AFPP – 7ème Conférence Internationale sur les Ravageurs en agriculture
Montpellier – 26 et 27 octobre 2005

Improving biological control of an invasive pest with molecular phylogeographic and population genetic approaches : the wheat stem sawfly, *Cephus cinctus* Norton, (Hymenoptera : Cephidae) as a case study.

MARIE-CLAUDE BON[1], THOMAS SHANOWER[2], WENDELL MORRILL[3], KIM HOELMER[1], CORINNE HURARD[1] AND JEAN-FRANCOIS MARTIN[4].

[1]EUROPEAN BIOLOGICAL CONTROL LABORATORY, USDA-ARS, CAMPUS INTERNATIONAL DE BAILLARGUET, 34980 MONTFERRIER LE LEZ, FRANCE.
[2]NORTHERN PLAINS AGRICULTURAL RESEARCH LABORATORY, USDA-ARS, SIDNEY, MONTANA, USA.
[3]DEPARTEMENT OF ENTOMOLOGY, MONTANA STATE UNIVERSITY, BOZEMAN, MONTANA, USA.
[4]CENTRE DE BIOLOGIE ET DE GESTION DES POPULATIONS, ECOLE NATIONALE SUPERIEURE AGRONOMIQUE DE MONTPELLIER, CAMPUS INTERNATIONAL DE BAILLARGUET, 34980 MONTFERRIER LE LEZ, FRANCE.

Summary:
The wheat stem sawfly (WSS), *Cephus cinctus* Norton (Hymenoptera: Cephidae) has become a chronic pest of wheat in the semi-arid steppe region of the North American Great Plains. To develop a more general and conceptual framework that would have predictive value for the biological control of this pest, it is important to clarify the geographical history and population structure of this species so that source areas for biological control agents can be more accurately predicted. In our present work, we conducted a phylogeographical study based on 104 samples from the North American Great Plains. Mitochondrial DNA COI gene sequences uncovered 25 haplotypes. Most populations of *C. cinctus* showed a high haplotype diversity except those from Canada, and the number of private haplotypes was higher than the number of shared haplotypes in most populations. Only one haplotype was shared by the three geographical areas that are represented by Canada, Montana and Wyoming, North Dakota, Idaho, Nebraska. The AMOVA revealed slight but significant genetic differences among the regions - Canada-Montana and the other states -. Our data suggest that the threat to wheat in the Northern Great Plains is not due to one uniform genetic entity *C. cinctus* and any biological control effort should be prepared to contend with a genetically diverse *C. cinctus* spread.

Résumé: Amélioration de la lutte biologique d'un ravageur envahissant grâce à des approches de Phylogéographie Moléculaire et de Génétique des populations: un cas d'étude, Le Cèphe du blé, *Cephus cinctus* Norton, (Hymenoptera: Cephidae).
Le cèphe du blé (WSS), *Cephus cinctus* Norton (Hymenoptera: Cephidae) est devenu un ravageur continuel du blé dans la région semi-aride des Grandes Plaines Nord Américaines. Afin de mettre en place un cadre plus conceptuel qui permettrait une meilleure prédiction pour la lutte biologique contre ce ravageur, il est important de clarifier l'histoire géographique et la structure des populations de cette espèce et ce, en vue d'identifier les zones géographiques « sources » de ses enemis naturels. Dans le présent travail, nous avons conduit une étude phylogéographique basée sur 104 individus, échantillonnés dans les grandes plaines Nord américaines. Les séquences du gène mitochondrial codant la Cytochrome Oxydase I recouvrent pas moins de 25 haplotypes. La plupart des populations de *C. cinctus* exhibe une importante diversité haplotypique à l'exception de celles du Canada, et le nombre d'haplotypes uniques est supérieur à celui des haplotypes partagés

entre populations. Seul un haplotype est commun aux trois régions géographiques qui sont représentées par le Canada, le Montana et l'ensemble des états du Wyoming, North Dakota, Idaho, Nebraska. L'AMOVA a révélé des différences faibles mais néanmoins significatives entre le Canada et le Montana, pris ensemble, et le reste des états américains. Nos données suggèrent que la menace pour le blé cultivé dans les Grandes Plaines du Nord n'est pas attribuable à une seule entité génétique de *C. cinctus* et que les efforts en matière de lutte biologique devraient prendre en compte cette dispersion génétiquement diverse de *C. cinctus*.

