GSA Data Repository 2018079
Kiessling et al., 2018, Pre–mass extinction decline of latest Permian ammonoids: Geology, https://doi.org/10.1130/G39866.1.

Item DR1: Supplementary methods

Maximum diameters of the conchs were measured in the lab. For incomplete specimens, the diameter was estimated from the height of the phragmocone in the case that the end of somatic growth is indicated by conspicuous septal crowding. In paratirolitids, there is an empirical relationship between the apertural height of the phragmocone ($ah_{phragmocone}$) and the diameter of the conch including the body chamber (dm_{conch}):

$dm_{conch} = ah_{phragmocone}*3.13$

We computed the volume body chamber with a simple frustum volume equation approximating unknown parameters from aperture shapes: $V = 1/3h\ (a^2 + ab + b^2)$, where a and b describe the base and top of the frustum, i.e. the whorl profile that may range from quadrate or circular to rectangular, oval or trapezoidal, and h describes the height of the frustum, i.e. the length of the body chamber, estimated by the circumference of the body chamber at half apertural height for a length of 270 degrees (the average body chamber length of paratirolitid ammonoids) (Fig. DR1).

Vertical positions of ammonoids in sections were measured as precisely as possible, but a measurement error of ±5 cm is admitted and hence all measurements were rounded to the next 5 cm prior to analysis.

Item DR2: Additional analyses

Prerequisites for classical confidence intervals

One prerequisite for calculating standard confidence intervals is that fossil horizons are randomly distributed in time (Strauss and Sadler, 1989; Marshall, 1990). We first tested this assumption with a Kolmogorov-Smirnov (KS) test on the distribution of fossil horizons in the two sections. Histograms of fossil horizons suggest that fossil recovery is sparse from six to four meters below the boundary clay and the null hypothesis of a uniform distribution is rejected for the entire thickness of the Paratirolites Limestone ($p = 1 \cdot 10^{-5}$ and $p = 6 \cdot 10^{-4}$ for Julfa and Baghuk, respectively) (Fig. DR2). The hypothesis of a uniform distribution is not rejected for horizons in the top four meters in both sections (p = 0.99 and 0.37, respectively). Because a uniform distribution is the null hypothesis in the KS test, a failure to reject the hypothesis does not necessarily mean that distributions are significantly uniform. We therefore also applied the method suggested by Solow et al. (2006) to check if a uniform distribution is true over the top four meters of the *Paratirolites* Limestone. Using the runif() function in R we simulated uniformly distributed horizons with the same number and range as the observed data in the top four meters of section. Repeating the procedure 1000 times allows us to estimate the 0.025 quantiles of the simulated ordered values and to check if systematic excursions from the simulation bands are evident in the observed data. As this is not the case (Fig. DR3), a uniform distribution of fossil horizons can be assumed.

Within the top 4 m of the *Paratirolites* Limestone, the null hypothesis of a uniform distribution was also not rejected for the vertical distribution of individual species, which further supports the use of first generation (Wang and Marshall, 2016) methods. There is also no significant tendency of gap sizes to either increase or decrease with stratigraphic position in both sections (Spearman's rho = -0.05, $p = 0.73$; rho = -0.14, $p = 0.32$, respectively). In summary, the assumptions underlying the classical methods to calculate stratigraphic confidence intervals are

met (random distribution of fossil horizons, uniform collection intensity, no increase in gap sizes towards the endpoint of interest).

Testing extinction scenarios

Springer's (1990) method uses the distribution of confidence levels of stratigraphic ranges to test whether species range data are compatible with a co-extinction event. An extinction or emigration pulse is likely to result in random range truncations. Therefore, in order to be compatible with an abrupt extinction event, the confidence levels should be independently uniformly distributed. This is assessed with a KS test of the confidence intervals on stratigraphic ranges. Several horizons may fail to reject the null hypothesis of a uniform distribution of confidence levels such that the best position is sought with the minimum D value. An extinction pulse above the last appearance of Iranian ammonoids is very unlikely, as the already high D values at the base of the boundary clay increase up sections. We applied Springer's method down the sections, consecutively excluding species, which make have their last appearance above a candidate extinction pulse. The D values in both composite sections show distinct minima. The local minima for which a uniform distribution of confidence levels cannot be rejected are presented in Figs. 2 and DR4 as candidate extinction pulses.

The possibility that the top 4 m of the *Paratirolites* Limestone simply record background turnover was addressed with another loop. This time, the hypothesized extinction level was held constant at the base of the boundary clay and occurrences were only considered down to a variable level ranging from -0.5 to -4m. Solow's (1996) likelihood-ratio statistic (Fig. DR5) was applied every 10cm moving down sections.

