BENTHIC DIATOM COMMUNITY STRUCTURE IN BOREAL STREAMS

DISTRIBUTION PATTERNS ALONG ENVIRONMENTAL AND SPATIAL GRADIENTS

Janne Soininen

Department of Biological and Environmental Sciences, P.O. Box 65, FIN-00014
University of Helsinki, Finland

Academic dissertation in limnology.

To be presented, with the permission of the Faculty of Biosciences, for public criticism in Auditorium XII, Unioninkatu 34, on 27th September, at 12 noon.

Helsinki 2004

Author`s address:
Department of Biological and Environmental Sciences
P.O. Box 65, FIN-00014 University of Helsinki, Finland
e-mail: janne.soininen@helsinki.fi

Supervisor:
Pertti Eloranta, Prof.
Department of Biological and Environmental Sciences
P.O. Box 65, FIN-00014 University of Helsinki, Finland
e-mail: pertti.eloranta@helsinki.fi

Reviewers:
Eugen Rott, Prof.
Institute for Botany
University of Innsbruck
A-6020 Innsbruck
Austria
e-mail: eugen.rott@uibk.ac.at

Sergi Sabater, Prof.
Department of Environmental Sciences
University of Girona
Campus Montilivi
17071 Girona
Spain
e-mail: sergi.sabater@udg.es

Opponent:
Helmut Hillebrand, Prof.
Institute for Botany
University of Cologne
Gyrhofstrasse 15
D-50931 Köln
Germany
e-mail: helmut.hillebrand@uni-koeln.de

ISBN 952-91-7426-8 (nid.)
ISBN 952-10-1937-9 (PDF)
http://ethesis.helsinki.fi
Yliopistopaino
Helsinki 2004

ABSTRACT

The past decade has seen growing appreciation of the role of regional influences in determining the structure of local communities. An emerging view among ecologists is that local community composition is controlled by acting of nested filters which select species with suitable traits for prevailing conditions, thus leading communities regulated by local environmental factors and regional, mainly historical or dispersal related factors. Running waters are naturally open, hierarchical and heterogeneous ecosystems. This heterogeneity prevails in physical, chemical and biological elements across multiple spatial and temporal scales. The growing and prospering of benthic algae in streams is the outcome of complex interactions between hydrological, chemical and biotic factors. Diatoms constitute a major part of the cell and species number in benthic algal communities offering the most useful algal community for studying large-scale ecological patterns in stream ecosystems.

The major aims of this thesis were (i) to find the main factors regulating benthic diatom community structure in boreal streams at different spatial scales, (ii) to test the correspondence between ecoregional delineations and spatial patterns in community structure, (iii) to assess seasonal community persistence and stability of benthic diatom communities and (iv) to investigate if benthic diatoms offer a usable tool for water quality assessment. Results of direct ordinations emphasized the predominance of chemical-constituent concentration and ion composition on structuring benthic diatom communities of running waters. Conductivity was the strongest environmental gradient explaining diatom distribution patterns in Finnish running waters at the national scale. The other important determinants of diatom community structure were latitude, pH, total P, and water colour.

Results of this thesis showed that diatom communities exhibit a rather strong spatial component especially at a national scale. This was shown both by variation partitioning and by a direct test of congruence between diatom community structure and the spatial coordinates of the sampling sites. The proportion of variation explained independently by spatial factors was quite large, ca. 25 %, at the largest, national, scale. Furthermore, it seems that even at rather small spatial scales (ca. 10^2 km), pure spatial component still plays an role in regulating benthic diatom community composition. Moreover, data of this thesis support also the view that beta-diversity of benthic diatoms might be higher than previously believed.

When studying temporal patterns of community structure, stability tended to be lowest among epiphytic communities. Moreover, species turnover seemed to be highest among epiphyton and lowest among epipelic communities. Although these differences could also result from lower diversity in epiphyton, they probably indicate lower persistence among epiphytic communities in boreal streams. For bioassessment needs, diatom-based weighted averaging models offer usable tool for water quality monitoring of boreal streams. Given the strong spatial patterns in community composition, it seems evident that bioassessment programs utilising lotic diatoms would benefit from geographical stratification, using e.g. ecoregions or subecoregions as regional delineations. However, since local in-stream factors were even more important than spatial factors in explaining diatom distributions, a combination of regional stratification and local environmental features might provide the most suitable framework for diatom-based bioassessment of boreal streams.

List of papers

This thesis is based on the following articles referred to in text by Roman numbers (I-VI).

I Soininen, J., Paavola, R. & Muotka, T. 2004: Benthic diatom communities in boreal streams: community structure in relation to environmental and spatial gradients. *Ecography* 27: 330-342.

II Soininen, J. 2004: Determinants of benthic diatom community structure in boreal streams: the role of environmental and spatial factors at different scales. *Int. Rev. Hydrobiol.* 89: 139-150.

III Soininen, J. 2002: Responses of epilithic diatom communities to environmental gradients in some Finnish rivers. *Int. Rev. Hydrobiol.* 87: 11-24.

IV Soininen, J. & Eloranta, P. 2004: Seasonal persistence and stability of diatom communities in rivers: are there habitat specific differences? *Eur. J. Phycol.* 39:153-160.

V Soininen, J. & Niemelä, P. 2002: Inferring the phosphorus levels of rivers from benthic diatoms using weighted averaging. *Arch. Hydrobiol.* 154:1-18.

VI Soininen, J. & Könönen, K. 2004: Comparative study of monitoring South-Finnish rivers and streams using diatom and macroinvertebrate community structure. *Aquat. Ecol.* 38: 63-75.

Author`s contribution

I Study was planned jointly. Janne Soininen and Timo Muotka wrote the paper jointly. Riku Paavola made the main statistical analyses.

IV Study was planned jointly. Janne Soininen wrote the paper and Pertti Eloranta planned and conducted the sampling.

V Janne Soininen planned the study and wrote the paper. Pirjo Niemelä provided part of the diatom data.

VI Study was planned jointly. Janne Soininen wrote main part of the paper.

INTRODUCTION

1.1 Benthic algae in streams

Running waters are naturally open, hierarchical and heterogeneous ecosystems. This heterogeneity prevails in physical, chemical and biological elements across multiple spatial and temporal scales. The hierarchical system of running water ecosystems consists of drainage systems, streams within drainage systems, stream segments and reaches within streams, pool-riffle sequences within reaches and microhabitats within pools or riffles (Frissel *et al.*, 1986). These spatial scales are linked to corresponding temporal scales of natural physical phenomena occurring in river ecosystems. Although many running waters receive their major energy inputs from allochtonous sources (see e.g. Fisher & Likens, 1973; Peterson *et al.*, 1986), autochtonous production often has a notable role in stream ecosystem energy budgets (Fisher & Carpenter, 1976; Fisher *et al.*, 1982). According to the River Continuum Concept (Vannote *et al.*, 1980), proportion of primary production to respiration attains its highest level in the middle reaches of the river continuum. Conditions (e.g. substratum type, current velocity and light) favor periphyton growth, or if epipelic habitats are present, the growth of benthic algae. Moreover, if substratum type and current velocity are suitable, vascular hydrophytes and aquatic bryophytes can be present as well. Phytoplankton maintains true potamoplanktonic communities only in the widest lowland rivers (Allan, 1995). In smaller streams, “phytoplankton” consists of drifting algae detached from the bottom substratum or from lakes and ponds upstream.

The primary groups of algae present in running waters are blue-green algae (Cyanophyta), green algae (Chlorophyta), diatoms (Bacillariophyta) and red algae (Rhodophyta). Their main growth form or morphology in stream benthos excluding benthic diatoms, is filamentous. Benthic diatoms are typically unicellular, but can form colonies or chain-like structures also in benthos (Stevenson *et al.*, 1996). Many of the green algal filaments (e.g. genera like *Cladophora*, *Spirogyra* and *Ulothrix*) are macroscopic even as individual filaments, whereas most of the other algae are macroscopic only in mass occurrences.

One of the most distinct features of benthic algae is a substantial heterogeneity of biomass and species composition prevailing at multiple spatial and temporal scales. Benthic algae typically form complex, multi-layered matrix of unicellular, colonial and filamentous morphologies entangled with a mixture of other organisms (e.g. microbes, meiofauna) living on the substrate (Stevenson *et al.*, 1996). Diatoms constitute a major part of the cell and species number in benthic algal communities. A significant part of algal primary production, and even biomass can be assigned to diatoms, if filamentous algae are scarce.

The growing and prospering of benthic algae in streams is the outcome of complex interactions between hydrological, chemical and biotic factors (Stevenson *et al.*, 1996, Fig. 1). Local “proximate” variables, like discharge regime, are controlled by regional “ultimate” factors like geology, topography or climate operating at spatial scales of catchments or even ecoregions. In addition, human activities act to change both proximate and ultimate variables in an increasing rate, leading towards variously impacted biological communities, e.g. algal communities with increased amount and biomass of nuisance species, or in general, impoverished biological communities.

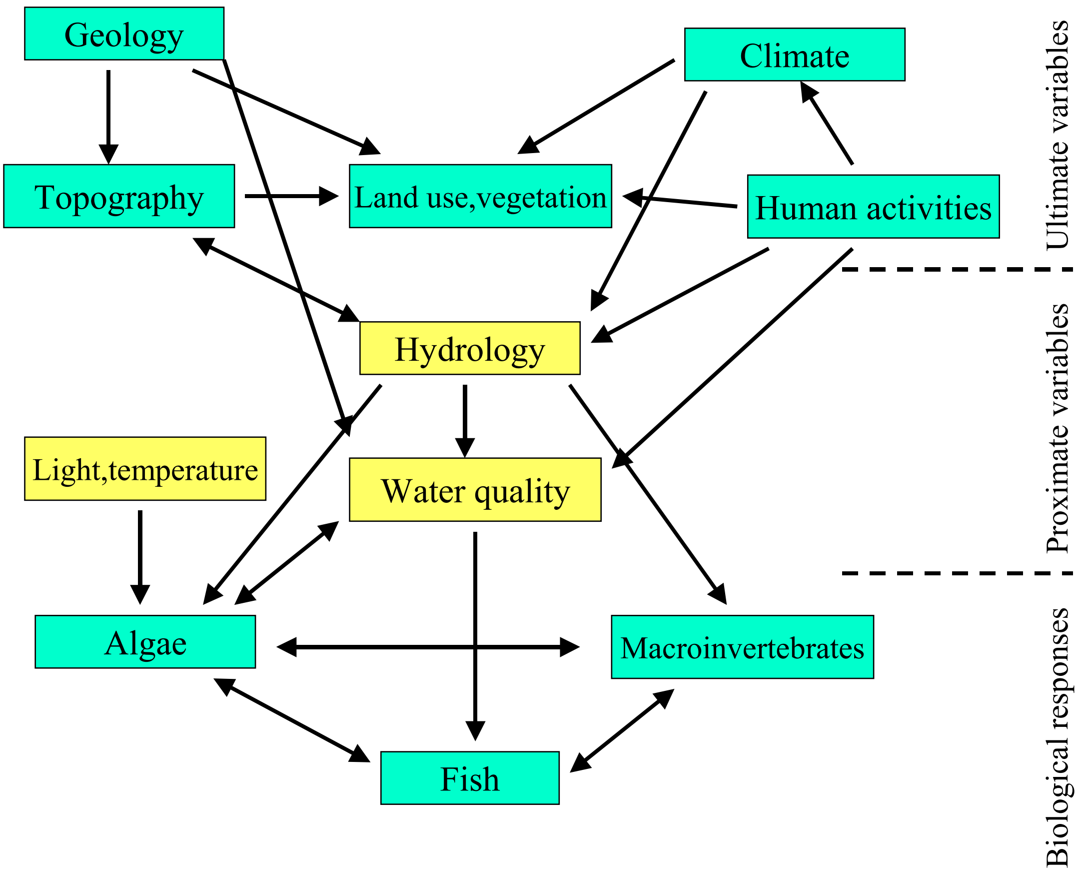

Fig 1. Diagram showing how ultimate landscape variables control the proximate physical and chemical variables of streams, which in turn control biological responses (modified from Biggs *et al.*, 1990).

Factors that potentially influence benthic algal communities include light, temperature, current, substrate, scouring effects of floods, water chemistry and grazing (Hynes, 1970; Whitton, 1975). Fluctuations in discharge cause changes in channel width, depth and current velocity. Therefore, discharge regime plays frequently an overriding role in the regulation of production, biomass and community composition of benthic organisms in general (e.g. Hart & Finelli, 1999).

1.2 Determinants of community structure: the role of local and regional factors

The structure of local communities has traditionally been considered to be regulated mainly by local physical and chemical factors. Recently it has been argued that community patterns are regulated by interacting local and regional factors, prevailing at multiple spatial and temporal scales (Ricklefs, 1987; Menge & Olson, 1990; Levin, 1992; Zobel, 1997; Lawton, 1999). To join a local community, every species in a regional pool must pass a series of nested “filters” (Poff, 1997; Zobel, 1997; Lawton, 1999) (Fig. 2). Filters are scaled habitat features that influence the probability that taxa with specified traits are able to join and persist as a member of a local community (Poff, 1997). All species are assumed to be capable of dispersing to all locales in a region. Therefore, the absence or very low abundance of a species reflects the action of selective forces or, in fact, habitat features prevailing at multiple scales (Tonn *et al*., 1990). To pass through a filter, a species must possess appropriate functional traits matching the selective characteristics of the filter.

The "history filter" determines the regional species pool, and it consists of large-scale historical, climatic and evolutionary factors such as migration and speciation (Hillebrand & Blenckner, 2002). In boreal areas, recurring glacial periods can be considered an important historical climatic factor influencing stream biota (Brown & Lomolino, 1998). Richness of the regional species pool, dispersal distance and the abundance of propagules are main factors determining the "dispersal filter". An "environmental filter" consists of habitat features, which affect e.g. to species adaptation to local abiotic conditions, their resistance to changes in physical and chemical conditions and grazing, and competitive ability in a local community.

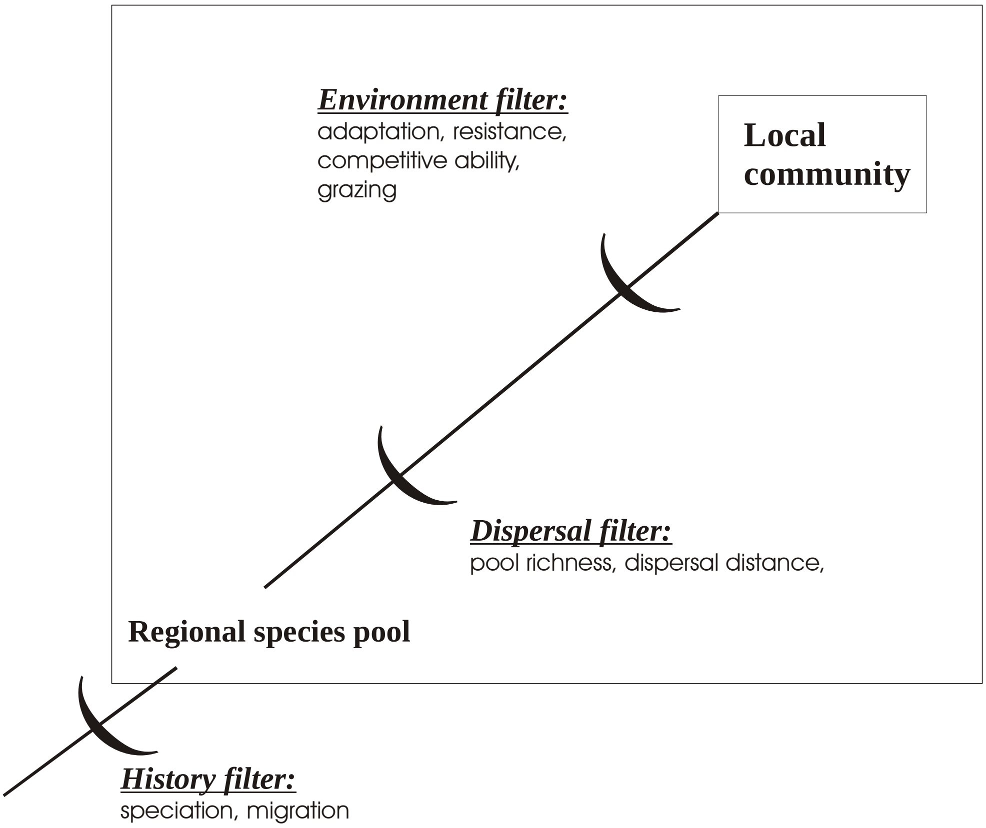

Fig 2. Conceptual model visualizing the assembly of local communities through series of nested filters (modified from Hillebrand & Blenckner, 2002).

