

Appendix 3

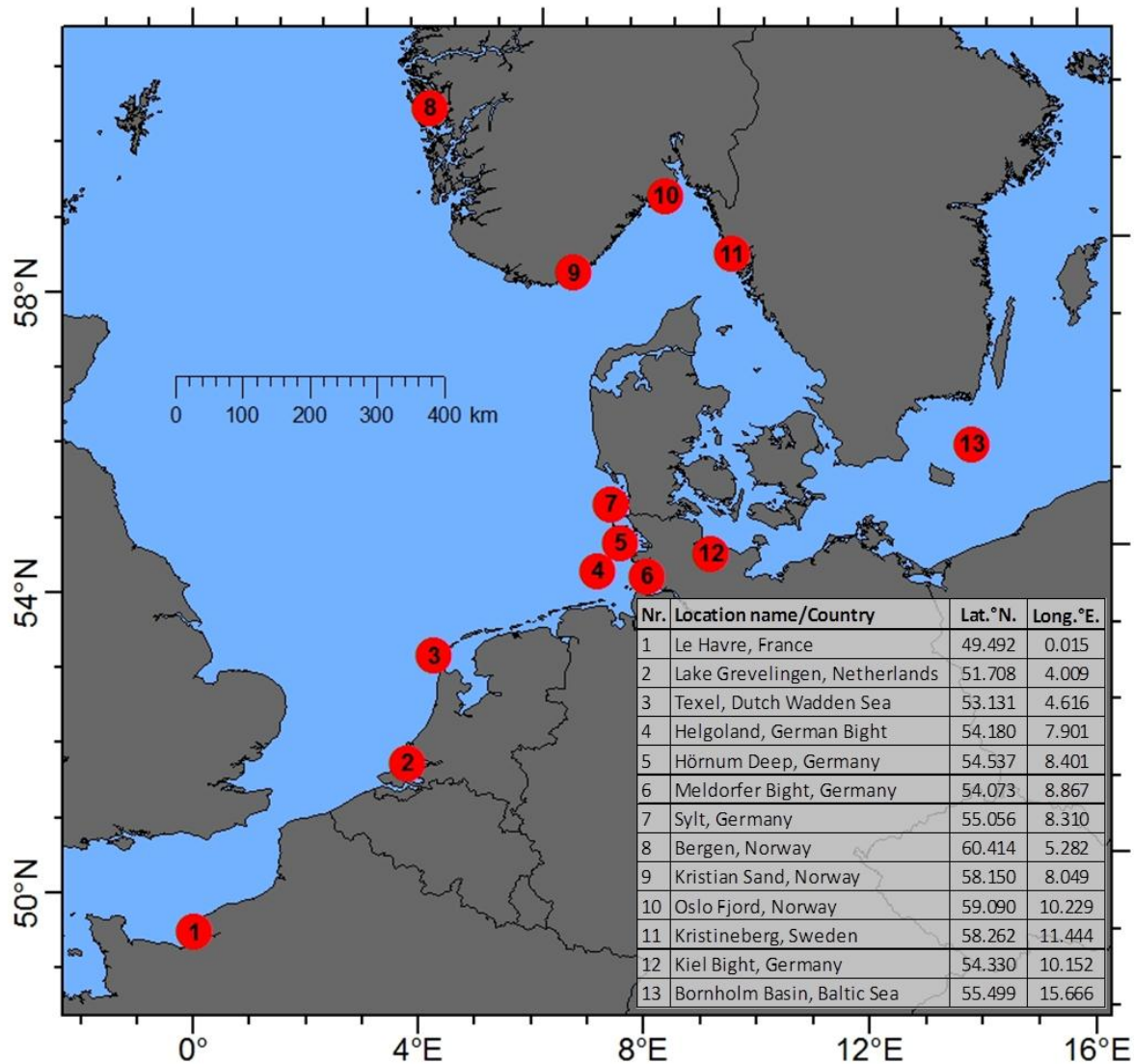


Figure S1 Location information of 13 monitoring stations across northern Europe (station number, locality name and decimal position), where presence/absence of the invasive comb jelly *Mnemiopsis leidyi* have been monitored from 2005 to 2015. Only stations with intensive continuous sampling effort have been considered to avoid bias due to low sampling frequency/low water volumes processed per station. Only presence or absence data for consecutive sampling events during its peak season (summer/autumn) have been considered (see Table S1 for reference). Station 8 and 13 represent a local monitoring region including several stations.