Key-words: Biological Control, Pest, Phylogeography, Wheat Stem Sawfly.

INTRODUCTION

The goal of classical biological control is to limit the population density of a pest organism through the use of non-indigenous predators, parasites, pathogens or herbivores, primarily found from the pest's area of origin. To develop a general and conceptual framework that would have predictive value for biological control, the determination of the origin of the pest that colonizes agricultural cropping systems as well as the knowledge of the composition, distribution, pattern of spread in new habitat and population dynamics of these insects are central to choosing appropriate biocontrol candidates (Roderick and Navajas, 2003; Brown, 2004; Wajnberg, 2004). This information is of interest both to researchers and to resource managers, whose goal it is to contain or control invasive species by understanding and optimizing the insect pest-plant host-natural enemy biosystems. All invasive species go through a similar series of steps as they invade new environments: (i) migration into the new habitat; (ii) initial colonization and establishment; and subsequent widespread dispersal (Sakai *et al.,* 2001). Throughout all these processes, there is a considerable potential for genetic changes to occur through founder effects, genetic drift and selection. Individuals that colonize a new habitat are, by definition, a genetic subset of their source population, hence involving a bottleneck relative to source population. Molecular genetic approaches have become increasingly used to address such situations (Roderick and Navajas, 2003; Brown, 2004). Molecular markers have been used successfully to study intraspecific genetic variation (genealogy) and its relationships to geographical and spatial distribution of species, for which the term of Phylogeography was coined (Avise, 2000). The first example in which Phylogeography was successfully applied to classical biological control was for *Bemisia tabaci* (Kirk *et al.,* 2000; Brown, 2004). In this case, relationships between biogeographic lineages of whiteflies and natural enemies were accurately estimated using mitochondrial markers allowing for a more "customized" biological control program.
In this paper, we report on the use of intraspecific phylogeograpy and its potential predictive value for a more customized classical biological control program of another pest.

The wheat stem sawfly, *Cephus cinctus* Norton, (Hymenoptera: Cephidae) is a major insect pest in dryland wheat (*Triticum* L. spp.: Poaceae) fields in the Northern Great Plains of the United States and in southern regions of the prairie provinces of Canada. Yield losses of up to 70% have been reported in parts of the USA and Canada, with annual losses exceeding US$50 million. First noticed in Canada in the late 1800s, wheat stem sawfly was considered as an agricultural pest by the early 20th century and has been the subject of a classical biological control program since several decades (Shanower and Hoelmer, 2004). Two hypotheses have been put forward to explain its origin and species status.
One hypothesis proposes that the wheat stem sawfly is native to North America, originally attacking sawflies in wild grasses and then adapting to spring wheat, and more recently to winter wheat. Although *C. cinctus* is currently considered to be a single species, Lou et al. (1998) found substantial genetic divergence using RAPDs to discriminate between Montana and North Dakota populations displaying a different duration of postdiapause development. Although wheat was readily infested by sawflies, parasitoids have been slower to utilize this

host in this new habitat. The primary natural enemies of the wheat stem sawfly are two closely related braconid wasps, *Bracon cephi* (Gahan) and *Bracon lissogaster* Muesebeck (Hymenoptera: Braconidae). Though parasitism levels in wheat fields have been found to be as high as 98% (Morrill, 1997), the overall impact of these natural enemies remains inconsistent, as infestation rates at some sites are high for unknown reasons.
However, this native American origin has been questioned by Ivie and Zinovjev's discovery (1996). These authors synonymized the Asian species *C. hyalinatus* Konow under *C. cinctus*, concluding that *C. cinctus* North American populations may have been the result of a recent introduction from Northeastern Asia starting from the late eighteenth century (Ivie, 2001). Taken as a whole, the nature of these contradictory data precludes the optimization of the host specificity testing and hence the current biological control program of this pest.
Because neither an extensive sampling of the Asian *C. hyalinatus* populations nor a sampling of the North American populations of *C cinctus* on wild grasses, has been achieved yet, comparisons of the genetic diversity between source populations and invasive populations on wheat are still to be undertaken. However, with the opportunity to examine samples from the main wheat producing areas in North America, our first goal was to focus on a general understanding of the genetic variation and structure of *C. cinctus* populations from wheat in the Northern Great Plains. To achieve this objective, we analyzed the mitochondrial Cytochrome Oxidase I (COI) gene sequences of 104 indivduals from 57 populations across the Northern Great Plains. The sequence level variation was used to establish diversity patterns of this pest and provide some preliminary guidance for the current biological control program.

