'analyse de l'architecture génétique de la tolérance, oscillant entre 1 gène (Macnair *et al.* 1999) et 3 régions QTL (Willems *et al.* soumis) à effets majeurs, il est fort possible que le polymorphisme observé fasse effectivement appel à des gènes secondaires. Si le ou les gènes majeurs de la tolérance au Zn d'*A. halleri* sont fixés, le polymorphisme quantitatif observé suggère au contraire que les mécanismes secondaires présentent eux-mêmes un polymorphisme importants. **Un tel polymorphisme atteste de l'existence d'un potentiel évolutif** vers plus ou moins de tolérance.

Des différences de tolérance moyenne significatives ont été notées entre populations NM, suggérant la présence de mécanismes secondaires de tolérance dans certaines de ces populations. Ces différences n'ont pu être liées ni au niveau de métal dans les sols, ni à aucune variable géographique. Le mode d'évolution de la tolérance en l'absence de pression métallique reste donc obscur. Si l'estimation du coût physiologique à la tolérance aux métaux chez *A. halleri* nécessitera des études délicates, rendues difficiles par l'impossibilité d'effectuer des comparaisons intraspécifiques (*e.g.* transplantations réciproques), il apparaît déjà néanmoins que la **tolérance constitutive n'est pas un handicap majeur au développement de l'espèce sur sol non pollué.**

L'évolution de la tolérance en milieu métallifère est plus intelligible. Des capacités de tolérance moyenne significativement plus élevées ont été enregistrées dans les populations M, attestant qu'en dépit de la constitutivité de la tolérance aux métaux, les populations M s'adaptent aux conditions métallifères et se différencient des populations NM. Il est donc notable pour les études de physiologie moléculaire que **les mécanismes de tolérance engagés dans les populations M peuvent différer de ceux présents au sein des populations NM.**

La différenciation phénotypique des populations M n'a pu être rapportée ni à une histoire commune de ces populations, ni à un isolement génétique avec les populations NM. L'analyse de la structure spatiale de la généalogie chloroplastique a montré que les populations M ne constituaient pas un groupe homogène de populations aux relations généalogiques étroites mais rassemblaient au contraire des populations fondées indépendamment dans les différentes régions polluées du Nord de l'aire de répartition. Ce résultat a été confirmé par l'analyse de la structure génétique de microsatellites nucléaires. Celle-ci n'a pas rassemblé les populations en fonction de leur affinité édaphique mais selon un patron général d'isolement par la distance, tout au moins dans l'ensemble de populations situées au Nord des Alpes. **L'évolution indépendante des populations M pose la question de l'identité des mécanismes secondaires localement sélectionnés pour augmenter les capacités de tolérance.** Etant donné le polymorphisme intraspécifique pour ces mécanismes, des différences sont en effet envisageables. En particulier, des populations M ont récemment été mentionnées en Italie (W. H. O. Ernst, com. pers.). Nous savons aujourd'hui grâce à l'analyse phylogéographique du génome chloroplastique que ces populations appartiennent à un groupe phylogéographique historiquement isolé de celui des populations M situées dans le groupe nord et utilisées jusqu'à présent. Les données microsatellites ont confirmé la divergence des génomes nucléaires des groupes nord et sud. Si les populations M italiennes montrent elles aussi des capacités accrues de tolérance, il sera intéressant de déterminer dans quelle mesure les mécanismes responsables diffèrent de ceux des populations du nord. Pour cela, des études QTL menées sur des descendances de croisements entre accessions M et NM du même groupe phylogéographique, devraient permettre de mettre en évidence des régions génomiques distinctes entre les populations M du nord et celles du sud.

Le corollaire de l'histoire indépendante des grands groupes de populations M est une **proximité génétique plus importante entre populations géographiquement proches qu'entre populations appartenant à un même type édaphique**. Dans la rare situation où les populations M et NM sont relativement proches géographiquement dans notre échantillonnage (*i.e.* populations M de Silésie et NM de Slovaquie), les données de génotypages à des locus neutres chloroplastiques et nucléaires ne permettent en effet pas de mettre en évidence de différenciation entre populations. Les résultats chloroplastiques témoignent de l'histoire phylogéographique commune des populations de Silésie et de Slovaquie et les données microsatellites identifient des pools génétiques communs largement partagés par ces populations. Ces résultats sont une donnée importante dans l'étude des interactions locales entre populations NM et M et des mécanismes évolutifs associés à la colonisation des sols pollués. Tout d'abord, ils soulignent **l'importance d'une étude beaucoup plus fine de la structure locale des populations et d'une estimation plus précise de l'intensité des flux de gènes entre populations M et NM**. En comparant un plus grand nombre de populations dans une même région, une telle étude sera plus à même de distinguer les mécanismes microévolutifs permettant le maintien local des différences phénotypiques observées entre types édaphiques sans différenciation des populations aux locus neutres. Ensuite, ces résultats attestent que les différences phénotypiques entre populations M et NM ne sont pas associées à une différenciation à des locus neutres. Il est donc théoriquement possible de mettre en évidence des régions du génome impliquées dans la divergence adaptative entre populations NM et M, en comparant les niveaux de différenciation des populations à divers locus répartis sur tout le génome. Une **approche de scan génomique** constituerait ainsi une analyse indépendante des analyses QTL, et permettrait d'identifier parmi les nombreux gènes candidats, ceux qui sont véritablement sélectionnés pour conférer des capacités accrues de tolérance aux métaux (Storz 2005).

IV Références

- Allen WR, Sheppard PM (1971) Copper tolerance in some Californian populations of the monkey flower, *Mimulus guttatus*. *Proceedings of the Royal Society of London, Series B Biological Sciences*, **177**, 177-196.
- Al-Shehbaz IA, O'Kane SLJ (2002) Taxonomy and Phylogeny of *Arabidopsis* (Brassicaceae). In: *The Arabidopsis Book* (eds. Somerville CR, Meyerowitz EM). American Society of Plant Biologist, Rockville.

- Antonovics J, Bradshaw AD, Turner RG (1971) Heavy metal tolerance in plants. *Advances in Ecological Research*, **7**, 1-85.
- Assunção AGL, Ten Bookum WM, Nelissen HJM, *et al.* (2003) A cosegregation analysis of zinc (Zn) accumulation and Zn tolerance in the Zn hyperaccumulator *Thlaspi caerulescens*. *New Phytologist*, **159**, 383-390.
- Baker AJM (1987) Metal tolerance. *New Phytologist*, **106(suppl.)**, 93-111.
- Baker AJM, Proctor J (1990) The influence of cadmium, copper, lead, and zinc on the distribution and evolution of metallophytes in the British Isles. *Plant Systematics and Evolution*, **173**, 91-108.
- Becher M, Talke IN, Krall L, Kramer U (2004) Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. *The Plant Journal*, **37**, 251-268.
- Bennett KD (2004) Continuing the debate on the role of Quaternary environmental change for macroevolution. *Philosophical Transactions of the Royal Society of London Serie B*, **359**, 295-303.
- Bert V, Bonnin I, Saumitou-Laprade P, de Laguerie P, Petit D (2002) Do *Arabidopsis halleri* from nonmetallicolous populations accumulate zinc and cadmium more effectively than those from metallicolous populations? *New Phytologist*, **155**, 47-57.
- Bert V, Macnair MR, de Laguerie P, Saumitou-Laprade P, Petit D (2000) Zinc tolerance and accumulation in metallicolous and non-metallicolous populations of *Arabidopsis halleri* (Brassicaceae). *New Phytologist*, **146**, 225-233.
- Bert V, Meerts P, Saumitou-Laprade P, *et al.* (2003) Genetic basis of Cd tolerance and hyperaccumulation in *Arabidopsis halleri*. *Plant and Soil*, **249**, 9-18.
- Boyd RS (1998) Hyperaccumulation as a plant defense strategy. In: *Plants that hyperaccumulate heavy metals* (ed. Brooks RR), pp. 181-201. CAB international, Oxford.
- Boyd RS (2004) Ecology of metal hyperaccumulation. *New Phytologist*, **162**, 563-567.
- Boyd RS, Martens SN (1998) The significance of metal hyperaccumulation for biotic interactions. *Chemoecology*, **8**, 1-7.
- Bradshaw AD (1970) Plants and industrial waste. *Trans. Bot. Soc. Edinb.*, **41**, 71-64.
- Clemens S (2001) Molecular mechanisms of plant metal tolerance and homeostasis. *Planta*, **212**, 475-486.
- Dahmani-Muller H, van Oort F, Balabane M (2001) Metal extraction by *Arabidopsis halleri* grown on an unpolluted soil amended with various metal-bearing solids: a pot experiment. *Environmental Pollution*, **114**, 77-84.
- Dahmani-Muller H, van Oort F, Gelie B, Balabane M (2000) Strategies of heavy metal uptake by three plant species growing near a metal smelter. *Environmental Pollution*, **109**, 231-238.
- Dräger DB, Desbrosses-Fonrouge A-G, Krach C, *et al.* (2004) Two genes encoding *Arabidopsis halleri* MTP1 metal transport proteins co-segregate with zinc tolerance and account for high MTP1 transcript levels. *The Plant Journal*, **39**, 425-439.

