

REVIEW AND
SYNTHESISQuantifying the evidence for biodiversity effects on
ecosystem functioning and services

Patricia Balvanera,^{1*} Andrea B. Pfisterer,² Nina Buchmann,³ Jing-Shen He,⁴ Tohru Nakashizuka,⁵ David Raffaelli⁶ and Bernhard Schmid²

¹Centro de Investigaciones en Ecosistemas, UNAM, Morelia, Mexico

²Institute of Environmental Sciences, Universität Zürich, Zürich, Switzerland

³Institute of Plant Sciences, ETH Zurich, Zurich, Switzerland

⁴Department of Ecology, Peking University, Peking, China

⁵Research Institute for Humanity and Nature, Kyoto, Japan

⁶Environment Department, University of York, York, UK

*Correspondence: E-mail: pbalvane@oikos.unam.mx

Abstract

Concern is growing about the consequences of biodiversity loss for ecosystem functioning, for the provision of ecosystem services, and for human well being. Experimental evidence for a relationship between biodiversity and ecosystem process rates is compelling, but the issue remains contentious. Here, we present the first rigorous quantitative assessment of this relationship through meta-analysis of experimental work spanning 50 years to June 2004. We analysed 446 measures of biodiversity effects (252 in grasslands), 319 of which involved primary producer manipulations or measurements. Our analyses show that: biodiversity effects are weaker if biodiversity manipulations are less well controlled; effects of biodiversity change on processes are weaker at the ecosystem compared with the community level and are negative at the population level; productivity-related effects decline with increasing number of trophic links between those elements manipulated and those measured; biodiversity effects on stability measures ('insurance' effects) are not stronger than biodiversity effects on performance measures. For those ecosystem services which could be assessed here, there is clear evidence that biodiversity has positive effects on most. Whilst such patterns should be further confirmed, a precautionary approach to biodiversity management would seem prudent in the meantime.

Keywords

Biodiversity–ecosystem functioning, diversity manipulations, ecosystem property, ecosystem services, ecosystem type, experimental design, meta-analysis, stability, trophic level.

Ecology Letters (2006) 9: 1146–1156

INTRODUCTION

Human needs have been, and continue to be, satisfied at the expense of altered land use, climate, biogeochemical cycles and species distributions (MA 2005). As a result, biodiversity is declining a thousand times faster now than at rates found in the fossil record (MA 2005), raising concerns about consequences of such loss for ecosystem functioning, the provision of ecosystem services and human well being (Schläpfer & Schmid 1999; Chapin *et al.* 2000; Loreau *et al.* 2001; Kinzig *et al.* 2002; Díaz *et al.* 2005; Hooper *et al.* 2005; MA 2005; Srivastava & Vellend 2005). Such concerns have moved beyond the science community to the global stakeholder and policy community with the publication of the Millennium Assessment (Díaz *et al.* 2005; MA 2005). That analysis acknowledges that biodiversity probably plays

a significant role in directly providing goods and services as well as regulating and modulating ecosystem properties (this term is used here to include 'processes' and 'functioning') that underpin the delivery of ecosystem services.

Considerable research has gone into teasing out the linkages between biodiversity, functioning and services (Naeem & Wright 2003), and experimental approaches now account for 40% of the publications in this area (Fig. 1). Most experiments have manipulated diversity or have assembled different diversities as a treatment variable and documented the response of ecosystem properties and processes, including modifying effects of environmental factors on such relationships (Naeem *et al.* 1994; Tilman 1996; McGrady-Steed *et al.* 1997; Hector *et al.* 1999). The experimental designs used, results obtained and interpretations made, have not been consistent and the field has been contentious and lively

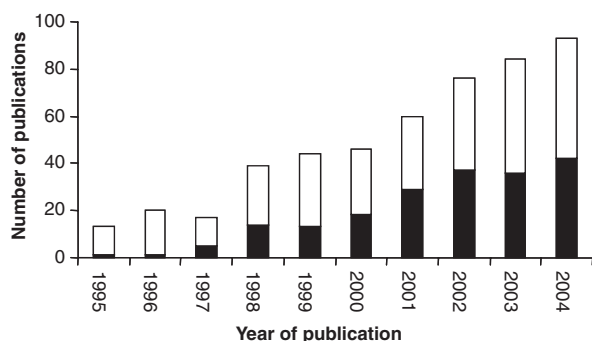


Figure 1 The number of biodiversity–ecosystem functioning articles published during the last decade is steadily growing (ISI Web of Science). Experimental work (filled section) has contributed around 40% of the total number of articles (total bar) since the beginning of this century.

(Grime 1997; Wardle *et al.* 1997; Huston *et al.* 2000; Lepš 2004). Attempts have been made to provide common frameworks, identify areas of consensus or future challenges, as well as potential management and policy implications (Schläpfer & Schmid 1999; Loreau *et al.* 2001; Kinzig *et al.* 2002; Schmid *et al.* 2002; Díaz *et al.* 2005; Hooper *et al.* 2005), but these syntheses have taken the form of largely subjective assessments through qualitative literature reviews. Such reviews provided an important foundation (in particular Schmid *et al.* 2002) for us to construct a more complete database using strict selection criteria (Schläpfer & Schmid 1999) for the formal meta-analysis presented here. Specifically, we pose the following questions: (i) what are the most commonly addressed relationships between biodiversity and ecosystem properties? (ii) How do the experimental designs used and the ecosystem properties measured affect the outcomes and interpretation of biodiversity–ecosystem functioning relationships? (iii) What can be learnt about biodiversity–ecosystem service relationships that could be useful for decision makers?

METHODS

Data collection

One hundred and three publications were included in our database, representing 446 ecosystem property measurements from 1954 to June 2004 (see Appendix S1 and Table S1). These publications were identified from the ISI Web of Science and Biological Abstracts database using criteria previously using the following search terms (Schläpfer & Schmid 1999): *biodiversity or species richness and stability or ecosystem function or productivity or yield or food web*. Where appropriate, we contacted authors of publications to obtain additional information and additional publications. Information about

specifics of experimental designs, the ecosystem properties measured and the significance and size of reported effects were entered into our database. We did not include duplicate records, for example, the same experiment and same measurement reported in a different publication or measured in a different year (repeated measures). If, however, the repeated measures were used to derive a new variable such as temporal variation in the ecosystem property, these data were included. We did not include studies that compared monocultures with mixtures of a single higher diversity level or single-species removal experiments. We used all records that reported effect sizes, allowing us to calculate correlation coefficients for the relationship between biodiversity and ecosystem property, but we excluded studies from our database, which reported only significance.