Fig. S2 Temperature anomalies and current characteristics in northern Europe for 2008. Temperature anomalies (red = above, blue = below 2007 – 2014 average), current velocities and directions (black arrows, in m s^{-1}) averaged for meteorological winter periods (January – March). Data source: CMEMS model.

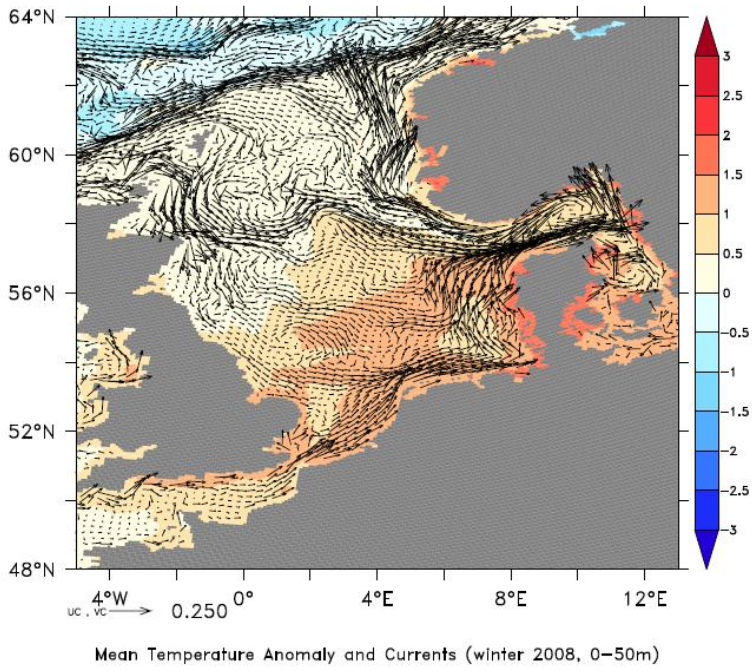


Table S1 Database with >12,400 *M. leidy* records – see Appendix 2, deposited on Pangaea, doi.org/10.1594/PANGAEA.884403.

Table S2 Sequence data of ctenophores caught in the North Sea during 2014 and 2015. Blast search yield 98 to 100% identity to the two published ITS sequences of *Bolinopsis infundibulum* from Europe (average $99.1 \pm 0.95\%$). Closest match with *Mnemiopsis leidyi* was considerably lower and showed an absolute difference in its identity score of 5-8% compared to *B. infundibulum*.

Location	Sequences	Accession number	species ID
Off NE UK coast	ITS, n=12	KY204070-81	<i>Bolinopsis infundibulum</i>
		KY204083	<i>Bolinopsis infundibulum</i>
Off DK coast	ITS, n=1	KY204082	<i>Mnemiopsis leidyi</i>

Table S3. Probability of *Mnemiopsis leidyi* occurrence as a function of average winter temperature (Jan.-Mar.) in the upper 50m of the water column for 13 stations throughout northern Europe from 2008 to 2015, based on two different generalised linear models: Model 1 considering average temperature in the previous winter and model 2 the cumulative effect as temperature average of the two preceding winters. The odds ratio estimates show that winter temperature average in model 2 has a stronger impact on the likelihood of occurrence (n = 104) and for this model critical temperatures for each station (Table S4) range between 1.3 to 2.7°C ($2.03 \pm .19^\circ\text{C}$; av. \pm SE). Station was included as a fixed factor (Model 1: $F_{12,90} = .18$, $p = .999$; Model 2: $F_{12,90} = .34$, $p = .979$).

Model	Estimate	SE	Odds ratio	t-value	DF	p
Model 1	1.59	.47	3.9	3.38	90	.0011
Model 2	8.97	3.88	7882	2.31	90	.0231

Table S4. Parameter estimates for station, intercept and winter temperature from generalized linear models using average winter temperature of the preceding winter (model 1) and cumulative effects as average winter temperature of the two preceding winters (model 2). Note that for both models, Texel (Dutch Wadden Sea, Fig. S1) was used as reference. Estimates from model 1 and 2 were used to calculate the respective critical winter temperatures for each station. Stations where both presence and absence had been observed are marked with *. For stations where absence had not been registered in the observation period, estimates of critical temperature is not a realistic scenario. We therefore only used estimates of critical temperature for the stations where both presence and absence had been registered (bold station names highlighted by*). Station name and number (see Fig. S1 for reference) are indicated.