- Escarré J, Lefèbvre C, Gruber W, *et al.* (2000) Zinc and cadmium hyperaccumulation by *Thlaspi caerulescens* from metalliferous and nonmetalliferous sites in the mediterranean area: implication for phytoremediation. *New Phytologist*, **145**, 429-437.
- Frérot H, Lefèbvre C, Petit C, *et al.* (2005) Zinc tolerance and hyperaccumulation in F1 and F2 offspring from intra and interecotype crosses of *Thlaspi caerulescens*. *New Phytologist*, **165**, 111-119.
- Ingrouille MJ, Smirnoff N (1986) *Thlaspi caerulescens* J. & C. presl. (*T. alpestre*) in Britain. *New Phytologist*, **102**, 219-233.
- Kadereit JW, Griebeler EM, Comes HP (2004) Quaternary diversification in European alpine plants: pattern and process. *Philos Trans R Soc Lond B Biol Sci*, **359**, 265-274.
- Koch M, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of Chalcone Synthase and Alcohol Dehydrogenase loci in *Arabidopsis*, *Arabis* and related genera (Brassicaceae). *Molecular Biology and Evolution*, **17**, 1483-1498.
- Koch M, Haubold B, Mitchell-Olds T (2001) Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. *American Journal of Botany*, **88**, 534-544.
- Kruckeberg AR (1986) An essay: the stimulus of unusual geologies for plant speciation. *Systematic Botany*, **11**, 455-463.
- Lexer C, Fay MF (2005) Adaptation to environmental stress: a rare or frequent driver of speciation? *Journal of Evolutionary Biology*, **18**, 893-900.
- Macnair M, Gardner M (1998) The Evolution of Edaphic Endemics. In: *Endless Forms, Species and Speciation* (eds. Howard DJ, Berlocher SH), pp. 157-171. Oxford University Press, New York.
- Macnair MR (1993) Tansley review N°49: the genetic of metal tolerance in vascular plants. *New Phytologist*, 541-559.
- Macnair MR, Bert V, Huitson SB, Saumitou-Laprade P, Petit D (1999) Zinc tolerance and hyperaccumulation are genetically independent characters. *Proceeding of the Royal Society of London, Biological Serie*, **266**, 2175-2179.
- Macnair MR, Tilstone GH, Smith SE (2000) The Genetics of Metal Tolerance and Accumulation in Higher Plants. In: *Phytoremediation of Contaminated Soil and Water* (eds. Terry N, Banuelos G, Vangronsveld J), pp. 235-250. CRC Press, Boca Raton.
- McCauley DE, Smith RA, Lisenby JD, Hsieh C (2003) The hierarchical spatial distribution of chloroplast DNA polymorphism across the introduced range of *Silene vulgaris*. *Molecular Ecology*, **12**, 3227-3235.
- Meerts P, van Isacker N (1997) Heavy metal tolerance and accumulation in metallicolous and nonmetallicolous populations of *Thlaspi caerulescens* from continental Europe. *Plant Ecology*, **133**, 221-231.
- Mitchell-Olds T (2001) *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends in Ecology & Evolution*, **16**, 693-700.
- Noret N, Meerts P, Tolra R, *et al.* (2005) Palatability of *Thlaspi caerulescens* for snails: influence of zinc and glucosinolates. *New Phytologist*, **165**, 763-772.

- O'Kane SL, Al-Shehbaz IA (2003) Phylogenetic position and generic limits of *Arabidopsis* (Brassicaceae) based on sequences of nuclear ribosomal DNA. *Annals of the Missouri Botanical Garden*, **90**, 603-612.
- Pollard AJ, Powell KD, Harper FA, Smith JAC (2002) The genetic basis of metal hyperaccumulation in plants. *Critical Reviews in Plant Sciences*, **21**, 539-566.
- Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671-688.
- Weber M, Harada E, Vess C, Roepenack-Lahaye E, Clemens S (2004) Comparative microarray analysis of *Arabidopsis thaliana* and *Arabidopsis halleri* roots identifies nicotianamine synthase, a ZIP transporter and other genes as potential metal hyperaccumulation factors. *The Plant Journal*, **37**, 269-281.
- Willems G (2006) *Characterisation of zinc and cadmium tolerance in Arabidopsis halleri*, Université de Sciences et Technologies de Lille 1 - Université Libre de Belgique.
- Willems G, Godé C, Dräger DB, *et al.* (submitted) Quantitative Trait Loci mapping of zinc tolerance in the metallophyte *Arabidopsis halleri* ssp. *halleri*. *Genetics*.
- Willis KJ, Niklas KJ (2004) The role of Quaternary environmental change in plant macroevolution: the exception or the rule? *Philos Trans R Soc Lond B Biol Sci*, **359**, 159-172; discussion 172.

Annexes

Numéro de la mutation *	Couple amorce/enzyme utilisé en PCR-RFLP*	Type de polymorphisme associé*	Polymorphisme de séquence associé	Codage*	séquence de l'amorce SNP	Nom de l'amorce	polymorphisme SNP attendu
1	K1K2 HpaII	restriction	ATCAGG ATCTGG	1 9	AAAATAAGGTTGTACCGATC	SNP_tmK_4133/4152	A T
2	K1K2 AcsI	restriction	AGGTAT AGGAAT	1 9	ACTGGTATGCCCAAAAAGG	SNP_tmK_3295/3314	T A
3	K1K2 tru9I	restriction	TTGAAT TTTAAT	1 9	TGTAAATTATAAGATTTTT	SNP_tmK_3179/3198	G T
4	CS Ahi I	restriction	TAGC CACTT TAGCTACTT	1 9	ACTTCCCACCTAGC	SNP_CS_33984/33998	C T
5	CS HinfI	restriction	AATA CATT AATA GATT	1 9	AGCCAATAAAGATAAAAATA	SNP_CS_35278/35297	C G
9	K1K2 HinfI	restriction	TTTTAT TTTGAT	1 9	AAAGACGAAGTATAT	SNP_tmK_2271/2289	T G
10	K1K2 HinfI	taille	ATATCTTATTCTTATTG A-----TATTCTTATTC	1 2	ATCTTGTTTGITATTCTAT	SNP_tmK_3774/3755	A T
11	K1K2 tru9I	taille	CTAAA - T CTAAAAT	2 1	ATTATTCATGATTGACTAAA	SNP_tmK_2055/2074	T A

* Pauwels M, Saumitou-Laprade P, Holl A-C, Petit D, Bormin I (2005) Multiple origin of metalicolous populations of the pseudometallophyte *Arabidopsis halleri* (*Brassicaceae*) in Central Europe: the cpDNA testimony. *Molecular Ecology*, 14, 4403-4414.

Les fragments K1K2 et CS ont été amplifiés séparément en suivant le protocole décrit dans Pauwels *et al.* (2005). La purification des produits PCR s'est effectué par digestion enzymatique dans un volume réactionnel total de 15µL (4,5µL H₂O ; 4,5µL de CS ; 6µL K1K2) auquel ont été ajoutés 2 U de SAP et 2 U Exo. L'incubation a été effectuée à 37°C pendant 1H30 ; puis à 75°C durant 15 minutes, et stocké à 4°C. Le génotypage SNP a été effectué à l'aide du kit ABI Prism® SNaPshot™ Multiplex. Une deuxième purification a été effectuée en ajoutant 1 U de SAP au mélange réactionnel, à 37°C pendant 1 heure, 75°C pendant 15 minutes, puis stockage à 4°C. Le génotypage a été effectué par électrophorèse sur ABI Prism 3100 Genetic Analyzer. (0,5µl de produit SNaPshot, 9µl de Hi-Di formamide du kit SNaPshot™ Multiplex et 0,25µl de GeneScan-120 LIZ size standard du kit SNaPshot™ Multiplex). Dénaturation préalable de 5 minutes à 95°C. La migration a été effectuée sur un capillaire de 36 cm avec le tampon de migration POP-4 de Applied Biosystems et le Dye Set E5, les données sont acquises sur le logiciel Run 3100 Data Collection et traitées sur Genemapper 3.5, Applied Biosystems.

