SUPPLEMENTARY INFORMATION

This material presents technical details to support the discussion in the main paper. We discuss here the Swift data analysis, the analysis of the photometric data, the details of the analysis of the TNG spectroscopic data and the modelling of the GRB redshift distribution.

1. Swift data analysis

Swift-BAT triggered on GRB 090423 at 07:55:19.35 UT on 23 April 2009. BAT data were analysed using the HEASOFT software package (version 6.6.2) with the Swift Calibration Database (CALDB) version BAT(20090130). Background-subtracted light curves in different energy channels, energy spectra and corresponding response functions were derived from the BAT event file as processed with the BAT software tool `batgrbproduct`, by using the mask-weighting technique for the BAT refined position[4], and by using standard and BAT-dedicated software tools. The mask-weighted light curve (Supplementary Figure 1) showed a couple of overlapping peaks starting at $T_0 - 2$ s, peaking at $T_0 + 4$ s, and ending at $T_0 + 15$ s. The estimated duration, T_{90}, was 10.3 ± 1.1 s for the mask-weighted light curve in the 15-150 keV band. T_{90} is defined as the duration 90% of the total prompt γ-ray fluence in the observer frame (i.e. the interval from 5% to 95% of the total fluence) and is estimated using the `battblocks` software tool. Noteworthy is the light curve of the hardest channel, from 100 to 150 keV, showing a very weak signal as compared with those of the other energy channels. This reflects the spectral softness of this GRB, as shown also by the total energy spectrum. The latter was accumulated from -0.7 to 11.7 s and is fitted with a cut-off power law, $N(E) \sim E^{-\Gamma} \exp[(2-\Gamma)E/E_p]$, with the best-fit value for the peak energy, $E_p = 48^{+6}_{-5}$ keV, and a photon index $\Gamma = 0.6^{+0.5}_{-0.6}$. This value of E_p is fully consistent with that determined by fitting the Fermi/Gamma-ray Burst Monitor spectrum[10] with the canonical "Band" function[31]. The corresponding total fluence in the 15-150 keV energy band is $(5.9 \pm 0.4) \times 10^{-7}$ erg cm^{-2}. The 1-s peak photon flux measured from 3.5 s in the 15-150 keV band is 1.7 ± 0.2 ph s^{-1} cm^{-2}. Uncertainties are given at 90% confidence. At $z = 8.1$, GRB 090423 is found to be consistent with the $E_{\rm p,rf} - E_{iso}$ correlation[13] within 0.5σ (Supplementary Figure 2). We note that, even considering the measured peak energy as obtained by a fit of the Fermi/GBM data with a cut-off power-law spectrum, i.e. $E_p = 82 \pm 15$ keV and thus $E_{p,rf} = 746 \pm 137$ keV (ref. [10]), GRB 090423 would still be consistent within 2σ with the $E_{p,rf} - E_{iso}$ correlation.

The BAT light curve shown in Supplementary Figure 3 is the mask-weighted curve extracted between 15 and 150 keV, binned so as to ensure S/N> 2 with a minimum binning time of 0.512 s. Extrapolation of the BAT flux down to the 0.3-10 keV band was performed by assuming the above spectral model. We note that the remarkable X-ray flare detected by XRT is seen in the BAT data as well. We note that assuming that the flare is still part of the prompt emission[32], the total duration of the prompt phase in the source rest frame might be $\sim$ 20s, similar to other long GRBs.

The XRT observations began 73 s after the trigger: up to $\sim$ 300 s the signal was dominated by a flare. As in many other GRBs, the light-curve then flattened to a shallow decay phase which could be well modelled by a power-law with index $\alpha_{X,1} = 0.13 \pm 0.11$. At $t \sim 4500$ s the X-ray afterglow steepened to $\alpha_{X,2} = 1.3 \pm 0.1$ (errors at 68% confidence level). The flare was modelled by a standard profile[34]: this is characterised by a $1/e$ rise-time $t_{rise} = 29.1 \pm 3.6$ s; $1/e$ decay-time $t_{decay} = 65.5 \pm 3.6$ s; $1/e$ width of $\Delta t = 94.6 \pm 7.3$ s, while the asymmetry parameter is $k = 0.38 \pm 0.03$. This implies a variability measure $\Delta t/t_{peak} = 0.66$ and a brightness contrast ΔFlux/Flux around 25. While the flare parameters are defined following ref. [34], the reported uncertainties are worked out by using the entire covariance matrix. At the redshift of the burst, the flare has an energy $E_{iso} = 3.6 \times 10^{51}$ erg in the redshifted 0.3-10 keV band of XRT, comparable to the energy released during the prompt emission of other GRBs.