MATERIALS AND METHODS

Fifty-seven populations of *C. cinctus* were collected, only from infested wheat host by numerous collaborators in two countries and seven states - USA (Montana, Wyoming, Idaho, North Dakota, Nebraska) and Canada (Alberta and Saskatchewan), were analyzed in this study (Figure 1). Samples from each locality were preserved in 96° ethanol and kept frozen at –20°C for genetic investigations. Ethanol-preserved voucher specimens are being maintained at the EBCL-ARS-USDA Laboratory at Montpellier, France.

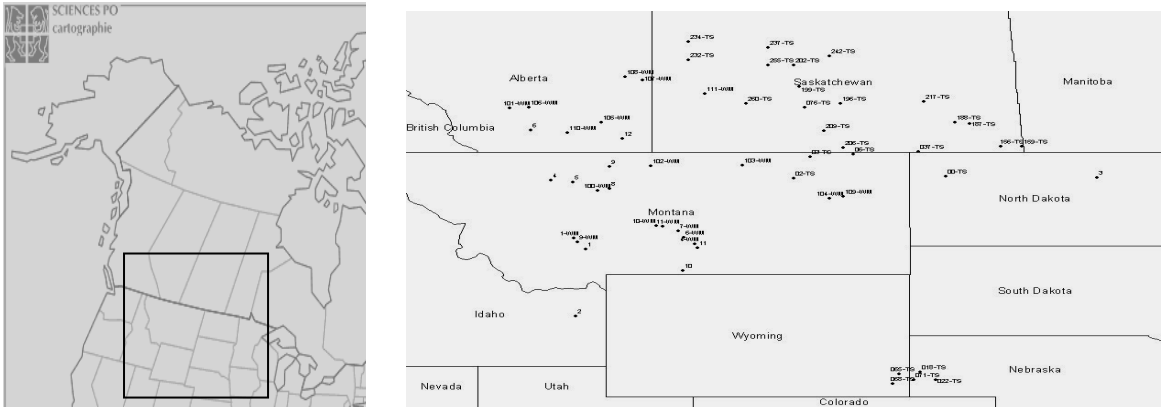

Figure 1. Maps showing the collection sites of *C. cinctus* in the Northern Great Plains.
Figure 1. Cartes décrivant les sites de collections de *C. cinctus* dans les Grandes Plaines Nord Américaines.

DNA analysis – Extraction, Amplification and Sequencing

Genomic DNA was extracted from the head or the whole body of one individual only, following a CTAB/phenol/chloroform-based procedure. The primers C1-J-2183 (5'-CAACATTTATTTTGATTTTTTGG-3') and TL2-N-3014 (5'-TCCAATGCACTAATCTGCCATATTA-3') as described in Simon *et al.* (1994) were used to amplify a ~830 bp fragment of the cytochrome oxidase gene. PCRs were carried out in 25 µl of 10X buffer, 1U Taq Polymerase (Qiagen), 0.2 mM of each dNTP, 0.5 µM of each oligonucleotide primer, and 2 µl of template DNA. Samples were denatured at 94°C for 3 min and then PCRs were carried out for 30 cycles of 30 s denaturation at 92°C, 30 s annealing at 52°C, 1 min elongation at 62°C and 7 min final extension at 62°C in a Perkin Elmer 9700 thermocycler. Both ends of the PCR products were sequenced using the same primers and the BigDye Terminator method (Perkin Elmer, Foster City, CA) on an Applied Biosystems 3730 XL ™ DNA Analyser (Genome express SA, France). Each individual sequence obtained will be deposited in the Genbank database.