Model	Station name	Nr.	Estimate	SE	df	t	P	Critical temperature
1	Intercept		8.49	1267	90	.01	.995	
1	Winter temperature		1.59	.47	90	3.38	.001	
1	Bergen*	8	-10.62	1267	90	-.01	.993	1.3
1	Kiel Bight*	12	-11.27	1267	90	-.01	.993	1.7
1	Kristineberg*	11	-11.26	1267	90	-.01	.993	1.7
1	Kristian Sand*	9	-11.99	1267	90	-.01	.993	2.2
1	Oslo Fjord*	10	-11.92	1267	90	-.01	.993	2.2
1	Bornholm Basin*	9	-12.62	1267	90	-.01	.992	2.6
1	Helgoland	4	1.59	1770	90	.00	.999	-6.3
1	Lake Grevelingen	2	-.65	1798	90	-.00	1.000	-4.9
1	Le Havre	1	-4.13	1836	90	-.00	.998	-2.7
1	Meldorfer Bight	6	4.56	1715	90	.00	.998	-8.2
1	Hörnum Deep	5	3.66	1734	90	.00	.998	-7.6
1	Sylt	7	3.64	1728	90	.00	.998	-7.6
1	Texel	3	0	-5.3
2	Intercept		-25.70	1087	90	-.02	.981	
2	Winter temperature		8.97	3.88	90	2.31	.023	
2	Bergen*	8	13.36	1086	90	.01	.990	1.4
2	Kiel Bight*	12	9.52	1086	90	.01	.993	1.8
2	Kristineberg*	11	9.97	1086	90	.01	.993	1.8
2	Kristian Sand*	9	5.16	1086	90	0	.996	2.3
2	Oslo Fjord*	10	6.28	1086	90	.01	.995	2.2
2	Bornholm Basin*	9	1.72	1086	90	0	.999	2.7
2	Helgoland	4	10.34	1469	90	.01	.994	1.7
2	Lake Grevelingen	2	-2.90	1511	90	0	.999	3.2
2	Le Havre	1	-24.56	1577	90	-.02	.988	5.6
2	Meldorfer Bight	6	25.95	1440	90	.02	.986	-.03
2	Hörnum Deep	5	21.48	1448	90	.01	.988	.5
2	Sylt	7	21.48	1445	90	.01	.988	.5
2	Texel	3	0	2.9

Supporting methods:

Statistical methods. Estimation of critical winter temperature. The effect of a 1°C temperature increase on the probability of occurrence has been analysed using odds ratios, while the critical winter temperature i.e. the temperature at which there is a higher than 50% likelihood of encountering *M. leidyi* has been analysed using critical winter temperature approach. For the latter, the parameter estimates for the logistic regression have been used to estimate the temperature at which an odds of 1, which is equivalent to a fifty-fifty probability of presence /absence , occurs.

$$(S1) \quad odds = \frac{p}{1-p} = \frac{\text{probability of presence of characteristic}}{\text{probability of absence of characteristic}}$$

Hence, finding the critical temperature, at which temperature is there a 50 % probability of presence i.e. at which temperature does it become more likely than fifty – fifty that presence will be encountered, we solved the following equations:

$$(S2) \quad \text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \text{intercept} + \beta_{\text{temp}} * \text{temp} + \beta_{\text{station}} * \text{station}$$

Where

$$(S3) \quad \log\left(\frac{p}{1-p}\right) = \log(1)$$

β_{station} is the parameter estimate found by the model for each station (see Table S4).

$$(S4) \quad \log(1) = \text{intercept} + \beta_{\text{temp}} * \text{temp} + \beta_{\text{station}} * \text{station}$$

$$(S5) \quad \text{temp} = \left(\frac{\log(1) - \text{intercept} - \beta_{\text{station}} * \text{station}}{\beta_{\text{temp}}} \right)$$

Winter temperature Baltic Sea.

To obtain average winter temperature data for the upper 50 m of the Baltic Sea area, we used the Kiel Baltic Sea Ice-Ocean Model (BSIOM) (Lehmann & Hinrichsen, 2000; Lehmann *et al.*, 2002) which has a higher precision than more general models which apply for larger areas. The horizontal resolution of the BSIOM coupled sea ice–ocean model is at present 2.5 km, and 60 levels in the vertical plane, leading to a vertical resolution of 3 m. The model is forced by low frequency sea level variations in the North Sea/Skagerrak calculated from the Baltic Sea Index (BSI) (Lehmann *et al.*, 2002) and meteorological forcing taken from the Swedish Meteorological and Hydrological Institute (SMHI).