Benthic diatom communities are traditionally considered as being regulated more by local environmental conditions than by broad-scale climatic, vegetational, and geological factors (Pan *et al.*, 1999, 2000). Water chemistry, in particular, has been considered to set strong local environmental filter regulating diatom communities. On a more general level, it has been argued that the species composition of small (especially unicellular) organisms is dominated by cosmopolitan species with high dispersal ability (e.g. Finlay *et al.*, 1996; Fenchel *et al.*, 1997). In addition, small organisms have long evolutionary history and perceive their environment with a fine resolution, but as homogeneous at macrospatial scales (Azovsky, 2002). Therefore, local factors should be much more important than regional ones, acting as strong environmental filter selecting species able to cope with the prevailing conditions. Consequently, communities of unicellular organisms should be characterized by a high local species richness compared to regional or global richness, that is, they should have low turnover (β) diversity (Finlay *et al.*, 1996; Fenchel *et al.*, 1997; Hillebrand & Azovsky, 2001; Azovsky, 2002).

Recently, however, this view have been challenged, and Hillebrand *et al.* (2001) found that although macroecological patterns documented for multicellular organisms differ from those reported for unicellular communities, there is in fact no strict evidence showing that unicellular organisms exhibit higher local species richness than metazoans. For freshwater diatoms in particular, the concept of predominantly cosmopolitan distribution has been strongly criticized by Kociolek & Spaulding (2000; see also Mann & Droop, 1996), who argued that a considerable proportion of diatoms seems to be endemic or at least show a regionally restricted distribution. Studies stressing macroecological questions among benthic diatoms are still very few (but see Potapova & Charles, 2002 and papers **I** and **II**) especially in boreal areas. In conclusion, the relative roles of local and regional factors as determinants of biological communities need to be thoroughly studied in the future, especially in aquatic ecosystems and among unicellular organisms.

Assembly rules for local communities, regulated by the interplay of local and regional factors, are macroecological questions studied intensively in different fields of ecology (e.g. Lawton, 1996; Gaston & Blackburn, 1999; Lawton, 1999; Blackburn & Gaston, 2003). Macroecology or ecological biogeography (see Brown & Lomolino, 1998) is primarily based on empirical data and therefore, hypotheses are not easily testable (see McGill, 2003). The strength of observed patterns depends on the extent to which various mechanisms act in concert; clear patterns arise when several processes act in the same direction (Gaston & Blackburn, 1999). Given their interdisciplinary nature, macroecological questions operate under multiple frameworks, and thus, observed patterns can have multiple explanations (Gaston & Blackburn, 1999). The central question becomes not which explanation is the correct one, but what are their relative roles.

1.3 Use of diatoms in bioassessment in streams

Biological indicators describe water quality and its changes over a long time scale more reliably than a few, discrete physicochemical analyses. Especially in running waters, where concentrations can fluctuate notably even within a few hours, biological monitoring has been proven to be useful (e.g. Whitton *et al.*, 1991; Prygiel & Coste, 1993; Rosenberg & Resh, 1993; Whitton & Rott, 1996). Benthic diatoms have been found to be practicable for river monitoring purposes in several European studies (see e.g. Whitton *et al*., 1991; Whitton & Rott, 1996; Prygiel *et al*., 1999). In Finland also, the applicability of diatoms in water quality assessments has been tested recently (e.g. Eloranta, 1995; Eloranta & Andersson, 1998; Eloranta & Soininen, 2002). Diatoms are very suitable bioindicators because their ecology is generally known rather well. In addition, diatom cell densities and number of local taxa are usually very high. Compared to benthic macroinvertebrates, diatoms are considered more sensitive indicators of water chemistry owing to their shorter life cycles and nature as primary producers (e.g. Steinberg & Schiefele, 1988).

Due to difficulties in monitoring rapidly fluctuating nutrient levels in rivers, diatom based tools are very useful in estimating the trophic status of a river (Kelly & Whitton, 1995; Kelly, 1998). The weighted averaging (WA) regression and calibration method dates back to Gause (1930) and it was reintroduced and developed by e.g. ter Braak & Looman (1986) and ter Braak & Prentice (1988). It is based on the theory and observation that the relationship between abundances or the probability of occurrence of the taxa and environmental variables is often unimodal (ter Braak & Looman, 1986; ter Braak & van Dam, 1989). A taxon will be most abundant in concentrations close to its phosphorus optimum, from which the expected abundance or probability of occurrence will gradually decrease. WA models have been widely used in paleoecological studies that infer the past water quality of lakes (usually phosphorus or pH), using e.g. sediment diatoms, chrysophytes or chironomids (e.g. Hall & Smol, 1992, 1996; Christie & Smol, 1993; Weckström *et al*., 1997; Korhola *et al*., 1999; Miettinen, 2003). There are only a few studies that have inferred total P concentrations of rivers using diatom communities (but see Winter & Duthie, 2000 and **V**).

When viewed across large areas, stream communities frequently exhibit a strong spatially-structured variation (Li *et al*., 2001; Heino *et al*., 2003a; Parsons *et al*., 2003). It is therefore important that the relative roles of local environmental variables vs. large-scale spatial factors be reliably identified. If such spatial structuring proves to be a rule, stream bioassessment programs may benefit from regional stratification, based on *a priori* delineations. Ecoregions provide a reasonable starting point for such spatial stratification. But because of their generally non-aquatic origin (e.g. climate, geology, vegetation cover, land use, etc.), they should be rigorously tested before accepted as an appropriate level of spatial resolution for long-term biomonitoring of freshwater communities. Ecoregion-level differences in freshwater communities have been mainly studied on macroinvertebrates (Hawkins & Vinson, 2000; Johnson, 2000; Sandin & Johnson, 2000; Heino *et al*., 2002) and fish (McCormick *et al*., 2000; Van Sickle & Hughes, 2000). Corresponding studies on benthic algae are rare, and they have shown subtle regional patterns in algal community structure (Whittier *et al*., 1988; Pan *et al*., 1999, 2000).

1.4 The main objectives of the study

Running waters in Finland typically have rather low conductivity and high humus content. Nevertheless, pristine or near-pristine streams in Finland exhibit distinct geographical, especially north-to-south, patterns in their physicochemical characteristics, largely paralleling regional trends in geology, soil type, topography, land use, and potential natural vegetation (Heino *et al*., 2002). The present thesis focuses on patterns of benthic diatom community structure in relation to environmental (chemical and physical) and spatial factors (latitude and longitude) in boreal streams. About half of the study sites are “reference” sites in near-pristine condition, being mostly small headwater streams. The rest of the sites are “impacted” by human activities, mainly by load of nutrients and suspended solids from agriculture, or nutrient and organic load by treated sewage.

The main objectives of this thesis are:

1. To find the main factors regulating benthic diatom community structure in boreal streams, and to assess the relative contributions of environmental and spatial factors as determinants of benthic diatom community structure at different spatial scales (**I-III**).

2. To describe the main diatom community types and their best indicator species in Finnish boreal streams (**I, III**), and to study how distinct these biologically defined community types are, and how well they can be predicted using environmental variables.

3. To test if a regional classification scheme based on terrestrial landscapes (ecoregions) provides a reasonable framework for a corresponding regional grouping of streams according to their benthic diatom communities (**I**).

4. To assess seasonal community persistence and stability of benthic diatom communities in different habitats (**IV**).

5. To investigate if benthic diatoms are a usable tool for water quality assessment, and in particular, in estimation of phosphorus concentrations of running waters (**III, V**).

6. To compare the performance of diatom and macroinvertebrate community structure as tools for water quality assessment (**VI**).

2. MATERIAL AND METHODS

2.1 Study area

Papers I-III, V:

The study area was composed of 197 (**I, II**) (Fig. 3), 146 (**III**) and 145 (**V**) stream sites in Finland. The sampled sites were chosen to cover long gradients in conductivity, pH, humus, and nutrient concentrations (see papers **I, III** and **V** for details). Diatoms were sampled in 1986 and between 1996-2001. A set of 56 sites sampled in 1986 was included, because most of them represent near-pristine conditions, being only slightly affected by agriculture and fish farming (Eloranta, 1995; see also Eloranta & Kwandrans, 1996). Furthermore, recent visits to these sites verified that they (stream channel + riparian zone) had not been modified to any noticeable degree between 1986 and 1996, so these samples were considered to be comparable with the rest of the material. In paper **I** and **II**, diatom material represented all the five ecoregions of Finland, i.e. hemiboreal, south boreal, middle boreal, north boreal, and arctic-alpine ecoregions (Fig. 3). Ecoregions were defined using the delineations of Alalammi & Karlsson (1988) based on climate, relief, vegetation, and land use. Since some of the ecoregions span large areas known to differ in many features important to freshwater biota (Heino *et al*., 2002), our data were further stratified according to subecoregions, based on major drainage systems and regional landscape characteristics within each ecoregion, mainly following Alalammi & Karlsson (1988). For a more detailed description of the five ecoregions and subecoregions, see paper **I,** and Heino *et al*. (2002). In papers **III** and **V**, sampled stations were primarily the same, with the exclusion of 47 near pristine streams in northern and eastern parts of Finland. In paper **I**, sampling stations were classified into reference sites (near-pristine or, especially in southern Finland, least impacted stream conditions) and impacted sites. For a reference site, the level of catchment disturbance (mainly forestry or agriculture) had to be less than 10 %, and the integrity of the riparian zone (% human disturbance in the water-riparian ecotone, assessed in situ) had to be more than 80 %.

In paper **II**, three hierarchical spatial scales were used in ordinations. The largest scale was the whole of Finland (scale ca. 10^3 km, 197 study sites) (Fig. 3), second largest was the ecoregion (three regions, scale ca. 10^2 km; 92, 47 and 33 study sites), and the smallest was the river system R. Vantaanjoki (scale 10 km-10^2 km, 21 study sites). In paper **V**, the calibration set was comprised of 97 sampling sites and the test set of 47 sites.

Papers IV, VI:

The study in paper **IV** was conducted in three boreal rivers in southern Finland. Epilithic, epiphytic and epipelic diatoms were sampled monthly at four sampling stations from June to October. Sampled rivers were impacted mainly by nutrients from agriculture and by nutrients and organic compounds from treated sewage. River Keravanjoki was sampled at two stations, River Porvoonjoki and River Mustijoki at one station.

In paper **VI**, the study sites consisted of eutrophic or moderately nutrient rich rivers and some smaller, less eutrophic streams situated in southern Finland. Most of the rivers drained cultivated land, and consequently were rather turbid. Epilithic diatoms were collected at 52 sampling stations between 1997 and 2000. Benthic macroinvertebrates were sampled at 22 stations in years 2000 and 2002.

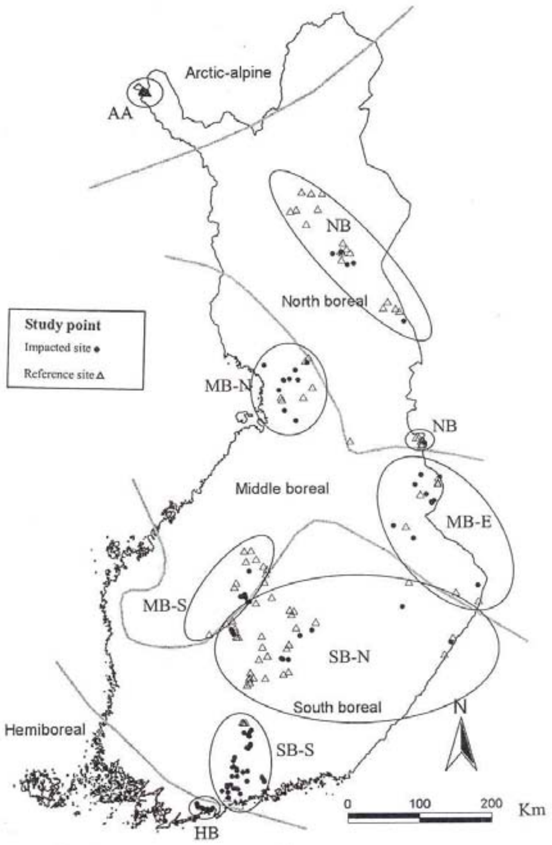

Fig. 3. Map of Finland showing the locations of the sampling sites within the five ecoregions of Finland (**I, II**). Middle boreal and South boreal ecoregions are further divided into subecoregions. Ecoregions and subecoregions were delineated according to Alalammi &Karlsson (1988). Abbreviations: AA = Arctic-alpine, NB = North boreal, MB-N = Middle boreal northern, MB-E = Middle boreal eastern, MB-S = Middle boreal southern, SB-N = South boreal northern, SB-S = South boreal southern and HB = Hemiboreal.

Simultaneous sampling of diatoms and benthic fauna was done at 12 stations in summer and autumn 2000. River Ingarskilaån, R. Siuntionjoki, R. Vantaanjoki and R. Keravanjoki are moderately eutrophic turbid rivers near the southern coast of Finland. They are influenced by agriculture, mainly cultivation. Stations at R. Vantaanjoki and R. Keravanjoki are impacted also by treated sewage. Nutrient concentrations, turbidity and conductivity are rather high (**VI**). Three of the sampled streams, Glomsån, Glimsån and Myllypuro are less impacted and especially Glomsån and Myllypuro are more oligotrophic.

2.2 Sampling and in-stream measurements

Diatoms were sampled by brushing stones with a toothbrush, following the recommendations of Kelly *et al.* (1998) (papers **I-III**, **V**, **VI**). At least five, pebble-to-cobble (5-15 cm) sized stones were collected from the stream bottom. They were brushed and the diatom suspension was put in a small plastic bottle. In R. Pikkujoki, diatom samples were taken above and below a sewage treatment plant for detecting the impact of effluents on diatom community structure (**III**). In paper **IV**, epilithic, epiphytic and epipelic diatoms were sampled at four sampling stations monthly from June to October. Epilithic diatoms were sampled following Kelly *et al.* (1998). Epiphytic samples were taken by brushing the undersurfaces and petioles of at least five *Nuphar lutea* leaves and the epipelic samples were taken from sediment surfaces using a pipette. In all studies, diatom samples were preserved in ethanol or formaldehyde. In papers **I-III** and **V-VI**, sampling was conducted during low flow conditions from June to August. Macroinvertebrates were sampled at shallow, fast flowing riffle sites using a standardized kick-method (SFS 5077, 1989; Kantola *et al*., 2001) (**VI**). The kicknet (net frame 25 x 30 cm, mesh size 0,5 mm) was placed on the streambed and the bottom material was kicked with a foot placed immediately upstream from the net for 30 seconds. All loose material (sand, cobbles, stones, macrophytes, bryophytes) from the upper 5-10 cm layer of the bottom substrate, carried by the current or kicking movement into the net, was included in the sample. Five (or three) samples from each site were taken from different types of microhabitats in the studied riffles to get a representative sample of the local species pool present at a site. Samples were preserved in 70 % ethanol and analysed separately.

At most of the stations, water samples were taken simultaneously with diatom samples (**I-VI**). Samples were analysed for at least water colour, conductivity, pH, and total phosphorus using national standards. Some of the sites are part of the national water quality database. For these sites water chemistry data were taken from the database, using results of the nearest sampling occasion. Current velocity was measured at each sampling site along transects (n = 5) and perpendicular to the flow, using a current meter (Seba 735) and covering the whole study section (ca. 20-30 m). Shading by the riparian canopy was visually estimated on a scale from zero to five. Stream width was also measured at each study site.

2.3 Laboratory analyses

Diatom samples were cleaned from organic material in the laboratory using wet combustion with acid (HNO_3: H_2SO_4; 2:1) and mounted in Dirax or Naphrax (**I-VI**). Two or three replicate slides of each sample were prepared. A total of 250-500 frustules per sample were identified and counted using phase contrast light microscopy (magnification 1000x). Species were identified according to Krammer & Lange-Bertalot (1986-1991) and Lange-Bertalot & Metzeltin (1996). In total, 430 diatom samples were counted for this thesis by the author during years 2000-2002. Benthic fauna was sorted in the laboratory on a white, shallow tray and identified to the lowest feasible taxonomic level (**VI**).

2.4 Data analyses

Diatom taxa occurring in at least two or three samples, with a relative proportion of 1 % or more in at least one sample were included in the statistical analyses. Species abundances were arcsine square root- or log-transformed. Major statistical analyses used in the papers are shown in Table 1.

The major patterns of community compositions and maximum amount of variation in the data were described using Detrended Correspondence Analysis (DCA) (Hill & Gauch, 1980) **(II, III, V, VI**). Rare species were downweighted in all DCA ordinations. DCA was performed using program PC-ORD version 4 (McCune & Mefford, 1999). In papers **I, II** and **IV** Non-Metric Multidimensional Scaling (NMDS) was used to describe major patterns in diatom community composition. NMDS is highly suitable for ecological data containing numerous zero values (Minchin, 1987). A three or two-dimensional solution was chosen depending on the strength of change in stress value on sequential dimensions. Sorensen`s distance measure was used. NMDS was performed using program PC-ORD version 4 (McCune & Mefford, 1999).

