Ann. Bot. Fennici 44: 213–218 ISSN 0003-3847
Helsinki 20 June 2007

Applying principal components analysis (PCA) for separating wingless birch fruits — a palaeoecological case study from northern Norway

Minna Väliranta[1,2] & Jan Weckström[1]

[1)] *Department of Biological and Environmental Sciences, P.O. Box 65, FI-00014 University of Helsinki, Finland (e-mails: minna.valiranta@helsinki.fi, jan.weckstrom@helsinki.fi)*
[2)] *Department of Geology, P.O. Box 64, FI-00014 University of Helsinki, Finland*

Received 9 May 2006, revised version received 1 Nov. 2006, accepted 8 Nov. 2006

Väliranta, M. & Weckström, J. 2007: Applying principal components analysis (PCA) for separating wingless birch fruits — a palaeoecological case study from northern Norway. — *Ann. Bot. Fennici* 44: 213–218.

We used principal components analysis (PCA) to distinguish wingless birch fruit bodies of different birch species from each other. Firts, we measured seven dimension variables from a modern fruit selection and analysed their explanatory power using PCA. Then, we included the subfossil birch fruits in PCA as passive variables and compared the result with the data gathered by visual analysis. PCA clearly separates the modern fruits of *Betula nana*, *B. pubescens* ssp. *czerepanovii*, and *B. pendula* from each other. PCA also provides an effective and objective tool for producing more detailed palaeoecological data of the occurrences of different birch species than does the traditional visual analysis only. Instead of classifying fossil birch fruits into rough type-classes at least *Betula nana*, *B. pubescens* ssp. *czerepanovii*, and *B. pendula* can be separated from each other on the basis of the morphological features.

Key words: modern birch fruits, principal components analysis (PCA), subfossil birch fruits

Introduction

Palaeobotanical research, especially in Finland, is largely based on pollen analysis (e.g. Hyvärinen 1975, Tolonen & Ruuhijärvi 1976, Seppä 1996), whereas plant macrofossil analysis is a rather neglected method (however *see* e.g. Vasari 1962, Vasari *et al*. 1996, Seppä *et al*. 2004). Pollen data can, in general, only provide a regional view of the past plant communities, whereas plant macrofossil analysis allows to reconstruct a local catchment-level plant history (Birks 2001, Birks 2003, Seppä *et al*. 2004, Väliranta 2005). This approach is especially valuable when reconstructing past environments in areas that nowadays are located beyond or above the tree line. For instance, subfossil birch fruits are one of the most abundant macroscopic plant remains preserved and found from lake sediments. When compared with information derived from past pollen assemblages, identification of subfossil birch fruits could provide important additional information of past environmental conditions of the catchment, due to the different ecologi-

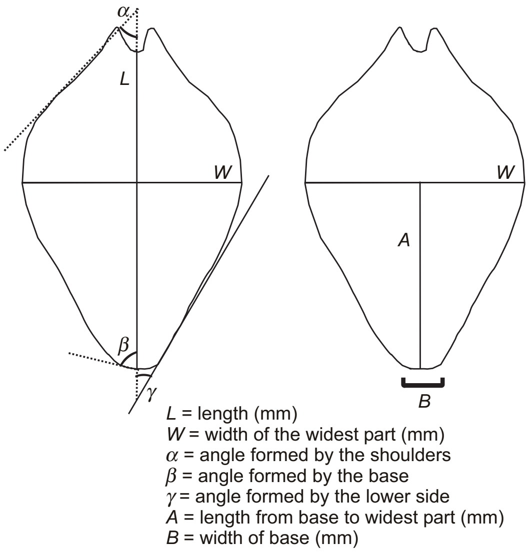

Fig. 1. The measured variables on *Betula* fruit bodies.