Molecular analyses expanded. Seven species-specific, highly polymorphic microsatellite loci (Reusch *et al.*, 2010) were used to genotype animals (n = 201) before and after the cold sweep in northern Europe. Dried tissue was extracted from filters using Qiagen blood and tissue kit with primer concentrations and PCR conditions as outlined in published protocols (Reusch *et al.*, 2010; Bolte *et al.*, 2013). Amplified fragments were analyzed on an ABI 3130 genetic analyzer using Rox-350 (Applied Biosystems) as internal size standard. Allele sizes were scored using the software Genemarker v1.91 (SoftGenetics, LLC). Arlequin v3.5.1.2 (Excoffier & Lischer, 2010) was used for population comparisons computing pairwise F_{ST} values using non-parametric permutation procedures. Type I errors were taken into account (Narum, 2006) by adjusting p values following the Benjamini-Hochberg False Discovery Rate (B-H FDR) procedure (Benjamini

& Yekutieli, 2001). The software Structure v2.3.4 (Pritchard *et al.*, 2000) was used to infer genetic clustering without *a priori* assumption about expected number of clusters. Structure implements a Bayesian inference algorithm for detecting the number of clusters (*k*) that best explain genetic variation within a multilocus dataset. We used 100,000 re-iterations for the burn-in and 1,000,000 Markov chain Monte Carlo (MCMC) repetitions. Probabilities were calculated for *k* ranging from 1 to 5 with five replicates for each *k* and the most likely number of *k*'s inferred through the Evanno method (Evanno *et al.*, 2005). Structure v2.3.4. analyses with 100,000 burnin periods and 1.000.000 MCMC repetitions after each burnin with *k* estimation from 1 to 5 with 5 iterations each.

Supporting references:

- Benjamini, Y. & Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, **29**, 1165-1188.
- Bolte, S., Fuentes, V., Haslob, H., Huwer, B., Thibault-Botha, D., Angel, D., Galil, B., Javidpour, J., Moss, A.G. & Reusch, T.B.H. (2013) Population genetics of the invasive ctenophore *Mnemiopsis leidyi* in Europe reveal source-sink dynamics and secondary dispersal to the Mediterranean Sea. *Marine Ecology Progress Series*, **485**, 25-36.
- Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.
- Excoffier, L. & Lischer, H.E.L. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567.
- Haraldsson, M., Jaspers, C., Tiselius, P., Aksnes, D.L., Andersen, T. & Titelman, J. (2013) Environmental constraints of the invasive *Mnemiopsis leidyi* in Scandinavian waters. *Limnology and Oceanography*, **58**, 37-48.
- Lehmann, A. & Hinrichsen, H.H. (2000) On the wind driven and thermohaline circulation of the Baltic Sea. *Physics and Chemistry of the Earth Part B-Hydrology Oceans and Atmosphere*, **25**, 183-189.
- Lehmann, A., Krauss, W. & Hinrichsen, H.H. (2002) Effects of remote and local atmospheric forcing on circulation and upwelling in the Baltic Sea. *Tellus Series a-Dynamic Meteorology and Oceanography*, **54**, 299-316.
- Lellouche, J.M., Le Galloudec, O., Drévilion, M., Regnier, C., Greiner, E., Garric, G., Ferry, N., Desportes, C., Testut, C.E., Bricaud, C., Bourdallé-Badie, R., Tranchant, B., Benkiran, M., Drillet, Y., Daudin, A. & De Nicola, C. (2013) Evaluation of global monitoring and forecasting systems at Mercator Ocean. *Ocean Science*, **9**, 57-81.
- Narum, S.R. (2006) Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics*, **7**, 783-787.
- Neumann, G. & Pierson, W. (1967) *Principles of Physical Oceanography*. Prentice-Hall Int. Inc., London.
- Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945-959.
- Reusch, T.B.H., Bolte, S., Sparwel, M., Moss, A.G. & Javidpour, J. (2010) Microsatellites reveal origin and genetic diversity of Eurasian invasions by one of the world's most notorious marine invader, *Mnemiopsis leidyi* (Ctenophora). *Molecular ecology*, **19**, 2690-9.