Two-way indicator species analysis (TWINSPAN) was used to define diatom community types (**I**). TWINSPAN is based on reciprocal averaging (Gauch, 1982), and it is widely used in freshwater ecology and bioassessment (e.g. RIVPACS, see Wright *et al*., 1984 and Wright *et al*., 1998). Despite its drawbacks (see Legendre & Legendre, 1998), TWINSPAN has been shown to perform well in the classification of benthic assemblages when compared with alternative clustering techniques (Moss *et al*., 1999). The statistical significance of differences between the community composition of different TWINSPAN groups **(I)** or in different habitats **(IV)** was tested using Multi-

Response Permutation Procedures (MRPP) (Berry *et al.*, 1983; Biondini *et al.*, 1985; Zimmermann, 1985). It is a non-parametric procedure for testing the significance of possible differences between *a priori* classified groups. MRPP has the advantage of not requiring assumptions like multivariate normality and homogeneity of variances and it is easily applied to multivariate space. MRPP was done using program PC-ORD version 4.

Possible indicator species of certain river groups **(I)**, or certain substrata **(IV),** were identified using Indicator Species Analysis (IndVal) (Dufrene & Legendre, 1997; McGeogh & Chown, 1998). The method combines information on the abundance and faithfulness of occurrence of species abundance in a particular group. IndVal is considered superior to more traditional methods of identifying indicators (e.g. TWINSPAN) on both statistical and practical grounds (Legendre & Legendre, 1998, McGeoch & Chown, 1998). The significance of the indicator value for each species was tested by a Monte Carlo randomization test. IndVal was performed using program PC-ORD version 4.

Discriminant Function Analysis (DFA) was used for interpreting the biological TWINSPAN groups (**I**), i.e. to examine which chemical and physical factors best discriminated among the site groups. In addition, DFA was used to study how well biologically defined community types (TWINSPAN groups) can be predicted using environmental variables, that is, how large a proportion of sites were classified into correct TWINSPAN groups using environmental data.

Table 1. Summary of the main statistical analyses and number of diatom samples and study sites in different articles. See text for abbreviations.

Paper	**Diatom samples**	**Study sites**	**Main statistical methods**
I	197	197	TWINSPAN, MRPP, IndVal, CCA, PCA, pCCA, CS, NMDS, ProTest
II	197	197	DCA, CCA, pCCA, NMDS, ProTest
III	294	146	DCA, CCA
IV	60	4	NMDS, MRPP, IndVal, linear regression, Spearman correlation
V	157	145	DCA, CCA
VI	108	60	DCA, CCA

After verifying that gradients were long enough, Canonical Correspondence Analysis was applied (CCA, ter Braak, 1986, ter Braak & Verdonschot, 1995) (**I-III, V, VI**). CCA is a direct gradient ordination method, which is appropriate for biological data having unimodal responses to the environmental gradients and containing many zeros (absences). In CCA, diatom data (relative abundances of taxa) were the response variables, constrained by the explanatory environmental variables. CCA was run using CANOCO version 4.0 with forward selection **(I)** (ter Braak & Smilauer, 1998) or using program PC-ORD version 4 (**II, III, V, VI**).

Partial CCA (pCCA) (Borcard *et al*., 1992, Økland & Eilertsen, 1994) was used to partition variation in species data into three components: (1) pure environmental (physical and chemical factors), (2) pure spatial (latitude and longitude), and (3) spatially structured environmental (the part explained jointly by the two groups of explanatory variables) (**I, II).** In paper **I**, Principal Component Analysis (PCA) was first performed on correlation matrix of the environmental variables to reduce the dimensionality of the original data into a few easily interpretable principal components (i.e. environmental gradients). By accepting only the three first components for subsequent analysis, it was ensured that the dimensionality of the environmental data matched closely that of the spatial data. Because the use of unexplained variation in pCCA has been recently questioned (Økland, 1999), only the amount of variation explained by the two sets of explanatory variables was discussed.

The classification strength (CS) of the ecoregions, subecoregions and TWINSPAN groups was tested using the randomization protocol of Van Sickle & Hughes (2000) (**I**). The mean of all between-class-similarities (B) and the within-class mean similarity (W) were first calculated using Sorensen similarity coefficient. CS is defined as the difference between these similarities (CS = W–B). Values of this measure range from 0 to 1, values near zero indicating that sites are randomly assigned to classes. The observed values of CS were compared to permutated values, obtained through 1000 random permutations.

Procrustes analysis was used to test whether the proximity of sites in a biological ordination could be explained by mere spatial distance between the sampling sites (**I, II**). Therefore, the strength of congruence between the spatial coordinates (longitude and latitude) of the study sites and a biotic ordination (Non-metric multidimensional scaling, NMDS) was tested. Procrustes analysis is used for testing the concordance between two ordinations, and it works by reflecting, rotating, translating and dilating one ordination and then superimposing it on a second one, minimizing the sum of the squared residual (m^2) between corresponding observation. The m^2 statistic is then used as a measure of association between the two ordinations; low values of m^2 indicate strong concordance (Digby & Kempton, 1987). Procrustean Rotation Test (ProTest) extends Procrustes analysis by providing a test to assess the statistical signifigance of the Procrustean fit using a permutation procedure (Jackson, 1995). Randomization procedure (9999 permutations) was used to determine whether the sum of residuals is less than expected by chance.

In paper **IV**, persistence was defined as the continuous presence of species populations in a community and stability as the degree of constancy in the relative abundance of organisms (Connell & Sousa, 1983; Scarsbrook, 2002). In addition to analyses mentioned above, the changes in community similarity were related with changes in environmental conditions using multiple linear regression. Stability between sampling months in the rank abundance of taxa was assessed using Spearman rank correlation (Townsend *et al.,* 1987; Weatherley & Ormerod, 1990). Furthermore, monthly changes in species composition or percentage turnover (T) were used to indicate community persistence. Turnover was calculated as $T = (G + L)/ (S_1 + S_2) \times 100$ where G and L are the number of taxa gained and lost between months, and S_1 and S_2 are number of taxa present in successive sampling months (Diamond & May, 1977; Brewin *et al*., 2000). Dominance and species richness of diatom communities on each substratum were assessed using rank-abundance diagrams.

In paper **V**, the estimate of each species` total phosphorus optimum was obtained using total P values of the river sites in the training set, weighted by the abundance of the taxa in these sites. The initial estimate of the total P was calculated as a weighted average using the inverses of the squared tolerances (range of their variation along the phosphorus gradient) as additional weights. Because averages are taken twice (regression and calibration), the range of the estimated phosphorus values shrinks (ter Braak & van Dam, 1989). To correct this, the initial site estimates were regressed both on the observed values (classical regression) and *vice versa* (inverse regression). With the obtained deshrinking parameters, bias of the estimates was corrected towards the observed total P values. An independent test set was used to cross-validate the Weighted Average (WA) model. To assess the performance of the model, root mean squared error of prediction (RMSEP, Wallach & Goffinet, 1989) and the Pearson correlation coefficient (r) between the observed and inferred total P values were used.

The ecological status of the rivers was evaluated using the pollution diatom index IPS (**III** and **VI**, Coste in CEMAGREF, 1982), updated version of the trophy index TDI (**III**, Kelly, 1998), the indicator list of Van Dam *et al*., (1994) and the macroinvertebrate pollution index ASPT (average score per taxon, family level identification) (**VI**, Armitage *et al*., 1983).

3. RESULTS

3.1 Regulating factors - role of environmental and spatial components (I, II)

The relative roles of environmental and spatial factors as determinants of diatom community structure were studied using direct ordination (CCA), variation partitioning (pCCA) and Procrustes analysis (ProTest). The diatom-environment correlations for CCA axis 1 (0.959) and 2 (0.926) were high, indicating a relatively strong relation between diatoms and the measured environmental variables in the whole data set (n = 197, I, II). The eigenvalues of the first two axes (0.435 and 0.227) were both significant ($p < 0.01$; Monte Carlo permutation test, 99 permutations), and they explained 10.2 % of the total variation (6.469) in the species data. Conductivity, total P, pH, and latitude were the most significant contributors to axis 1 (**I, II**). This axis mainly separated electrolyte poor, soft waters in central and northern Finland from southern enriched, hard waters (**I**). The second CCA-axis primarily separated humic or turbid streams from clear-water streams; colour and pH were the most important variables along this axis. At the ecoregion and river system scales, eigenvalues of the first two CCA-axes were lower but significant excluding the R. Vantaanjoki due primarily to lower inertia of the data (**II**). Like in the whole station set, conductivity, pH, total P, latitude and colour were primarily regulators of the diatom distribution patterns at these smaller scales. Partial CCA revealed that pure spatial component explained ca. 20 % of variation in diatom data at each three spatial scales (Fig. 4). Pure spatial component was slightly more important at the largest scale, explaining almost 25 % of variation. On the other hand, environmental factors captured over 70 % of explained variation in species data in the North boreal ecoregion and in the R. Vantaanjoki drainage system. Proportion of variation explained by the combined effect of environmental and spatial factors was at the largest scale almost 40 %, indicating that the diatom communities of boreal streams incorporate a rather strong spatial component. This also implied that the environmental gradients had a strong spatial structure at that scale. At ecoregion and river system scale, environmental variables had a smaller spatially structured component due to a smaller spatial extent of the study area.

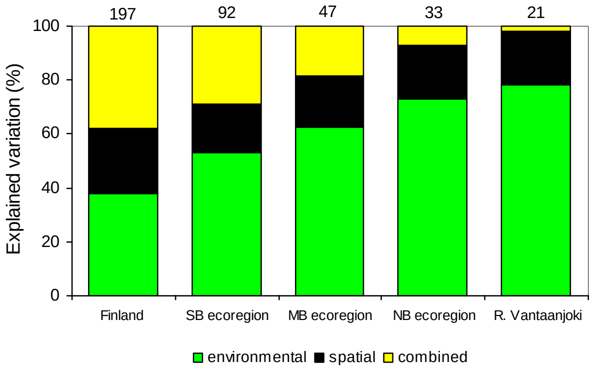

Fig 4. Variation partitioning (Borcard *et al.* 1992) of diatom data at three spatial scales using partial Canonical Correspondence Analysis. Only explained variation is shown.

According to Procrustes Analysis and subsequent Procrustean Rotation Test (ProTest), spatial coordinates of the study sites and patterns in diatom community structure, as summarized by Non-metric Multidimensional Scaling (NMDS) ordination axes, were strongly concordant ($m^2 = 0.862$, $p = 0.001$) across the largest spatial scale (**II**). For impacted sites, diatom data and spatial coordinates were, surprisingly, more strongly concordant than for reference sites at the largest spatial scale ($m^2 = 0.809$, $p = 0.0001$; $m^2 = 0.915$, $p = 0.001$, respectively). At ecoregional scales, concordance was especially strong in south boreal ecoregion ($m^2 = 0.841$, $p = 0.001$) and in middle boreal ecoregion ($m^2 = 0.779$, $p = 0.001$).

However, in north boreal ecoregion, concordance was low ($m^2 = 0.992$, n.s.). At a river scale, the degree of congruence were rather high ($m^2 = 0.887$). However, biotic and spatial data matrices did not show significant concordance in R. Vantaanjoki due to smaller number of sampling sites.

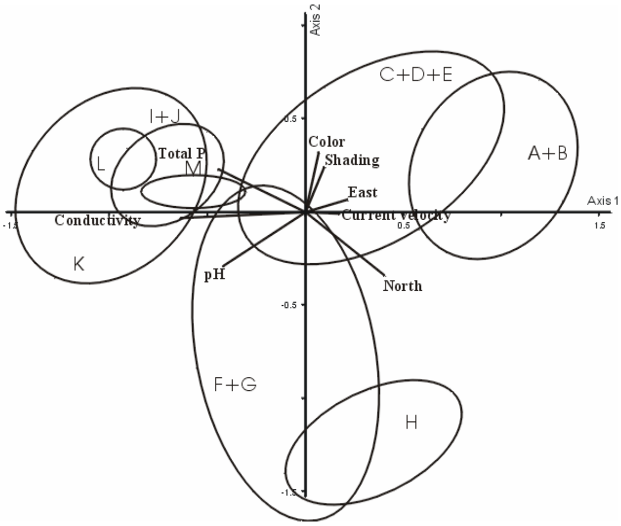

Fig. 5. Ordination diagram showing the distributions of the TWINSPAN site groups (denoted by capital letters) and relative contributions of environmental variables in the CCA space. Ellipses encircle 90 percent of sites belonging to a given group.

3.2 Diatom community types and indicator species (I, III)

Using TWINSPAN grouping and Indicator Species Analysis (IndVal) based on diatom species composition, 13 distinct river groups were found in Finland, which all were statistically significant (all $p< 0.0001$) according to Multi-Response Permutation Procedures (MRPP) (Fig. 6) (**I**). TWINSPAN groups were rather well separated from each other in CCA-space (Fig. 5). Discriminant Function Analysis (DFA) was used for interpreting the TWINSPAN groups. First four functions explained 94.2 % of variance. The river groups were mainly separated by their chemical characteristics, yet they were spatially structured as well. The first two functions were mainly related to conductivity (eigenvalue 0.928) and water colour (eigenvalue 0.627), respectively. Subsequent gradients were primarily related to physical factors (current velocity, shading, width) and pH. Although the number of groups was high, 68 % of original biological groups were predicted correctly using the four discriminant functions based on physical and chemical data.

Using only four categories, Finnish running waters might be classified into “clearwater neutral” (groups F-H, mostly in central and northern Finland), “humic acid” (groups A-E, mostly in eastern and northern Finland), “eutrophic polluted” (groups I-K, in southern Finland) and “meso-eutrophic” rivers (groups L-M, mostly in southern Finland) (Fig. 6). These river classes can be characterized by indicator species (IndVal) with statistically significant abundances and faithfulness of occurrence in these river groups (**I**). Communities described by indicator species (IndVal) naturally differ from the ones characterized by dominant species (Table 2, paper **I**; Table 2, paper **III**). In clear oligotrophic streams, strong indicator species represented e.g. genera *Achnanthes*, *Cymbella* and *Gomphonema*. In humic, acid streams, most of the strongest indicators were acidophilic species of the genus *Eunotia (*e.g. *E. incisa* and *E. rhomboidea)*. In eutrophic, polluted southern Finnish streams most of the strong indicators were biraphid motile species of genera *Navicula (*small *Navicula*-species like *N. agrestis* and *N. saprophila*), *Nitzschia* (*N. palea*) and *Surirella* (*S. brebissonii*) indicating low water quality as well as features of the habitats, being mainly soft bottoms. In meso-eutrophic streams most of the strongest indicators were of genera *Diatoma, Navicula* and *Nitzschia*.

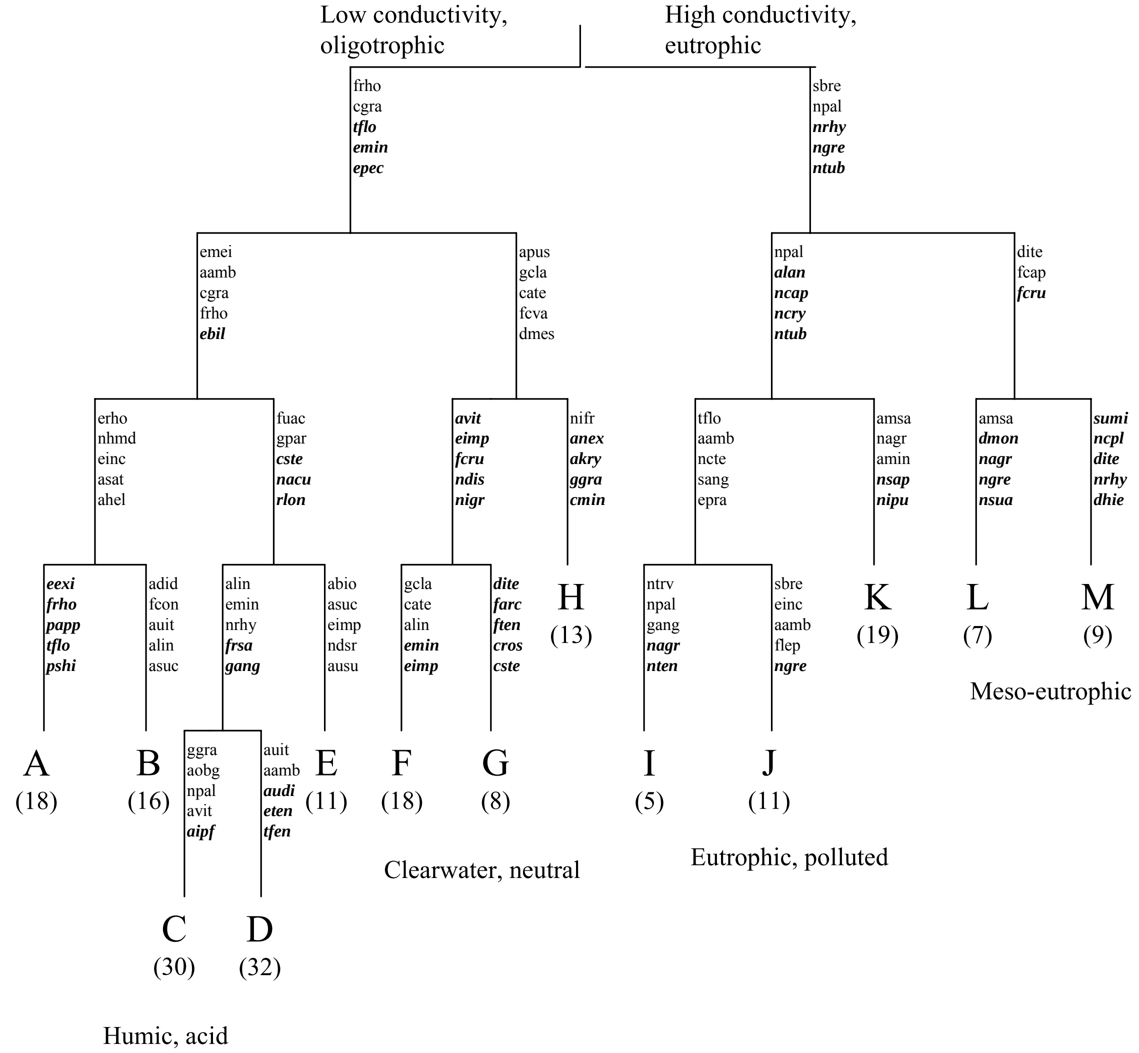

Fig. 6. TWINSPAN classification of the study streams. Figures in parentheses refer to the number of sites in each TWINSPAN group (A-M). Taxa in bold italics were identified as indicators only by the Indicator Value method (IndVal), while all others were identified by both TWINSPAN and IndVal. See Appendix 1 and Table 1 in paper **I** for species abbreviations.