cal needs for growing conditions, reproducing and germination of different birch species (e.g. Kallio & Mäkinen 1978). Finds of subfossil tree-type birch fruits from locations above or beyond modern tree line refer, by themselves, to better growing conditions in the past. In addition, however, a reliable and precise identification of fossil birch-fruit material, covering the time period after the last deglaciation, could provide an added value to the debate of the history of different birch species, and especially of the history, origin and taxonomic status of the mountain birch *Betula pubescens* ssp. *czerepanovii* (synonym for *B. pubescens* ssp. *tortuosa*) (Vaarama & Valanne 1973, Kallio & Mäkinen 1978, Hämet-Ahti 1987, Aas & Faarlund 2001 and references therein). Accordingly, palaeoecological interpretation should be based on a trustworthy separation of pubescent and white birches (*B. pubescens* and *B. pendula*), mountain birch (*B. pubescens* ssp. *czerepanovii*), and dwarf birch (*B. nana*). It is commonly recognised among palynologists that it is extremely difficult, if not impossible, to identify different birch species based on pollen analysis only (however *see* Mäkelä 1996). By contrast, in optimal circumstances tree-type birch fruits are relatively easy to distinguish from dwarf birch fruits. Yet, a lot of morphological variation exists (Białobrzeska & Truchanowiczówna 1960) and hybridization between different birch species complicates the visual based identification further due to the transmission of varying amount of genetic information to fruits of a new hybrid (Valanne & Sulkinoja 1991).

As shown in studies of van Dinter and Birks (1996), and Freund *et al.* (2001) principal components analysis (PCA) is a useful tool to separate wingless birch fruit bodies of different birch species from each other. As a complement to the previous studies carried out in relatively southern latitudes, we here apply PCA to a small subarctic lake in northern Norway (69°45´N, 23°50´E) to provide new data of the past presence of different birch species in an area that currently is located at the subarctic tree line. Moreover, we will emphasize the potential of the objective PCA method to distinguish fossil birch fruits from each other in a more reliable way than the traditional, more subjective, visual method does.

Methods

Fruit measurements

Modern fruits of five taxonomic birch groups were selected: *Betula pubescens*, *B. pendula*, *B. nana*, *B. pubescens* ssp. *czerepanovii*, and a hybrid *B. nana* × *pubescens*. Fruits were collected from the botanical garden of the University of Turku, Finland. The fruits represent a selection of birches that originate from northern provenances but have been transplanted/germinated relating to propagation studies carried out at the University of Turku. Nomenclature follows that of Hämet-Ahti *et al.* (1998). Thirty fruits of each group were copy-drawn using a binocular and a camera Lucida. Each fruit was drawn three times and seven different size and shape variables were measured from each drawing according to van Dinter and Birks (1996) (Fig. 1). The average of three separate measurements was used for statistical analysis. The measured variables were: length (L, mm), width of the widest part (W, mm), angle formed by the shoulders (α), angle formed by the base (β), angle formed by the lower side (γ), length from base to the widest part (A, mm), and width of base (B, mm). Instead of using only one morphometrical ratio (cf. van Dinter & Birks 1996, Freund *et al*. 2001) we used three different ratios, *A*:*L*, *A*:*W* and *B*:*W* (× 100), to measure the symmetry of the fruit body. The mean, median, minimum, and maximum values of the measurements are shown in Table 1. All ten morphological variables were used in the PCA.

Subfossil fruits (220 individuals) were at first identified visually to as low a taxonomic level as possible from a 185-cm-long sediment core at 5 cm intervals. Their dimensions were then measured and compared with the modern birch fruit data using PCA.

Principal components analysis

Principal components analysis (PCA) was used to interpret and summarise the major patterns of variation within the birch fruit data. PCA is an indirect ordination technique for obtaining a low-dimensional representation of multivariate data so that the data can be explored visually in a two dimensional PCA correlation biplot and any structure in the data identified (Everitt 1978). In the correlation biplot, variables with high positive correlations have acute angles between their vectors. The length indicates the strength of the variable in relation to the displayed ordination (ter Braak 1994). The fossil birch fruits are included in the PCA as supplementary (passive) variables. These variables do not affect the analysis but are positioned subsequently of the ordination biplot as if they had taken part in it (ter Braak & Šmilauer 1998). PCA is performed on untransformed and centred and standardised species data using the program CANOCO 4.5. (ter Braak & Šmilauer 2002).

Results and discussion

PCA ordination is shown as a two-dimensional correlation biplot (Fig. 2a). The eigenvalues for PCA axes 1 and 2 are 0.44 and 0.19, respec-

Table 1. Summary of measurements of modern birch fruits.