Data analyses

Biodiversity effects were measured as simple or multiple correlation coefficients, r . Using r instead of r^2 (the coefficient of determination) had the advantage that we could assign negative and positive signs to effects. Maintaining negative and positive effects and using a Z -transformation (see below) allowed us to test the overall distribution for normality and to obtain normally distributed error terms after fitting explanatory terms.

Simple correlation coefficients (365 records) were only available where biodiversity was treated as an independent continuous variable or where a linear or log-linear contrast was made for the factor biodiversity. When biodiversity was analysed as a factor with more than one level (or as a polynomial), we calculated multiple correlation coefficients from the entries in the analysis of variance tables (81 records). We used adjusted r^2 values to derive correlation coefficients because these correct for the degrees of freedom used to fit a model (Sokal & Rohlf 1995). When the relationship between the levels of the biodiversity factor and the response variable was generally negative, we gave the multiple correlation coefficient a minus sign. In addition to the sign, we also noted the shape of the relationship (see below). To simultaneously analyse simple and multiple correlation coefficients we normalized them using Fisher's z -algorithm (Rosenberg *et al.* 2000)

$$Z_r = 0.5 \times \ln \left(\frac{1+r}{1-r} \right) \quad (1)$$

and analysed these Z_r -values as a new dependent variable. We did all analysis with all 446 correlation coefficients and with the subset of the 365 simple coefficients. Because the results were the same, we only present those from the full analysis.

The common, normalized effects measure allowed us to analyse all data together with a single general-linear modelling framework, despite the overwhelming heterogen-

city of studies. Based on major controversies as well as areas of consensus identified in previous qualitative synthesis (Schläpfer & Schmid 1999; Loreau *et al.* 2001; Kinzig *et al.* 2002; Schmid *et al.* 2002; Diaz *et al.* 2005; Hooper *et al.* 2005), a set of hypothesis were constructed about possible effects of the specifics of experimental designs and the ecosystem properties measured on the biodiversity effects observed (Table 1). The studies were classified into groups using a separate explanatory factor for each of the hypotheses (Table 1). The significance and explanatory power of these factors and of interactions was then assessed in mixed-model analyses of variance (ANOVA). Study site and reference were random terms in the model.

We compared a small number of alternative models for the fixed terms using adjusted r^2 values (which gave the same model ranking as Akaike and Bayesian information criteria). The selected final model contained only main effects but no interactions of fixed terms. Due to correlations between fixed terms, we assessed their explanatory power in two ways if they were entered: (i) first into the model or (ii) in a sequence of decreasing order of their F -values when entered first. The random effects were added after the fixed effects in the sequence study site/reference, imposing a nesting of these terms. In one case, a single publication reported results from two study sites and in

another case, a single publication reported results from two separate experiments. In these two cases, we gave each publication two reference IDs to ensure full nesting. To avoid weak pseudo-replication due to measurements of multiple ecosystem properties in single experiments, terms referring to specifics of experimental design and study site could be tested against the reference ID instead of the residual mean square as error term. We used this very strict test but list the mean squares in the ANOVA table so that readers can calculate the more liberal F -test as well. The reciprocal of the variance in the individual Z_r values, based on the individual study sizes, was used as a weighting factor in the ANOVA (Crawley 1993). This ensured that studies with small sample sizes were not over-rated in comparison with studies with large sample sizes. Throughout the paper, we report result in terms of these weighted average normalized effect sizes Z_r and their standard errors.

Ecosystem properties that could unequivocally be related to ecosystem services (MA 2003; Diaz *et al.* 2005), and thus that could be assigned a positive (or negative) value for human well being, were further analysed based on mean values and standard errors of effect sizes. Some judgment is involved in the assignment of positive or negative value, because a particular ecosystem property may not be seen as the same benefit by all stakeholders of biodiversity

Table 1 Hypotheses tested in the meta-analysis and corresponding explanatory terms in ANOVA

Explanatory term	Null hypothesis
Type of diversity measure	H_1 , biodiversity effects are independent of type of diversity measure used to estimate relationship (e.g. species vs. functional diversity)
Type of experimental system	H_2 , biodiversity effects are independent of type of experimental system (e.g. bottle, field)
Ecosystem type	H_3 , biodiversity effects are independent of ecosystem type (e.g. grassland, forest)
Main cause of diversity changes	H_4 , biodiversity effects are independent of main cause of diversity changes (direct vs. indirect manipulation of diversity)
Design for direct species diversity manipulations	H_5 , biodiversity effects are the same whether total density is held constant (substitutive designs) or not (additive or designs without control of total density)
Type of indirect species diversity gradients	H_6 , biodiversity effects are independent of the type of indirect species diversity gradients [natural variation vs. gradient (e.g. nitrogen addition)]
Maximum species number	H_7 , biodiversity effects are independent of maximum species number in most diverse treatment
Trophic-level manipulated	H_8 , biodiversity effects are independent of trophic level manipulated
Trophic level measured	H_9 , biodiversity effects are independent of trophic level measured
Number of trophic links between them	H_{10} , biodiversity effects are independent of number of trophic links between level manipulated and level measured
Ecosystem property	H_{11} , biodiversity effects are independent of the ecosystem property measured
Organization level of ecosystem property	H_{12} , biodiversity effects are independent of the level of organization at which the ecosystem property was measured (population- vs. community- vs. ecosystem-level)
Biotic vs. abiotic ecosystem properties	H_{13} , biodiversity effects are independent of whether ecosystem property is biotic or abiotic
Dominant cycle to which ecosystem property belongs	H_{14} , biodiversity effects are independent of whether ecosystem property is associated to water, nutrient, energy or biotic dynamics
Nature of ecosystem property	H_{15} , biodiversity effects are independent of whether ecosystem property is a stock or a rate
Study site	H_{16} , biodiversity effects are independent of location of study site

Listed are the null hypotheses we tried to reject.

(Srivastava & Vellend 2005). Only those ecosystem properties for which at least five effect size measurements were available were included in the analysis.