Data analysis

Sequences were checked and corrected by eye, then aligned with ClustalX (Thompson *et al.,* 1997). Genetic variation was analysed for the entire set of samples and for subsets consisting of three major geographical areas: Canada, Montana and Wyoming plus Idaho plus North Dakota plus Nebraska. Standard sequence diversity indices computed included: *A* (number of alleles = variable haplotypes), *S* (number of segregating sites = variable nucleotide positions), gene diversity, and nucleotide diversity, π. These were calculated using the ARLEQUIN 2.0 Software (Schneider *et al.*, 2000). Tajima's test was computed using DNASP 3.53 (Rozas & Rozas 1999) to determine whether the mutations in the data set varied significantly from a neutral pattern.
The average number of pairwise nucleotide differences (*K*) was calculated in DNASP v. 3.53. Inference on the distribution of the genetic diversity in the North Great Plains was determined using the analysis of molecular variance (AMOVA) method Escoffier *et al.* (1992) as implemented by ARLEQUIN. Analysis of molecular variance was used to descipher the components of genetic diversity imputable to the variance among the two geographical areas sampled in the North Great Plains (A: Canada and Montana) and B: Wyoming plus Idaho plus North Dakota plus Nebraska relative to that observed within each of them. The significance of the resultant F_{St} estimates and variance component were tested with 1,023 permutations.
The evolutionary relationships between *C. cinctus* samples were investigated using PAUP* version 4.0b10 (Swofford 2002). A Neighbor-Joining (NJ) tree was computed using Kimura 2-parameter model as the distance measure. The reliability of each node in the phylogeny was assessed using the bootstrap method with 1000 pseudo-replications.

RESULTS

The alignment of the COI sequences for the 104 specimens of *C. cinctus* contained 797 positions, revealing 24 parsimony informative sites. Tajima's test indicated that there is no significant departure from neutrality for the data set (Tajima's D=-1,007, P>0.1).
Diversity indices are summarized in Table 1. Analysis yielded a total number of 25 observed haplotypes which were not uniformly distributed in the geographical range sampled (Table 1). The most common haplotype was found in all three geographical groups and was represented in 34% of the sequenced individuals. The highest diversity was found in the group including Wyoming, Idaho, North Dakota and Nebraska, with nucleotide diversity of 0.94 compared to 0.60 in Montana and 0.27 in Canada. Likewise gene diversities and mean pairwise differences were higher for Wyoming, Idaho, North Dakota and Nebraska together compared to the two other groupings.

Geographical origin	Canada	Montana	Wyoming & Idaho & North Dakota & Nebraska
Haplotype			
H1	10		
H2		1	
H3	17		
H4	1		
H5	1	2	
H6		1	
H7		1	
H8	18	15	2
H9		1	
H10			1
H11		1	
H12			1
H13		2	
H14			1
H15			1
H16		1	
H17		1	
H18			1
H19			4
H20			1
H21			1
H22		2	
H23		1	
H24			1
H25	9	5	
Diversity indices			
N	6	13	10
A	3	10	8
S	7	23	26
π	0,27%	0.60%	0.94%
Gene diversity	0.75±0.02	0.81±0.06	0.92±0.06
Mean Pairwise differences	0.28±0.11	0.61±0.14	0.96±0.20

Table 1. Geographic distribution of mtDNA haplotypes and molecular diversity indices estimated from COI sequences obtained from samples of *Cephus cinctus* collected across the Northern great Plains.
N: Number of total haplotypes; *A*: number of alleles (unique halotypes); *S*: number of segregating sites ; π: nucleotide diversity (see text).
Table 1. Distribution géographique des haplotypes mitochondriaux et indices de diversité moléculaire estimés à partir des séquences de COI obtenues d'échantillons de *Cephus cinctus* collectés dans les Grandes Plaines Nord Américaines.