		L (mm)	*W* (mm)	*A* (mm)	*B* (mm)	α	β	γ	*A*:*L* (%)	*A*:*W* (%)	*B*:*W* (%)
B. nana (number on fruits = 30)											
	mean	1.74	1.23	0.87	0.33	44.37	81.47	29.93	49.97	72.04	26.76
	median	1.75	1.19	0.87	0.30	45.00	81.00	28.50	50.01	71.59	22.31
	min	1.43	0.96	0.54	0.14	28.00	70.00	16.00	37.50	46.94	13.33
	max	1.90	1.65	1.09	0.95	55.00	98.00	45.00	63.01	94.00	64.63
B. pubescens ssp. *czerepanovii* (number on fruits = 30)											
	mean	2.77	1.25	1.61	0.22	39.73	68.15	22.07	57.69	130.89	17.99
	median	2.78	1.25	1.59	0.19	39.00	68.50	21.00	58.16	125.51	16.29
	min	2.03	0.88	0.79	0.14	24.00	48.00	11.00	35.11	73.38	9.62
	max	3.53	1.80	2.37	0.49	56.00	84.00	46.00	76.95	216.67	46.08
B. pubescens (number on fruits = 30)											
	mean	2.01	1.10	1.15	0.21	37.73	72.87	23.03	57.21	106.34	19.39
	median	2.04	1.11	1.17	0.18	36.50	74.00	22.00	57.11	103.67	17.43
	min	1.43	0.87	0.72	0.11	23.00	54.00	12.00	40.22	63.72	9.38
	max	2.49	1.33	1.62	0.42	58.00	88.00	31.00	71.43	157.47	38.00
B. nana × *pubescens* (number on fruits = 30)											
	mean	1.94	1.19	1.06	0.25	38.30	73.12	27.63	54.51	89.94	21.40
	median	2.03	1.17	1.09	0.22	38.50	73.50	27.50	54.78	84.48	20.00
	min	1.04	0.70	0.49	0.11	32.00	57.00	13.00	41.48	53.33	7.75
	max	2.84	1.44	1.55	0.47	50.00	94.00	40.00	73.46	163.16	41.43
B. pendula (number on fruits = 30)											
	mean	2.03	0.94	1.08	0.17	35.32	64.43	23.67	53.16	115.36	17.91
	median	2.08	0.95	1.04	0.16	35.00	64.00	23.00	53.23	115.36	17.36
	min	1.49	0.75	0.75	0.11	21.00	44.00	14.00	40.81	72.12	11.97
	max	2.49	1.17	1.32	0.30	51.00	80.00	38.00	61.50	155.42	30.61

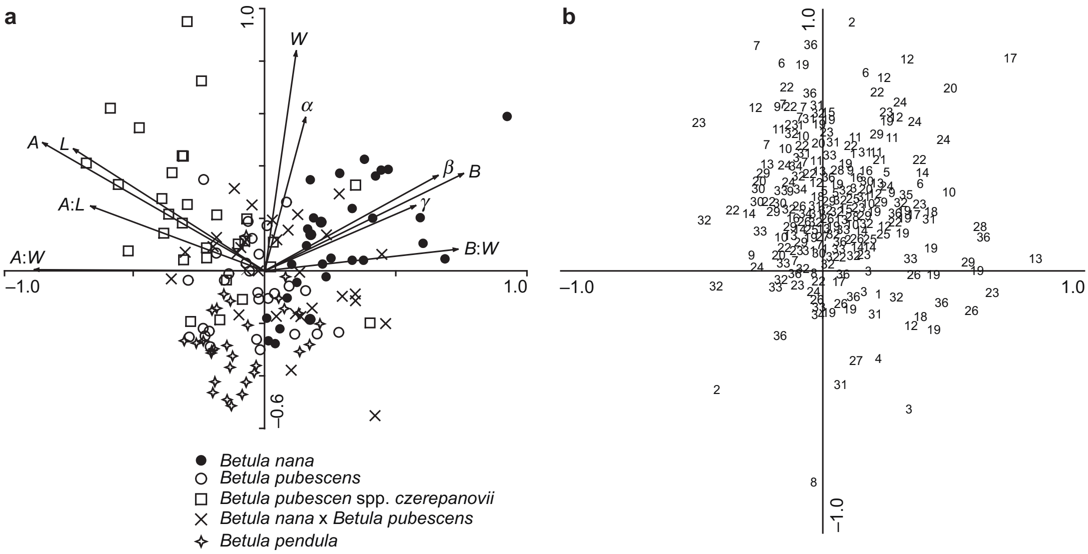

Fig. 2. Principal components analysis (PCA) ordination diagram showing the position of (**a**) modern *Betula* fruits and (**b**) subfossil *Betula* fruits in relation to the measured morphological variables. The numbers in panel **b** correspond to the depths in Fig. 3a.