3.3 Ecoregions as classification units (I)

The possible differences in diatom community structure among three ecoregions and eight subecoregions were tested using a randomization protocol (**I**).

The classification strength (CS: similarity within ecoregions – similarity between regions) of ecoregions was 0.090. It was only slightly improved by including only sites at “purified” ecoregions, i.e. sites with at least 25 km to the nearest ecoregion boundary (CS = 0.107). At the level of subecoregions, classification was almost equally strong for all sites (CS = 0.107) as for near-pristine reference sites only (CS =

0.123). Finally, CS for the biologically-defined TWINSPAN typology (division level three), which here served as a CS benchmark, was 0.127 and it only slightly exceeded that of the subecoregions (0.127 vs 0.107), both having eight site groups. All CS values were higher than expected by chance (Monte Carlo randomisation test with 1000 permutations, all $p < 0.001$).

3.4 Seasonal community persistence and stability on three substrata (IV)

Seasonal community persistence (continuous presence of species populations in a community) and stability (degree of constancy in the relative abundance of taxa) were assessed using several statistical methods (**IV**). Monthly changes in species composition or percentage turnover (T) were highest among epiphytic communities, indicating the lowest persistence (Fig. 7). The differences in species turnover among habitats were significant at two stations (ANOVA; $p<0.05$ and $p<0.01$, respectively). Epipelic communities were more persistent than epilithic communities at three sampling stations.

The variation of successive samples in the ordination (Non-Metric Multidimensional Scaling (NMDS) space was substantial and without an evident pattern; there were no clear differences between the stabilities of communities on stones, plants or sediment (Fig. 8). Ordinations showed that in R. Porvoonjoki, epipelic diatom community was clearly the most instable, indicated by a large deviation of successive samples in the ordination space. At the other stations, however, epipelic communities were rather stable compared to communities in other habitats. At two stations (R. Mustajoki and R. Keravanjoki), Multi-Response Permutation Procedures (MRPP) showed significant among-group differences ($p<0.0001$), indicating that diatom communities were distinctly different between the three substrata. The other stations did not show any significant among-group differences. The stability between sampling months in the rank abundance of diatom taxa was lowest among epiphytic communities (**IV**). Especially in the R. Mustijoki, correlation between rank abundances of epiphytic taxa for successive months was lower than for stones or sediment. At each station, the Spearman correlation between constitutive sampling months decreased substantially until August indicating a distinct change in rank abundances. Communities in September and October, however, tended to be more similar to communities in June. A major spate at the end of August coincided with the change towards higher similarity with the June communities. However, there were no significant correlations between changes in community similarity and changes in water chemistry or discharge conditions using multiple linear regression. Using simple linear regression, however, changes in total P were significantly ($p<0.05$) related to community stability (Pearson correlation between consecutive sampling months) in the epilithon.

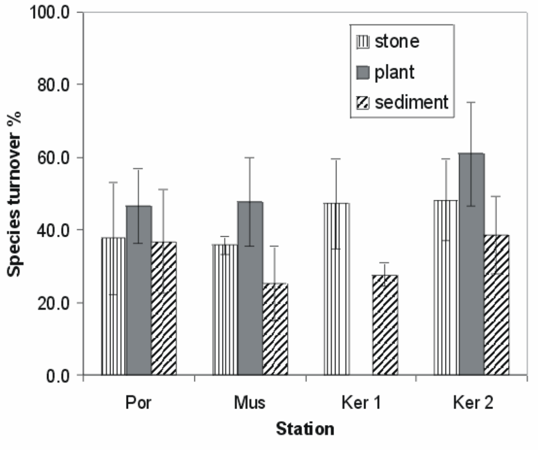

Fig 7. Average monthly species turnover (± SD) indicating community persistence at four sampling stations. Turnover is expressed as a proportion of species gained and lost to total species number. Por = R. Porvoonjoki, Mus = R. Mustijoki, Ker 1 = R. Keravanjoki station 1 and Ker 2 = R. Keravanjoki station 2.

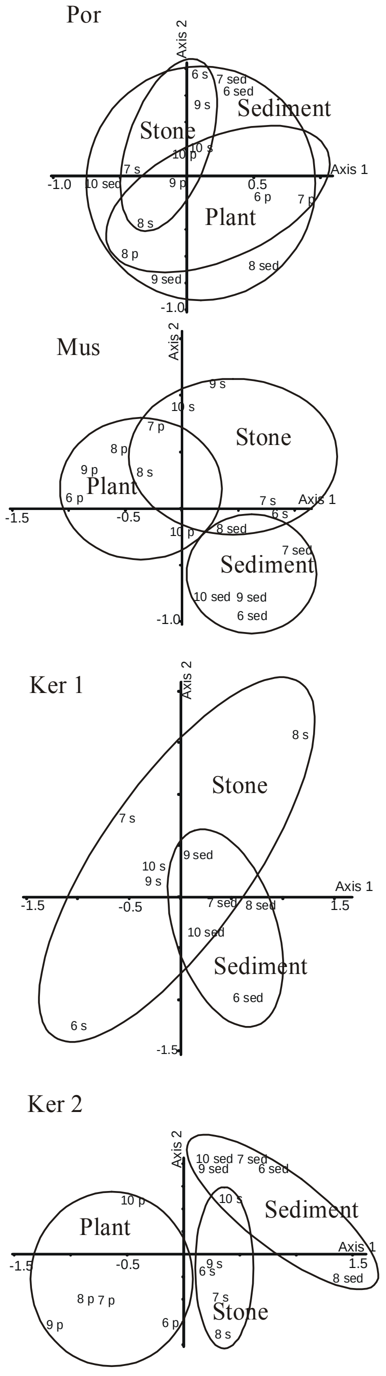

Fig 8. Ordination diagrams for Non-metric Multidimensional Scaling (NMDS) analyses of diatom communities on three substrata. Numbers refer to successive sampling months. s = stone, p = plant, sed = sediment. See Fig. 7 for station abbreviations.

Epiphytic communities had notably higher dominance and fewer species than communities on stones or sediment, as clearly illustrated by monthly rank-abundance diagrams (**IV**). Species richness was highest among epipelic samples. Community dominance among epiphytic and epilithic samples was highest in September when the abundance of the dominant species (*Cocconeis placentula*) on plant surfaces reached 70 %. Overall, species richness was lowest in August.

3.5 Inferring the phosphorus levels of running waters using diatoms (V)

Direct ordination (CCA) was first used to study if total P contributed significantly to diatom distribution patterns at the study sites in the test set (**V**). The eigenvalues of the first two CCA axes (0.55 and 0.18, respectively) were both significant ($p < 0.01$; Monte Carlo permutation test, 99 permutations). They explained 12.4 % of the total variance in the diatom community. The canonical coefficients and intraset correlations indicated that conductivity, total P and pH made the most significant contribution to axis 1, and colour and stream width to axis 2. According to the constrained CCA analysis, total P as a sole influencing factor had a significant ($p<0.01$, Monte Carlo permutation test, 99 permutations) effect on the diatom community structure. The ratio of the constrained axis (λ_1) and first unconstrained axis (λ_2) for total P was 0.81. Based on literature (e.g. Hall & Smol, 1992, 1996; Winter & Duthie, 2000), it was considered high enough for modelling total P.

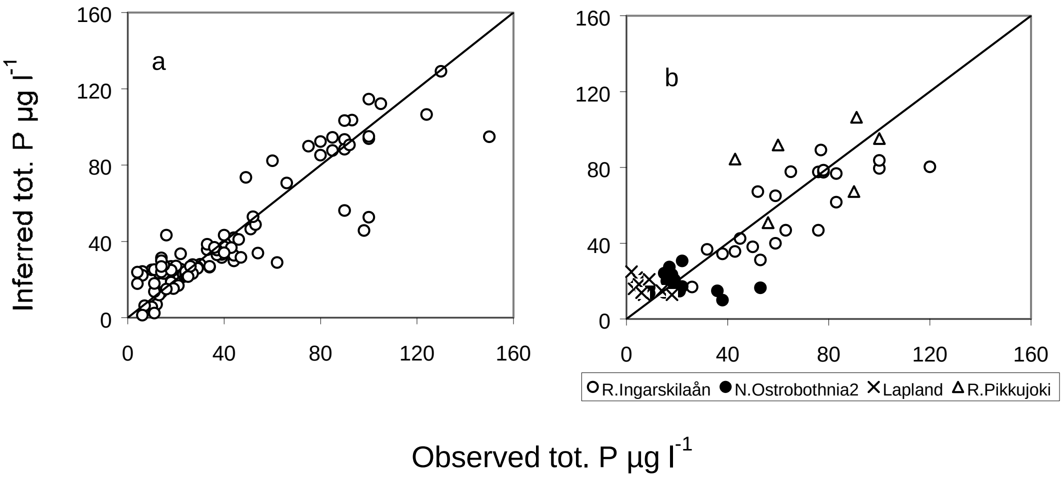

Fig. 9. Relationship between observed total P and diatom inferred total P using a) weighted averaging with tolerance weighting and inverse deshrinking in the training set, b) weighted averaging without tolerance weighting and with inverse deshrinking in the test set.

Training set

The total P optima and tolerances were calculated using equation 1 (**V**) for 120 diatom taxa using the training set. Some of the species abundances did not show clear unimodal or linear relationships with the total P concentration, and species distribution patterns along the trophic gradient were usually rather noisy. The species were segregated, however, along the total P gradient with different optima. The species indicating oligotrophy (e.g. genus *Eunotia*) were restricted to low concentrations. By contrast, species indicating eutrophy (e.g. *Navicula cryptocephala* and *Nitzschia palea*) had higher optima and usually larger tolerances (**V**).

The correlations between the observed and diatom inferred total P concentrations were high. The highest correlation ($r = 0.91$) was found using weighted averaging regression with species tolerances as additional weights (Fig 9a). The smallest prediction error (RMSEP 14 µg P l^{-1}) was obtained using inverse deshrinking. As a whole, inverse deshrinking performed much better yielding prediction errors notably smaller than classical deshrinking. Logarithmic transformation did not significantly improve the performance of the WA inferences. Consequently, WA models were presented based on untransformed total P data.

Test set

The correlation between the observed and inferred total P was only slightly lower in the independent test set than in the training set (Fig. 9b). The correlation was highest ($r = 0.87$) and the prediction error smallest (RMSEP 16 µg P l^{-1}) when using inverse deshrinking without tolerances as additional weights. As in the training set, classical deshrinking yielded clearly larger prediction errors than inverse deshrinking, especially in weighted averaging without tolerance weighting. The calculated total P concentrations were slightly higher than observed in very nutrient poor stations (observed total P < 15 µg l^{-1}), especially in oligotrophic northern rivers (North Ostrobothnia and Lapland).

3.6 Diatom and macroinvertebrate based bioassessment tools (III, VI)

At R. Pikkujoki, the purified sewage degraded the water quality clearly. Diatom inferred saprobity (IPS-index) and trophy (TDI-index) increased below the sewage load point (Fig. 10) (**III**). The changes were significant at $p < 0.05$ and $p < 0.01$, respectively. The community dominated by *Achnanthes minutissima, Fragilaria capucina* and *Melosira varians* changed into a community dominated primarily by *Gomphonema parvulum, Navicula*-species (e.g. *Navicula cryptocephala*, *N. gregaria* and *N. minima*) and *Nitzschia palea* below the sewage load. The biologically-inferred water quality improved farther downstream due to self-purification. The recovery zone was ca. 10 km long. In May, the sewage treatment plant did not have as large an effect on the water quality as it did in August, because of higher discharge and, therefore, dilution of sewage. The IPS index proved to be more stable than TDI, with smaller standard deviations between the parallel composite samples.

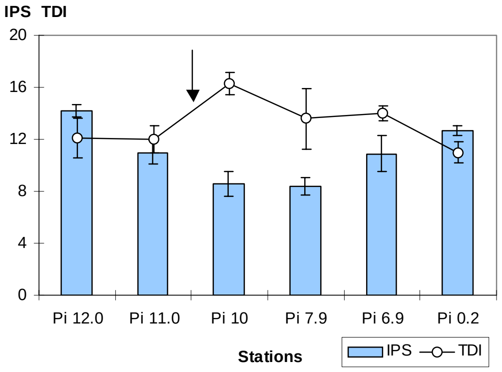

Fig 10. The values of IPS and TDI indices in R. Pikkujoki in August 2000. The site loaded by sewage is marked by an arrow. Numbers in the station names refer to the distance from the Finnish south-coast.

In a comparative study monitoring both diatoms and macroinvertebrates, structures of both biotic communities were first related to environmental factors using direct ordination (**VI**). For diatoms, the first CCA-axis of was primarily related to total P and conductivity, separating the river sites with the highest trophy levels and electrolyte concentrations from the other sites. Axis 2 was mainly related to latitude, longitude and current velocity. As a whole, diatom community structure was most affected by conductivity, total P and latitude. For macroinvertebrates, the first CCA-axis primarily separated the most electrolyte-rich, wide river sites from the others; while conductivity, pH and stream width contributed most importantly to this axis. The second axis was a gradient of humus and total P concentrations, yet stream width contributed as well. As a whole, macroinvertebrate community structure was most affected by stream width, conductivity and pH.

The Detrended Correspondence Analysis (DCA) for diatom data indicated that communities differed clearly between the sampling stations (**VI**). The eigenvalues of the first two axes for diatom DCA were 0.454 and 0.214 and together accounted for 24 % of the cumulative variance. The first axis was primarily a gradient of conductivity and phosphorus concentrations. Eutrophic stations had rather similar diatom communities. Variation between replicate samples was notably smaller than between the sites, but was highest in R. Ingarskilaån.

The eigenvalues for the macroinvertebrate DCA were 0.296 for the first axis and 0.150 for the second axis explaining 18 % and 9 % of variation, respectively. The analysis mainly separated sampling stations in R. Vantaanjoki (V 44) and R. Glimsån from the other sites along axis 1 (**VI**).

The values of diatom index IPS varied from 9.5 (station K) to 16.3 (station Glo) (Fig. 11), reflecting poor to good quality (Eloranta & Soininen, 2002). According to diatoms, water quality was expectedly clearly highest at the most oligotrophic stations in R. Glomsån and R. Myllypuro. Diatom community composition at these stations differed from the other sites notably according to DCA analysis as well (**VI**). IPS index values of most of the stations indicated that water quality was moderate. The values of ASPT index varied rather little, from 6.2 to 7.2. According to macroinvertebrates, water quality was best in R. Myllypuro and at one station in R. Vantaanjoki (V 24).

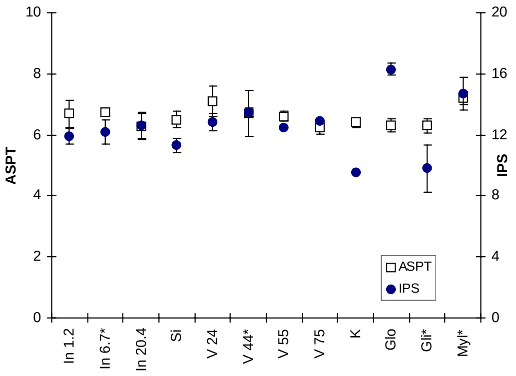

Fig. 11. The values (± SD) of the macroinvertebrate (ASPT) and the diatom index (IPS) at 12 sampling stations. See paper **VI** for station abbreviations. Stations having significantly different variances in index values, assessed by paired F-tests, are marked by asterisks.