Annexe 2: Spatial genetic structure within a metallicolous population of *Arabidopsis halleri*, a clonal, self-incompatible and heavy-metal-tolerant species. *Molecular Ecology*, **13**, 2959-2967.

Fabienne van Rossum, Isabelle Bonnin, Stéphane Fenart, Maxime Pauwels, Daniel Petit, Pierre Saumitou-Laprade, P.

Abstract *Arabidopsis halleri*, a close wild relative of *A. thaliana*, is a clonal, insect-pollinated herb tolerant to heavy metals (Zn, Pd, Cd) and a hyperaccumulator of Zn and Cd. It is of particular interest in the study of evolutionary processes and phytoremediation. However, little is known about its population gene flow patterns and the structure of its genetic diversity. We used five microsatellite loci to investigate the genetic structure at a fine spatial scale (10 cm to 500 m) in a metallicolous population of *A. halleri*. We also studied the contributions made by clonal propagation and sexual reproduction (seed and pollen dispersal) to the genetic patterns. Clonal diversity was high ($DG > 0.9$). Clonal spread occurs only at short distances (< 1 m). Both clonal spread and limited dispersal, associated with sexual reproduction, contribute to the significant spatial genetic structure revealed by spatial autocorrelation analysis. The shape of the autocorrelogram suggests that seed dispersal is restricted and pollen flow extensive, which may be related to intense activity by insect pollinators. Clonal spread was more extensive in the lowly polluted zone than in the highly polluted zone. This cannot be interpreted as a strategy for promoting the propagation of adapted genotypes under the harshest ecological constraints (highest heavy metal concentrations). The higher fine-scale spatial genetic structure found in the lowly polluted zone can be ascribed to plant densities that were lower than in the highly polluted zone. No evidence of genetic divergence due to spatial heavy metal heterogeneity was found between lowly and highly polluted zones.

Keywords: *Arabidopsis halleri*, clonal spread, heavy metal tolerance, pollen dispersal, seed dispersal, spatial autocorrelation

Spatial genetic structure within a metallicolous population of *Arabidopsis halleri*, a clonal, self-incompatible and heavy-metal-tolerant species

FABIENNE VAN ROSSUM,* ISABELLE BONNIN,† STÉPHANE FÉNART, MAXIME PAUWELS, DANIEL PETIT and PIERRE SAUMITOU-LAPRADE

Laboratoire de Génétique et Evolution des Populations Végétales, UMR CNRS 8016, FR CNRS 1818, Université de Lille 1, Bâtiment SN2, F-59655 Villeneuve d'Ascq Cedex, France

Abstract

Arabidopsis halleri, a close wild relative of *A. thaliana*, is a clonal, insect-pollinated herb tolerant to heavy metals (Zn, Pd, Cd) and a hyperaccumulator of Zn and Cd. It is of particular interest in the study of evolutionary processes and phytoremediation. However, little is known about its population gene flow patterns and the structure of its genetic diversity. We used five microsatellite loci to investigate the genetic structure at a fine spatial scale (10 cm to 500 m) in a metallicolous population of *A. halleri*. We also studied the contributions made by clonal propagation and sexual reproduction (seed and pollen dispersal) to the genetic patterns. Clonal diversity was high ($D_G > 0.9$). Clonal spread occurs only at short distances (< 1 m). Both clonal spread and limited dispersal, associated with sexual reproduction, contribute to the significant spatial genetic structure revealed by spatial autocorrelation analysis. The shape of the autocorrelogram suggests that seed dispersal is restricted and pollen flow extensive, which may be related to intense activity by insect pollinators. Clonal spread was more extensive in the lowly polluted zone than in the highly polluted zone. This cannot be interpreted as a strategy for promoting the propagation of adapted genotypes under the harshest ecological constraints (highest heavy metal concentrations). The higher fine-scale spatial genetic structure found in the lowly polluted zone can be ascribed to plant densities that were lower than in the highly polluted zone. No evidence of genetic divergence due to spatial heavy metal heterogeneity was found between lowly and highly polluted zones.

Keywords: *Arabidopsis halleri*, clonal spread, heavy metal tolerance, pollen dispersal, seed dispersal, spatial autocorrelation

Received 9 March 2004; revision received 7 June 2004; accepted 6 July 2004

Introduction

Contamination of soils by heavy metals is an important environmental and health problem because of the high toxicity of heavy metals (e.g. Salt *et al.* 1998). Some plant species can grow and reproduce on these soils; they are defined as heavy metal tolerant (Macnair 1987). These

plants are of great interest for phytoremediation, a technique that uses heavy-metal-tolerant plants to extract metals from polluted soils or to limit the erosion of toxic elements (Salt *et al.* 1998). Because heavy metals represent highly constraining selective pressures, heavy metal tolerance also provides an excellent model system to study recent adaptive evolutionary processes (e.g. Macnair 1987). Metallicolous populations occurring in areas polluted by heavy metals due to industrial human activities are of particular interest; the sites often constitute contaminated islands in noncontaminated areas, and the amount of heavy metals in the soil is very high, which means there are very strong ecological constraints (e.g. Macnair 1987).

Correspondence: Pierre Saumitou-Laprade. Fax: +33 320 436 979; E-mail: pierre.saumitou-laprade@univ-lille1.fr

*Present address: Plant Science and Nature Management, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium.

†Present address: UMR de Génétique Végétale (CNRS-INRA-UPS-INAPG), Ferme du Moulon, F-91190 Gif sur Yvette, France.

2960 F. VAN ROSSUM ET AL.

Molecular tools and genomics offer new perspectives to study evolutionary processes, including the function and evolution of genes. But because of the considerable investments they require, such investigations have been carried out on only a few model species (Mitchell-Olds 2001). In this context, *Arabidopsis halleri* (Brassicaceae) is a particularly interesting species; this clonal perennial herb is a close wild relative of the model species *Arabidopsis thaliana*, for which a vast array of genomic tools and molecular markers is available (e.g. Mitchell-Olds 2001; Clauss *et al.* 2002). Unlike *A. thaliana*, *A. halleri* is tolerant to high soil concentrations of Zn, Cd and Pb, and hyperaccumulates Zn and Cd. It has been shown that tolerance to Zn is constitutive of the species, but variance for the tolerance has been observed among populations and individuals; metallicolous populations demonstrate increased tolerance to Zn compared with nonmetallicolous populations (Bert *et al.* 2000; Fournier *et al.* unpublished). It is, therefore, a good candidate species for studying adaptive processes, but also for phytoextraction (Bert *et al.* 2002). However, little is known about gene flow patterns and the structure of genetic diversity in *A. halleri* (Clauss *et al.* 2002). To our knowledge, very few studies have explored gene flow patterns within metallicolous populations (e.g. McNeilly 1967; Antonovics & Bradshaw 1970; Ducouso *et al.* 1990). Also, genetic structure at a fine spatial scale has not been investigated for metallicolous populations of insect-pollinated plant species.