Wang et al. (2012) proposed a simulation to put maximum constraints on the duration of a mass extinction. This simulation uses the observed number of species and their stratigraphical distribution to check under which hypothesis for the true duration of a mass extinction would the observed duration of a mass extinction (e.g., the stratigraphic difference between the highest and lowest upper endpoint of species) comply with observed pattern of fossil horizons. The method is very sensitive to the number of fossil horizons, allowing for a broad range of possible durations when sampling is limited. The apparent extinction duration spans 2.95 m at Julfa and 3.3 m at Baghuk, whereas the simulations allow for a true duration between 0 and more than 3.2 m, when all non-singleton species are included. However, the hypothesis of a sudden extinction (duration of 0) is readily rejected when the simulation is constrained to species occurring in at least four horizons (Tables DR3-4).

When applied to maximum intervals over which a simultaneous extinction at the boundary clay cannot be refuted (Fig. DR5), the simulation of Wang et al. (2012) allows for a true duration over the entire interval when the minimum number of horizons is not constrained (but singletons omitted). A true duration of zero is excluded for Baghuk when only species occurring in at least 5 horizons are included in the simulation over the top 1.3 m (4 species). The three species in the top 0.5m at Julfa occur each in only two horizons such that the true duration cannot be constrained further.

Trends in ammonoid diameters

The conch diameters yield a similar pattern the logarithm of volume of ammonoid body chambers (Fig. DR6). The nearly steady decline starts after a plateau (Julfa) of even peak (Baghuk) at ca. 2.5 m, that is, earlier than the decline of body chamber volume (Fig. 3).

References cited in the appendix

Marshall, C. R., 1990, Confidence-intervals on stratigraphic ranges: Paleobiology, v. 16, p. 1-10.

Solow, A. R., 1996, Tests and confidence intervals for a common upper endpoint in fossil taxa: Paleobiology, v. 22, p. 406-410.

Springer, M. S., 1990, The effect of random range truncations on patterns of evolution in the fossil record: Paleobiology, v. 16, p. 512-520.

Strauss, D., and Sadler, P. M., 1989, Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges: Mathematical Geology, v. 21, p. 411-427.

Wang, S. C., and Marshall, C. R., 2016, Estimating times of extinction in the fossil record: Biology Letters, v. 12, p. 20150989.

Wang, S. C., Zimmerman, A. E., McVeigh, B. S., Everson, P. J., and Wong, H., 2012, Confidence intervals for the duration of a mass extinction: Paleobiology, v. 38, p. 265-277.