To evaluate the intrinsic column density absorbing the GRB 090423 spectrum, we extracted data in the 3900–21568 s time interval (observer frame). This interval was selected in order to avoid the bright X-ray flare whose variable spectrum might alter the fit and in order to have sufficient signal in the extraction region which we define as a count rate of more than 0.01 counts s^{-1}. The resulting 7984 s exposure contains 680 counts in the range between 0.3-10 keV. The ancillary response file (arf) was created with the task `xrtmkarf` (within `heasoft` v.6.2.2) using the relevant exposure file and the latest v.11 reponse matrix function (rmf). The spectrum was binned to 20 counts per bin in order to assure a reasonable χ^2 statistic.

We fit the X-ray spectrum with a composite absorption model consisting of a Galactic contribution and an intrinsic absorption fixed to z=8.1 using the `tbabs` model within the XSPEC (v12.5.0aa) package. We left the Galactic value free to vary in the $2.9 - 3.2 \times 10^{20}$ cm^{-2} range (based on the absorption maps by [35] and [36]). The X-ray continuum was modeled with a power law, as is customary for the afterglow spectra of GRBs. The overall fit is good with a reduced $\chi^2_{\rm red} = 1.12$ (28 degrees of freedom, corresponding

to a null hypothesis probability of 30%). The resulting power law photon index is $\Gamma_X = 1.97^{+0.15}_{-0.16}$. For the intrinsic column density, we get a value of $N_H(z) = 6.8^{+5.6}_{-5.3} \times 10^{22}$ cm^{-2} (90% confidence level), among the highest of all Swift GRBs[37]. The results refer to a solar composition and metallicity. Assuming that the medium is not Thomson thick, a lower limit of the metallicity can be obtained by $N_H(z) \lesssim (1/\sigma_T)(Z/Z_\odot)^{-1}$, where σ_T is the Thomson cross-section (e.g. ref. [38]). We find $Z > 0.043\ Z_\odot$. A lower limit on the value of $N_H(z) > 6 \times 10^{21}$ cm^{-2} is found at 95% confidence level corresponding to a lower limit on the metallicity of $Z > 0.004\ Z_\odot$.

2. Analysis of the photometric data

We analyzed all the available photometric data[7] by using the /it zphrem code**??**, in order to determine the photometric redshift and spectral properties of the afterglow. Our code fits a model of functional form $f_\nu \propto \nu^{-\beta} t^{-\alpha_O}$, including dust extinction (by an SMC-type extinction law) as a free parameter.

We find that the data are best fit within the time range $4.2 \times 103 < t < 6.6 \times 104$ s) with a model characterised by a temporal decay with a power law index $\alpha_O = 0.50 \pm 0.05$ (we quote hereafter 95% confidence intervals). The dust content is constrained to be $E(B-V) < 0.15$, and the photometric redshift is $z_{\rm phot} = 8.3 \pm 0.3$, consistent with the spectroscopic results. We caution that the rest-frame wavelength observed extends only out to 2500Å, and only the three JHK filters do indeed measure any flux redwards of Lyman-α. That is the reason why the spectral index is only loosely constrained and its error is asymmetric ($\beta = 0.4^{+0.2}_{-1.4}$), although its relatively blue color still enables us to put stringent limits to the possible dust content in the afterglow environment. Supplementary Figure 4 shows the projections of the $(z, \alpha_O, \beta, E(B-V))$ four-dimensional confidence intervals on the different bidimensional planes.

Extending the analysis to the whole available temporal window ($2.5 \times 10^2 < t < 1.4 \times 10^6$ s) renders impossible to find a good fit with a single temporal power-law, because of the different decay regimes that the afterglow goes through.