Replicate macroinvertebrate samples tended to have slightly lower variation in index values. The average variation in ASPT index values was 0.31 SD among replicate samples and coefficient of variation (CV) was on average 4.7 %. Among the IPS index, average variation was 0.72 SD and coefficient of variation was 6.0 %. According to F-test, variances were significantly different at four stations (In 6.7, V 44, Gli and Myl). Correlation between diatom and macroinvertebrate indices was rather low (r = 0.29; n.s.). Finally, the community similarity of five replicate samples at eight stations was assessed using Pearson correlation. For diatoms, correlation varied from 0.803 to 0.931 and on average it was 0.874. As a whole, community similarity between the replicate samples was slightly lower among macroinvertebrates; correlation varied from 0.480 to 0.964, and on average correlation was 0.770.

4. DISCUSSION

4. 1 Determinants of benthic diatom community structure in boreal streams (I-III)

The results of direct ordinations in this thesis (**I-III**) are well congruent with previous works emphasizing the predominance of chemical-constituent concentration and ion composition on structuring benthic algal communities of running waters (Biggs, 1990; Leland & Porter, 2000; Winter & Duthie, 2000; Griffith *et al.*, 2002). The variable identified with having the strongest association with periphytic, and especially diatom communities has rather often been conductivity (Biggs, 1990, 1995; Pan *et al.*, 1999; Munn *et al.*, 2002). Conductivity indicates primarily the concentrations of main cations (Ca, Mg), and is closely related to water pH, an important variable for aquatic biota, regulating many physiological processes. Conductivity is considered an easy, conservative surrogate for stream enrichment, because major ions are not intensively involved in biological processes, and relative fluctuations in conductivity are smaller than, for example, for nutrients (Biggs, 1990, 1995). Conductivity integrates several important watershed processes indicating strongly the nature of geology and, therefore, is the strongest environmental gradient in explaining diatom distribution patterns in Finnish running waters at the national scale. The other important determinants of diatom community structure were latitude, pH, total P, and water colour. Similarly, in a study of diatom distributions in Labrador lakes, water colour emerged as one of the key determinants of diatom communities, leading Fallu *et al.* (2002) to suggest that the amount of humic substances should be generally important in explaining diatom distributions across broad geographic regions in electrolyte-poor, oligotrophic freshwater systems. Water colour has also been identified as one of the main correlates of macroinvertebrate assemblage structure in boreal streams (Malmqvist & Mäki, 1994; Paavola *et al.*, 2000; Heino *et al.*, 2003a). Humic substances react with other dissolved substances, such as phosphorus, reducing their bioavailability (Jones, 1990). Additionally, humus may lower the toxic effects of heavy metals on some algae (e.g. Gensemer *et al.*, 1993). The results of this thesis show also that water colour is generally important, often a clearly distinguishable second gradient (**I**) in explaining diatom distributions across broad geographic regions in electrolyte-poor, oligotrophic freshwater systems.

The set of environmental variables was somewhat limited due to heterogenous sources of the data. If important variables are missing, the variation explained by pure environmental component will be underestimated (Økland & Eilertsen, 1994). Inclusion of nitrogen results for instance, would have probably increased somewhat the proportion of variation explained solely by environmental variables. Moreover, particle size at the benthos and iron concentrations may be important determinants of diatom community composition especially at the level of ecoregions (Soininen *et al.*, unpublished). Large-scale spatial factors, such as climate, geology and vegetation also influence diatom species distributions. According to Leland (1995), both temperature and basin geology (reflected in conductivity and pH), strongly influence phytobenthos communities. Temperature can influence diatoms directly by affecting metabolic processes and cell division (Bothwell, 1988), and indirectly via changes in physical, chemical and biological properties of rivers and streams. Potapova & Charles (2002) reported that mean annual air temperature affected significantly the diatom distribution patterns in USA rivers. However, temperature is not only a regional variable, but can vary locally, depending on riparian characteristics and possible ground water influx. I did not address the question whether grazing could have some major effect on lotic diatom community structure in this summary or in independent papers due to lack of data on benthic fauna. However, that issue will be addressed in forthcoming paper concerning community concordance in boreal headwater streams (see future research directions).

4.2 Spatial scale, organism body size and taxonomy (I, II)

Biotic communities in streams are controlled by multiple factors prevailing at different temporal and spatial scales (Biggs, 1995; Stevenson, 1997; Angermeir & Winston, 1998; Sandin, 2003). For unicellular, and small-sized organisms in general, local factors should be much more important than regional ones, thus setting a strong environmental filter (sensu Poff, 1997). This filter selects species able to cope with the conditions prevailing at a site. Benthic diatom communities are considered as being regulated primarily by local environmental conditions while broad-scale climatic, vegetational, and geological factors have a minor role (Pan *et al*., 1999, 2000). However, results of this thesis (**I, II**) show that diatom communities exhibit a rather strong spatial component especially at a national scale. This was shown both by variation partitioning (partial CCA) and by a direct comparison of the NMDS ordinations of diatom communities and the spatial coordinates of the sampling sites.

The proportion of variation explained independently by spatial factors was quite large, ca. 25 %, at the largest spatial scale (**I, II**). Corresponding figures (23-31 % of explainable variation) were reported by Potapova & Charles (2002) for the whole of USA and for Omernik's (1987) level 1 ecoregions. Furthermore, it seems that even at small scales (ca. 10-10^2 km), pure spatial component still plays an role (ca. 20 % of explained variation) in regulating benthic diatom community composition. However, the spatially structured environmental component (combined effect) was small especially in North boreal ecoregion and at river system scale. Diatom communities seemed to be more strongly spatially structured in southern Finland; this was clearly seen in variation partitioning and direct test of congruence between diatom community structure and spatial coordinates. This might reflect the bias in the number of sampling sites (more southern sites) or the fact that most of the impacted sites are situated near the southern coast of Finland.

When very small spatial scales (1 m - 10 m) are concerned, pure spatial component arises from natural spatial autocorrelation and patchiness of benthic biotic communities. According to Passy (2001), space alone contributed 10 % of explainable variance in the diatom data at a riffle scale. In general, spatial variation in algal communities is the result of physical and biological factors prevailing at multiple scales. The heterogeneity of the communities is induced at a local scale primarily by differences in light and current regimes, intensity of grazing, stages of succession, and variation in substratum (Peterson & Stevenson, 1989, 1990; Ledger & Hildrew, 1998; Sommer, 2000). Large heterogeneity prevails among benthic diatom communities in scales from meters to tens of meters, especially in varying current regimes, and it appears that this scale can be rather important in diatom distribution patterns (Soininen, 2003).

It seems evident that decreasing body size is correlated with decreasing influence of regional processes on community structure (Hillebrand & Azovsky, 2001) and therefore, spatial structure among diatoms should be weaker than e.g. among benthic fauna. However, Hillebrand *et al*., (2001) stated that although the species composition of unicellular organisms is less influenced by geographic and dispersal related factors, diatoms lack strictly ubiquitous dispersal. Data of this thesis support also the view that turnover diversity (ß-diversity) of benthic diatoms might be much higher than previously believed (**I, II**). It might be that diatoms disperse along a continuum from endemic species to cosmopolitans depending e.g. on ecological tolerances, body size and life form. The frequent disturbances in running waters may lessen the effects of some processes on the community structure of organisms. These patterns and processes are, however, still inadequately known in running waters. In conclusion, community concordance and diversity patterns among multiple aquatic organism groups with different dispersal capacity needs rigorous testing using extensive data from streams (but see Paavola *et al*., 2003; Heino *et al*. 2004).

Kociolek & Spaulding (2000) argued that the importance of geographical factors in explaining diatom distribution has previously been underestimated. They further claimed that in explanations of diatom distributions, more emphasis should be given to broad-scale historical factors, (e.g. glacial period in boreal areas) than to explanations stressing the role of present-day dispersal capacity. Given the rather strong spatial structure of benthic diatoms in studies of this thesis, it might be that endemism or rather restricted geographical distribution is more common in benthic diatoms than in planktonic diatoms or phytoplankton in general. In benthos, cells are attached to or living on the bottom or other substrata, and not that susceptible to wind or other physical forces. In general, dispersal of viable algal cells might be less efficient than has been previously believed; airborne dust often contains diatom frustules, which are mostly dead (Round, 1981).

Spatial structure of the data is strongly affected by the species concept used. A fine-scale taxonomy tends to lead to discovery of more taxa with narrower geographical and ecological distributions. Therefore, it is essential that both the species concept and species identification are congruent throughout a study. Furthermore, the use of the same identification keys by all researchers might contribute to the perception that most freshwater diatoms seem cosmopolitan (Mann & Droop, 1996; Kociolek & Spaulding, 2000). Mann & Droop (1996) further emphasize the fact that the prevailing diatom species concept hides diversity, endemism and spatial structure of diatom communities. If dispersal is lower than previously believed, it should increase the speciation, which is, however stated to be rather low among microbial organisms due to pervasive gene flow (Godfray & Lawton, 2001). Large-scale patterns in diversity and dispersal of microbial eukaryotes are under a strong debate among ecologists, and opposite view of endemism is that distribution is governed by ubiquitous dispersal and the spatial distribution of suitable habitats (e.g. Finlay, 2002; Finlay et al., 2002). In the future, comparable data sets of organisms with varying dispersal capacity collected within large geographical areas should further reveal true ecological patterns concerning this issue.

4.3 Ecoregions as classification units (I)

Given the strong latitudinal patterns in community composition, it seems evident that bioassessment programs utilising lotic diatoms would benefit from geographical stratification, using e.g. ecoregions or subecoregions (**I**). The spatial patterns exhibited by benthic diatoms in this study corresponded fairly closely with those documented for stream macroinvertebrates in Finland by Heino *et al*. (2002, 2003a).

The level of classification strengths using ecoregions (CS = 0.090) and subecoregions (CS = 0.107) were rather similar to those obtained for macroinvertebrates (CS = 0.096 and CS = 0.138) in boreal streams (Heino *et al*., 2002). However, as noted by Van Sickle & Hughes (2000), the classification strength obtained using ecoregional delineations may partly result from spatial autocorrelation, rather than ecological factors that determine the ecoregional boundaries. Subecoregional differences were slightly stronger among reference sites (**I**). This is to be, however, expected; human disturbance is likely to reduce spatial heterogeneity and thus mask ecoregional differences. However, the hypothesis that spatial structure *per se* would be more evident among near-pristine reference sites was not supported; according to ProTest, location of the study sites and diatom community structure were more related among impacted sites (**II**). By contrast, Pan *et al*., (2000) noted that in their diatom data from Mid-Atlantic Highlands, ecoregional differences were more evident among randomly selected sites than reference sites.

Stream ecologists seem to share the view that ecoregional classifications should not be used alone to partition variance in community composition (Hawkins & Vinson, 2000; Hawkins *et al*., 2000; Sandin & Johnson, 2000; Van Sickle & Hughes, 2000). Since local in-stream factors were even more important than spatial factors in explaining diatom distributions (see also Potapova & Charles, 2002), a combination of regional stratification and local environmental features might provide the most robust framework for diatom-based bioassessment of boreal streams, as previously proposed for benthic fauna (Hawkins *et al*., 2000; Sandin & Johnson, 2000; Heino *et al*., 2002).

4.4 Diatom community types and indicator species (I, III)

Classification reduces or partitions natural variation of biological data into classes and is considered as a necessary first step in biological assessment (Gerritsen *et al*., 2000; Sandin & Johnson, 2000). Indicator species have a key role; they add ecological meaning to the clusters derived from data, and help to identify where to stop dividing clusters further into subsets (Dufrene & Legendre, 1997). Although the number of significant TWINSPAN groups was high, meaningful ecological interpretations were found for most of them (**I**). According to DFA-analysis, groups were primarily separated by chemical variables (mainly conductivity and water colour), yet physical factors also contributed to site classification. Most of the sites in each group were located within a restricted geographical area, demonstrating the tight relation between chemical and regional factors in Finnish streams (Heino *et al*., 2002). In this thesis, 68 % of sites were predicted into the correct TWINSPAN group using physicochemical factors. This percentage is slightly higher than that reported by Heino *et al*. (2003a), and clearly higher than by Sandin (2003), when testing TWINSPAN typologies based on benthic macroinvertebrate assemblages in boreal streams. This higher congruence between physicochemical factors and the biological classification indicates probably the fact that local small-scale factors are more important determinants of diatom community composition than for benthic fauna.

The most important indicator species characterizing each TWINSPAN group differed morphologically and ecologically. The indicator species for groups I to M were mainly motile biraphid taxa representing genera *Navicula*, *Nitzschia* and *Surirella* indicating high sedimentation and low current velocity at the sampling sites. The group J had several planktonic species, e.g., *Aulacoseira ambigua* and *Cyclotella meneghiniana*, as indicators, showing the importance of species origin, as well as features of the habitat. The TWINSPAN method has been criticized by Belbin & McDonald (1993), for failing to identify secondary gradients underlying one strong dominant gradient. Moreover, TWINSPAN always produces a hierarchical structure, even if this structure is subtle or nonexistent (Dufrene & Legendre, 1997). In this thesis, TWINSPAN nevertheless succeeded in producing ecologically meaningful groups that were explainable by several important gradients (levels of conductivity, water colour, total P, pH, and current regime),

IndVal proved to be more sensitive than TWINSPAN, as expected (Dufrene & Legendre, 1997; McGeogh & Chown, 1998), in finding indicator species with high specificity and fidelity for the groups concerned (**I**). Most of the detected indicator species occurred, expectedly, in low numbers, e.g. *Achnanthes kryophila*, *Navicula trivialis*, *Cymbella affinis* and *Navicula densestriata*. It is known that rare species can be important, even critical, in community ecology and bioassessment in detecting primary or secondary environmental gradients or impacts (Cao *et al*., 1998, 2001). Species which tend to occur locally in low abundance, tend also to be more narrowly distributed (e.g. Hanski, 1982; Brown, 1984; Hanski & Gyllenberg, 1997; Gaston, 1998 and see Fig. 12), thus increasing the spatial structure or variation of the data. In this sense, rare species are an important part of biological communities, as shown also in this thesis.

Although some of the taxa in these data were almost ubiquitous, some species exhibited regionally restricted distributions (**I**). For example, *Achnanthes biasolettiana, A. carissima, A. didyma* and *Cymbella affinis* had a distinctly northern distribution, whereas other species, e.g. *Navicula gregaria, N. reichardtiana, N. tenelloides* and *Surirella minuta*, occurred mainly or exclusively in southern, often eutrophic and turbid streams. A corresponding latitudinal gradient has been previously described for stream macroinvertebrates by Sandin & Johnson (2000), Heino *et al*. (2002) and Sandin (2003). Pienitz *et al*. (1995) reported a rather strong latitudinal gradient in diatom distribution patterns in boreal areas. These patterns of spatial variability probably have been further accentuated by covariation of geographical location and water chemistry across the study area (see Heino *et al*., 2002). In this thesis, the phenomenon was seen in variation partitioning analyses, where a large proportion of variation was explained by the spatially structured environmental component.

4.5 Seasonal community persistence and stability (IV)

Concepts of stability or persistence refer to discrete disturbance events, such as changes in discharge levels that do not have long-term effects (Connell & Sousa, 1983). Disturbances (sensu Poff, 1992) should have discrete effects on biological communities, and should not to be defined only on statistical grounds (see Resh *et al*., 1988). Spates can potentially have different effects on benthic algal communities, depending on the initial taxonomic composition, as well as features of the habitat. In this respect, the major spate that occurred in August can probably be considered as a disturbance event (**IV**). No continuous data on current velocities were available, but it is likely that any discharge almost a magnitude higher than the “normal” discharge regime, will produce clearly higher current velocities. Consequently, it will increase shear stress on the stream bottom (Hart & Finelli, 1999). The presumed effects of increased shear stress and possible changes in water chemistry were reflected in the instability of rank abundances of taxa at all four sampling stations. Communities tended to revert to their pre-disturbance state towards the autumn. Another possible explanation is a seasonal pattern in water temperature, because communities in mid-summer seemed to be different from those in June or autumn. The most probable explanation is that these factors acted jointly to influence the community composition. Changes in community stability could not be related to fluctuations in environmental conditions. Lack of any significant correlation between seasonal community structure or its fluctuations and environmental factors has been noted in some other studies (see e.g. Duncan & Blinn, 1989). To rigorously prove a dependency of seasonal community structure on environmental factors would need very intensive sampling.

Intuitively, epilithic diatom communities should be more stable than epiphytic and epipelic communities because of inherent stability of stony substrata. For example, fast-growing macrophytes can be rather unstable substrata for algal colonization and growth (Burkholder & Wetzel, 1989). Community persistence, indicated by species turnover, tended to be highest among epipelic, and lowest among epiphytic samples. According to the rank-abundance diagrams, a characteristic feature of the epiphytic communities was high dominance and low species richness. Consequently, a gain or a loss of a single species in epiphytic communities results in a higher turnover percentage than in communities with higher species richness. On the other hand, epipelic samples were the most species rich, being often a collection of species from multiple habitats. This perhaps explains the relatively low species turnover. Measures of community persistence are, of course, strongly influenced by rare species (Robinson *et al*., 2000). Despite the fact that rare species can be critical in detecting the primary or secondary environmental gradients or impacts regulating the community composition and diversity (Cao *et al*., 1998, 2001), their removal from analyses of persistence is important when strictly quantitative sampling or laboratory methods are not used.