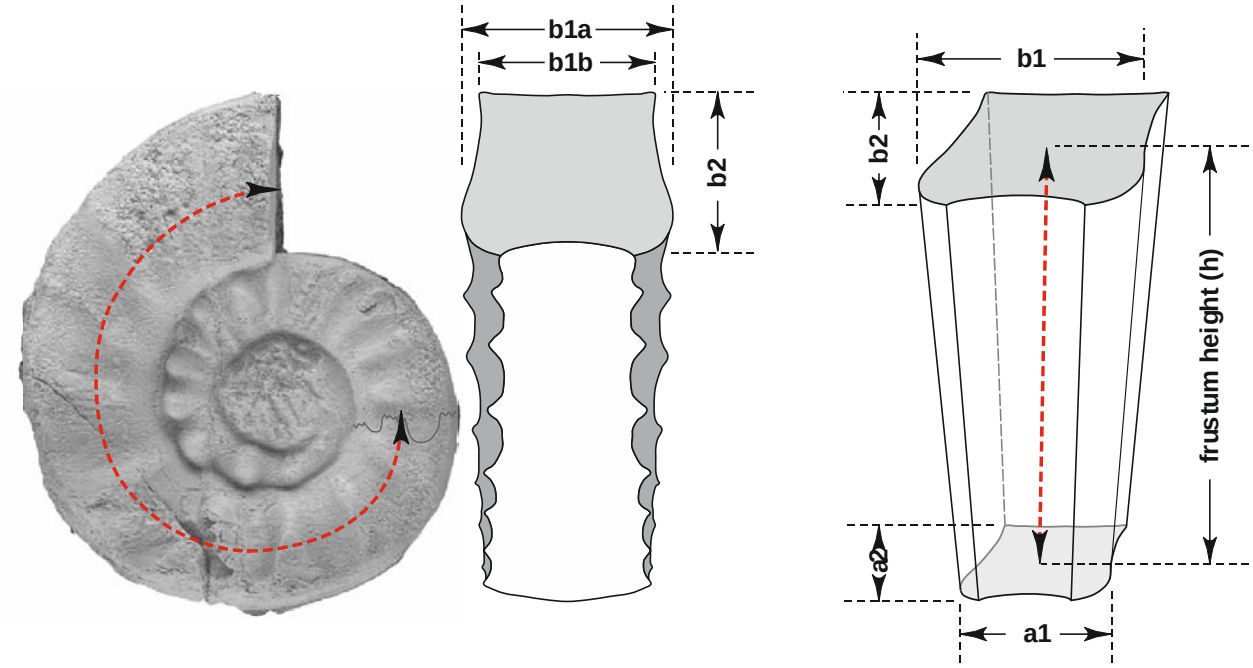


Figure DR1. Measurements for calculating the volume of the dzhulifitid phragmocone.

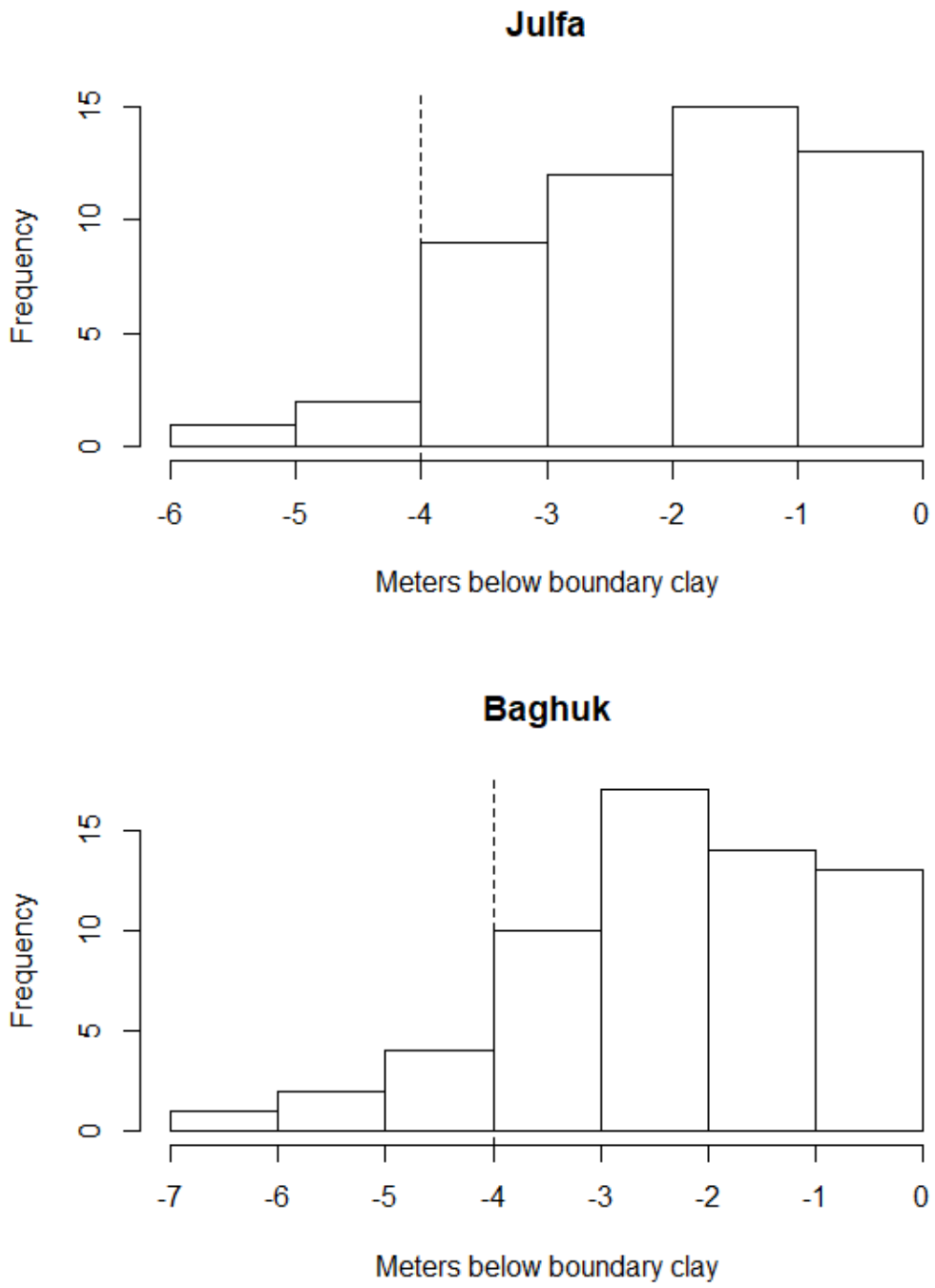
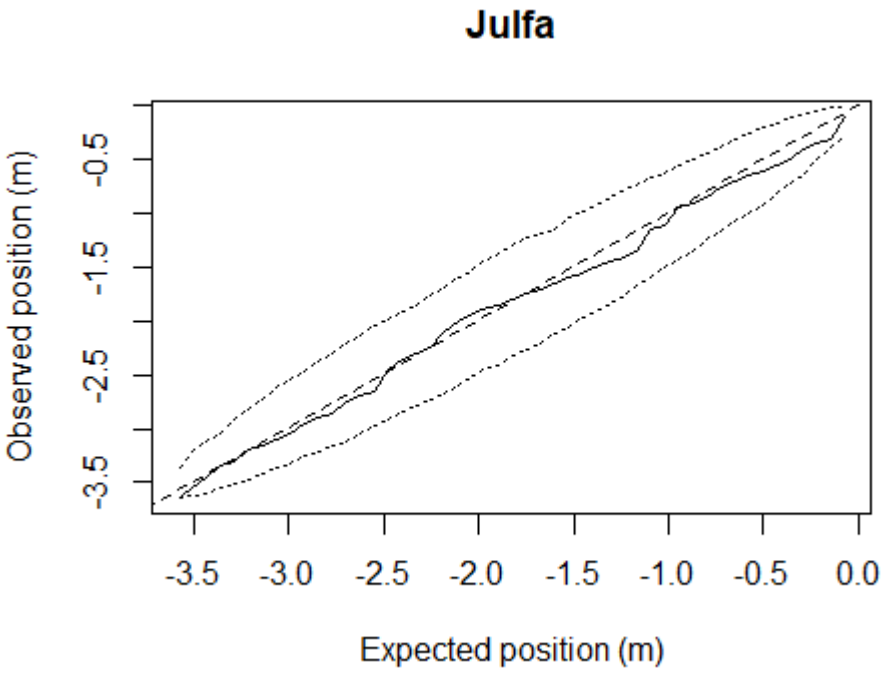