tively, thus capturing 62.7% of the total variance in the data. The third (λ_3 = 0.13) and fourth (λ_4 = 0.10) PCA axes show low eigenvalues and are not discussed further. Three groups with positive correlations can be identified. The first group consists of variables indicating the shape of the fruit base (β, *B*, γ and *B*:*W*). Variables describing the length characteristics (*A*, *L*, and *A*:*W*) form the second group. These are strongly correlated to PCA axis 1. The third group is highly correlated with PCA axis 2 and comprises variables describing the width of the fruit (*W*, α). Small and round fruits of *B*. *nana* (mainly upper right quarter) and longer, slimmer and more steep-sided *B*. *pubescens* ssp. *czerepanovii* fruits (upper left quarter) are distinctively separated by PCA axis 1 with only two fruits of *B*. *pubescens* ssp. *czerepanovii* being plotted among the fruits of *B*. *nana*. These outliers may be the result of natural variation in shape (Białobrzeska & Truchanowiczówna 1960). Width (*W*) and α seem not to be important variables to separate dwarf birch from mountain birch. Instead, these features separate white birch, *B*. *pendula*, from the other birch taxa. *B*. *pendula* fruits seem to, in general, have distinctly narrow shoulders and are thus negatively correlated with the PCA axis 2. *Betula pubescens* and the hybrid *B*. *nana* × *pubescens* are plotted around the origo and cannot reliably be separated from each other based on the wingless fruits only.

The fossil *Betula* fruits are plotted as numbers (Fig. 2b), which correspond to the depths shown in Fig. 3a. At first sight it appears that the fossil fruits are scattered all around the PCA biplot. However, this exercise provides interesting and useful information when converted as a stratigraphical diagram and compared with the results of visual birch-fruit identification (Fig. 3). Based on visual analysis (Fig. 3a), subfossil fruits are categorized into three groups: *Betula nana*, tree-type *Betula*, and *Betula* sp. Most of the fruits could not be identified to a species level and are thus marked as *Betula* sp. Moreover, it is also impossible to separate the different tree-type birch fruits (*B*. *pubescens*, *B*. *pendula*, and *B*. *pubescens* ssp. *czerepanovii*) from each other. The PCA analysis, instead, separates four groups of birch fruits (Fig. 3b) providing more detailed information about the occurrences and amounts of different birch species. Despite the assumption that pure *B*. *nana* fruits should easily be distinguishable, the amount of *B*. *nana* clearly increases as a result of PCA (Fig. 3). The high and frequent abundance is in accordance with the fact that dwarf birch is one of the most dominant vascular plant inhabiting northern fjeld areas (Kallio & Mäkinen 1978). In addition, PCA also suggests that *B. pubescens* ssp. *czerepanovii* has been present in the catchment almost throughout the time under examination. This fits the conception of Hämet-Ahti (1987), who suggests that mountain birch actually originates from isolated *B. pubescens* populations that survived in northwestern Europe through the last glaciation. PCA also suggests that *B. pendula* has been present in the area. This is interesting, since according to Kallio and Mäkinen (1978) in Finland the modern northern limit of *B. pendula* is ca. 69°N. The fruits mainly occur at the lower part of the core that most probably represents the time period of early to mid Holocene. The sediment core is not dated, but based on ^{14}C radiocarbon dates derived from three lakes being situated in close vicinity, with sediment thicknesses of 209 cm, 240 cm and 140 cm, respectively, and representing at least the last 7000 years (R. Bradshaw pers. comm.) this sediment sequence (185 cm) likely covers at least the same time span. Furthermore, since the long-term average sedimentation rate in subarctic/arctic lakes is shown to be ca. 0.1–0.2 mm a^{-1} (e.g. Barnekow 1999, Sorvari 2001, Korhola & Weckström 2004), the sedimentation probably started during the early Holocene. Another evidence for early Holocene inception is the finds of *Nuphar* sp. from the bottom part of the core. This species cannot survive in subarctic lakes, thus, its presence indicates warmer than present climate conditions. According to Mäkelä and Hyvärinen (2000), *B. pendula* had a much wider northern distribution during the Holocene thermal optimum 8000–5000 years ago but the distribution decreased towards the present due to the slowly cooling climate. Thus, in theory it is possible that *B. pendula* colonised the catchment after deglaciation. Alternatively, the fruits might represent the northern form of the silver birch *B. pendula* var. *lapponica*. Unfortunately, this could not be verified, since no modern reference fruit-material was available. A possible misinterpretation of PCA, resulting from the internal variation in fruit morphology reported e.g. by Białobrzeska and Truchanowiczówna (1960) must also be considered. However, PCA seems to separate the modern *B. pendula* fruits fairly well from the other birch fruits (Fig. 2a).