Spatial genetic structure within plant populations is primarily determined by the balance and interplay between local genetic drift and gene flow through sexual reproduction, i.e. through seed and pollen dispersal (e.g. Wright 1943; Heywood 1991). The magnitude of structuring depends on the relative contribution of pollen and seed dispersal to gene flow, the level of self-fertilization and the availability of compatible mates, plant density, and pollinator abundance and foraging behaviour (e.g. Williams 1994; Richards 1997; Vekemans & Hardy 2004). For instance, a decrease in pollinator numbers may contribute to the disruption of pollination processes, and therefore reduce intrapopulation gene flow but increase local genetic drift (e.g. Kwak *et al.* 1998). If habitat deterioration due to human activities has been reported to negatively affect pollinators (e.g. Kwak *et al.* 1998; Kevan 1999), little is known about the effects of heavy metal contamination of soils on pollinator abundance (Westerbergh & Saura 1994; Leita *et al.* 1996). Clonal propagation can also contribute to the patterns of genetic structure. A nonrandom spatial arrangement of clones can substantially increase geitonogamy and/or local consanguineous matings, local genetic drift and isolation-by-distance (Heywood 1991; Charpentier 2002). Spatial genetic structure may also be influenced by the spatial variation of selection through environmental heterogeneity (Linhart & Grant 1996). Several studies on plant

populations reveal a fine-scale, intrapopulation genetic structure based on neutral genetic markers that is associated with habitat variation, e.g. with edaphic factors for the wind-pollinated *Festuca ovina* (e.g. Prentice *et al.* 2000), with plant community composition for the outcrossing *Gypsophila fastigiata* (Lönn *et al.* 1996) and with aridity for *Triticum dicooides* (Li *et al.* 2000). Genetic adaptation to heterogeneous environments may result in the formation of distinct ecotypes, which may or may not continue to exchange effective migrants.

In this study, we use microsatellite markers to investigate the spatial genetic structure at a microgeographical scale (ranging from 10 cm to 500 m) in a metallicolous population of *A. halleri*. This large, continuous population is located in an anthropogenic calaminary site surrounded by industrial and urban areas. The soil shows a gradient of heavy metal concentrations (mainly Zn, Pb, Cd). We address the following questions:

- 1 How is genetic diversity structured within the whole population, and at a very fine spatial scale?
- 2 What are the relative contributions of sexual reproduction and clonal propagation to the genetic pattern?

These results are analysed and discussed in light of spatial environmental heterogeneity of the site, i.e. by comparing zones that vary in the amounts of heavy metals in the soil.

Materials and methods

The species

Arabidopsis halleri (Brassicaceae) is a diploid, self-incompatible, insect-pollinated, perennial, rosette-forming herb, which propagates by stolons. Plants produce large numbers of seeds. In Europe, it is distributed among continental and mountainous areas, occurring on acidic, fresh and oligotrophic soils, but also on soils with high heavy metals content (Clapham & Akeroyd 1993; Bert *et al.* 2002). In France, it is non-native and only found in heavy metal-contaminated sites due to human activities (e.g. Berton 1946).

Study site and sampling

'Bois des Asturies' in the town of Aubry in northern France, comprises a very large (several thousands of individuals), continuous population of *A. halleri*, described since 1944 and probably introduced in 1920–25 (Berton 1946). This site was downwind of a zinc smelter and also served as a waste deposit site for the metallurgical factory Umicor. A gradient of heavy metal concentrations (mainly Zn, Pb and Cd) can be found in the soil, increasing from a lowly polluted (LP) zone (extractable [Zn] < 3000 µg/g, [Pb] < 1000 µg/g, [Cd] < 100 µg/g) to a highly polluted (HP)

GENETIC STRUCTURE IN *ARABIDOPSIS HALLERI* 2961

zone (extractable [Zn] > 30 000 µg/g, [Pb] > 5000 µg/g, [Cd] > 300 µg/g) (Van Rossum *et al.* unpublished).

Leaf material was sampled from 612 individual rosettes (= ramets) of *A. halleri*, following two patterns: (i) a minimal distance of every 3 m along a transect covering the population area (232 ramets); and (ii) near the transect (0.5–3 m), using a grid, every 10 cm within four plots of a 0.5 × 3 m area ($n = 59$ –133 for a total of 380 ramets). Two plots (LP1 and LP2) were located in the LP zone (extractable [Zn] = 2792 and 5091 µg/g, respectively). They were 8 m apart. The other two (HP1 and HP2) were located 47 m apart in the HP zone (extractable [Zn] = 29 706 and 32 792 µg/g, respectively). Plots between zones were 295–318 m apart. Because we could detect the presence or absence of rosettes with the quadrat size (10 × 10 cm) of the grid, the number of ramets (= n) and genets (distinct multilocus genotypes, = G) can be considered as an estimate of plot ramet and genet density, respectively.

Microsatellite analysis

Leaves were dried at 55 °C for 24 h. DNA extractions were performed on dried leaf material (10–15 mg dry weight) using the extraction kit Dneasy® from Qiagen®. Polymorphism was assayed on each DNA sample at five microsatellite loci transferred from *A. thaliana* using the resources of the *Arabidopsis* community. Three loci were kindly provided by Thomas Mitchell-Olds (*LYR132*, *LYR133*, *LYR417*); *GC16* (= *ATTS0392*) was described in Clauss *et al.* (2002), and *ATH* (= *ATHCTR1A*) in Bell & Ecker (1994). In order for all loci polymerase chain reaction (PCR) products to be labelled with a single labelled primer, each forward primer contained a 5'-tail of 19 bp homologous to the universal consensus M13 forward primer sequence, followed by the specific locus sequence (Oetting *et al.* 1995).

PCR were carried out in a total volume of 15 µL containing 10 ng of template DNA, 3.5 mM MgCl₂, 200 µg/mL of BSA (for *LYR132*, *LYR133* and *GC16*), 100 µM dNTP, 0.2 µM each primer (the M13 forward primer was fluorescence-labelled with either IRD-700 or IRD-800), 20 mM Tris-HCl (pH 8.3), 50 mM KCl and 0.4 units of AmpliTaq DNA polymerase (Perkin-Elmer). PCR was carried out on a Perkin-Elmer Gene-Amp system 9700 (94 °C for 5 min, followed by locus-specific amplification: 94 °C for 30 s, annealing temperature for 35 s, 72 °C for 30 s, for eight cycles, followed by M13 labelling amplification: 94 °C for 30 s, 56 °C for 30 s, 72 °C for 30 s, for 32 cycles, and a final extension 72 °C for 2 min). The annealing temperature of the locus-specific cycles varied between 50 and 60 °C, depending on the locus.

Amplification products were analysed on a Li-Cor automated DNA sequencer 4200 (Li-Cor-ScienceTec). PCR fragments were separated on 7% denaturing polyacrylamide gels (Long ranger, FMC) of 25 cm (for *LYR417*) and

33 cm, sizes were assessed with Li-Cor's BASE IMAGER software v. 4.1 by comparison with an appropriate labelled molecular mass marker (Li-Cor-ScienceTec).

Data analysis

Extent of donality. In order to assess the independence of the loci, a test for genotypic disequilibrium between pairs of loci and sequential Bonferroni-type correction (Rice 1989) was performed using *FSTAT* (Goudet 2001) for the transect (all sampled individuals are distinct genets, see Results).

To estimate the extent of clonality, we tested to see whether clustered individuals of the same multilocus genotype could be clones generated by asexual propagation or whether they could be identical genotypes (for the markers used) produced via sexual reproduction. The probability of obtaining a second encounter of the same multilocus genotype in the sampled individuals (p_{gen}), under the assumption of linkage equilibrium, was calculated for the whole population (= transect + plots) according to Parks & Werth (1993). It may be assumed with over 95% confidence that the multiple occurrence of a multilocus genotype with $p_{\text{gen}} < 0.05$ is likely to be generated by clonal propagation of a single genet. If $p_{\text{gen}} > 0.05$, ramets are likely to be accounted for by sexual reproduction (Parks & Werth 1993). This analysis was performed for the 574 ramets showing no missing data.

Genetic variation within population. In order to analyse genets (sexually reproduced genotypes) only, distinct multilocus genotypes were identified for each plot (no clone for the transect), and one ramet per genotype was randomly chosen. The following measures of genetic variation were calculated for each locus using *GEN-SURVEY* (Vekemans & Lefèbvre 1997): mean number of alleles (A), observed heterozygosity (H_O) and expected heterozygosity (H_E) corrected for small sample size (Nei 1978) for the entire (transect + plots) 'genet' population (491 genets). The average measures were also computed for all loci. Wright's inbreeding coefficient (F_{IS}) corrected for small sample size (Kirby 1975) was estimated for the entire 'ramet' and 'genet' populations (612 ramets and 491 genets, respectively). The mean F_{IS} over all loci was also determined for the four plots with ramets and with genets, and the standard errors estimated using a jackknife procedure over the loci (Sokal & Rohlf 2000). The significance of F_{IS} for each locus, and over all loci for the entire population and for each plot was tested by randomization tests using *FSTAT*.