Figure DR2. Histograms of the stratigraphic distribution of fossil horizons at Julfa (top) and Baghuk (bottom). Dashed vertical lines mark the levels above which the assumption of a random (uniform) distribution of fossil horizons cannot be rejected.



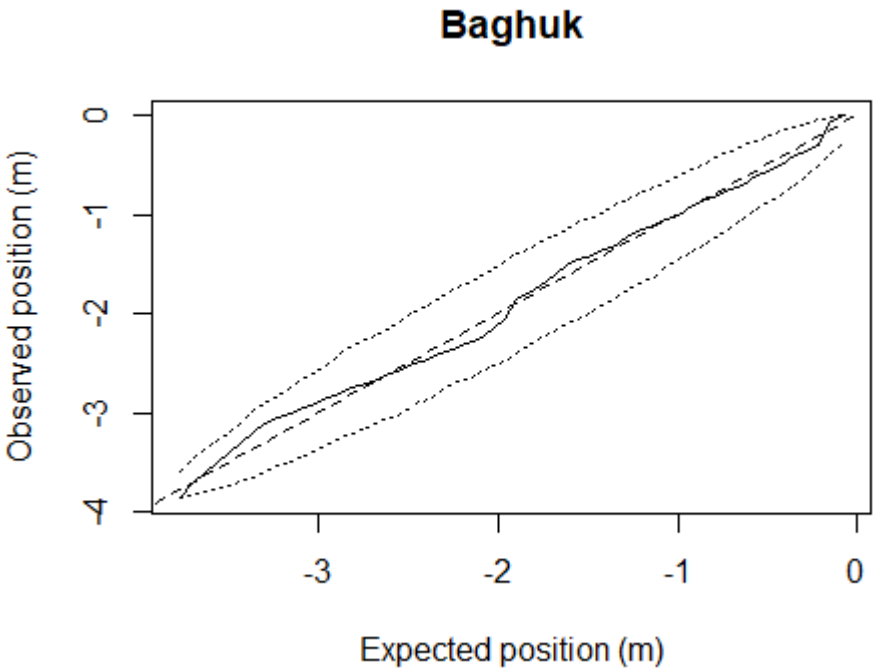


Figure DR3. Observed ordered positions of ammonoid-bearing horizons plotted against expected positions (solid line) predicted from simulated uniform distributions (n = 1000). Also shown are the 2.5 and 97.5 percentiles of expected ordered positions (dotted) and the 1:1 line (dashed).