Groupings for specifics of experimental design and ecosystem properties (number of records in parentheses)

Type of diversity measure

These included species richness (393), functional group richness (23), evenness (11) and diversity indices (19). Although we aimed to include diversity effects in the broadest sense of the word, the majority of studies examined species richness effects only. Some studies reported effects of functional group richness, but only a few of these were intentionally designed from the start to examine effects of varying functional diversity.

Type of experimental system

System types were bottle (microcosm studies) or pot (111), greenhouse, including climate chambers (62) and field (273). Pot and greenhouse systems differ from field systems in that the latter experience natural climate and light regimes. Field systems included studies that directly and indirectly manipulated species diversity.

Main cause of diversity change

Direct manipulations (398) of diversity were distinguished from indirect ones (48). Indirect manipulations were found only in field studies and were further categorized as follows.

Type of indirect species diversity gradients

Indirect manipulations of diversity were divided into natural variation (39) and gradient (9). In the first category, naturally varying diversity levels were constructed. In the second category, a natural (succession) or experimental gradient in environmental conditions (nutrient application or multiple factors) generated the differences in diversity levels.

Design of direct species diversity manipulation experiments

Direct manipulations of diversity were subdivided into those which were set up so that total density remained constant, i.e. substitutive experiments (357), and others, mostly additive experiments (41).

Maximum species number

Three levels of maximum diversity were recognized: low (≤ 10 species, $n = 211$), intermediate (11–20 species, $n = 104$) and high (> 20 species, $n = 131$).

Ecosystem type

These encompassed forest (43), grassland (258), marine (32), freshwater (68), bacterial microcosm (seven), soil commu-

nity (15), crop/successional (10) and ruderal/salt marsh (13).

Trophic level manipulated and trophic level measured

Studies that manipulated diversity and/or measured diversity effects at different trophic levels were categorized into: primary producer (319 manipulated and 241 measured), primary consumer (30 and 91), secondary consumer (four and 13), detritivores (15 and 38), mycorrhiza (47 and 15), multitrophic (31 and five) and ecosystem level (0 and 43). 'Multitrophic' refers to studies where diversity was manipulated on more than one trophic level or where the ecosystem property involves more than one trophic level (e.g. total macrofaunal biomass). Ecosystem level refers to properties measured in the entire ecosystem within the abiotic compartment (e.g. nutrient loss from the system).

Number of trophic links

We counted the number of trophic links between the trophic level manipulated and the level at which the property was measured (Fig. 2).

Effect form

The shapes of the biodiversity–ecosystem property relationships were classified into negative (40), negative linear (92), negative log-linear (41), idiosyncratic (113), positive (70), positive linear (56), positive log-linear (34). This classification was performed independently of significance or size of biodiversity effects simply by inspecting results presented in the text and figures of the publications analysed. This variable is similar to the effect size itself and could be used as an alternative dependent variable in log-linear analysis of deviance. We include this variable in the supplementary online material but except for a single case (see below) the only reported dependent variable in the present paper is effect size *per se*.

Ecosystem properties measured

We included any physical characteristics of the ecosystems, including process rates of energy and nutrient flow. To simplify comparisons, we grouped similar ecosystem properties (EP), which resulted in 28 groups; an additional group was used to collect those measures that could not be assigned. We distinguished between properties of the ecosystem and those of an invader (defined as any species added after the establishment of a community) and we also distinguished between effects on means of properties measured and those that relate to their variances.

Organizational level of the ecosystem property measured

We distinguished between population-level properties, recorded for individual target species, such as density, cover

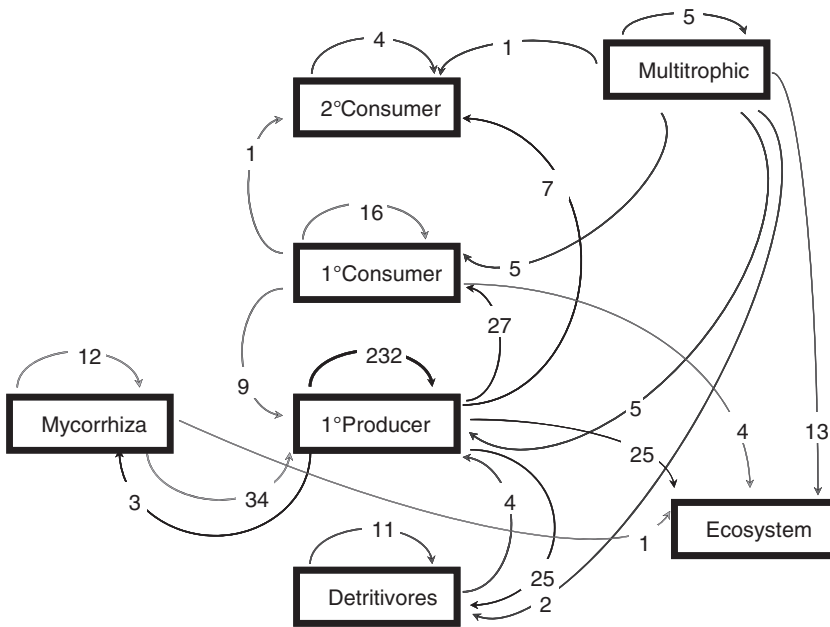


Figure 2 Number of measurements in published biodiversity–ecosystem functioning experiments for different trophic levels manipulated (base of arrow) and trophic levels measured (end of arrow). A dominance of measurements and manipulations of primary producers was observed.

or biomass, and their temporal variance; community-level properties, recorded for multispecies assemblages, such as density, biomass, consumption, diversity and their temporal variance; and ecosystem-level properties, recorded for abiotic components, such as nutrient, water or CO₂ and their temporal variance.

Dominant dynamic of ecosystem property

Properties were assigned to the ecosystem cycle in which they predominate: water, nutrient, energy or biotic dynamics.

Nature of ecosystem property

Stock vs. rate measurements of ecosystem properties were distinguished.

Ecosystem service

Ecosystem services are the benefits people obtain from ecosystems. Our classification followed that of the Millennium Ecosystem Assessment (MA 2003; Díaz et al. 2005). A list of ecosystem properties considered to underpin each ecosystem service, as well as the directionality of expected benefits to human well being, is provided below in the Results section.

Groupings according to place of study and identity of experiment (number of groups in parentheses)

Location of study site (60)

Site location of an experiment ranged from a precise place to a broad region, depending on the extent of the study.