3. Analysis of the TNG spectroscopic data

We observed the afterglow of GRB090423 with the near-IR camera and spectrograph NICS[40] on the Italian 3.6m Telescopio Nazionale Galileo (TNG) at La Palma. We used the the lowest spectral resolution mode, offered by the Amici prism[41]. This prism provides a simultaneous spectral coverage over a wide wavelength range, between 0.8 and 2.4 μm, and has a high efficiency. It yields a constant

spectral resolution $R \approx 50$ over the whole wavelength range. These characteristics make the instrument especially well-suited for studying the spectral distribution of faint objects.

We obtained 128 minutes of on-target spectroscopy. The afterglow was positioned in the slit using as reference a nearby star approximately 30 arcseconds away (at J2000 coordinates 09:55:35.31, +18:09:03.9). We used a dithering mosaic of 8 cycles, each including two coadds of single 120s exposures, repeated 4 times. The mean time of our observations was Apr 23.98, approximately 15.5 hours after the burst detection. The 2-dimensional spectrum is shown in Supplementary Figure 5.

Standard reduction tasks for NIR spectroscopy were performed independently by four different groups in our team, all of them reaching consistent results. Wavelength calibration was obtained by using a standard calibration table provided by the TNG and matching the deep telluric absorption bands. This method allows for wavelength calibrations better than 0.005 μm at 1.1 μm, and its contribution ($\Delta z = \pm 0.04$) to the final error budget on redshift is negligible.

Relative flux calibration was performed by using the observed spectral shape of the reference star. Its optical (SDSS) and near-IR (2MASS) colors are consistent with those of an M3-III star. The absolute calibration of the spectrum was obtained from the comparison with the simultaneous photometric measurements obtained by GROND (H=19.94 (Vega), ref. [7]). We estimate the slit losses to be less than 30%.

The observed flux is compatible with zero below a wavelength of 1.1 μm, while a significant flux ($> 99\%$ confidence level) is measured redwards of this limit. Assuming that this is due to hydrogen absorption by a virtually completely thick Lyman-α forest, then the redshift at which the GRB occurred is $z = 8.1^{+0.1}_{-0.3}$. The quoted error includes the uncertainties on the wavelength calibration and on the estimate of the break position. This value makes GRB090423 the most distant object spectroscopically identified to date. By using a standard cosmology with $\Omega_\Lambda = 0.73$, $\Omega_M = 0.27, H_0 = 71\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$, we find that GRB 090423 was detected at a lookback time of greater than 13 Gyrs.

We tentatively identified two absorption features at 1.3 and 2.2 μm. These would be consistent with blends of Si IV and Fe II at 1400Å and 2400Å, $z = 8.1$ rest-frame, respectively. The detection, however, has a low confidence level due to the low S/N of the spectrum.

4. Modelling the GRB redshift distribution

We compute the probability of detecting of GRB 090423 in three different scenarios for the formation

and cosmic evolution of long GRBs: (i) no evolution model, where GRBs follow the cosmic star formation and their luminosity function (LF) is constant in redshift; (ii) luminosity evolution model, where GRBs follow the cosmic star formation but the LF varies with redshift; (iii) density evolution model, where GRBs form preferentially in low–metallicity environments. In the first two cases, the GRB formation rate is simply proportional to the global cosmic star formation rate as computed by [42]. For the luminosity evolution model, the typical burst luminosity is assumed to increase with redshift as $(1+z)^{\delta}$. Finally, for the density evolution case, the GRB formation rate is obtained by convolving the observed SFR with the fraction of galaxies at redshift z with metallicity below Z_{th} using the expression computed by [43]. In this scenario, the LF is assumed to be constant.

The computation works as follows (see also [21,22,23,24,26,43]). The observed photon flux, P, in the energy band $E_{\rm min} < E < E_{\rm max}$, emitted by an isotropically radiating source at redshift z is

$$P = \frac{(1+z)\int_{(1+z)E_{\rm min}}^{(1+z)E_{\rm max}} S(E)dE}{4\pi d_L^2(z)}, \qquad (1)$$

where $S(E)$ is the differential rest–frame photon luminosity of the source, and $d_L(z)$ is the luminosity distance. To describe the typical burst spectrum we adopt the functional form proposed by [31], i.e. a broken power–law with a low–energy spectral index α, a high–energy spectral index β, and a break energy $E_b = (\alpha-\beta)E_p/(2+\alpha)$, with $\alpha = -1$ and $\beta = -2.25$ (ref. [45]). In order to broadly estimate the peak energy of the spectrum, E_p, for a given isotropic–equivalent peak luminosity, $L = \int_{1\,{\rm keV}}^{10000\,{\rm keV}} ES(E)dE$, we assumed the validity of the correlation between E_p and L (ref. [46]).