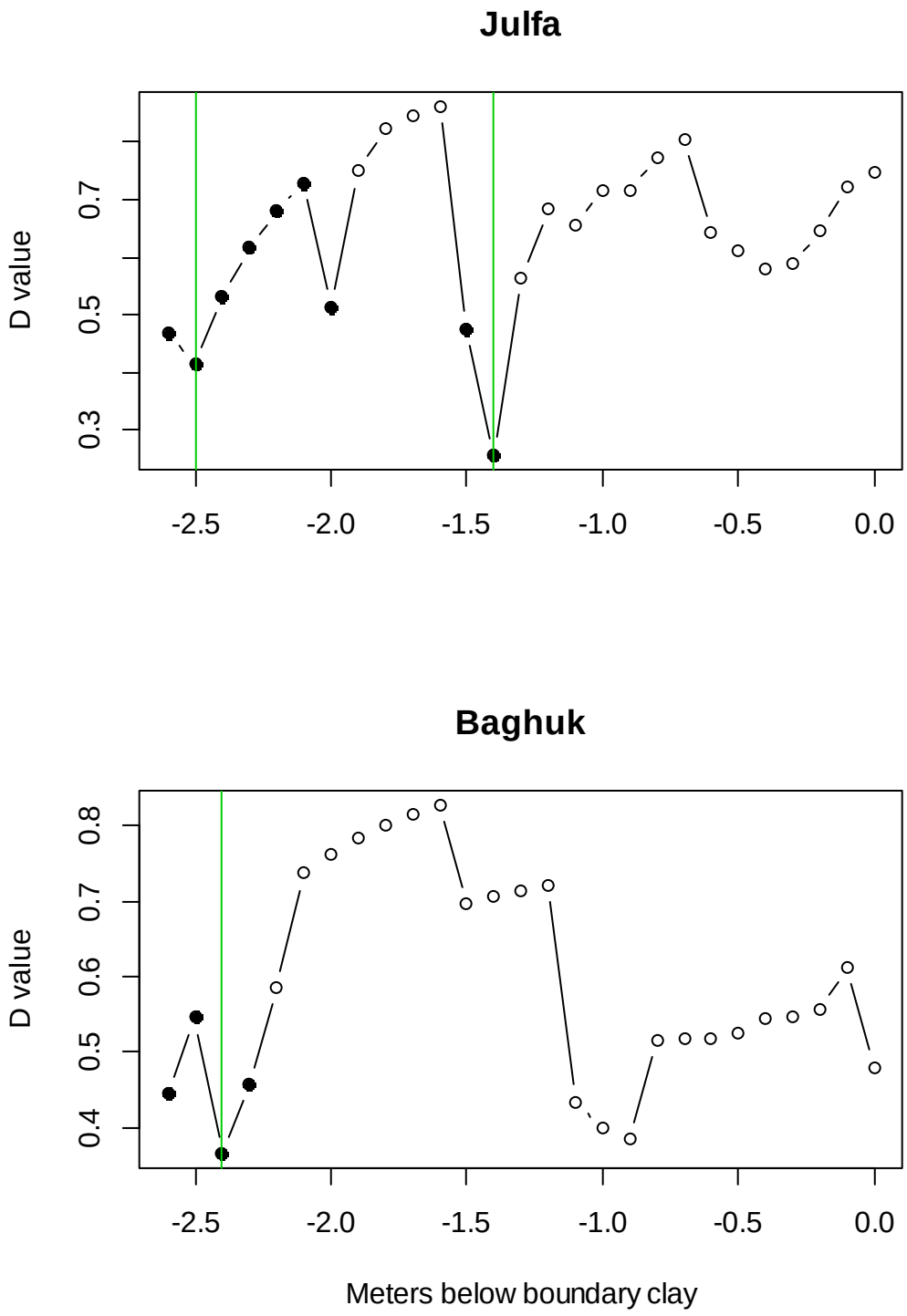


Figure DR4. D-values of the KS test applied down the sections at Julfa (top) and Baghuk (bottom). Solid points mark intervals below which the hypothesis of a uniform distribution of confidence levels cannot be rejected. Green vertical lines correspond to the green lines in Fig. 2.