Study site (75)

Generally equivalent to location, this term was used to distinguish different studies within a single location. Study site reflects a set of environmental conditions particular to that experiment.

Reference ID (105)

This corresponded to individual publications, except where a single publication reported results from more than one study, in which case this publication received two reference IDs. This ID is used to distinguish between groups of potentially non-independent measurements in order to avoid pseudo-replication.

RESULTS

The overall mean of the standardized effect sizes Z_r (weighted by the reciprocal of the variance of the individual Z_r -values) was significantly positive ($\bar{X} = 0.101 \pm 0.028$, $t = 3.57$, d.f. = 445, $P < 0.001$), indicating that negative responses of ecosystem properties to biodiversity manipulations are less frequent or less strong than positive ones. Nevertheless, the reported effect sizes varied greatly, ranging from -2.71 to 2.39 . In the following sections, we explore the sources of this variation.

Effects of specifics of experimental design and study site

Some specifics of the experimental design which we originally expected to have an influence on effect sizes in fact could not be included in the final analysis model,

suggesting that they need not be a concern when designing future biodiversity experiments. For instance, there was only a weak influence of the type of diversity measure on measured effect sizes (Table 2). Of particular note is that effect sizes were only slightly larger when functional-group rather than species richness was manipulated (adjusted mean values \pm SE of Z_r -values: 0.191 ± 0.103 vs. 0.116 ± 0.030).

In contrast, the type of experimental system employed (bottle vs. greenhouse vs. field) strongly modified biodiversity effects (Table 2). More positive effects were found where environmental variables could be controlled best, such as in greenhouses and climate chambers (0.467 ± 0.084) compared with bottle/pot experiments (0.100 ± 0.051) or field experiments (0.007 ± 0.033).

Effect sizes also varied markedly between different types of ecosystem (Table 2). For the four ecosystem types which were represented most frequently in the data set, average

effect sizes were close to zero (grassland 0.039 ± 0.038 , freshwater -0.010 ± 0.065 , marine -0.006 ± 0.109 , forest -0.116 ± 0.076), whereas average effect sizes were larger and positive for the ecosystem types with fewer records (ruderal/salt marsh, 1.058 ± 0.154 ; bacterial, 0.317 ± 0.095 ; crop/successional, 0.245 ± 0.052 ; soil, 0.094 ± 0.086). This could imply that the research community's perception of the magnitude and direction of biodiversity effects may be biased by the focus to date on relatively few ecosystem types that included measures of negative impacts on properties. There was considerable variation among study sites, but this was not significant in the multiway ANOVA using the strict F -test with reference ID as error term (Table 2). In other words, effect sizes varied as much between references within study sites as between study sites.

Although average effect sizes were practically identical for studies that manipulated biodiversity directly or indirectly

Table 2 Results from one-way analyses of variance (ANOVA)s in the sequence of decreasing F -values and multiway ANOVA using this sequence for fitting the corresponding fixed terms (see Methods for details)

<i>H</i> no.	Variable	d.f.	Sum of squares	Mean squares	<i>F</i>	<i>P</i> -value	% Explained variance
One-way ANOVA							
12	Organization level EP	2	2031.7	1015.9	40.27	<0.001	15.4
5	Type direct manipulations*	2	1802.5	901.2	35.00	<0.001	13.6
7	Maximum species number	2	1319.0	659.3	24.57	<0.001	10.0
2	Experimental system	2	1071.0	535.3	19.54	<0.001	8.1
3	Ecosystem type	7	2255.8	322.3	12.89	<0.001	17.1
11	Ecosystem property	28	3241.7	115.8	4.83	<0.001	24.5
16	Study site	74	6168.6	83.4	4.39	<0.001	46.7
1	Type diversity measure	3	377.2	125.7	4.33	0.005	2.9
15	Nature of EP	1	86.5	86.5	2.92	n.s.	0.7
8	Trophic-level manipulated	5	305.1	61.0	2.08	n.s.	2.3
9	Trophic-level measured	6	295.2	49.2	1.67	n.s.	2.2
10	Number of links	1	37.4	37.4	1.28	n.s.	0.3
14	Cycle type EP	4	143.9	36.0	1.21	n.s.	1.1
13	Biotic vs. abiotic EP	1	27.3	27.3	0.93	n.s.	0.2
6	Type indirect gradient*	2	14.1	7.1	0.24	n.s.	0.1
4	Direct vs. indirect	1	2.2	2.2	0.07	n.s.	0.0
ANOVA for selected model							
12	Organization level EP	2	2031.9	1016.0	83.69	<0.001	15.38
5	Type direct manipulations†	2	1295.5	647.4	18.19	<0.001‡	9.81
7	Maximum species number	2	349.3	174.7	4.91	<0.05‡	2.64
2	Experimental system	2	485.0	242.5	6.81	<0.01‡	3.67
3	Ecosystem type	7	660.3	94.3	2.65	<0.05‡	5.00
11	Ecosystem property	28	1196.6	42.7	3.52	<0.001	9.06
16	Study site	65	2501.7	38.5	1.08	n.s.‡	18.94
	Reference (within study site)	26	925.5	35.6	2.93	<0.001	7.01
	Residual	337	3762.4	12.0			28.49
	Total	444	13208.1	29.8			100.00

H no., hypothesis number (see Table 1); n.s., not significant ($P > 0.05$).

*These two terms include the last term (direct vs. indirect) as a category 'none'.

†This term includes the term 'direct vs. indirect' as a category 'none'.

‡ F -test using reference ID as error term.

(hypothesis 4), and between versions of indirect manipulations (hypothesis 6), average effect sizes were smaller if direct manipulations maintained total density constant (substitutive designs, 0.031 ± 0.030) than if they did not (0.868 ± 0.102) (Table 2). This confirms something which has long been known to agricultural scientists and plant ecologists using substitutive designs (Harper 1977), the importance of not confounding increasing species richness and total density in experiments.

Average effect sizes were positive if the maximum species richness was larger than 20 species (0.344 ± 0.052) and close to zero for the other two categories (two to 10 species: -0.049 ± 0.030 ; 11–20 species: -0.034 ± 0.081) (Table 2). Yet only 33 of 105 experiments (reference IDs) employed more than 20 species at the highest diversity level. With respect to effect form there was an indication that the odds ratio between linear and log-linear-negative or -positive relationships was greatest in experiments where maximum species richness was lowest ($P < 0.05$), but even where maximum species richness was high, this ratio was > 1 .