Frustules were cleaned from organic material in order to reliably identify diatoms at species level. Cleaning of diatoms makes it impossible to distinguish cells that were originally alive from those that were dead, which can have some influence on results. It is probable that among epipelic diatoms, the proportion of dead cells was higher than in other habitats (especially among epiphyton), thus artefactually increasing the species number in epipelon.

These clear between-habitat differences in community persistence were not seen in the NMDS analyses, where community stability did not show any clear patterns between the three habitats. According to NMDS analyses, diatom communities were significantly different in the three habitats at two stations. There was no clear trend indicating that any particular habitat would possess a highly specific diatom flora. Winter & Duthie (2000) identified no clear habitat preferences, nor seasonality among stream epilithic, epipelic and epiphytic diatoms, although epilithic communities tended to be slightly different from diatoms on plants or sediment in two Canadian streams. On the other hand, there are studies documenting specific habitat preferences among diatoms (Reavie & Smol, 1997; Lim *et al.,* 2001; Antoniades & Douglas, 2002). Species significantly confined to plant surfaces were more or less similar to typical epilithic species (e.g. *Achnanthes lanceolata*), except *Cocconeis placentula*, which is generally assumed to grow preferentially on plants (Krammer & Lange-Bertalot, 1986-1991). Many of the species confined to epipelon were, expectedly, motile biraphid diatoms (e.g. species of the genera *Nitzschia* and *Surirella*) suggesting motility is an important species trait in the epipelic habitats (see also Pan *et al*., 1999).

The studied patterns in persistence and stability described the diatom communities in three eutrophic and turbid South-Finnish streams. Patterns may differ in rivers with different physical and chemical characteristics. However, it can be assumed that main processes and results apply to other boreal rivers or are valid in larger geographical areas, as long as suitable habitats (epilithon, epiphyton and epipelon) are present. In future studies, more and different kinds of river sites (impacted and near-pristine sites) should be sampled, preferable at multiple temporal scales ranging from days to years.

4.6 Diatoms in bioassessment of rivers (V, VI)

In running waters, the fluctuation of phosphorus concentrations is rapid and the amplitude is larger than in lakes. Therefore a single analysis of concentration does not have the same reliability of “true” trophic status as in standing waters. Using a WA model for river diatoms, it was found that correlations between the observed and inferred total P in the training set were high ($r = 0.91$) (**V**). Correlation in stream epilithon found earlier by Winter & Duthie (2000) was lower ($r = 0.72$). Compared to linear regression applied to the same data set (Eloranta & Soininen, 2002), WA modelling yielded higher correlation with measured total P. The correlation was notably higher than for total P and TDI index (Kelly & Whitton, 1995) applied to Finnish river material (Eloranta, 1999). These facts imply the superiority of WA modelling and the importance of indicator values obtained from a local calibration set. The correlation found for river waters is comparable to results published by Agbeti (1992) and Hall & Smol (1992, 1996) for sediment diatom communities in lakes and is higher than that found for epilithic algae in lakes (King *et al*., 2000). These facts further prove the potential of applying WA models for assessment of running water ecosystems.

Due to the different scales and the range of variation of observed concentrations, comparing the present prediction errors to previous studies is even more difficult than comparing the level of correlations. In lakes, prediction errors (RMSEP) have been smaller (Agbeti, 1992; Hall & Smol, 1992) due to lower phosphorus concentrations in general and to the different nature of physicochemical conditions in standing waters. The RMSEP in the training set is an underestimate of the true prediction error (Wallach & Goffinet, 1989; Stoermer & Smol, 1999). On the contrary, ter Braak & van Dam (1989) point out that the prediction error in the test set can be overly pessimistic due e.g. to different methods used in the analyses. However, the correlations were surprisingly high in the independent test set. In the test set, three to four parallel diatom samples were taken, which could have slightly stabilized the results compared to the training set and the studies mentioned above.

Inverse deshrinking yielded smaller prediction errors than classical deshrinking (**V**). In most studies, inverse deshrinking has yielded smaller prediction errors (Hall & Smol, 1992; Weckström *et al*., 1997; King *et al*., 2000). However, classical deshrinking has in some cases been preferred (Agbeti, 1992; Christie & Smol, 1993; Hämäläinen & Huttunen, 1996). The model generally performed very well in oligotrophic waters (Lapland, central Finland). At very nutrient-poor sites, however, the inferred total P concentrations were higher than the observed ones. In running waters, saturation of nutrients is obtained presumably at lower concentrations than in standing waters, due to the replenishment of the nutrient resources by water turbulence (Bothwell, 1988; Stevenson *et al.*, 1996). By contrast, in high concentrations (> 100 µg l^{-1} total P), the models predictions were lower than the observed concentrations. The biases at the ends of the gradient are caused by "edge-effects"; species responses are truncated at the gradient edges (Stoermer & Smol, 1999). A notable part of the phosphorus in river water is fixed in particulate matter. Thus, the concentration of total phosphorus alone does not accurately indicate the trophic conditions of a river (Ekholm, 1998). In four humic, nutrient rich, but electrolyte poor, northern sampling stations the inferred concentrations were substantially lower than the observed ones (**V**). Despite of elevated phosphorus levels, low electrolyte and high humus concentrations are unfavourable to species normally found in eutrophic rivers (especially genera *Navicula* and *Nitzschia*). In very turbid waters, model yielded clearly lower values than observed. In low light intensities, primary production may be limited by the level of available radiation, which might lead to even free reactive phosphorus being accumulated in the water.

In conclusion, the WA modelling provides a tool for evaluating trophic conditions, also delivering indicator species suitable for the prevailing conditions and water quality. In WA modelling it is assumed that the variability cannot be reduced greatly by using more complex response curves than the Gaussian curve; it combines simplicity with a good performance.

In a comparative study of river monitoring using diatoms and macroinvertebrates, the observed rather low correlation between ASPT- and IPS-index was quite expected. This is probably due to the basic differences in the two organism group`s roles and function in food webs (**VI**). Stream width was the most important factor regulating the macroinvertebrate community structure, followed by chemical factors like conductivity and pH. Similarly, according to Paavola *et al.* (2000), stream size is an important regulating factor for macroinvertebrate communities in northern Finland, along with water colour and nutrient concentrations. Stream size is usually strongly connected with discharge conditions and light regime. Contrary to macroinvertebrates, diatoms were mainly regulated by chemical factors, especially by nutrient and electrolyte concentrations. The predominance of total P as a regulating factor for diatoms was highlighted. Likewise, Triest *et al.* (2001) pointed out that primary producers are stronger and more straightforward indicators for river`s trophic status than are macroinvertebrates. Diatoms are probably less sensitive to changes in river habitat and its heterogeneity owing to their higher density and smaller spatial extent. Microscopic organisms perceive the world with a very fine resolution, but for them the environment may seem rather homogeneous at larger spatial scales (Azovsky, 2002). Furthermore, a valid time scale for water quality indication is directly dependent on the life cycles of organisms concerned. Diatom community reacts to changes in water quality within a few days – weeks (Eloranta, 1999), whereas macroinvertebrates integrate water quality for much longer period, for some months – couple of years (Skriver, 2000).

For replicate samples, community composition was more similar among diatoms than among macroinvertebrates (**VI**). This can be due to diatom sampling being limited to epilithon. Furthermore, only the diatom samples were considered as true replicates due to the fact that macroinvertebrate samples were taken from different habitats. In addition, for diatoms, cell numbers per cm^2 can be in the millions (Eloranta & Kunnas, 1979; Blinn *et al*., 1980) whereas densities of benthic macroinvertebrates are limited to some hundreds to thousands benthic animals per square meter (Laasonen *et al*., 1998). It is intuitively clear that a higher cell density and higher number of species per area lessen the community variation at small (1-10 m) spatial scales. In fact, smaller organisms have lower turnover diversity, microscopic communities being thus more diverse at small spatial scales (Finlay *et al*. 1996; Fenchel *et al*., 1997; Lawton, 1998).

IPS index values for replicate samples tended to vary more than the ASPT index. This may result from the rather different premises of the indices. In the IPS index, every species has its own sensitivity and indicator value (Coste in CEMAGREF, 1982). Thus, slight changes in the species dominance or introduction of a new species will directly affect the index value. On the other hand, the fact that every species has its own sensitivity and indicator value gives a rather stable base, and the functioning of the IPS index does not depend on whether certain indicator species occur or not. By contrast, in the ASPT index, only family-level identification is needed. Therefore, abundance changes in closely related species, as long as they are in the same families or in the same family group with similar index value, do not affect the outcome.

The objectives of the study should determine the sampling strategy used. In the case of water quality assessment, it is reasonable to minimize natural variation in the algal or faunal communities, due e.g. to substrate type, to be able to focus on the effects of water quality on the biotic communities. Habitat stratification (Norris *et al*., 1996) is a sampling strategy recommended for such an approach. In the case of monitoring e.g. effects of channelization or other structural degradation on benthic communities, samples are needed from many types of habitats: riffles, pools and other available habitats. Macroinvertebrate sampling was conducted in riffles, but in many types of microhabitats to assess the overall riffle community and the local species pool. Diatom sampling concentrated more strictly on assessing water quality, thus samples were collected only from stones, following recommendations of Kelly *et al*. (1998).

Due mainly to reasons based on tradition, macroinvertebrates have a leading role in stream bioassessment in northern Europe, as also in many other parts of the world. Yet, many recent studies have shown that community concordance, i.e. similarity in patterns of community structure among major organism groups (e.g. Jackson & Harvey, 1993), is often rather low in freshwater systems, especially at small (e.g. within-watershed) spatial scales (Allen *et al*., 1999a and b; Paavola *et al*., 2003). Therefore, because freshwater biomonitoring requires considerable effort and resources, it may be advisable and, ultimately, cost-effective to base stream biomonitoring on multiple taxonomic groups, e.g. macroinvertebrates and benthic diatoms.

The use of functional groups or genus-level identification instead of species might alleviate some of the problems introduced by uncertain and variable taxonomy, at least for cost-effective monitoring purposes. Aggregating biota into functional groups is consistent with the use of species traits, instead of taxonomic identities, for fishes or macroinvertebrates (Townsend & Hildrew, 1994; Poff & Allan, 1995), and it has been successfully applied also for benthic algae (Kutka & Richards, 1996; Pan *et al*., 1999; Leland & Porter, 2000). However, too coarse-level identification or species grouping might obscure some important environmental gradients or impact and weaken the spatial structure of the data.

In future studies of diatom community structure, and especially in basic research, semi-quantitative field methods are needed if species richness is being assessed (Paavola et al., unpublished; Soininen et al., unpublished). Samples can be taken along transects, e.g. as ten subsamples, which are then pooled into containers. The subsamples can be taken with a toothbrush and scalpel using a plastic template of a predefined sampling area (ca. 5 cm^2). For comparability of samples, similar number of diatom valves (e.g. 500) should be counted. These methods ensure that sampling effort is consistent, a fact that has been often neglected among diatomists due mainly to methodological traditions.

5. CONCLUSIONS

- Results emphasized the predominance of chemical-constituent concentration and ion composition in structuring benthic diatom communities of boreal streams. Conductivity integrates several important watershed processes and, therefore, is the strongest environmental gradient in explaining diatom distribution patterns in Finnish running waters.

- This thesis showed that diatom communities exhibit a rather strong spatial component especially at a national scale. The proportion of variation explained independently by spatial factors was ca. 25 %, at that scale. Furthermore, it seems that even at rather small spatial scales, pure spatial component still plays an role (ca. 20 % of explained variation) in regulating benthic diatom community composition.

- Given the strong spatial patterns in community composition, it seems evident that bioassessment programs utilising lotic diatoms would benefit from geographical stratification. A combination of regional stratification and local environmental features might provide the most suitable framework for diatom-based bioassessment of boreal streams.

- Diatom species turnover seemed to be highest among epiphyton and lowest among epipelic communities. Although these differences can partly result from lower diversity in epiphyton, they probably indicate lower persistence among epiphytic communities in studied boreal rivers. There were, however, no clear differences in community stability among epiphytic, epilithic and epipelic diatoms.

- The predictive power of the WA model was good in the training set. The model also performed well in the independent test set. Therefore, diatom based weighted averaging is a useful tool for evaluating trophic conditions of running waters in different ecoregions.

- Given the rather low correlation between the diatom and macroinvertebrate indices, it seems advisable and cost-effective to base stream biomonitoring of boreal streams on multiple taxonomic groups, e.g. macroinvertebrates and benthic diatoms.

6. FUTURE RESEARCH DIRECTIONS

As was expected, this thesis raised several important questions, which need thorough studying in the near future. Especially the rather strong spatial structure detected for diatom communities in boreal streams (**I, II**) opened up some new avenues for testing macroecological theories with benthic diatoms as model organism:

(i) Community concordance in boreal streams

Aim of this study is to assess community concordance (Jackson & Harvey, 1993; Allen et al., 1999a and b; Paavola et al., 2003) among multiple freshwater organisms (diatoms, macroalgae, macroinvertebrates, fishes and bryophytes) in boreal headwater streams. All these biotic groups have been sampled semi-quantitatively in 47, mostly near-pristine streams to assess whether these groups show significant congruence in community composition and diversity patterns. Several sets of background data (water chemistry, in-stream physical measurements and catchment variables) are available and these data will be analysed using e.g. several multivariate statistical methods.

(ii) Determining the regional species pool

The aim of this study is to determine the number of local sites that need to be sampled in order to obtain a reliable estimate of the number of species in the regional (within-watershed) species pool (e.g. Dupré, 2000). The regional species richness will be estimated using species-accumulation curves (e.g. Ugland *et al.*, 2003). Semi-quantitative methods will be used during the field and laboratory work.

(iii) Local vs. regional diversity of stream organisms

The main objective of this study is to determine the local vs. regional species richness of benthic diatoms in boreal streams. Seven to eight regions (watersheds) (see Heino *et al.*, 2003b) will be sampled using semi-quantitative field methods. The number of local sites to be sampled per region will be estimated using species accumulation curves (ii). Communities in which local diversity is linearly dependent on regional diversity have been referred to as "unsaturated" or "Type I" communities indicating that they are mainly under regional control (Cornell & Lawton, 1992; Caley & Schluter, 1997; Griffiths, 1999). By contrast, if local species richness saturates at high levels of regional species richness, then communities are regarded as "saturated" or "Type II" communities and controlled by interspecific competition and other local processes. These data will enable a comparison of diversity patterns and their determinants with a comparable data set of benthic macrofauna and aquatic bryophytes sampled at the same sites (Heino *et al.*, 2003b). Several spatial scales (stone, riffle, stream, watershed) will be used in this study.

(iv) Relationship between distribution and abundance of species

This study focuses to explore the relationships between distribution and abundance of species. Locally abundant species tend to be widely distributed, while species which are locally few in number (locally rare) show often only narrow distribution (e.g. Hanski, 1982; Brown, 1984; Hanski & Gyllenberg, 1997; Gaston, 1998). This relationship has been found for several, mostly terrestrial groups of organisms embracing from avifauna to beetles, and this study will test this hypothesis on stream algae, invertebrates and aquatic mosses, i.e. organisms with different dispersal capacity. Preliminary results (Soininen, unpublished) show that distribution and abundance of benthic diatoms are strongly positively correlated (Fig. 12) and clearly the most diatom species show satellite mode (sensu Hanski, 1982) in their distribution further implying that benthic diatom communities are rather strongly spatially structured.

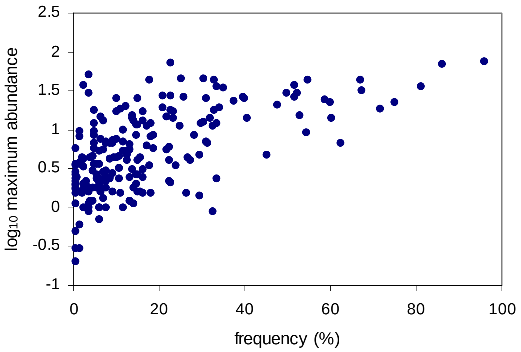

Fig. 12. Relationship between maximum abundance and distribution of benthic diatom species ($r^2 = 0.36$, $p < 0.001$, $n = 212$) sampled at 197 Finnish stream sites.

(v) Relationship between current velocity, turbidity and diversity

Preliminary results show that in turbid rivers diatom species richness is quite high even at high current velocities (Soininen, 2003, 2004). In clear water rivers, diversity is notably lower in fast-flowing part of the riffle. The aim of this study is to explore whether the phenomenon is consistent.