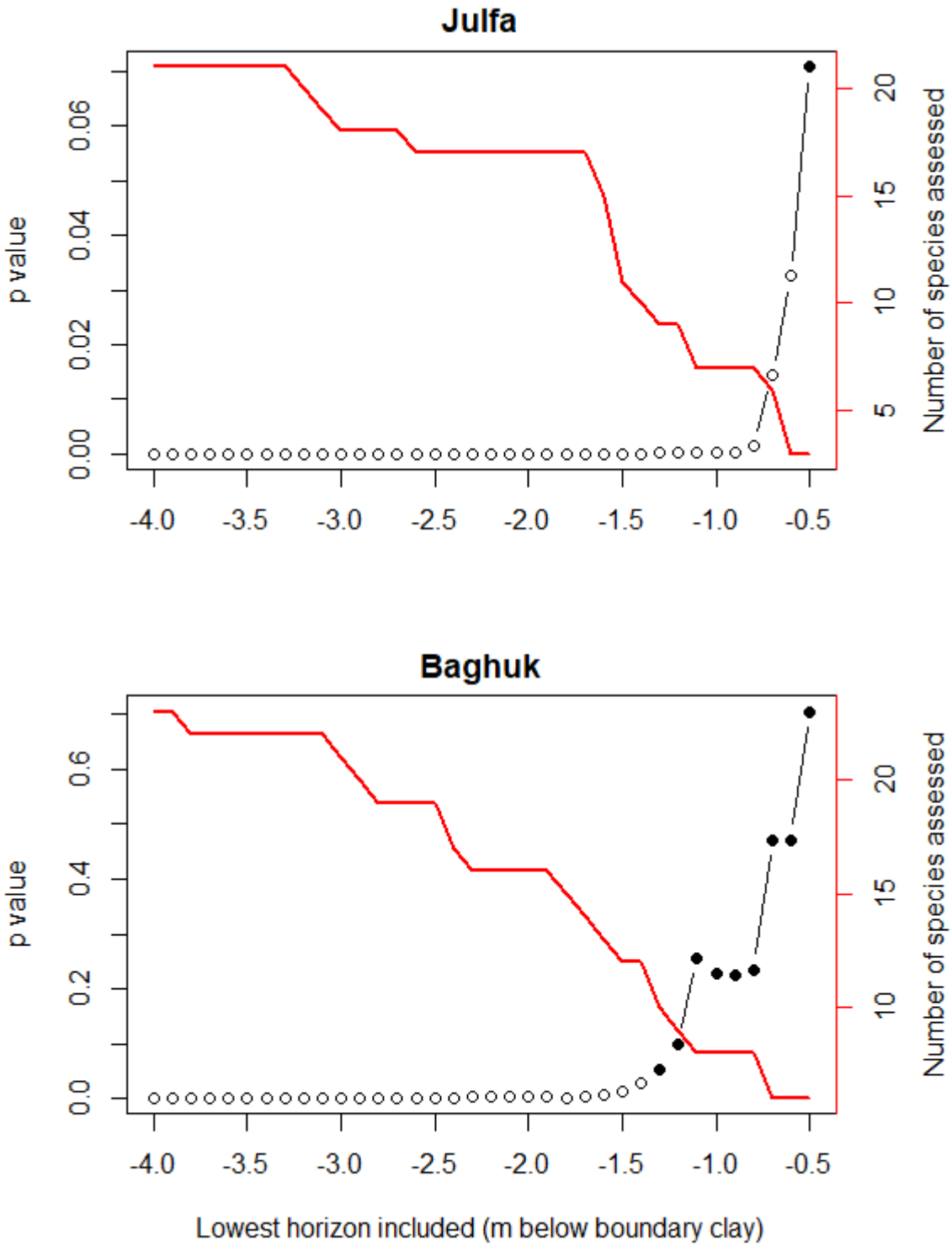


Figure DR5. P-values for testing the null hypothesis of a common upper endpoint of species at the base of the boundary clay. Tests were iteratively performed on occurrences down to a horizon applying the likelihood ratio statistic of Solow (1996). Solid points mark horizons above which the hypothesis of common upper endpoint at the base of the boundary clay cannot be rejected.