There were no overall effects of trophic level manipulated, trophic level measured or number of trophic links between manipulated and response trophic levels (Table 2). Nevertheless, productivity-related effect sizes did significantly decline with increasing number of trophic links ($F_{1,140} = 5.74$, $P < 0.05$).

Effects of ecosystem properties measured

Biodiversity effects differed significantly among the 29 different groups of ecosystem properties (Table 2). A large fraction of the variance in effect sizes was explained by comparing population-, community- and ecosystem-level measures of ecosystem properties (Organization level EP in Table 2). Biodiversity negatively affected population-level measures (-0.332 ± 0.053), but positively affected community-level measures (0.270 ± 0.036). Ecosystem-level measures showed an intermediate response (0.066 ± 0.046). In contrast, no differences were found between biotic and abiotic ecosystem properties, stocks and rates, nor between those more related to carbon, nutrient, water or biotic cycles (terms 'biotic vs. abiotic EP', 'nature of EP' and 'cycle type EP', respectively, in Table 2).

Biodiversity–ecosystem service relationships

Biodiversity effects were explored in more detail by plotting mean values and SE for groups of ecosystem properties in Fig. 3 and relating these groups to ecosystem services.

Productivity is a fundamental supporting ecosystem service that underpins the provision of services such as food or wood (MA 2003; Díaz *et al.* 2005). Generally, increasing biodiversity at one trophic level increased productivity at the

same trophic level (Fig. 3). Plant diversity also appeared to enhance belowground plant and microbial biomass (Fig. 3), indicating positive biodiversity effects on the regulating ecosystem service of erosion control, as large root and mycorrhizal networks are expected to reduce soil erosion.

Positive biodiversity effects (Fig. 3) were found for most ecosystem properties associated with nutrient cycling services. Plant diversity had positive effects on decomposer activity and diversity, and both plant and mycorrhizal diversity increased nutrients stored in the plant compartment of the ecosystem. It is unclear whether plant or detritivore diversity has a general effect on soil nutrient supply.

Increasing the diversity of primary producers contributed to a higher diversity of primary consumers, which we consider here as a supporting service (Fig. 3). Our results also suggest positive effects of biodiversity on the closely related regulating service of pest control; higher plant diversity contributed to lowering plant damage (Fig. 3). The effects of plant diversity on the performance and diversity of predatory insects or other animals that control pests require further investigation. In the case of the regulation of invasive species, a service of economic significance and an area of considerable debate (Levine & D'Antonio 1999; Fargione *et al.* 2003), we found reduced invader abundance, survival, fertility and diversity when plant diversity was higher (Fig. 3).

Temporal stability is directly linked to reliability of service delivery (Díaz *et al.* 2005). Our analysis indicates that more diverse systems have greater temporal stability, as well as greater resistance to external forces such as nutrient perturbations and invading species (Fig. 3). However, this was not the case for other stressors such as warming, drought or a high variance in other environmental conditions. In contrast to the suggestion of qualitative reviews (e.g. Srivastava & Vellend 2005), portfolio and insurance effects of biodiversity (Tilman 1996; Naeem & Li 1997; Yachi & Loreau 1999), i.e. effects on variances or disturbance responses of ecosystem properties, are not more common than performance effects of biodiversity, i.e. effects on means of ecosystem properties ($F_{1,444} = 0.09$, $P = 0.75$).

DISCUSSION

The database assembled here clearly contains an over-representation of some ecosystem types and ecosystem properties, especially grasslands and primary production measures. It is not surprising that experimental grassland plots are often used as model systems in biodiversity studies, because grassland is a widespread system, experiments can be relatively easily set up at constant total density (as opposed to microcosms with strong population dynamics),