The analysis of molecular variance performed with grouping sample sites according to major geographical areas in Northern Great plains (Canada and Montana) versus Wyoming, Idaho, North Dakota and Nebraska showed that 88.9% (F_{st}=0.11, P<0.001) of the variation (Table 2) was distributed within populations, but there remained a residual and significant portion (7.2%) as a result of differences among regions.

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation	P-value
Between regions 1 and 2	1	1.50	0.032	7.2	<0.05
Between populations within 1 and 2	1	1.14	0.017	3.8	<0.05
Within populations	101	40.42	0.401	88.9	<0.001

Table 2. AMOVA testing population structure between the two regions 1) Canada and Montana and 2) Wyoming, Idaho, North Dakota and Nebraska. See test for details.
Table 2. Test AMOVA de structure génétique des populations entre les deux régions 1) Canada et Montana et 2) Wyoming, Idaho, North Dakota et Nebraska. Voir texte pour les détails.

The Neighbor-Joining tree obtained with the COI sequences (Figure 2) exhibited a relatively shallow genealogy. Samples with identical geographic origin did not necessarily group together.

DISCUSSION

In the past decades, there has been a rapid development in new molecular techniques and evolutionary approaches for analyzing the population structure of a species (Sakai *et al.*, 2001, Brown, 2004). These can be used to directly measure the genetic divergence between populations of a species. Genetic distances have been previously estimated in a variety of species of Hymenoptera, with values ranging from 0.003 to 0.020 for local populations and from 0.059 to 0.190 for cryptic species (as reported by Unruh and Wooley,1999). In *C. cinctus,* we found that they could range from 0.0013 at the population level to 0.0188 between two populations (for example between Montana and Idaho or Wyoming). This level of genetic diversity is similar to that observed in the COI of other insects (Unruh and Wooley,1999).
Most populations of *C. cinctus* showed high haplotype diversity except for those from Canada, and the number of private haplotypes was higher than the number of shared haplotypes in most populations. Our results indicate also that within a population, there can be multiple and divergent haplotypes while widespread populations may be monotypic for this marker (all individuals sharing the same haplotype). This pattern of genetic variation is highly favorable for colonizing species, providing the genetic material enabling rapid response to selection pressure (Lou *et al.,* 1998; Roderick and Navajas, 2003).
Genetic variation was mostly observed within populations but nevertheless there was measurable divergence among the populations within each geographical region and among the three geographical regions. The results are consistent with those obtained by Lou *et al.,* (1998) even though the amplitude of the among-population variation observed is lower when using the mtDNA data than for the RAPD data. The genetic structure of most species seems to be governed by several components, including the dispersal ability of the species. *C cinctus* is a weak flier and long-distance dispersal has not been documented for this species. The low mobility of this species would strongly limit the rate of gene flow between populations and affect its genetic structure. Nevertheless, the large scale distribution of some haplotypes suggests that there could have been a recent, long dispersal event(s) contrary to the expectations for a low mobility. *C cinctus* may have dispersed into US and began to subdivide in response to the changing environmental conditions. In addition, human intervention can affect dispersal of organisms, and this has the potential to drastically alter the genetic structure of natural populations.
At this point of the genetic study, these data already provide insights on how much genetic diversity is represented in the North Great Plains locations. These data suggest that the threat to wheat in the Northern Great Plains is not due to one uniform genetic entity of *C.*

cinctus and any biological control effort should be prepared to contend with a genetically diverse *C. cinctus* spread.

ACKNOWLEDGMENTS

We thank Gaetan Gotanegre for his technical assistance, Drs. Beyer and Weaver for their valuable help in the samples collection and anonymous reviewers for comments on the manuscript.

REFERENCES

Avise, J.C. 2000. Phylogeography: the History and formation of species. Harvard University Press, Cambridge, Massachusetts. 447 p.

Brown, J.K. 2004. Tracing the origin of cryptic insect Pests and vectors, and their natural enemies. In: Ehler, L.E., Sforza, R., Mateille, T. (Eds), Genetics, Evolution, and Biological Control. CABI Publishing, Wallingford, UK, pp. 113-135.