Given a normalized GRB LF, $\phi(L)$, the observed rate of bursts with $P_1 < P < P_2$ is

$$\frac{dN}{dt}(P_1 < P < P_2) = \int_0^{\infty} dz \frac{dV(z)}{dz}\frac{\Delta\Omega_s}{4\pi}\frac{\Psi_{\rm GRB}(z)}{1+z}\int_{L(P_1,z)}^{L(P_2,z)} dL'\phi(L'), \qquad (2)$$

where $dV(z)/dz$ is the comoving volume element, $\Delta\Omega_s$ is the solid angle covered on the sky by the survey, and the factor $(1+z)^{-1}$ accounts for cosmological time dilation. $\Psi_{\rm GRB}(z)$ is the comoving burst formation rate and the GRB LF is described by a power law with an exponential cut–off at low luminosities[47], i.e. $\phi(L) \propto (L/L_{\rm cut})^{-\xi}\exp(-L_{\rm cut}/L)$.

For the three scenarios, we optimize the model free parameters (GRB formation efficiency, burst typical luminosity at $z = 0$ and the power index ξ of the LF) by fitting the differential number counts observed by BATSE (see ref. [24,26] for a detailed description of the models and of the analysis). We

find that it is always possible to find a good agreement between models and data. Moreover, we can reproduce also the differential peak flux count distribution observed by Swift in the 15-150 keV band without changing the best fit parameters. On the basis of these results, we compute the probability to detect with Swift a GRB at $z \geq 8$ with photon flux P. The results are plotted in Supplementary Figure 6 (top panels) together with the cumulative number of GRBs at $z \geq 8$ expected to be detected by Swift in one year of observations (bottom panels). From the plot it is clear that the no evolution model fails to account for the observation of GRB 090423, since only $\sim 4 \times 10^{-4}$ GRBs are expected to be detected at $z \geq 8$ in ~ 4 years of Swift observations. Evolutionary models (both in luminosity or in density) can easily account for the discovery of GRB 090423. We note that the results confirm the need for cosmic evolution in the GRB luminosity function and/or in the GRB density obtained by recent analysis of the whole Swift GRB dataset. Indeed, both the large number of $z \geq 2.5$ bursts[24] and the number of bright (i.e. with peak luminosity $L \geq 10^{53}$ erg s^{-1}) bursts[26] strongly require the existence of evolution.

Moreover, we want to stress here that our conclusions are conservative. First of all, many biases can hampered the detection of GRB at very high redshift. Indeed, a few very high-z bursts may be hidden among the large sample of Swift bursts that lack of an optical detection. Thus, the discovery of a single event at $z > 8$ in 4.5 yrs of Swift operation can be treated as a lower limit on the real number of high-z detection. Moreover, our choice of the GRB LF is also conservative, since the existence of large population of faint GRBs (i.e. for an LF with a more gentle decline or a rise in the faint end) would lead to a decrease of the expeceted number of GRBs at $z > 8$ strenghtening our conclusions.

31. Band, D.L. *et al.* BATSE observations of gamma-ray burst spectra. I - Spectral diversity. *Astrophys. J.* **413**, 281-292 (1993).

32. Zhang, B.-B. & Zhang, B. GRB 090423: pseudo burst at z=1 and its relation to GRB 080913. *GCN Circ.* 9216 (2009).

33. Amati, L. *et al.* On the consistency of peculiar GRBs 060218 and 060614 with the $E_{p,i} - E_{iso}$ correlation. *Astron. Astrophys.* **463**, 913-919 (2007).