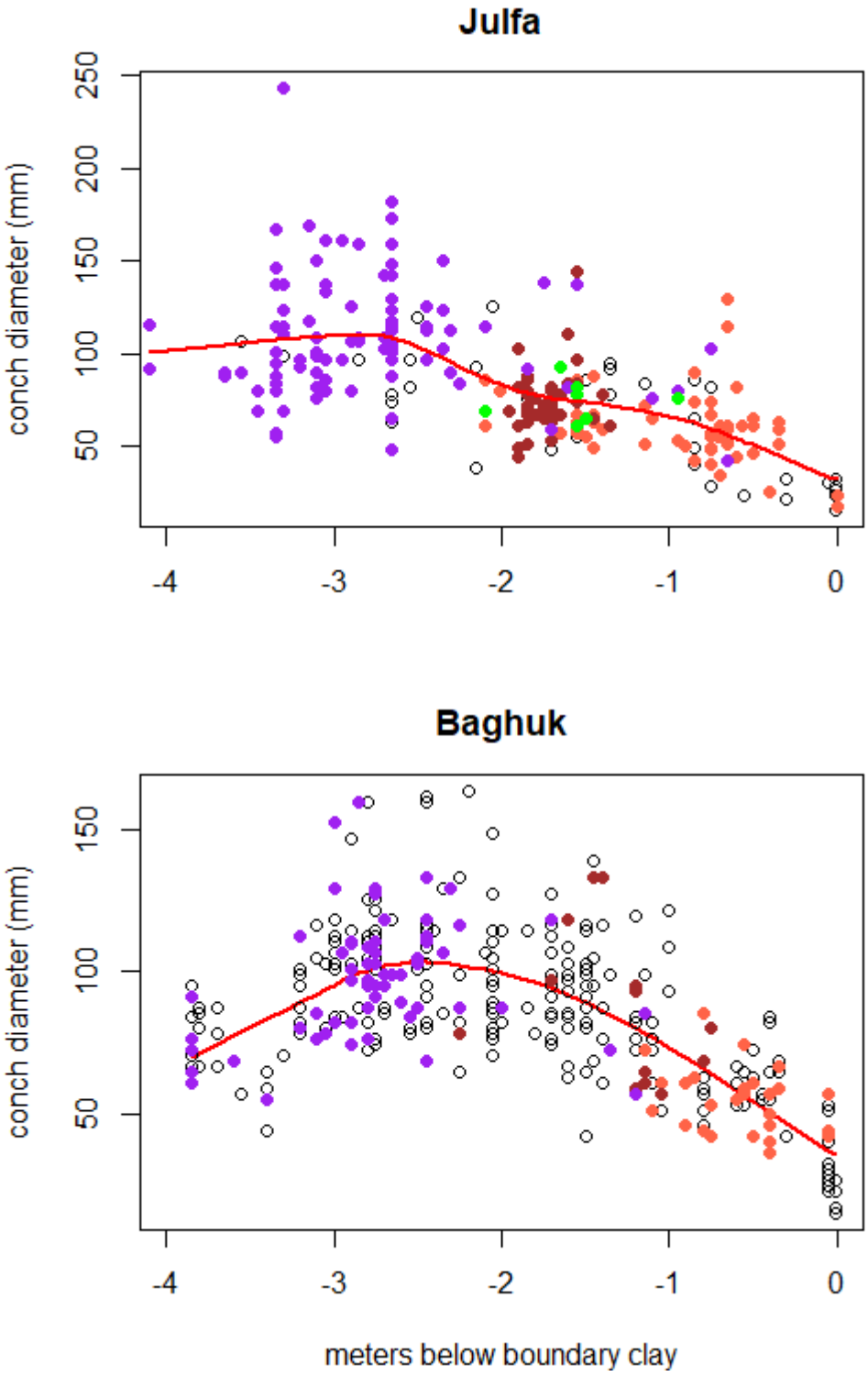


Figure DR6. Stratigraphic trends in conch diameters. Ammonoid biozones are shaded as in Fig. 2. The dominant and identified genera are colored: Purple = *Paratirolites*; brown = *Alibashites*, orange = *Abichites*, green = *Stoyanowites*.

Table DR1. Stratigraphic ranges and confidence intervals of ammonoids at Julfa. Order of species corresponds to Fig. 2A. Species with a last appearance datum (LAD) below -4m are not considered in hypothesis testing (grey marker).

Species	FAD (meters below boundary clay)	LAD (meters below boundary clay)	Number of horizons	Stratigraphic range in meters	Confidence of true extinction below the boundary clay	Top of 50% confidence interval (m)	Top of 95% confidence interval (m)
Dzhulfites zalensis	-5.3	-4.65	2	0.65	0.877	-4.00	7.70
Paratirolites trapezoidalis	-4.1	-3.05	6	1.05	0.999	-2.89	-2.19
Paratirolites coronatus	-3.55	-2.9	3	0.65	0.966	-2.63	-0.64
Paratirolites birunii	-3.35	-2.7	3	0.65	0.962	-2.43	-0.44
Paratirolites vediensis	-3.3	-2.1	10	1.2	1.000	-2.00	-1.63
Paratirolites kittli	-3.35	-1.55	17	1.8	1.000	-1.47	-1.18
Stoyanowites dieneri	-2.3	-1.55	4	0.75	0.965	-1.36	-0.26
Abichites subtrapezoidalis	-2.1	-1.45	5	0.65	0.991	-1.33	-0.73
Alibashites ferdowsii	-1.85	-1.45	4	0.4	0.990	-1.35	-0.76
Alibashites stepanovi	-1.9	-1.45	7	0.45	1.000	-1.39	-1.16
Alibashites uncinatus	-1.95	-1.45	6	0.5	0.999	-1.38	-1.04
Alibashites mojsisovicsi	-1.9	-1.15	11	0.75	1.000	-1.10	-0.89
Paratirolites multiconus	-1.15	-1.1	2	0.05	0.957	-1.05	-0.15
Stoyanowites aspinosus	-1.55	-0.95	3	0.6	0.850	-0.70	1.13
Abichites ariaeii	-1.15	-0.7	4	0.45	0.940	-0.58	0.07
Abichites abichi	-1.5	-0.65	5	0.85	0.897	-0.49	0.30
Neoaganides ultimus	-1.15	-0.65	2	0.5	0.565	-0.15	8.85
Paratirolites serus	-0.95	-0.55	5	0.4	0.969	-0.47	-0.10
Abichites alibashiensis	-1.4	-0.5	6	0.9	0.890	-0.37	0.24
Abichites paucinodus	-0.75	-0.35	4	0.4	0.848	-0.25	0.34
Abichites stoyanowi	-0.95	-0.3	10	0.65	0.967	-0.25	-0.04