Spatial genetic structure within population: contributions of clonal growth and sexual reproduction. We assessed spatial genetic structure at a microgeographical scale using spatial autocorrelation analyses. These were performed with kinship coefficients (Loiselle *et al.* 1995) using *SPAGED1* (Hardy & Vekemans 2002). In order to test for isolation-by-distance,

2962 F. VAN ROSSUM ET AL.

the multilocus kinship coefficient for each pair of individuals was plotted against the logarithm of the geographical distance separating them. The slope (b) of this linear regression provides a good estimator of the extent of gene dispersal at this scale, under the assumption that population genetic structure has reached equilibrium (Vekemans & Hardy 2004). For the graphical representation of kinship, average multilocus kinship coefficients per distance intervals (F_j) were computed for the following distance classes (upper bound distance in metres): $j = 22$ distance classes for the entire (transect + plots) 'ramet' and 'genet' population (0.2, 0.3, 0.4, 0.5, 0.75, 1, 2, 4, 8, 12, 16, 25, 50, 75, 100, 125, 150, 200, 250, 300, 350, 600), $j = 13$ for the four plots with ramets (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.5, 2, 3), and $j = 8$ for the four plots with genets (0.2, 0.4, 0.6, 0.8, 1, 1.5, 2, 3). Standard errors for the multilocus estimates of the kinship coefficients per distance class and the slope were estimated using a jackknife procedure over the loci. We tested the significance of the kinship coefficients and of the b estimates by comparing the observed values with those obtained after 2000 random permutations of the individuals among positions. We also quantified spatial genetic structure using the S_p statistics, calculated as $-b/(1 - F_{(1)})$ where $F_{(1)}$ is the mean F_j for the first distance interval. $F_{(1)}$ can be considered as an approximation of the kinship between pairs of neighbours, provided the first distance class contains enough pairs of individuals to get a reasonably precise $F_{(1)}$ value (Vekemans & Hardy 2004). $1/S_p$ can be an estimate of the neighbourhood size, $Nb = -(1 - F_{(1)})/b$ (Vekemans & Hardy 2004).

In case of a significant structuring effect of clonal growth, a higher spatial genetic structure can be expected for the ramets than the genets (e.g. Heywood 1991). Therefore, pairwise Wilcoxon matched pairs tests between ramets and genets were performed for the entire population and for the four plots on the slopes (b) of the regression with the five polymorphic loci as replicates.

The relative contributions of seed and pollen dispersal to total gene flow were inferred by analysing, for the entire 'genet' population, the shape of the regression between the kinship coefficients and the logarithm of the distances obtained from the spatial autocorrelation analysis, using the methods described in Heuertz *et al.* (2003). The shape of the regression can be described by the k -value, calculated using the coefficients of the term of second and third power of a cubic regression between the residuals of the regression and the logarithm of the distance (for details, see Vekemans & Hardy 2004). A concave shape ($k > 0$) indicates leptokurtic gene flow, suggesting more restricted seed dispersal than pollen dispersal ($\sigma_s \ll \sigma_p$). For a convex curvature ($k < 0$), $\sigma_s \geq \sigma_p$.

Clonality and sexual reproduction in the context of heterogeneous spatial distribution of heavy metals. To compare clonal

growth patterns between the LP and HP zones, the amount of multilocus genotypic variation was estimated for each of the four plots. The proportion of distinguishable multilocus genotypes was calculated as G/N (Ellstrand & Roose 1987), where G is the number of genets and N the number of ramets. Multilocus genotypic diversity (D_G) was calculated as a modification of the Simpson index (Pielou 1969). Ramets with missing data were excluded.

To test for the difference in genetic structure and sexual reproduction between the LP and HP zones, pairwise comparisons between plots (with genets) were performed on the slope (b) and on F_{IS} values using nonparametric Wilcoxon matched pairs tests (replicates = five polymorphic loci).

To detect genetic divergence between the LP and HP zones as a result of heavy metals heterogeneity, an analysis of molecular variance was performed on the four plots with ramets and with genets using ARLEQUIN (Schneider *et al.* 2000). The between-plot component of the total genetic diversity is given by F_{ST} . F -statistics were also used to partition the genetic diversity into its between-zone (LP and HP) component ($= F_{ZT}$) and its between-plot (within-zone) component ($= F_{SZ}$). Their relationship is described by $(1 - F_{ST}) = (1 - F_{SZ})(1 - F_{ZT})$ (Weir 1990). The significance of the F_{ST} , F_{SZ} and F_{ZT} values was tested using a nonparametric permutation approach (Excoffier *et al.* 1992).

Results

Extent of clonality

We did not find any linkage disequilibrium for the transect between the 10 pairs of loci after sequential Bonferroni correction ($P < 0.005$). In total, 459 (of 574 genotyped rosettes) distinct multilocus genotypes were identified for the entire population. The 232 ramets from the transect had distinct multilocus genotypes and can therefore be regarded as distinct genets. Five distinct multilocus genotypes occurred at least twice in different plots, four of which were located in LP and HP plots. At the plot scale, 38 distinct multilocus genotypes occurring at least twice within the same plot were identified; 30 of them (78.9%) were likely to be accounted for by clonal propagation ($p_{sc} < 0.05$). These could, therefore, be considered as putative clones.

Genetic variation within population

At the 'genet' population level (transect + plots), A varied from 3 to 8, H_G varied from 0.428 to 0.710, H_E varied from 0.540 to 0.738 and F_{IS} varied from 0.013 to 0.208 (Table 1). Two of five loci had significant positive F_{IS} values (heterozygote deficiency). Similar F_{IS} values were found for the ramets.

GENETIC STRUCTURE IN *ARABIDOPSIS HALLERI* 2963

Table 1 Estimates of within-population genetic variation (per locus and mean over all loci) for the entire population (transect + plots)

Locus	Range in allele size	A	Genets				Ramets	
			H_O	H_E	F_{IS}	b	F_{IS}	b
ATH	171–206	5	0.428	0.540	0.208**	-0.0021*	0.226**	-0.0027**
GC16	158–183	8	0.709	0.719	0.013 ns	-0.0027***	0.036 ns	-0.0039**
LYR132	229–238	3	0.562	0.610	0.079†	-0.0014***	0.040 ns	-0.0019**
LYR133	159–170	5	0.710	0.722	0.017 ns	-0.0042*	0.017 ns	-0.0063***
LYR417	185–244	6	0.683	0.738	0.074**	-0.0019***	0.090**	-0.0045***
Mean		5.4	0.618	0.666	0.071**	-0.0025***	0.075**	-0.0038***

For the genets: number of alleles (A), observed heterozygosity (H_O) and expected heterozygosity (H_E). For the genets and the ramets: Wright's inbreeding coefficient (F_{IS}) and slope (b) of the regression of pairwise kinship coefficients on the logarithm of geographical distance with significance level.

ns, not significant ($P > 0.10$); † $0.10 > P > 0.05$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

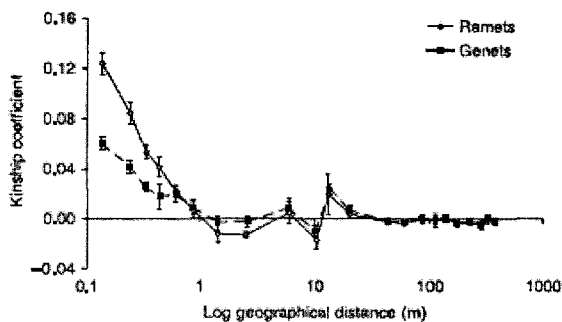


Fig. 1 Correlograms showing the spatial genetic structure for the entire (transect + plots) 'ramet' and 'genet' population, with mean Loiselle kinship coefficients over all loci (\pm SE) as a function of the geographical distance in metres (log-scale).

Spatial genetic structure within population: contributions of clonal growth and sexual reproduction

We found a significant linear relationship between decreasing pairwise kinship coefficients and the logarithm of increasing geographical distance in the entire (transect + plots) 'ramet' population (Fig. 1). The slopes (b) of the regression between pairwise kinship coefficients and

spatial distances were significant ($P < 0.05$) for all five loci, with a mean b -value of -0.0038 (Table 1).