7. ACKNOWLEDGEMENTS

I was not trained for this. I mean that, mainly for historical reasons, limnologists in Finland are experts of mostly or solely the ecology of standing waters. However, with the help of many wonderful persons I managed to finish my thesis. First of all, I would like to thank my supervisor, Prof. Pertti Eloranta, for guiding me into the interesting, but strange world of diatoms. It was year 1997 when I attended a diatom course in Lammi, and I got immediately fascinated. Thanks to Pertti especially for sharing his huge knowledge in algal ecology and taxonomy, as well as spurring me during the final phases of my work. In that diatom course, I also got to know Dr. Janina Kwandrans, a very important and lovely person who has been teaching me diatom identification and ecology through all these years. In year 2000 I attended a benthic algae course at Lake Erken, Sweden. I would like to show my gratitude to Dr. Maria Kahlert for organizing several nice international algae course there, and especially for the possibility to meet many wonderful young algologists. At Erken Lab I met also Riku Paavola, who has ever since been a very important person, especially because of his great wisdom and expertise in multivariate statistical methods and stream ecology in general. Riku also taught me some of his uncompromising attitude towards science during a sampling trip to North Finland. Thanks to Riku, I got to know Prof. Timo Muotka. He has been kindly co-working with me and has taught me a huge amount of things in stream ecology and science in general. I`m grateful also to Pirjo Niemelä for sharing the first steps in diatom ecology and identification with me. Katriina Könönen also kindly worked with me during these years.

I show my deepest gratitude to my opponent Prof. Helmut Hillebrand and my reviewers Prof. Eugen Rott and Prof. Sergi Sabater for spending their time and energy reading and commenting this study. This thesis was supported by grants from Maj and Tor Nessling Foundation, University of Helsinki, Niilo Helander Foundation and Emil Aaltonen Foundation. Timo Muotka, Jyrki Lappalainen, Riku Paavola and Jukka Horppila gave constructive comments on summary of this thesis.

I am indebted to Kimmo Kahilainen, my closest companion in arms, for sharing great moments in our cozy room at the back of the corridor, and discussions about new visions and trends in ecology, as well as for his constant, invisible mental whipping to keep working. I admit that without his company, I would have been lazier during these years. With Kimmo and

Pörje Nyberg, we had great moments during several long nights touring in idyllic bars of Kallio. I was honoured to get to know older scientists, especially Jukka Horppila, Jyrki Lappalainen, Hannu Lehtonen, Tommi Malinen, Leena Nurminen, Arto Tolonen and Mika Vinni, and to have fruitful discussions about freshwater biology. I have also been happy to get to know some nice junior-scientists like Antti Tuomaala, Laura Uusitalo, Anne Liljendahl-Nurminen and Zeynep Pekcan. Discussions with Jani Heino brought up some new visions on how to test ecological theories using stream organisms. And with Mikolaj Kokocinski, I always had so great time during congresses or field courses. I would like also to thank all of my friends, my parents Irma and Pentti, my sisters Helena and Henna and my brother Jussi for keeping my spirit high during this work. A big thanks goes especially to my friend Laura Saikku for wonderful time in Viikki ("Oh no, where is my primary target?"). Finally, I`m deeply grateful to lovely Lotta Hirvenkari for supporting me during the final phases of this work.

8. REFERENCES

Agbeti, M. D. 1992. Relationship between diatom assemblages and trophic variables: a comparison of old and new approaches. - Can. J. Fish. Aquat. Sci. 49: 1171-1175.

Alalammi, P. & Karlson, K.P. 1988. Atlas of Finland, Folio 141-143. Biogeography and nature conservation. - National Board of Survey and Geographical Society of Finland.

Allan, J.D. 1995. Stream ecology. The structure and function of running waters. Chapman & Hall, London.

Allen, A.P., Whittier, T.R., Larsen, D.P., Kaufmann, P.R., O`Connor, R.J., Hughes, R.M., Stemberger, R.S., Dixit, S.S., Brinkhurst, R.O., Herlihy, A.T. & Paulsen, S.G. 1999a. Concordance of taxonomic richness patterns across multiple assemblages in lakes of the northeastern United States. - Can. J. Fish. Aquat. Sci. 56: 739-747.

Allen, A.P., Whittier, T.R., Larsen, D.P., Kaufmann, P.R., O`Connor, R.J., Hughes, R.M., Stemberger, R.S., Dixit, S.S., Brinkhurst, R.O., Herlihy, A.T. & Paulsen, S.G. 1999b. Concordance of taxonomic composition patterns across multiple lake assemblages: effects of scale, body size, and land use. - Can. J. Fish. Aquat. Sci. 56: 2029-2040.

Angermeier, P.L. & Winston, M.R. 1998. Local vs. regional influences on local diversity in stream fish communities of Virginia. - Ecology 79: 911-927.

Antoniades, D. & Douglas, M.S.V. 2002. Characterization of high arctic stream diatom assemblages from Cornwallis Island, Nunavut, Canada. - Can. J. Bot. 80: 50-58.

Armitage, P.D., Moss, D., Wright, J.F. & Furse, M.T. 1983. The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. - Water Res. 17: 333-347.

Azovsky, A.I. 2002. Size-dependent species-area relationship in benthos: is the world more diverse for microbes? - Ecography 25: 273-282.

Belbin, L. & McDonald, C. 1993. Comparing three classification strategies for use in ecology. - J. Veg. Sci. 4: 341-348.

Berry, K. J., Kvamme, K. L. & Mielke, P. W. 1983. Improvements in the permutation test for the spatial analysis of the distribution of artifacts into classes. - Am. Antiq. 48: 547-553.

Biggs, B.J.F. 1990. Periphyton communities and their environments in New Zealand rivers. - N. Z. J. Mar. Freshw. Res. 24: 367-386.

Biggs, B.J.F. 1995. The contribution of flood disturbance, catchment geology and land use to the habitat template of periphyton in stream ecosystems. - Freshwater Biol. 33: 419-438.

Biggs, B.J.F., Duncan, M.J., Jowett, I.G., Quinn, J.M., Hickey, C.W., Davies-Colley, R.J. & Close, M.E. 1990. Ecological characterisation, classification and modelling of New Zealand rivers: an introduction and synthesis. - N. Z. J. Mar. Freshwater. Res. 24: 277-304.

Biondini, M.E., Bonham, C.D. & Redente, E.F. 1985. Secondary successional patterns in a sagebrush (*Artemisia tridentata*) community as they relate to soil disturbance and soil biological activity. - Vegetatio, 60: 25-36.

Blackburn, T.M. & Gaston, K.J. 2003. Macroecology. Concepts and consequences. Blackwell Publishing, Cornwall.

Blinn, D.W., Fredericksen, A. & Korte, V. 1980. Colonization rates and community structure of diatoms on three different rock substrata in a lotic system. - Br. Phycol. J. 15: 303-310.

Borcard, D., Legendre, P. & Drapeau, P. 1992. Partialling out the spatial component of ecological variation. - Ecology 73: 1045-1055.

Bothwell, M.L. 1988. Growth rate responses of lotic periphytic diatoms to experimental phosphorus enrichment: the influence of temperature and light. - Can. J. Fish. Aquat. Sci. 45: 261-269.

Brewin, P.A. Buckton, S.T. & Ormerod, S.J. 2000. The seasonal dynamics and persistence of stream macroinvertebrates in Nepal: do monsoon floods represent disturbance? - Freshwat. Biol. 44: 581-594.

Brown, J.H. 1984. On the relationship between abundance and distribution of species. - Am. Nat. 124: 255-279.

Brown, J.H. & Lomolino, M. 1998. Biogeography. Sinauer, Sunderland.

Burkholder, J.M. & Wetzel, R.G. 1989. Microbial colonization on natural and artificial macrophytes in phosphorus-limited, hardwater lake. - J. Phycol. 25: 55-65.

Caley, M.J. & Schluter, D. 1997. The relationship between local and regional diversity. - Ecology 78: 70-80.

Cao, Y. Williams, D.D. & Williams, N.E. 1998. How important are rare species in aquatic community ecology and bioassessment? - Limnol. Oceanogr. 43: 1403-1409.

Cao, Y., Larsen, D.P. & Thorne, R.St-J. 2001. Rare species in multivariate analysis for bioassessment: some considerations. - J. N. Am. Benthol. Soc.

CEMAGREF 1982. Etude des méthodes biologiques quantitatives d`appréciation de la qualité des eaux. Rapport Division Qualité des Eaux Lyon. Agence financière de Bassin Rhône – Méditerranée – Corse, Pierre-Bénite.

Christie, C.E. & Smol, J.P. 1993. Diatom assemblages as indicators of lake trophic status in southeastern Ontario lakes. - J. Phycol. 29: 575-586.

Connell, J.H. & Sousa, W.P. 1983. On the evidence needed to judge ecological stability or persistence. - Am. Nat. 121: 789-824.

Cornell, H.V. & Lawton, J.H. 1992. Species interactions, local and regional processes, and limits to the richness of ecological communities: a theoretical perspective. - J. Anim. Ecol. 61: 1-12.

Diamond, J.M. & May, R.M. 1977. Species turnover rates on islands: dependence on census interval. - Science 197: 266-270.

Digby, P.G.N. & Kempton, R.A. 1987. Multivariate Analysis of Ecological communities. Chapman and Hall.

Dufrene, M. & Legendre, P. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. - Ecol. Monogr. 67: 345-366.

Duncan, S.W. & Blinn, D.W. 1989. Importance of physical variables on the seasonal dynamics of epilithic algae in a highly shaded canyon stream. - J. Phycol. 25: 455-461.

Dupré, C. 2000. How to determine a regional species pool: a study in two Swedish regions. - Oikos 89: 128-136.

Ekholm, P. 1998. Algal-available phosphorus originating from agriculture and municipalities. - Monographs of the Boreal Environment Research 11: 7-60.

Eloranta, P. 1995. Type and quality of river waters in central Finland described using diatom indices. - In: Marino, D. & Montresor, M. (eds.), Proceedings of the 13th International Diatom Sympsium. Biopress, pp. 271-280.

Eloranta, P., 1999. Applications of diatom indices in Finnish rivers. In J. Prygiel, B.A. Whitton & Bukowska J. (eds.), Use of algae in monitoring rivers III. - Agence de l´Eau Artois-Picardie, Douai. pp. 138-144.

Eloranta, P. & Andersson, K. 1998. Diatom indices in water quality monitoring of some South-Finnish rivers. - Verh. int. Verein. Limnol. 26: 1213-1215.

Eloranta, P. & Kunnas, S. 1979. The growth and species communties of the attached algae in a river system in Central Finland. - Arch. Hydrobiol. 86: 27-44.

Eloranta, P. & Kwandrans, J. 1996. Distribution and ecology of freshwater red algae (Rhodophyta) in some central Finnish rivers. - Nord. J. Bot. 16: 107-117.

Eloranta, P. & Soininen, J. 2002. Ecological status of some Finnish rivers evaluated using benthic diatom communities. - J. Appl. Phycol. 14: 1-7.

Fallu, M.A., Allaire, N. & Pienitz, R. 2002. Distribution of freshwater diatoms in 64 Labrador (Canada) lakes: species:environment relationships along latitudinal gradients and reconstruction models for water colour and alkalinity. - Can. J. Fish. Aquat. Sci. 59: 329-349.

Fenchel, T., Esteban, G.F. & Finlay, B.J. 1997. Local versus global diversity of microorganisms: cryptic diversity of ciliated protozoa. - Oikos 80: 220-225.

Finlay, B.J. 2002. Global dispersal of free-living microbial eukaryote species. - Science 296: 1061-1063.

Finlay, B.J., Esteban, G.F. & Fenchel, T. 1996. Global diversity and body size. - Nature 383: 132-133.

Finlay, B.J., Monaghan, E.B. & Maberly, S.C. 2002. Hypothesis: the rate and scale of dispersal of freshwater diatom species is a function of their global abundance. Protist 153: 261-273.

Fisher, S.G. & Likens, G.E. 1973. Energy flow in Bear Brook, New Hampshire: an integrative approach to stream ecosystem metabolism. - Ecol. Monogr. 43: 421-439.

Fisher, S.G. & Carpenter, S.R. 1976. Ecosystem and macrophyte primary production of the Fort River, Massachusetts. - Hydrobiologia 47: 175-187.

Fisher, S.G., Gray, L.J., Grimm, N.B. & Busch, D.E. 1982. Temporal succession in a desert stream ecosystem following flash flooding. - Ecol. Monogr. 52: 93-110.

Frissell, C.A., Wiss, W.J., Warren, C.E. & Huxley, M.D. 1986. A hierarchical framework for stream classification: viewing streams in a watershed context. - Env. Manag. 10: 199-214.

Gaston, K.J. 1998. Rarity as double jeopardy. - Nature 394: 229-230.

Gaston, K.J. & Blackburn, T.M. 1999. A critique for macroecology. - Oikos 84: 353-368.

Gauch, H. G. 1982. Multivariate analysis in community ecology. Cambridge University Press.

Gause, G. F. 1930. Studies on the ecology of the Orthoptera. - Ecology 11: 307-325.

Gensemer, R.W., Smith, R.E.H., Duthie, H.C. & Schiff, S.L. 1993. pH tolerance and metal toxicity in populations of the planktonic diatom Asterionella: influences of synthetic and natural dissolved organic carbon. - Can. J. Fish. Aquat. Sci. 50: 121-132.

Gerritsen, J., Barbour, M.T. & King, K. 2000. Apples, oranges, and ecoregions: on determining pattern in aquatic assemblages. - J. N. Am. Benthol. Soc. 19: 487-496.

Godfray, H.C.J. & Lawton, J.H. 2001. Scale and species numbers. - Trends Ecol. Evol. 16: 400-404.

Griffith, M.B., Hill, B.H., Herlihy, A.T. & Kaufmann, P.R. 2002. Multivariate analysis of periphyton assemblages in relation to environmental gradients in Colorado Rocky Mountain streams. - J. Phycol. 38: 83-95.

Griffiths, D. 1999. On investigating local-regional species richness relationship. - J. Anim. Ecol. 68: 1051-1055.

Hall, R.I. & Smol, J.P. 1992. A weighted-averaging regression and calibration model for inferring total phosphorus concentration from diatoms in British Columbia (Canada) lakes. - Freshwat. Biol. 27: 417-434.

Hall, R.I. & Smol, J.P. 1996. Paleolimnological assessment of long-term water-quality changes in south-central Ontario lakes affected by cottage development and acidification. - Can. J. Fish. Aquat. Sci. 53: 1-17.

Hanski, 1982. Dynamics of regional distribution: the core and satellite species hypothesis. - Oikos 38: 210-221.

Hanski, I. & Gyllenberg, M. 1997. Uniting two general patterns in the distribution of species. - Science 275: 397-400.

Hart, D.D. & Finelli, C.M. 1999. Physical-biological coupling in streams: the pervasive effects of flow on benthic organisms. - Annu. Rev. Ecol. Syst. 30: 363-395.

Hawkins, C.P. & Vinson, M.R. 2000. Weak correspondence between landscape classifications and stream invertebrate assemblages: implications for bioassessment. - J. N. Am. Benthol. Soc. 19: 501-517.

Hawkins, C.P., Norris, R.H., Gerritsen, J., Hughes, R.M., Jackson, S.K., Johnson, R.K. & Stevenson, R.J. 2000. Evaluation of the use of landscape classifications for the prediction of freshwater biota: synthesis and recommendations. - J. N. Am. Benthol. Soc. 19: 541-556.

Heino, J., Muotka, T., Paavola, R., Hämäläinen, H. & Koskenniemi, E. 2002. Correspondence between regional delineations and spatial patterns in macroinvertebrate assemblages of boreal headwater streams. - J. N. Am. Benthol. Soc. 21: 397-413.

Heino, J., Muotka, T., Mykrä, H., Paavola, R., Hämäläinen, H. & Koskenniemi, E. 2003a. Defining macroinvertebrate assemblage types of headwater streams: implications for bioassessment and conservation. - Ecol. Appl. 13: 842-852.

Heino, J., Muotka, T. & Paavola, R. 2003b. Determinants of macroinvertebrate diversity in headwater streams: regional and local influences. - J. Anim. Ecol. 72: 425-434.

Heino, J.H, Paavola, R.P., Virtanen, R. & Muotka, T. 2004. Searching for biodiversity indicators in running waters: do bryophytes, macroinvertebrates, and fish show congruent diversity patterns? - Biodiv. Cons., in press.

Hill, M.O. & Gauch, H.G. 1980. Detrended correspondence analysis: an improved ordination technique. - Vegetatio 42: 47-58.

Hillebrand, H., Watermann, F., Karez, R. & Berninger, U.G. 2001. Differences in species richness patterns between unicellular and multicellular organisms. - Oecologia 126: 114-124.

Hillebrand, H. & Azovsky, A. 2001. Body size determines the strength of the latitudinal diversity gradient. - Ecography 24: 251-256.

Hillebrand, H. & Blenckner, T. 2002. Regional and local impact on species diversity - from pattern to processes. - Oecologia 132: 479-491.

Hynes, H.B.N. 1970. The Ecology of Running Waters. Liverpool University Press, Liverpool.

Hämäläinen, H. & Huttunen, P. 1996. Inferring the minimum pH of streams from macroinvertebrates using weighted averaging regression and calibration. - Freshwat. Biol. 36: 697-709.

Jackson, D.A. 1995. PROTEST: A Procrustes Randomization TEST of community-environment concordance. - Ecoscience 2: 297-303.