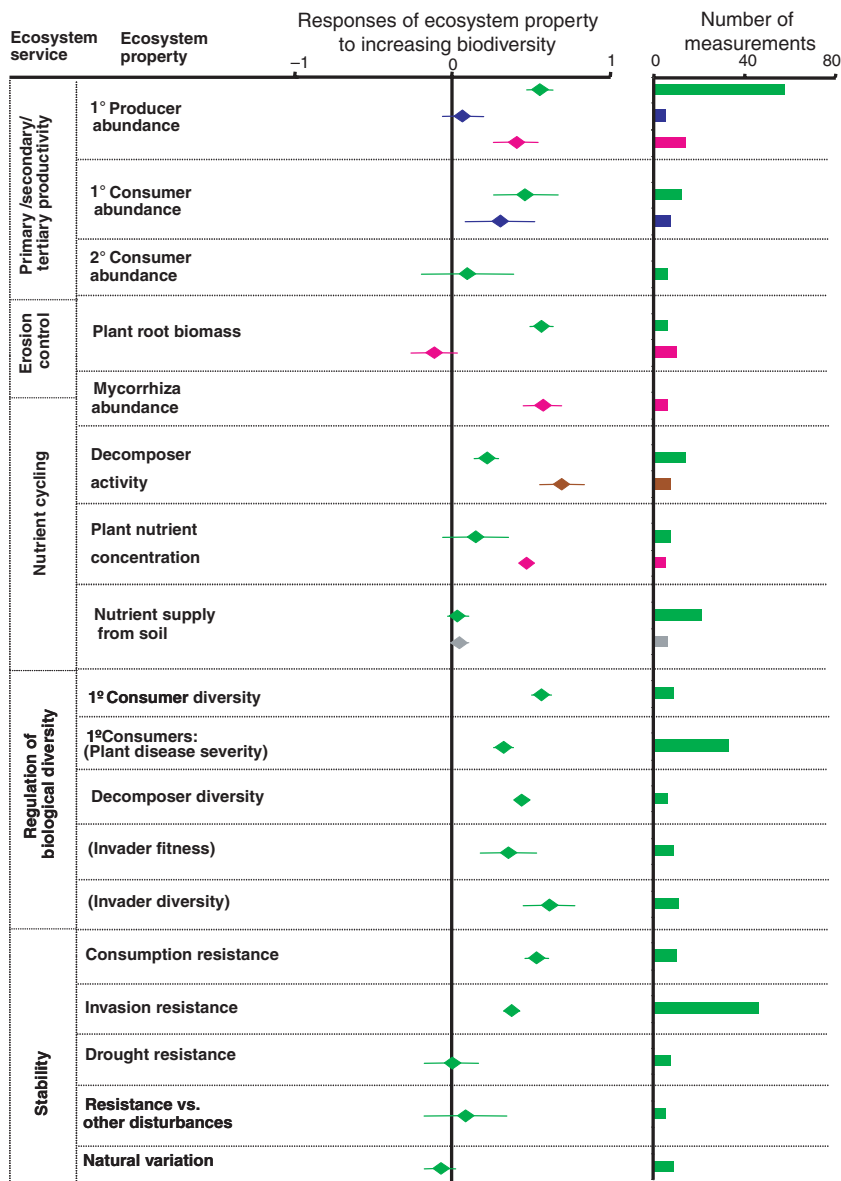


Figure 3 Magnitude and direction of biodiversity effects (shown are mean values and SE of normalized effect sizes Z_n , weighted by the reciprocal of the variance of the individual Z_n -values) and number of measurements available for ecosystem properties organized into ecosystem services. Coloured bars show differential effects of trophic level manipulated: green, primary producers; blue, primary consumers; pink, mycorrhiza; brown, decomposer; grey, multitrophic (multiple levels simultaneously manipulated). Ecosystem properties shown in parentheses were considered of negative value for human well being, and thus opposite of effect sizes are shown.

yet they do not require very large areas (as opposed to forests). In addition, primary productivity plays a major role in delivering a wide range of ecosystem services. Nevertheless, future biodiversity experiments should embrace a broader range of systems, properties and trophic levels if the generality of these relationships is to be established. In particular, a recent experiment that came to light after our analysis was carried out (Bell *et al.* 2005), suggests that bacterial systems hold great promise for future research of biodiversity effects on ecosystem functioning.

Notwithstanding this heterogeneity in the database, our analyses indicate an overall significant positive effect of biodiversity on ecosystem processes. We do not believe that this represents a publication bias towards positive effects,

because finding a significantly negative effect would be just as interesting and just as likely to be reported. Nevertheless, there was significant variation between studies in the magnitude and direction of biodiversity effects, attributable mainly to specifics of experimental design and the ecosystem properties measured, as also argued in qualitative reviews (Hooper *et al.* 2005).

Specifics of experimental design and ecosystem properties

A large number of negative effects were associated with population-level measures, whilst positive effects were associated with community-level measures. This result provides perhaps the strongest empirical evidence to date

for the prediction that individual populations are expected to fluctuate more with increasing biodiversity, but the community stability and productivity should be enhanced (May 1981; Tilman 1996).

In contrast to the outcomes of qualitative reviews (Hooper *et al.* 2005), we could not find a simple dependence of biodiversity effects on the trophic levels manipulated or measured. However, we did find productivity-related biodiversity effects that declined with increasing number of trophic links between those trophic levels which were manipulated and those at which the property was measured. This intuitively compelling result has never been reported before. It is clear that experiments need to be extended beyond the single trophic level approach to better understand such variations in biodiversity effects across an ecosystem (Petchey *et al.* 2002; Raffaelli *et al.* 2002).

Variation in biodiversity effects among study sites and references suggest that local environmental or specific unrecognized experimental factors may either increase or decrease biodiversity effects. Previous work (Hector *et al.* 1999) had already indicated important influences of location on biodiversity effects. The additional variation among references within study sites, which actually made the variation between sites non-significant, is reported here for the first time.

Sufficient information is not available to permit analysis of biodiversity-modifying factors, such as nutrient levels or elevated CO₂ (Hooper *et al.* 2005), but it is clear that biodiversity effects are significantly weaker in less-controlled experimental systems. Indeed, it is much more difficult to maintain diversity treatments on open field plots than in closed bottles; environmental heterogeneity, unpredictable biotic and abiotic environmental fluctuations and sampling variances are greater in the former. Thus, while our results would suggest that further research under controlled conditions is needed to improve our understanding of biodiversity effects on ecosystem functioning, extrapolation of those results to the larger landscape scale is likely to be hindered by the greater environmental heterogeneity and its effects on ecosystem functioning (Loreau *et al.* 2001; Hooper *et al.* 2005). In this respect, field experiments are likely to be more meaningful for extrapolation to the landscape scales at which humans impact on biodiversity and hence service delivery. On the other hand, in a recently constructed grassland experiment in Jena, Germany, Rosher *et al.* (2005) found a similar plant diversity–productivity relationship in small plots of 12.25 m² and in plots more than 30 times larger (400 m²).

The effect on our understanding of the relationship between biodiversity and ecosystem functioning of differences in the way biodiversity is manipulated, how experiments are set up, and how response variables are measured in such experiments has been much debated (Schmid *et al.* 2002; Lepš 2004). Different experimental designs and setups

are acknowledged to have their own advantages and shortcomings; but the present analysis has allowed a formal assessment of the degree to which these really are important. Surprisingly, we found no significant differences between those experiments where diversity was manipulated directly and those involving indirect manipulations by altering environmental conditions. However, there was clear evidence in favour of substitutive designs with control for constant total density of individuals at the start of an experiment. If total density is allowed to vary, in most cases in parallel with species richness, larger effects are seen, but one cannot unequivocally attribute them to biodiversity or density. In other words, such experiments are confounded.

Using a large number of species at the highest diversity levels of an experiment increases the chances of detecting biodiversity effects, although this must be weighed up against the increased work involved in setting up such an experiment. Nevertheless, there is a clear need to include higher levels of species richness in experiments. Unfortunately, interesting new simulation and empirical studies which used non-random extinction scenarios (Raffaelli 2004; Solan *et al.* 2004; Zavaleta & Hulvey 2004; Bunker *et al.* 2005; Schläpfer *et al.* 2005; Srivastava & Vellend 2005) could not be included in our analysis because they were published after our analyses were complete.

An important question when designing a biodiversity–ecosystem functioning experiment is what expression of diversity to manipulate: richness, evenness or functional groups? The literature is somewhat divided on this issue (Díaz & Cabido 2001; Loreau *et al.* 2001; Hooper *et al.* 2005; Petchey & Gaston 2006; Wright *et al.* 2006), but the predominant view is that functional groups may be more important than species richness, consistent with our own findings.