The entire 'genet' population (transect + plots) also showed a significant ($P < 0.05$) spatial genetic structure (mean b -value = -0.0025) (Fig. 1, Table 1). The slope of the 'ramet' population was, however, significantly steeper than that of the 'genet' population (Wilcoxon matched-pairs test $Z = 2.02$, $P < 0.05$). This indicated a more pronounced genetic structure for the 'ramet' population. For the genets, mean F_{IS} value (= 0.071) was similar to the average kinship coefficient between rosettes separated by the first distance interval ($F_{(1)} = 0.060$). This was not the case for the ramets ($F_{IS} = 0.075$ and $F_{(1)} = 0.124$) (Fig. 1). The estimate of neighbourhood size (Nb) for the genets was ~ 372 ; the S_p statistic was 0.0027. The shape of the linear regression between the kinship coefficients and the logarithm of the distance obtained for the spatial autocorrelation for the entire 'genet' population was found to be concave ($k = 2.19$; cubic regression $R^2 = 0.713$, $P < 0.001$), suggesting that seed dispersal was more restricted than pollen dispersal ($\sigma_s \ll \sigma_p$) according to Heuertz *et al.* (2003).

At the plot scale, the slopes (b) of the regression between pairwise kinship coefficients and spatial distances were significant ($P < 0.05$) for the four plots with genets and with ramets (Table 2). For LP1, HP1 and HP2, the slopes

Table 2 Mean Wright's inbreeding coefficient (F_{IS}) and mean slope over all loci (b) of the regression of pairwise kinship coefficients on the logarithm of geographical distance for the four plots with ramets and with genets, with significance level and standard error (SE), and for each plot, Wilcoxon matched pairs tests for b between ramets and genets

Plot	Ramets			Genets			SE	Wilcoxon test b	
	F_{IS}	SE	b	SE	F_{IS}	SE			b
LP1	-0.063 ns	0.082	-0.0890***	0.0181	-0.021 ns	0.078	-0.0482***	0.0072	*
LP2	-0.012 ns	0.099	-0.0763***	0.0184	-0.011 ns	0.064	-0.0393***	0.0050	†
HP1	0.149**	0.049	-0.0502***	0.0071	0.115**	0.041	-0.0333***	0.0087	*
HP2	0.075**	0.051	-0.0102***	0.0058	0.066*	0.059	-0.0063***	0.0056	*

ns, not significant ($P > 0.05$); † $P < 0.10$; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

2964 F. VAN ROSSUM ET AL.

Table 3 Estimates of multilocus genotypic variation for the four plots: number of distinct multilocus genotypes (G), proportion of distinct multilocus genotypes (G/N) and multilocus genotypic diversity (D_G)

Plot	n	N	G	G/N	D_G
LP1	75	74	26	0.351	0.906
LP2	59	59	32	0.542	0.958
HP1	133	124	103	0.831	0.996
HP2	112	104	98	0.942	0.999

n = number of ramets sampled, N = with no missing data.

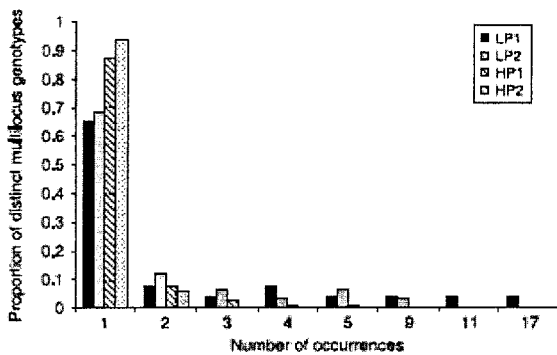


Fig. 2 Frequency-distribution diagram of the number of occurrences (= ramets) of the distinct multilocus genotypes for each of the four plots.

were significantly steeper ($P < 0.05$) for the ramets than for the genets, and marginally significant ($P = 0.08$) for LP2.

Clonality and sexual reproduction in the context of heterogeneous spatial distribution of heavy metals

Estimates of clonal diversity for each plot are given in Table 3. The proportion of distinct multilocus genotypes (G/N) varied from 0.351 to 0.942. Multilocus genotype diversity (D_G) was high, ranging from 0.906 to 0.999. Multiple occurrences of genotypes were mainly found in the LP1 and LP2 plots (Fig. 2). The distance between two identical multilocus genotypes did not exceed 1 m. These putative clones tended to aggregate in clumps, although there was some mixing of genets (Fig. 3). Ramet (n) and genet (G) density, G/N and D_G were lower for the LP1 and LP2 plots than for HP1 and HP2 (Table 3).

Pairwise comparison (Wilcoxon matched-pairs tests) between LP and HP plots (with genets) indicated that the slope was significantly ($Z = 2.02$, $P < 0.05$) lower for HP2 ($b = -0.0063$) than for LP1 and LP2 ($b = -0.0482$ and -0.0393 , respectively), and not significant ($P > 0.10$) for the other pairwise comparisons. There was only a tendency for higher F_{IS} values in the HP plots than in the LP plots

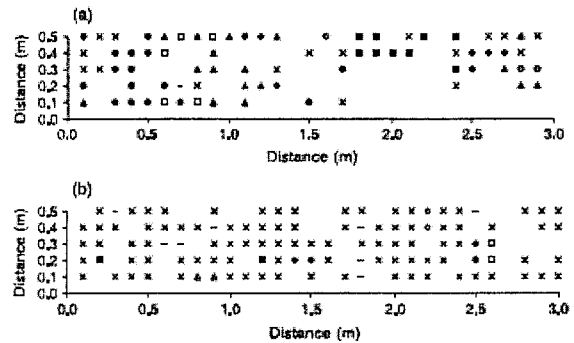


Fig. 3 Map of the distribution of ramets and their multilocus genotypes within (a) the LP1 plot ($G = 26$) and (b) the HP2 plot ($G = 98$). Symbols: X, multilocus genotypes with a single occurrence; -, incomplete genotype (missing data); others, multilocus genotypes with multiple occurrences (≥ 2). There is no correspondence between symbols in the two plots.

($Z = 1.75$, $P = 0.08$ between HP1 and LP1, HP1 and LP2, and HP2 and LP1) (Table 2). The other pairwise comparisons were not significant ($P > 0.10$).

Analysis of molecular variance on the plots with ramets and genets gave a value of -0.004 and -0.008 for the between-zone component of genetic diversity (F_{ZT}) and 0.031 and 0.019 for F_{SZ} (between-plots within LP and HP zones), respectively. The overall F_{ST} value was 0.027 for the ramets and 0.018 for the genets. The permutation test showed significance ($P < 0.001$) for the F_{SZ} and F_{ST} values, whereas F_{ZT} was nonsignificant ($P > 0.10$), indicating that the LP and HP zones were not genetically differentiated.

Discussion

Arabidopsis halleri has a typical outcrosser behaviour on the metalliferous site of the Bois des Asturies

Few studies have explored gene flow patterns and sexual reproduction in metalliferous plant populations. To our knowledge, this is the first study of genetic structure at a fine spatial scale conducted in a metalliferous population of an insect-pollinated plant species. The diversity of the *Arabidopsis halleri* population of 'Bois des Asturies' is high ($H_E = 0.666$; range: 0.540–0.738), and falls within the range of H_E values (0.11–0.75; mean 0.445) reported for 24 polymorphic loci by Clauss *et al.* (2002) in a metalliferous population of *A. halleri* in Germany. The slight, but significant, heterozygote deficiency ($F_{IS} = 0.07$, $P < 0.01$) is typical of an outcrossing species. However, a high F_{IS} value (0.208) was detected at one locus (*ATH*). A possible explanation might be the presence of null alleles, which can arise in microsatellite loci (Pemberton *et al.* 1995). Similar low values ($F_{IS} = 0.077$, $P = 0.10$) were reported in a German anthropogenic metalliferous population of *A. halleri* (M Clauss,

GENETIC STRUCTURE IN *ARABIDOPSIS HALLERI* 2965

pers. commun.) and in the dioecious *Silene dioica* on serpentine ($F_{IS} = 0.055$, based on allozymes) (Westerbergh & Saura 1992). In contrast, high F_{IS} values were reported for metallicolous populations of the self-compatible *Thlaspi caerulescens* ($F_{IS} = 0.107\text{--}0.510$, based on allozymes) (Dubois *et al.* 2003). So despite the constrained ecological conditions in soils contaminated by heavy metals, which might affect genetic patterns, the breeding system appears to be a major determining factor of genetic structure.