Escoffier, L., P.E. Smouse, and Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

Kirk, A.A., L.A. Lacey, J.K. Brown, M.A. Ciomperlik, J.A. Goolsby, D.C. Vacek, L.E. Wendel and Napompeth, B. 2000. Variation with the *Bemisia tabaci* species complex (Hemiptera: Aleyrodidae) and its natural enemies leading to successful bological control of Bemisia biotype B in the USA. *Bulletin of Entomological Research* 90: 317-327.

Ivie, M.A., and Zinovjev, A.G. 1996. Discovery of the Wheat Stem Sawfly (*Cephus cinctus* Norton) (Hymenoptera: Cephidae) in Asia, with the proposal of a new synonymy. *The Canadian Entomologist* 128: 347-348.

Ivie, M.A. 2001. On the geographic origin of the Wheat Stem Sawfly ((Hymenoptera: Cephidae): A new hypothesis of introduction from Northeastern Asia. *American Entomologist* 47: 84-97.

Lou, K.F., M.J. Weiss, P.L. Bruckner, W.L. Morrill, L.E. Talbert, and Martin, J.M. 1998. RAPD variation within and among geographic populations of Wheat Stem Sawfly (*Cephus cinctus* Norton). *The Journal of Heredity* 89: 329-335.

Morrill, W. 1997. The Wheat Stem Sawfly (*Cephus cinctus* Norton) (Hymenoptera: Cephidae), and associated parasitoids in the northern Great Plains of North America. *Trends in Entomology* 1: 171-174.

Roderick, G.K. and Navajas, M. 2003. Genes in novel environments: Genetics and evolution in biological control. *Nature Reviews Genetics* 4: 889-899.

Rozas, J., and Rozas, R. 1999. DNASP, Version 3: an integrated program for molecular popluation genetics and molecular evolution analysis. *Bioinformatics* 15: 174-175.

Sakai, A.K., F.W. Allendorf, J.S. Holt, D.M. Lodge, J. Molofsky, K.A. With, S., Baughman, R.J. Cabin, , J.E. Cohen, N.C. Ellstrand, D.E. Mc Cauley, P. O'Neil, I.M. Parker, J.N. Thompson, and Weller, S.G. 2001. The population biology of invasive species. *Annual review of Ecology and systematics* 32: 305-332.

Schneider, S., D. Roessli, and Excoffier, L. 2000. Arlequin ver. 2.000: A Software for Population Genetics Data Analysis, Version 2.000. Genetics and Biometry Laboratory, Dept. of Anthropology, University of Geneva, Switzerland.

Shanower, T.G. and Hoelmer, K.A. 2004. Biological control of wheat stem sawflies: past and future. *J. Agricultural & Urban Entomol*. 21: In press.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and Flook, P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Entomology Society of America* 87: 651-701.

Swofford, D.L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and Higgins, D.G. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 24: 4876-4882.
Unruh T.R. and Wooley J.B. 1999. Molecular methods in Classical Biological Control. In: Bellows T.S. and Fisher T.W. (Eds), Handbook of Biological Control: Principles and applications of Biological Control. Academic Press, San Diego, pp. 57-85.
Wajnberg, E. 2004. Measuring genetic variation in natural enemies used for biological control: Why and how?. In: Ehler, L.E., Sforza, R., Mateille, T. (Eds), Genetics, Evolution, and Biological Control. CABI Publishing, Wallingford, UK, pp. 19-37.