Biodiversity–ecosystem service relationships

Where ecosystem properties could be related to ecosystem services (Srivastava & Vellend 2005), clear positive effects of biodiversity were found, for both regulating and supporting services. Nevertheless, our ability to make these linkages at spatial (landscape) scales relevant to the human enterprise is limited at present (Kremen 2005). There is an urgent need to extend experimental, observational and theoretical work on biodiversity effects for an array of ecosystem functions that can be linked to ecosystem services, such as water quantity and quality, pollination, regulation of pests and human diseases, carbon storage and climate regulation, waste management and cultural services, and to evaluate biodiversity–ecosystem service relationships at the larger spatial scales relevant to management (Kremen *et al.* 2004; Balvanera *et al.* 2005).

The role of biodiversity in buffering environmental variation and thus providing consistent service delivery

has received extensive theoretical treatment (Tilman 1996; Yachi & Loreau 1999; Hooper *et al.* 2005). In general, a positive effect of biodiversity is expected on the stability of ecosystem properties (Tilman 1996; Naeem & Li 1997; Yachi & Loreau 1999; Hooper *et al.* 2005), and qualitative reviews have suggested that such effects on the variance in processes (stability) may be stronger than the effects on means (stocks and fluxes; Srivastava & Vellend 2005). The quantitative results from our meta-analysis do not support this view, rather indicating that biodiversity effects on disturbance buffering are dependent on the nature of the disturbance. Thus, while biodiversity effects on buffering of nutrient perturbations and invading species were positive, biodiversity effects on buffering influences of warming, drought or high environmental variance were neutral or slightly negative.

CONCLUSIONS

Whilst there are many qualitative reviews and position statements about the effects of biodiversity on ecosystem properties and services, our analysis provides the first extensive quantitative meta-analysis of this relationship. This analysis suggests that simple generalizations among ecosystem types, ecosystem properties or trophic level manipulated or measured will be difficult to sustain. Considerations of the way in which biodiversity is defined and manipulated, and disentangling the many separate effects and the interactions between them, as well as those with environmental heterogeneity, will be a major challenge for the next generation of experiments. We offer our database (*Supplementary material*) as a building block for continued synthesis attempts. The advantages of a formal meta-analysis are illustrated by the following novel contributions we have been able to bring to the synthesis: (i) biodiversity effects are weaker if biodiversity manipulations are less well controlled (e.g. field vs. greenhouse or climate chamber); (ii) biodiversity effects are weaker if the highest diversity levels in an experiment are lower (e.g. ≤ 10 vs. > 10 species); (iii) biodiversity experiments should avoid confounding diversity and total density (they should use a substitutive design); (iv) biodiversity effects are weaker at the ecosystem than the community level and negative at the population level; (v) productivity-related biodiversity effects decline with increasing number of trophic links between level manipulated and level measured; (vi) biodiversity effects on stability measures are not obviously stronger than biodiversity effects on performance measures.

There are clear messages for policy makers from these analyses. First, for those ecosystem services that could be assessed in the present study, there is clear evidence that biodiversity has positive effects on the provision of those services and that further biodiversity loss can only be

expected to compromise service delivery. Secondly, whilst further research is needed to confirm such linkages, in particular to extend the work to a broader range of systems and properties, society in the meantime should proceed in a precautionary manner in its use and management of biodiversity.

ACKNOWLEDGEMENTS

We thank the Swiss Biodiversity Forum for administrative support and guidance and SCOPE, with Chris Field, Carlo Heip and Osvaldo Sala for suggestions about the project design. We acknowledge helpful comments from three anonymous referees and S. Naeem which improved the manuscript. We thank the Swiss Agency for the Environment, Forests and Landscape (SAEFL) and the University of Zurich for financial support. We thank Manuel Maass for suggestions and Alberto Valencia and Heberto Ferreira for technical support.

REFERENCES

- Balvanera, P., Kremen, C. & Martínez-Ramos, M. (2005). Applying community structure analysis to ecosystem function: examples from pollination and carbon storage. *Ecol. Appl.*, 15, 360–375.
- Bell, T., Newman, J.A., Silverman, B.W., Turner, S.L. & Liley, A.K. (2005). The contribution of species richness and composition to bacterial services. *Nature*, 436, 1157–1160.
- Bunker, D.E., DeClerck, F., Bradford, J.C., Colwell, R.K., Perfecto, I., Phillips, O.L. *et al.* (2005). Species loss and aboveground carbon storage in a tropical forest. *Science*, 310, 1029–1031.
- Chapin, F.S., Zavaleta, E.S., Eviner, V.T., Naylor, R.L., Vitousek, P.M., Reynolds, H.L. *et al.* (2000). Consequences of changing biodiversity. *Nature*, 405, 234–242.
- Crawley, M.J. (1993). *Glim for Ecologists*. Blackwell, Oxford.
- Díaz, S. & Cabido, M. (2001). Vive la différence: plant functional diversity matters to ecosystem functioning (review article). *Trends Ecol. Evol.*, 16, 646–655.
- Díaz, S., Tilman, D., Fargione, J., Chapin, F.S. III, Dirzo, R., Kitzberger, T. *et al.* (2005). Biodiversity regulation of ecosystem services. In: *Trends and Conditions* (ed. MA). Island Press, Washington, DC, pp. 279–329.
- Fargione, J., Brown, C.S. & Tilman, D. (2003). Community assembly and invasion: an experimental test of neutral versus niche processes. *Proc. Natl Acad. Sci. USA*, 100, 8916–8920.
- Grime, J.P. (1997). Biodiversity and ecosystem function: the debate deepens. *Science*, 277, 1260–1261.
- Harper, J.L. (1977). *Plant Population Biology*. Academic Press, London.
- Hector, A., Schmid, B., Beierkuhnlein, C., Caldeira, M.C., Diemer, M., Dimitrakopoulos, P.G. *et al.* (1999). Plant diversity and productivity experiments in European grasslands. *Science*, 286, 1123–1127.
- Hooper, D.U., Chapin, F.S., Ewell, J.J., Hector, A., Inchausti, P., Lavorel, S. *et al.* (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecol. Monogr.*, 75, 3–35.