Gene flow appears to be very extensive

We detected a significant pattern of fine-scale spatial genetic structure consistent with the model of isolation-by-distance: ramets separated by short distances are more likely to be genetically related than those that are farther apart. Sexual reproduction contributes significantly to the spatial genetic structure. Analysis of the relative contributions of seed and pollen dispersal to total gene flow indicates that seed dispersal is more restricted than pollen dispersal ($\sigma_s \ll \sigma_p$). Seeds of *A. halleri* have no particular mechanism for long-distance dispersal (Clapham & Akeroyd 1993). Hence, seed dispersal is expected to be restricted. By contrast, the flowering period is long (from April to September). Insect visitors (bumblebees, small bees, syrphids, butterflies), which may be potential pollinators, are very abundant in the 'Bois des Asturies' during the flowering season of *A. halleri* (F Van Rossum, pers. observ.). Bumblebees and butterflies can fly over long distances. As a result, long-distance pollen dispersal may be favoured and extensive pollen flow expected (Kwak *et al.* 1998). This is consistent with the particularly high estimate of neighbourhood size ($Nb = 372$) and low Sp value (0.0027) found for the genets. Such values are quite surprising, for an insect-pollinated herbaceous plant species like *A. halleri*, most Nb estimates do not exceed 200 (e.g. Richards 1997; Hardy *et al.* 2000), and mean Sp value is -0.0171 (Vekemans & Hardy 2004). Such high Nb and low Sp values are generally found in wind-pollinated species ($Nb > 200$, mean $Sp = 0.0064$) (Richards 1997; Heuertz *et al.* 2003; Vekemans & Hardy 2004). Thus, our finding for *A. halleri* suggests extensive gene flow through pollen dispersal, revealing intense activity by insect pollinators. This also means that sites contaminated by heavy metals, although highly toxic, may be potentially important pollen and nectar providers for pollinators, especially in industrial, highly urban landscapes like in northern France, where suitable habitats for pollinators are few and far between.

Restricted seed dispersal possibly explains the steep slope of the spatial autocorrelation analysis at short distances (Fig. 1) and constitutes the main contribution of sexual reproduction to spatial genetic structure. The mean F_{IS} value is similar to the average kinship coefficient

between genets of the first distance interval ($F_{IT} = 0.060$), indicating that it might partly reflect biparental inbreeding (Vekemans & Hardy 2004).

Our finding of extensive gene flow by pollen suggests that adjacent metallicolous and nonmetallicolous populations of *A. halleri* might exchange genes. In such a situation, Antonovics (1968) hypothesized that outcrossers on soils contaminated by heavy metals might evolve toward self-fertilization, so that gene flow from a nontolerant population can be prevented, and the breakdown of coadapted traits therefore avoided. Tolerance to Zn has been shown to be constitutive of the species, but variance for the tolerance has been observed among populations and individuals, metallicolous populations being more tolerant than nonmetallicolous populations (Bert *et al.* 2000). Therefore, it would prove very interesting to compare our results on genetic structure and gene flow patterns observed in an isolated metallicolous population with those found in metallicolous populations adjacent to nonmetallicolous populations of *A. halleri*.

Extent of clonality and sexual reproduction in the context of heterogeneous spatial distribution of heavy metals

A. halleri is reported to be a clonal species (e.g. Clapham & Akeroyd 1993). In the population we studied, most genets (80%) were represented by one ramet, and genotypic diversity ($D_C > 0.90$) was higher than the average (0.61) reported by Ellstrand & Roose (1987) in their review of clonally reproducing species. There is some spatial mixing of the genets (Fig. 3, LP plots). This can be explained by the development and rooting of vegetative rosettes on the creeping flowering stems, which can remain after the senescence of the stems and become separated from the 'mother' ramet. Clonal growth contributes significantly to within-population spatial genetic structure, as indicated by the significantly higher b -values for the ramets than for the genets, and by the higher F_{ST} value for the plots with ramets. However, this only happens at a small spatial scale (< 1 m) (Fig. 1). Although *A. halleri* is potentially clonal, we suggest that clonality remains a marginal mode of species' propagation.

Clonal ability can contribute to propagate or perpetuate-adapted genotypes (e.g. Abrahamson 1980; Salemaa & Sievanen 2002). For instance, clonal spread was reported to be higher under harsh ecological conditions, e.g. at range margins (Eckert & Barrett 1993), increasing latitude (Stenström *et al.* 2001) and altitude (Young *et al.* 2002). A more intensive branching of the shoots was reported for *Arctostaphylos uva-ursi* and *Vaccinium uliginosum* plants growing on a copper-nickel polluted soil (Salemaa *et al.* 1999; Salemaa & Sievanen 2002). In our study, contrasted clonal growth patterns are observed between LP and HP zones. But as clonal spread occurs more readily in the LP

2966 F. VAN ROSSUM ET AL.

zone than in the HP zone (Table 3), it cannot be interpreted as a strategy for propagating or perpetuating adapted genotypes under the harshest ecological constraints (highest heavy metals concentrations). The contrasting clonal patterns observed for *A. halleri* seem to be a response to the variation in other environmental conditions in the site, such as humus forms (mull and mor in the LP and HP zones, respectively) or moisture levels (Gillet & Ponge 2002). This may result in locally different competition or nutrient availability (Lehman 1997).

The HP plots show significantly (or a trend for) weaker *b*-values than the LP plots. The LP and HP plots differ in clonal growth patterns but also in plant densities, as estimated by both genet and ramet numbers (Table 3). Plant density can affect the intensity of local genetic drift (Vekemans & Hardy 2004). With low plant density, genetic structure is expected to be more pronounced, as a result of a lower number of potential mates (Heywood 1991; Williams 1994). The higher plant density observed in the HP plots may reflect higher seed production and/or seedling recruitment and thus more marked sexual reproduction in this zone.

No evidence of genetic divergence due to heavy metal heterogeneity was found between LP and HP zones, as the F_{ST} value (between zones) was not significant. Three hypotheses can be considered to explain this lack of divergence. First, within-population gene flow is too extensive to allow divergent selection processes. Second, the lowest concentrations of heavy metals (Zn, Pb, Cd) in the gradient may already correspond to very constraining conditions ($[Zn] > 2792 \mu\text{g/g}$) and therefore may not have resulted in differential selection. Third, our study is based on a small set of randomly chosen microsatellites (only five). These are generally expected to be neutral markers, but may also be in disequilibrium with selected genes. A larger array of loci should, therefore, be investigated. In a further study on the *A. halleri* population of 'Bois des Asturies', we will examine whether ecotypic divergence based on quantitative trait (survival, growth, metal accumulation) variation can be found between the LP and HP zones, using a reciprocal transplants experiment.

Acknowledgements

The authors thank Cécile Godé, Bertrand Spilliaert and Anne Petit for technical advice and support in genotyping and chemical analysis. They are very grateful to Myriam Heuertz for advice on analysis for relative contributions of seed vs. pollen dispersal, to Xavier Vekemans, Olivier Hardy, Jean-François Arnaud and three anonymous referees for helpful comments on the manuscript, and to Lycia Touzet for correction of the English. This study was supported by funding from the Contrat de Plan Etat/Région Nord-Pas de Calais (PRC), from the European FEDER, and from the 'Programme Environnement Vie et Société' (PEVS) from the CNRS (contract N°00 N55/0083).