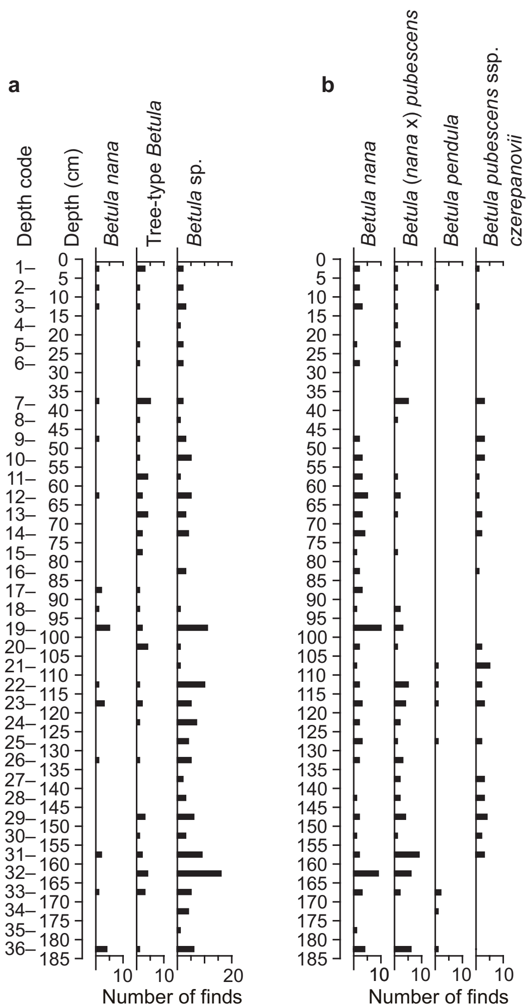

Fig. 3. The stratigraphical distribution of subfossil *Betula* fruit finds from a northern Norwegian lake based on (**a**) visual interpretation and (**b**) PCA.

PCA proves to be a helpful tool to separate subfossil birch fruits from each other and is a recommendable aid to use together with the visual analysis. In this study, the drawings of fruits as well as the variable measurements were done manually but modern computer-aided drawing and measuring techniques can provide a less time-consuming way to copy-draw fossil fruits. In order to verify the possible occurrence of *B. pendula* in northern Norway during the Holocene thermal optimum more studies are needed. In the near future DNA analysis might provide new insights of the phytogeographical history of different birch species in Fennoscandia.

Acknowledgements

This research was financed by the DART-project (ENV4-CT97-0586), the Academy of Finland (Grant 204291) and HERC (Helsinki University Environmental Research Centre) (Grant 470034467). We thank two anonymous reviewers for valuable comments.

References

Aas, B. & Faarlund, T. 2001: The Holocene history of the Nordic mountain birch belt. — In: Wielgolaski, F. E. (ed.), *Nordic mountain birch ecosystems*: 5–22. The Parthenon Publ. Group, New York, London.

Barnekow, L. 1999: Holocene tree-line dynamics and inferred climatic changes in the Abisko area, northern Sweden, based on macrofossil and pollen records. — *Holocene* 9: 253–265.

Białobrzeska, M. & Truchanowiczówna, J. 1960: The variability of shape of fruits and scales of the European birches (*Betula* L.) and their determination in fossil materials. — *Monogr. Botanicae* 9(2): 87–92.

Birks, H. H. 2001: Plant macrofossils. — In: Smol, J. P., Birks, H. J. B. & Last, W. M. (ed.), *Tracking environmental change using lake sediments*, vol. 3: 49–74. Kluwer Acad. Publ., Dordrecht.

Birks, H. H. 2003: The importance of plant macrofossils in the reconstruction of late-glacial vegetation and climate: examples from Scotland, western Norway, and Minnesota, USA. — *Quat. Sci. Rev.* 22: 453–473.

Everitt, B. S. 1978: *Graphical techniques for multivariate data*. — London, Heinemann Educ. Books Ltd.

Freund, H., Birks, H. H. & Birks, H. J. B. 2001: The identification of wingless *Betula* fruits in Weichselian sediments in the Gross Todtshorn borehole (Lower Saxony, Germany) — the occurrence of *Betula humilis* Schrank. — *Veget. Hist. Archaeobot.* 10: 107–115.