Table DR2. Stratigraphic ranges and confidence intervals of ammonoids at Baghuk. Order of species corresponds to Fig. 2B. Species with a last appearance datum (LAD) below -4m are not considered in hypothesis testing (grey marker).

Species	FAD (meters below boundary clay)	LAD (meters below boundary clay)	Number of horizons	Stratigraphic range in meters	Confidence of true extinction below the boundary clay	Top of 50% confidence interval (m)	Top of 95% confidence interval (m)
Dzhulfites 35	-7	-4.7	3	2.3	0.892	-3.75	3.29
Dzhulfites 36	-4.8	-4.7	2	0.1	0.979	-4.60	-2.80
Shevyrevites shevyrevi	-5.1	-4	6	1.1	1.000	-3.84	-3.10
Pseudogastrioceras relicuum	-4.9	-3.3	5	1.6	0.989	-3.00	-1.52
Paratirolites 5	-3.2	-2.9	3	0.3	0.991	-2.78	-1.86
Paratirolites 3	-3.85	-2.8	9	1.05	1.000	-2.70	-2.32
Paratirolites trapezoidalis	*-3.05*	*-2.8*	*2*	*0.25*	*0.918*	*-2.55*	*1.95*
Paratirolites vediensis	*-3.5*	*-2.45*	*9*	*1.05*	*1.000*	*-2.35*	*-1.97*
Paratirolites 2	-3.1	-2.35	8	0.75	1.000	-2.27	-1.95
Paratirolites kittli	-3.5	-2.3	12	1.2	1.000	-2.22	-1.92
Stoyanowites aspinosus	-1.8	-1.6	2	0.2	0.889	-1.40	2.20
Alibashites 10	-1.6	-1.2	3	0.4	0.938	-1.03	0.19
Alibashites 9	-1.7	-1.2	2	0.5	0.706	-0.70	8.30
Paratirolites 21	-2.45	-1.2	3	1.25	0.740	-0.68	3.14
Paratirolites 1	-2.65	-1.15	10	1.5	0.994	-1.03	-0.56
Paratirolites 6	-2.05	-1.1	6	0.95	0.979	-0.96	-0.32
Alibashites 8	-1.6	-1	4	0.6	0.947	-0.84	0.03
Abichites mojsisovicsi	-1.35	-0.75	2	0.6	0.556	-0.15	10.65
Alibashites ferdowsii	-2.25	-0.7	9	1.55	0.949	-0.56	0.00

Paratirolites serus	-0.75	-0.6	2	0.15	0.800	-0.45	2.25
Abichites alibashiensis	-1.15	-0.4	7	0.75	0.923	-0.31	0.09
Abichites 15	-0.75	-0.3	4	0.45	0.784	-0.18	0.47
Abichites 18	-0.4	-0.05	2	0.35	0.125	0.30	6.60

Table DR3. Simulation results of expected true extinction durations (Delta) at Julfa

Threshold of horizons	Observed duration (m)	Minimum Delta (m)	Maximum Delta (m)
2	2.95	0	3.25
3	2.75	0.65	3.45
4	2.75	2.05	3.55
5	2.75	2.45	3.55

Table DR4. Simulation results of expected true extinction durations (Delta) at Baghuk

Threshold of horizons	Observed duration (m)	Minimum Delta (m)	Maximum Delta (m)
2	3.3	0	3.65
3	2.85	0	3.65
4	2.75	1.75	3.55
5	2.75	2.35	3.45

Table DR5. Raw data of ammonoid distributions at Julfa and Baghuk alongside morphological data of measured specimens