References

- Abrahamson WG (1980) Demography and vegetative reproduction. In: *Demography and Evolution in Plant Populations* (ed. Solbrig OT), pp. 89–106. Blackwell, Oxford.
- Antonovics J (1968) Evolution in closely adjacent populations V. Evolution of self-fertility. *Heredity*, **23**, 219–238.
- Antonovics J, Bradshaw AD (1970) Evolution in closely adjacent populations VIII. Clinal patterns at a mine boundary. *Heredity*, **25**, 349–362.
- Bell CJ, Eckert JR (1994) Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics*, **19**, 137–144.
- Bert V, Bonnin I, Saumitou-Laprade P, de Laguérie P, Petit D (2000) Do *Arabidopsis halleri* from non-metallicolous populations accumulate zinc more effectively than those from metallicolous populations? *New Phytologist*, **155**, 47–57.
- Bert V, Macnair MR, de Laguérie P, Saumitou-Laprade P, Petit D (2000) Zinc tolerance and accumulation in metallicolous and nonmetallicolous populations of *Arabidopsis halleri* (Brassicaceae). *New Phytologist*, **146**, 225–233.
- Berton A (1946) Présentation de plantes: *Arabidopsis halleri*, *Armeria elongata*, *Oenanthe fluviatilis*, *Galinsoga parviflora dicoides*. *Bulletin de la Société Botanique Du Nord de la France*, **93**, 139–145.
- Charpentier A (2002) Consequences of clonal growth for plant matings. *Evolutionary Ecology*, **15**, 521–530.
- Clapham AR, Akeroyd JR (1993) *Cardaminopsis*. In: *Flora Europaea*, Vol. I (eds Tutin TG, Heywood VH, Burges NA, Valentine DH, Walters SM, Webb DA), p. 290. Cambridge University Press, Cambridge.
- Clauss MJ, Cobban H, Mitchell-Olds T (2002) Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaceae). *Molecular Ecology*, **11**, 591–601.
- Dubois S, Cheptou P-O, Petit C *et al.* (2003) Genetic structure and mating systems of metallicolous and nonmetallicolous populations of *Thlaspi caerulescens*. *New Phytologist*, **157**, 633–641.
- Ducouso A, Petit D, Valero M, Vernet P (1990) Genetic variation between and within populations of a perennial grass: *Arrhenatherum elatius*. *Heredity*, **65**, 179–188.
- Eckert CG, Barrett SCH (1993) Clonal reproduction and patterns of genotypic diversity in *Decodon verticillatus* (Lythraceae). *American Journal of Botany*, **80**, 1175–1182.
- Ellstrand NC, Roose ML (1987) Patterns of genotypic diversity in clonal plant species. *American Journal of Botany*, **71**, 123–131.
- Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.
- Gillet S, Ponge JF (2002) Humus forms and metal pollution in soil. *European Journal of Soil Science*, **53**, 529–539.
- Goudet J (2001) *ESTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices, Version 2.9.3.2*. <http://www.unil.ch/izea/software/fstat.html> [updated from Goudet 1995].
- Hardy OJ, Vanderhoeven S, Meerts P, Vekemans X (2000) Spatial autocorrelation of allozyme and quantitative traits within a natural population of *Centaurea jacea* (Asteraceae). *Journal of Evolutionary Biology*, **13**, 656–667.
- Hardy OJ, Vekemans X (2002) SPAGEDE: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

GENETIC STRUCTURE IN *ARABIDOPSIS HALLERI* 2967

- Heuertz M, Vekemans X, Hausman JF, Palada M, Hardy OJ (2003) Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. *Molecular Ecology*, **12**, 2483–2495.
- Heywood JS (1991) Spatial analysis of genetic variation in plant populations. *Annual Review of Ecology and Systematics*, **22**, 335–355.
- Kevan PG (1999) Pollinators as bioindicators of the state of the environment: species, activity and diversity. *Agriculture, Ecosystems and Environment*, **74**, 373–393.
- Kirby GC (1975) Heterozygote frequencies in small subpopulations. *Theoretical Population Biology*, **8**, 31–48.
- Kwak MM, Velterop O, van Andel J (1998) Pollen and gene flow in fragmented habitats. *Applied Vegetation Science*, **1**, 37–54.
- Lehmann C (1997) Clonal diversity of populations of *Calamagrostis epigejos* in relation to environmental stress and habitat heterogeneity. *Ecography*, **20**, 483–490.
- Leita L, Muhlbachova G, Cesco S, Barbattini R, Mondini C (1996) Investigation of the use of honey bees and honey bee products to assess heavy metals contamination. *Environmental Monitoring and Assessment*, **43**, 1–9.
- Li Y-C, Röder MS, Fahima T *et al.* (2000) Natural selection causing microsatellite divergence in wild emmer wheat at the ecologically variable microsite at Anmiad, Israel. *Theoretical and Applied Genetics*, **100**, 985–999.
- Linhart YB, Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*, **27**, 237–277.
- Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.
- Lönn M, Prentice HC, Bengtsson K (1996) Genetic structure, allozyme-habitat associations and reproductive fitness in *Gypsophila fastigiata* (Caryophyllaceae). *Oecologia*, **106**, 308–316.
- Macnair MR (1987) Heavy metal tolerance in plants: a model evolutionary system. *Trends in Ecology and Evolution*, **2**, 354–359.
- McNeilly T (1967) Evolution in closely adjacent plant populations III. *Agrostis tenuis* on a small copper mine. *Heredity*, **23**, 99–108.
- Mitchell-Olds T (2001) *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. *Trends in Ecology and Evolution*, **16**, 693–700.
- Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.
- Oetting WS, Lee HK, Flanders DJ *et al.* (1995) Linkage analysis with multiplexed short tandem repeat polymorphism using infrared fluorescence and M13 tailed primers. *Genomics*, **30**, 450–458.
- Parks JC, Werth CR (1993) A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). *American Journal of Botany*, **80**, 537–544.
- Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, **4**, 249–252.
- Pielou EC (1969) *An Introduction to Mathematical Ecology*. Wiley Interscience, New York.
- Prentice HC, Lönn M, Lager H, Rosén E, van der Maarel E (2000) Changes in allozyme frequencies in *Festuca ovina* populations after a 9-year nutrient/water experiment. *Journal of Ecology*, **88**, 331–347.
- Rice WR (1989) Analysing tables of statistical tests. *Evolution*, **43**, 223–225.
- Richards AJ (1997) *Plant Breeding Systems*, 2nd edn. Chapman & Hall, Cambridge.
- Salemaa M, Sievanen R (2002) The effects of apical dominance on the branching architecture of *Arctostaphylos uva-ursi* in four contrasting environments. *Flora*, **197**, 429–442.
- Salemaa M, Vanha-Majamaa I, Gardner PJ (1999) Compensatory growth of two clonal dwarf shrubs, *Arctostaphylos uva-ursi* and *Vaccinium uliginosum* in a heavy-metal polluted environment. *Plant Ecology*, **141**, 79–91.
- Salt DE, Smith RD, Raskin I (1998) Phytoremediation. *Annual Review of Plant Physiology and Plant Molecular Biology*, **49**, 643–668.
- Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN, Version 2000: A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Sokal RR, Rohlf FJ (2000) *Biometry: The Principles and Practice of Statistics in Biology Research*, 3rd edn. W.H. Freeman, New York.
- Stenström A, Jonsson BO, Jónsdóttir IS, Fagerstöm T, Augner M (2001) Genetic variation and clonal diversity in four clonal sedges (*Carex*) along the Arctic coast of Eurasia. *Molecular Ecology*, **10**, 497–513.
- Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 912–935.
- Vekemans X, Lefèbvre C (1997) On the evolution of heavy-metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. *Journal of Evolutionary Biology*, **10**, 175–191.
- Weir BS (1990) *Genetic Data Analysis*. Sinauer Associates, Sunderland, MA.
- Westerbergh A, Saura A (1992) The effect of serpentine on the population structure of *Silene dioica*. *Evolution*, **46**, 1537–1548.
- Westerbergh A, Saura A (1994) Gene flow and pollinator behaviour in *Silene dioica* populations. *Oikos*, **71**, 215–224.
- Williams CF (1994) Genetic consequences of seed dispersal in three sympatric forest herbs. II. Microspatial genetic structure within populations. *Evolution*, **48**, 1959–1972.
- Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.
- Young AG, Hill JH, Murray BG, Peakall R (2002) Breeding system, genetic diversity and clonal structure in the sub-alpine forb *Rutidosia leirolepis* F. Muell. (Asteraceae). *Biological Conservation*, **106**, 71–76.

The authors of the laboratory [Génétique et Evolution des Populations Végétales (GEPV)] are involved in population genetics of plant species with a special attention given to the genetics and evolution of zinc tolerance and zinc hyperaccumulation in *Arabidopsis halleri*. F. Van Rossum's research interests are centred around a range of questions in plant evolutionary biology and conservation genetics related to fragmentation of natural plant populations and spatial environmental heterogeneity. I. Bonnin is population geneticist particularly interested in evolution of selected characters and relationships between migration, diversity and adaptation in experimental populations of wheat and *Arabidopsis*.