Hämet-Ahti, L. 1987: Mountain birch and mountain birch woodland in NW Europe. — *Phytocoenologia* 15: 449–453.

Hämet-Ahti, L., Suominen, J., Ulvinen, T. & Uotila, P. (eds.) 1998: *Retkeilykasvio*, 4th ed. — Finnish Mus. Nat. Hist., Bot. Mus., Helsinki.

Hyvärinen, H. 1975: Absolute and relative pollen diagrams from northernmost Fennoscandia. — *Fennia* 142: 1–23.

Kallio, P. & Mäkinen, Y. 1978: Vascular flora of Inari Lapland. 4. Betulaceae. — *Rep. Subarctic Res. Station Kevo* 14: 38–63.

Korhola, A. & Weckström, J. 2004: Paleolimnological studies in Arctic Fennoscandia and the Kola Peninsula (Russia). — In: Pienitz, R., Douglas, M. S. V. & Smol, J. P. (eds.), *Long-term environmental change in Arctic and Antarctic lakes*, vol. 8: 381–418. Kluwer Academic Publishers, Dordrecht.

Mäkelä, E. 1996: Size distinctions between *Betula* pollen types — a review. — *Grana* 35: 248–256.

Mäkelä, E. & Hyvärinen, H. 2000: Holocene vegetation history at Vätsäri, Inari Lapland, northeastern Finland, with special reference to *Betula*. — *Holocene* 10: 75–85.

Seppä, H. 1996: Post-glacial dynamics of vegetation and tree-lines in the far north of Fennoscandia. — *Fennia* 174: 1–96.

Seppä, H., Hannon, G. & Bradshaw R. H. W. 2004: Holocene history of alpine vegetation and forestline on Pyhäkero Mountain, northern Finland. — *Arct. Antarc. Alp. Res.* 36(4): 593–600.

Sorvari, S. 2001: *Climate impacts on remote subarctic lakes in Finnish Lapland: limnological and palaeolimnological assessment with a particular focus on diatoms and Lake Saanajärvi*. — Kilpisjärvi Notes 16, Yliopistopaino, Helsinki.

ter Braak, C. F. J. 1994: Canonical community ordination. Part I: Basic theory and linear methods. — *Écoscience* 1: 127–140.

ter Braak, C. J. F. & Šmilauer, P. 1998: *CANOCO reference manual and user's guide to Canoco for Windows: software for canonical community ordination*, ver. 4. — Microcomputer Power, Ithaca, NY.

ter Braak, C. J. F. & Šmilauer, P. 2002: *CANOCO Reference manual and CanoDraw for Windows User's guide: Software for Canonical Community Ordination*, ver. 4.5. — Microcomputer Power, Ithaca, NY.

Tolonen, K. & Ruuhijärvi, R. 1976: Standard pollen diagrams from the Salpausselkä region of southern Finland. — *Ann. Bot. Fennici* 13: 155–196.

Vaarama, A. & Valanne, T. 1973: On the taxonomy, biology and origin of *Betula tortuosa* Ledeb. — *Rep. Subarctic Res. Station Kevo* 10: 70–84.

Valanne, T. & Sulkinoja, M. 1991: Lapin koivujen risteytymisestä. — *Luonnon Tutkija* 95: 83–87.

Väliranta, M. 2005: *Plant macrofossil evidence of changes in aquatic and terrestrial environments in north-eastern European Russia and Finnish Lapland since the late Weichselian*. — Publ. Dept. Geol. 2, Univ. Helsinki, Yliopistopaino, Helsinki.

Van Dienter, M. & Birks, H. H. 1996: Dishtinguishing fossil *Betula nana* and *B. pubescens* using their wingless fruits: implications for the late-glacial vegetational history of western Norway. — *Veget. Hist. Archaeobot.* 5: 229–240.

Vasari, Y. 1962: A study of the vegetational history of the Kuusamo district (north east Finland) during the late-Quaternary period. — *Ann. Bot. Soc.* "*Vanamo*" 33(1): 1–140.

Vasari, Y., Tonkov, S., Vasari, A. & Nikolova, A. 1996: The Late-quaternary history of the vegetation and flora in northeastern Finland in the light of a re-investigation of Aapalampi in Salla. — *Aquilo Ser. Bot.* 36: 27–41.