- Huston, M.A., Aarssen, L.W., Austin, M.P., Cade, B.S., Fridley, J.D., Garnier, E. *et al.* (2000). No consistent effect of plant diversity on productivity. *Science*, 289, 1255.
- Kinzig, A., Pacala, S.W. & Tilman, D. (2002). *Functional Consequences of Biodiversity: Empirical Progress and Theoretical Extensions*. Princeton University Press, Princeton, NJ.
- Kremen, C. (2005). Managing ecosystem services: what do we need to know about their ecology? *Ecol. Lett.*, 8, 468–479.
- Kremen, C., Williams, N.M., Bugg, R.L., Fay, J.P. & Thorp, R.W. (2004). The area requirements of an ecosystem service: crop pollination by native bee communities in California. *Ecol. Lett.*, 7, 1109–1119.
- Lepš, J. (2004). Diversity and ecosystem function. In: *Vegetation Ecology* (ed. Maarel, E.v.d.). Blackwell, Oxford, pp. 199–237.
- Levine, J.M. & D'Antonio, C.M. (1999). Elton revisited: a review of evidence linking diversity and invasibility. *Oikos*, 87, 15–26.
- Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J.P., Hector, A. *et al.* (2001). Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science*, 294, 804–808.
- MA (2003). *Ecosystems and Human Well-being: A Framework for Assessment*. Island Press, Washington, DC.
- MA (2005). *Ecosystems and Human Well-Being: Synthesis*. Island Press, Washington, DC.
- May, R.M. (1981). *Theoretical Ecology: Principles and Applications*, 2nd edn. Blackwell Scientific Publications, Oxford.
- McGrady-Steed, J., Harris, P.M. & Morin, P.J. (1997). Biodiversity regulates ecosystem predictability. *Nature*, 390, 162–163.
- Naeem, S. & Li, S. (1997). Biodiversity enhances ecosystem reliability. *Nature*, 390, 507–509.
- Naeem, S. & Wright, J.P. (2003). Disentangling biodiversity effects on ecosystem functioning: deriving solutions to a seemingly insurmountable problem. *Ecol. Lett.*, 6, 567–579.
- Naeem, S., Thompson, L.J., Lawler, S.P., Lawton, J.H. & Woodfin, R.M. (1994). Declining biodiversity can alter the performance of ecosystems. *Nature*, 368, 734–736.
- Petchey, O.L. & Gaston, K.J. (2006). Functional diversity: back to basics and looking forward. *Ecol. Lett.*, 9, 741–758.
- Petchey, O., Morin, P.J., Hulot, F.D., Loreau, M., McGrady-Steed, J. & Naeem, S. (2002). Contributions of aquatic model systems to our understanding of biodiversity and ecosystem functioning. In: *Biodiversity and Ecosystem Functioning. Synthesis and Perspectives* (eds Loreau, M., Naeem, S. & Inchausti, P.). Oxford University Press, Oxford, pp. 127–138.
- Raffaelli, D. (2004). How extinction pattern affect ecosystems. *Science*, 306, 1141–1142.
- Raffaelli, D., van der Putten, W.H., Persson, L., Wardle, D.A., Petchey, O., Koricheva, J. *et al.* (2002). Multi-trophic dynamics and ecosystem processes. In: *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives* (eds Loreau, M., Naeem, S. & Inchausti, P.). Oxford University Press, Oxford, pp. 147–154.
- Rosenberg, M.J., Adams, D.C. & Gurevitch, J. (2000). *Metawin 2.0 User's Manual: Statistical Software for Meta-Analysis*. Sinauer Associates, Sunderland, MA.
- Rosher, C., Temperton, V.M., Scherer-Lorenzen, M., Schmitz, M., Schumacher, J., Schmid, B. *et al.* (2005). Overyielding in experimental grassland communities- irrespective of species pool or spatial scale. *Ecol. Lett.*, 8, 419–429.
- Schläpfer, F. & Schmid, B. (1999). Ecosystem effects of biodiversity – a classification of hypotheses and exploration of empirical results. *Ecol. Appl.*, 9, 893–912.
- Schläpfer, F., Pfisterer, A.B. & Schmid, B. (2005). Non-random species extinction and plant production: implications for ecosystem functioning. *J. Appl. Ecol.*, 42, 13–24.
- Schmid, B., Joshi, J. & Schläpfer, F. (2002). Empirical evidence for biodiversity–ecosystem functioning relationships. In: *Functional Consequences of Biodiversity: Experimental Progress and Theoretical Extensions* (eds Kinzig, A., Tilman, D. & Pacala, P.). Princeton University Press, Princeton, NJ, pp. 120–150.
- Sokal, R.R. & Rohlf, F.J. (1995). *Biometry. The Principles and Practice of Statistics in Biological Research*. Freeman, New York.
- Solan, M., Cardinale, B.J., Downing, A.L., Engelhardt, K.A.M., Ruesink, J.L. & Srivastava, D.S. (2004). Extinction and ecosystem function in the marine benthos. *Science*, 306, 1177–1180.
- Srivastava, D.S. & Vellend, M. (2005). Biodiversity–ecosystem function research: is it relevant to conservation? *Annu. Rev. Ecol. Evol. Syst.*, 36, 267–294.
- Tilman, D. (1996). Biodiversity: population versus ecosystem stability. *Ecology*, 77, 350–363.
- Wardle, D.A., Booner, K.I. & Nicholson, K.S. (1997). Biodiversity and plant litter: experimental evidence which does not support the view that enhanced species richness improves ecosystem function. *Oikos*, 79, 247–258.
- Wright, J.P., Naeem, S., Hector, A., Lehman, C., Reich, P.B., Schmid, B. *et al.* (2006). Conventional functional classification schemes underestimate the relationship with ecosystem functioning. *Ecol. Lett.*, 9, 111–120.
- Yachi, S. & Loreau, M. (1999). Biodiversity and ecosystem functioning in a fluctuating environment: the insurance hypothesis. *Proc. Natl Acad. Sci. USA*, 96, 1463–1468.
- Zavaleta, E.S. & Hulvey, K.B. (2004). Realistic species losses disproportionately reduce grassland resistance to biological invaders. *Proc. Natl Acad. Sci. USA*, 306, 1175–1177.

SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

Appendix S1 Literature used for the meta-analysis.

Table S1 Database used in the meta-analysis.

This material is available as part of the online article from: <http://www.blackwell-synergy.com/doi/full/10.1111/j.1461-0248.2006.00963.x>

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Editor, Shahid Naeem

Manuscript received 22 March 2006

First decision made 15 May 2006

Manuscript accepted 3 July 2006