Jackson, D.A. & Harvey, H.H. 1993. Fish and benthic invertebrates: community concordance and community-environment relationships. - Can. J. Fish. Aquat. Sci. 50: 2641-2651.

Johnson, R. K. 2000. Spatial congruence between ecoregions and littoral macroinvertebrate assemblages. - J. N. Am. Benthol. Soc. 19: 475-468.

Jones, R.I. 1990. Phosphorus transformations in the epilimnion of humic lakes: biological uptake of phosphate. - Freshwat. Biol. 23: 323-337.

Kantola, L., Koskenniemi, E., Paavola, R. & Heikkinen, M. 2001. Ohjeita järvien ja jokien pohjaeläimistöseurannan näytteenottoon ja raportointiin (in finnish). Pohjois-Pohjanmaan ympäristökeskus. Ympäristöopas 87.

Kelly, M.G. 1998. Use of the trophic diatom index to monitor eutrophication in rivers. - Water Res. 32: 236-242.

Kelly, M.G. & Whitton, B.A. 1995. The Trophic Diatom Index: a new index for monitoring eutrophication in rivers. - J. Appl. Phycol. 7: 433-444.

Kelly, M.G., Cazaubon, A., Coring, E., Dell`Uomo, A., Ector, L., Goldsmith, B., Guasch, H., Hürlimann, J., Jarlman, A., Kawecka, B., Kwandrans, J., Laugaste, R., Lindstrøm, E.-A., Leitao, M., Marvan, P., Padisák, J., Pipp, E., Prygiel, J., Rott, E., Sabater, S., Van Dam, H. & Vizinet J. 1998. Recommendations for the routine sampling of diatoms for water quality assessments in Europe. - J. Appl. Phycol. 10: 215-224.

King, L., Barker, P. & Jones, R. I. 2000. Epilithic algal communities and their relationship to environmental variables in lakes of the English Lake District. - Freshwat. Biol. 45: 425-442.

Kociolek, J.P. & Spaulding, A.A. 2000. Freshwater diatom biogeography. - Nova Hedw. 71: 223-241.

Korhola, A., Weckström, J. & Nyman, M. 1999. Predicting the long-term acidification trends in small subarctic lakes using diatoms. - J. Appl. Ecol. 36: 1021-1034.

Krammer, K. & Lange-Bertalot, H. 1986-1991. Bacillariophyceae. Süβwasserflora von Mitteleuropa, 2 (1-4). Fischer.

Kutka, F.J. & Richards, C. 1996. Relating diatom assemblage structure to stream habitat quality. - J. N. Am .Benthol. Soc. 15: 469-480.

Laasonen, P., Muotka, T. & Kivijärvi, I. 1998. Recovery of macroinvertebrate communities from stream habitat restoration. - Aquat. Conserv. Mar. Freshwat. Ecosyst. 8: 101-113.

Lange-Bertalot, H. & Metzeltin, D. 1996. Indicators of oligotrophy. 800 taxa repsesentative of three ecologically distinct lake types: carbonate buffered, oligodystrophic, weakly buffered soft water. Iconographia Diatomologica vol.2.

Lawton, J.H. 1996. Patterns in ecology. - Oikos 75: 145-147.

Lawton, J.H. 1998. Small is beautiful, and very strange. - Oikos 81: 3-5.

Lawton, J.H. 1999. Are there general laws in ecology? - Oikos 84: 177-192.

Ledger, M.E. & Hildrew, A.G. 1998. Temporal and spatial variation in the epilithic biofilm of an acid stream. - Freshwat. Biol. 40: 655-670.

Legendre, P. & Legendre, L. 1998. Numerical ecology. Elsevier.

Leland, H.V. 1995. Distribution of phytobenthos in the Yakima River basin, Washington, in relation to geology, land use, and other environmental factors. - Can. J. Fish. Aquat. Sci. 52: 1108-1129.

Leland, H.V. & Porter, S.D. 2000. Distribution of benthic algae in the upper Illinois River basin in relation to geology and land use. - Freshwater Biol. 44: 279-301.

Levin, S.A. 1992. The problem of pattern and scale in ecology. - Ecology 73: 1943-1967.

Li, J., Herlihy, A., Gerth, W., Kaufmann, P., Gregory, S., Urquhart, S. & Larsen, D.P. 2001. Variability of stream macroinvertebrates at multiple spatial scales. - Freshwat. Biol. 46: 87-97.

Lim, D.S.S. Kwan, C. & Douglas, M.S.V. 2001. Periphytic diatom assemblages from Brathurst Island, Nunavut, Canadian high arctic: an examination of community relationships and habitat preferences. - J. Phycol. 37: 379-392.

Malmqvist, B. & Mäki, M. 1994. Benthic macroinvertebrate assemblages in north Swedish streams: environmental relationships. - Ecography 17: 9-16.

Mann, D.G. & Droop, S.J.M. 1996. Biodiversity, biogeography and conservation of diatoms. - Hydrobiologia 336: 19-32.

McCormick, F.H., Peck, D.V. & Larsen, D.P. 2000. Comparison of geographic classification schemes for Mid-Atlantic stream fish assemblages. - J. N. Am. Benthol. Soc. 19: 385-404.

McCune, B. & Mefford, M. J. 1999. PC-ORD. Multivariate Analysis of Ecological Data, Version 4. MjM Software Design.

McGeoch, M. & Chown, S.L. 1998. Scaling up the value of bioindicators. - Trends Ecol. Evol. 13: 46-47.

McGill, B. 2003. Strong and weak tests of macroecological theory. - Oikos 102: 679-685.

Menge, B.A. & Olson, A.M. 1990. Role of scale and environmental factors in regulation of community structure. - Trends Ecol. Evol. 5: 52-57.

Miettinen, J. 2003. A diatom-total phosphorus transfer function for freshwater lakes in southeastern Finland, including cross-validation including independent test lakes. - Boreal Env. Res. 8:

Minchin, P.R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. - Vegetatio 69: 89-107.

Moss, D., Wright, J.F., Furse, M.T. & Clarke, R.T. 1999. A comparison of alternative techniques for prediction of the fauna of running water sites in Great Britain. - Freshwater Biol. 41: 167-181.

Munn, M.D., Black, R.W. & Gruber, S.J. 2002. Response of benthic algae to environmental gradients in an agriculturally dominated landscape. - J. N. Am. Benthol. Soc. 21: 221-237.

Norris, R.H., McEltavy, E.P. & Resh, V.H. 1996. The sampling problem. In Petts, G. & Calow, P. (eds.) River biota. Diversity and dynamics.

Økland, R. H. 1999. On the variation explained by ordination axes and constrained ordination axes. - J. Veg. Sci. 10: 131-136.

Økland, R. H. & Eilertsen, O. 1994. Canonical correspondence analysis with variation partitioning: some comments and an application. - J. Veg. Sci. 5: 117-26.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. - Ann. Ass. Am. Geogr. 77: 118-125.

Paavola, R., Muotka, T. & Tikkanen, P. 2000. Macroinvertebrate community structure and species diversity in humic streams of Finnish Lapland. - Verh. Int. Verein. Limnol. 27: 2550-2555.

Paavola, R., Muotka, T., Virtanen, R., Heino, J. & Kreivi, P. 2003. Are biological classifications of headwater streams concordant across multiple taxonomic groups? - Freshwat. Biol. 48: 1912-1923.

Pan, Y., Stevenson, R.J., Hill, B.H., Kaufmann, P.R. & Herlihy, A.T. 1999. Spatial patterns and ecological determinants of benthic algal assemblages in Mid-Atlantic streams, USA. - J. Phycol. 35: 460-468.

Pan, Y., Stevenson, R.J., Hill, B.H. & Herlihy, A.T. 2000. Ecoregions and benthic diatom assemblages in Mid-Atlantic Highlands streams, USA. - J. N. Am. Benthol. Soc. 19: 518-540.

Parsons, M., Thoms, M. C. & Norris, R. H. 2003. Development of stream macroinvertebrate models that predict watershed and local stressors in Wisconsin. - J. N. Am. Benthol. Soc. 22: 105-122.

Passy, S.I. 2001. Spatial paradigms of lotic diatom distribution: a landscape ecology perspective. - J. Phycol. 37: 370-378.

Peterson, C.G. & Stevenson, J.R. 1989. Substratum conditioning and diatom colonization in different current regimes. - J. Phycol. 25: 790-793.

Peterson, C.G. & Stevenson, J.R. 1990. Post-spate development of epilithic algal communities in different current environments. - Can. J. Bot. 68: 2092-2102.

Peterson, B.J., Hobbie, J.E. & Corliss, T.L. 1986. Carbon flow in a tundra stream ecosystem. - Can. J. Fish. Aquat. Sci. 43: 1259-1270.

Pienitz, R., Douglas, M.S.V., Smol, J.P., Huttunen, P. & Meriläinen, J. 1995. Diatom, chrysophyte and protozoan distributions along a latitudinal transect in Fennoscandia. - Ecography 18: 429-439.

Poff, N.L. 1992. Why disturbances can be predictable: a perspective on the definition of disturbance in streams. - J. N. Am. Benthol. Soc. 11: 86-92.

Poff, N.L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. - J. N. Am. Benthol. Soc. 16: 391-409.

Poff, N.L. & Allan, J.D. 1995. Functional organization of stream fish assemblages in relation to hydrological variability. - Ecology 76: 606-627.

Potapova, M.G. & Charles, D.F. 2002. Benthic diatoms in USA rivers: distributions along spatial and environmental gradients. - J. Biogeogr. 29: 167-187.

Prygiel, J., & Coste, M. 1993. The assessment of water quality in the Artois-Picardie water basin (France) by the use of diatom indices. - Hydrobiologia 269/270: 343-349.

Prygiel, J., Whitton, B.A. & Bukowska, J, (eds.) 1999. Use of algae for monitoring rivers III. Agence de l´Eau Artois-Picardie, Douai.

Reavie, E.D. & Smol, J.P. 1997. Diatom-based model to infer past littoral habitat characteristics in the St. Lawrence River. - J. Great Lakes Res., 23: 339-348.

Resh, V.H., Brown, A.V., Covich, A.P., Gurtz, M.E., Li, H.W., Minshall, G.W., Reice, S.R., Sheldon, A.L., Wallace, J.B. & Wissmar, R. 1988. The role of disturbance in stream ecology. - J. N. Am. Benthol. Soc. 7: 433-455.

Ricklefs, R.E. 1987. Community diversity: relative roles of local and regional processes. - Science 235: 167-171.

Robinson, C.T., Minshall, G.W. & Royer, T.V. 2000. Inter-annual patterns in macroinvertebrate communities of wilderness streams in Idaho, USA. - Hydrobiologia 421: 187-198.

Rosenberg, D.M. & Resh,V.H. (eds.) 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman and Hall. New York.

Round, F.E. 1981. The ecology of algae. Cambridge University Press, Cambridge.

Sandin, L. 2003. Benthic macroinvertebrates in Swedish streams: community structure, taxon richness, and environmental relations. - Ecography 26: 269-282.

Sandin, L. & Johnson, R. K. 2000. Ecoregions and benthic macroinvertebrate assemblages of Swedish streams. - J. N. Am. Benthol. Soc. 19: 462-474.

Scarsbrook, M.R. 2002. Persistence and stability of lotic invertebrate communities in New Zealand. - Freshwat. Biol. 47: 417-431.

SFS 5077 1989. Water quality. Sampling the bottom fauna in running waters.

Skriver, J. 2000. Biological monitoring in Nordic Rivers and lakes. Report to the Nordic Council of Ministers.

Soininen, J. 2003. Heterogeneity of benthic diatom communities in different spatial scales and current velocities in a turbid river. - Arch. Hydrobiol. 156: 551-564.

Soininen, J. 2004. Assessing the current related heterogeneity and diversity patterns of benthic diatom communities in a turbid and a clear water river. - Aquat. Ecol. (in press).

Sommer, U. 2000. Benthic microalgal diversity enhanced by spatial heterogeneity of grazing. - Oecologia 122: 284-287.

Steinberg, C. & Schiefele, S. 1988. Biological indication of trophy and pollution of running waters. - Z Wasser-Abwasser Forsch 21: 227-234.

Stevenson, R.J. 1997. Scale-dependent determinants and consequences of benthic algal heterogeneity. - J. N. Am. Benthol. Soc. 16: 248-262.

Stevenson, R. J., Bothwell, M. L. & Lowe, R. L. (eds.). 1996. Algal ecology. Academic Press, Inc. San Diego.

Stoermer, E.F. & Smol, J.P. (eds.) 1999. The Diatoms: Applications for the environmental and earth sciences. Cambridge University Press, Cambridge.

ter Braak, C. J. F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. - Ecology 67: 1167-1178.

ter Braak, C.J.F. & Looman, C.W.N. 1986. Weighted averaging, logistic regression and the Gaussian response model. - Vegetatio 65: 3-11.

ter Braak, C. J. F. & Prentice, I.C. 1988. A theory of gradient analysis. - Adv. Ecol. Res. 18: 271-317.

ter Braak, C. J. F. & Smilauer, P. 1998. Program CANOCO, version 4.0. Centre for Biometry.

ter Braak, C. J. F. & van Dam, H. 1989. Inferring pH from diatoms: a comparison of old and new calibration methods. - Hydrobiologia 178: 209-223.

ter Braak, C. J. F. & Verdonschot, P. 1995. Canonical correspondence analysis and related multivariate methods in aquatic ecology. - Aquatic Sciences 57: 255-289.

Tonn, W.M., Magnuson, J.J., Rask, M. & Toivonen, J. 1990. Intercontinental comparison of small-lake fish assemblages: the balance between local and regional processes. - Am. Nat. 136: 345-375.

Townsend, C.R. & Hildrew, A.G. 1994. Species traits in relation to a habitat templet for river systems. - Freshw. Biol. 31: 265-275.

Townsend, C.R. Hildrew, A.G. & Schofield, K. 1987. Persistence of stream invertebrate communities in relation to environmental variability. - J. Anim. Ecol. 56: 597-613.

Triest, L., Kaur, P., Heylen, S. & De Pauw, N. 2001. Comparative monitoring of diatoms, macroinvertebrates and macrophytes in the Woluwe River (Brussels, Belgium). - Aquat. Ecol. 35: 9-17.

Ugland, K.I., Gray, J.S. & Ellingsen, K.E. 2003. The species-accumulation curve and estimation of species richness. J. Anim. Ecol. 72: 888-897.

Van Dam, H., Mertens, A. & Sinkeldam, J. 1994. A coded checklist and ecological indicator values of freshwater diatoms from the Netherlands. - Neth. J. Aquat. Ecol. 28: 117-133.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. & Cushin, C.E. 1980. The river continuum concept. - Can. J. Fish. Aquat. Sci. 37: 130-137.

Van Sickle, J. & Hughes, R.M. 2000. Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon. - J. N. Am. Benthol. Soc. 19: 370-384.

Wallach, D. & Goffinet, B. 1989. Mean squared error of prediction as a criterion for evaluating and comparing system models. - Ecol. Model. 44: 299-306.

Weatherley, N.S. & Ormerod, S.J. 1990. The constancy of invertebrate assemblages in soft-water streams: implications for the prediction and detection of environmental change. - J. Appl. Ecol. 27: 952-964.

Weckström, J., Korhola, A. & Blom, T. 1997. Diatoms as quantitative indicators of pH and water temperature in subarctic Fennoscandian lakes. - Hydrobiologia 347: 171-184.

Whittier, T.R., Hughes, R.M. & Larsen, D.P. 1988. Correspondence between ecoregions and spatial patterns in stream ecosystems in Oregon. - Can. J. Fish. Aquat. Sci. 45: 1264-1278.

Whitton, B.A. (ed.) 1975. River ecology. University of California Press, Berkeley CA.

Whitton, B.A. & Rott, E. (eds.) 1996. Use of algae for monitoring rivers II. Institüt für Botanik, Universität Innsbruck, Innsbruck.

Whitton, B.A. E. Rott, & Friedrich, G. (eds.) 1991. Use of algae for monitoring rivers. Institüt für Botanik, Universität Innsbruck, Innsbruck.

Winter, J.G. & Duthie, H.C. 2000. Epilithic diatoms as indicators of stream total N and total P concentration. - J. N. Am. Benthol. Soc. 19: 32-49.

Wright, J.F., Armitage, P.D. & Furse, M.T. 1984. A preliminary classification of running water sites in Great Britain based on macroinvertebrate species and the prediction of community types using environmental variables. - Freshw. Biol. 14: 221-256.

Wright, J. F., Furse, M. T. & Moss, D. 1998. River classification using invertebrates: RIVPACS application. - Aquat. Cons. Mar. Freshwat. Ecosyst. 8: 617-631.

Zimmermann, G. M., Goetz, H. & Mielke, P. W. 1985. Use of an improved statistical method for group comparisons to study effects of prairie fire. - Ecology 66: 606-611.

Zobel, M. 1997. The relative role of species pools in determining plant species richness: an alternative explanation of species coexistence? - Trends Ecol. Evol. 12: 266-269.