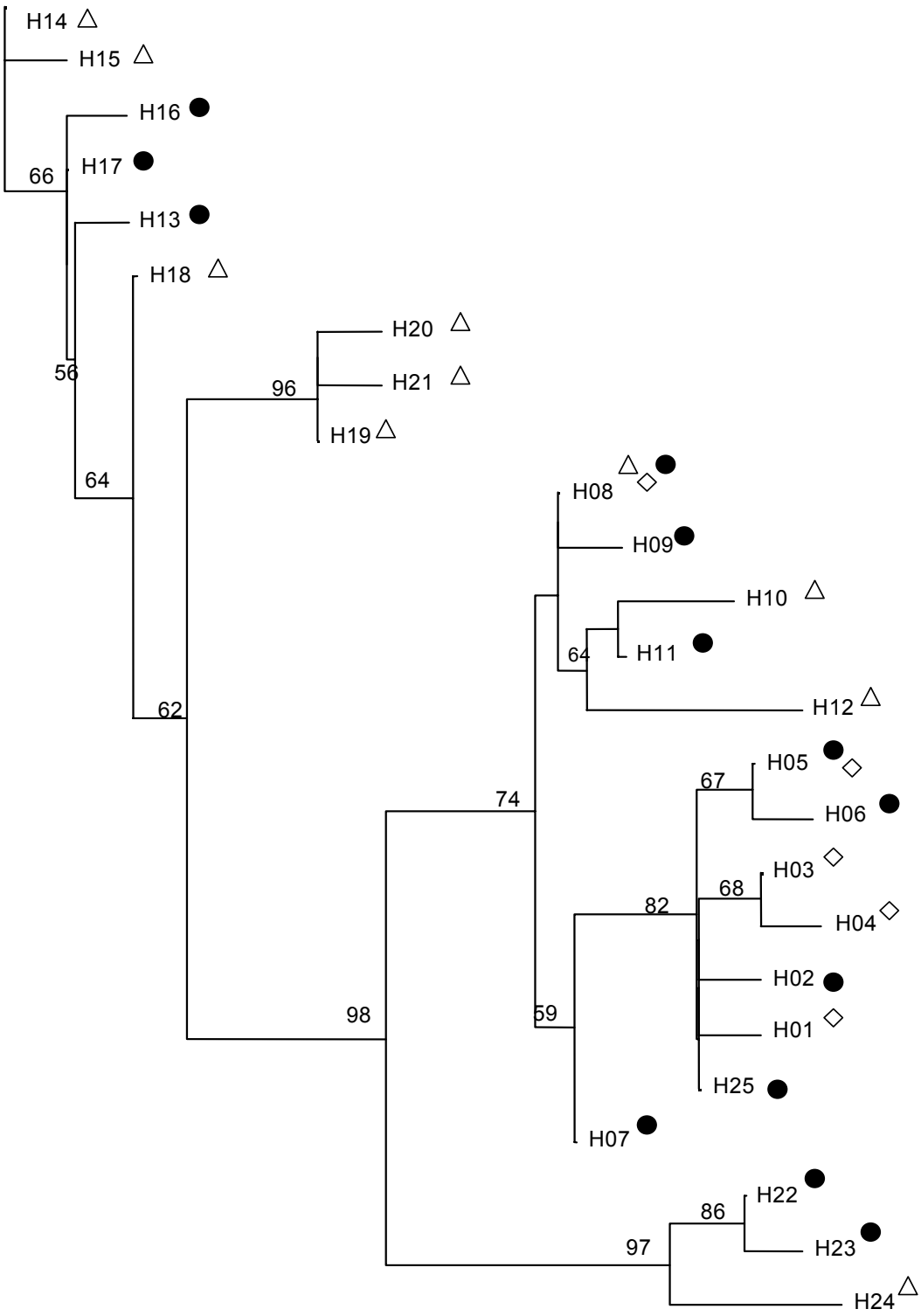

Figure 2. Neighbor-joining tree with Kimura- 2 distance of the 25 *C. cinctus* haplotypes. Numbers on the branches are bootstrap values (1 000 replicates). The three symbols refer to the three geographical areas – Montana● - Wyoming & Nebraska & Idaho & North Dakota△ Canada respectively ◇ .

Figure 2. Phylogramme construit selon la methode du Neighbor-Joining (distance de Kimura-2) à partir des 25 haplotypes observés chez *C. cinctus.* Les chiffres sur les branches correspondent aux valeurs de Bootstrap (1000 réplications). Les trois symboles réfèrent aux trois zones géographiques – Montana● , Wyoming & Nebraska & Idaho & North Dakota △ - Canada respectively ◇ .