34. Norris J.P. *et al.* Long-Lag, Wide-Pulse Gamma-Ray Bursts. *Astrophys. J.* **627**, 324-345 (2005).

35. Dickey, J.M. & Lockman F.J. H I in the Galaxy. *Annu. Rev. Astron. Astrophys.* **28**, 215-261 (1990).

36. Kalberla, P.M.W. *et al.* A New Whole HI Sky Survey. Proceedings of ASP Conference **317**, 13 (2004).

37. ev. sostituire con Campana in prep. Evans, P.A. *et al.* Methods and results of an automatic analysis of a complete sample of Swift-XRT observations of GRBs. ArXiv e-print:0812.3662 (2008).

38. Campana, S. *et al.* A Metal-rich Molecular Cloud Surrounds GRB 050904 at Redshift 6.3. *Astrophys. J. Lett.* **654**, 17-20 (2007).

39. Fernández-Soto A. *et al.* z-ph-REM: A photometric redshift code for the REM telescope. arXiv:astro-ph/0309492 (2003).

40. Baffa, C. *et al.* NICS: The TNG Near Infrared Camera Spectrometer. *Astron. Astrophys.* **378**, 722-728 (2001).

41. Oliva E. Infrared instrumentation for large telescopes : an alternative approach. MemSAIt **71**, 861 (2000)

42. Hopkins, A. M. & Beacom, J. F. On the Normalization of the Cosmic Star Formation History. *Astrophys. J.* **651**, 142-154 (2006).

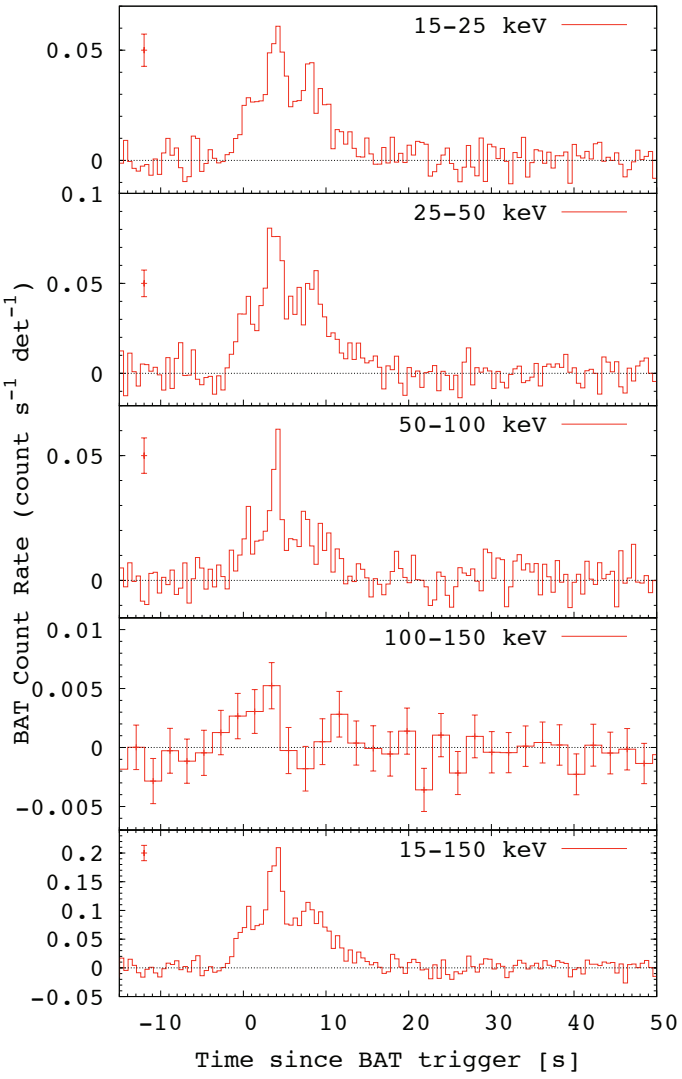
43. Langer, L. & Norman, C. A. On the Collapsar Model of Long Gamma-Ray Bursts:Constraints from Cosmic Metallicity Evolution. *Astrophys. J. Lett.* **638**, 63-66 (2006).

44. Natarajan, P. *et al.* The redshift distribution of gamma-ray bursts revisited. *Mon. Not. R. Astron. Soc.* **364**, L8-L12 (2005).

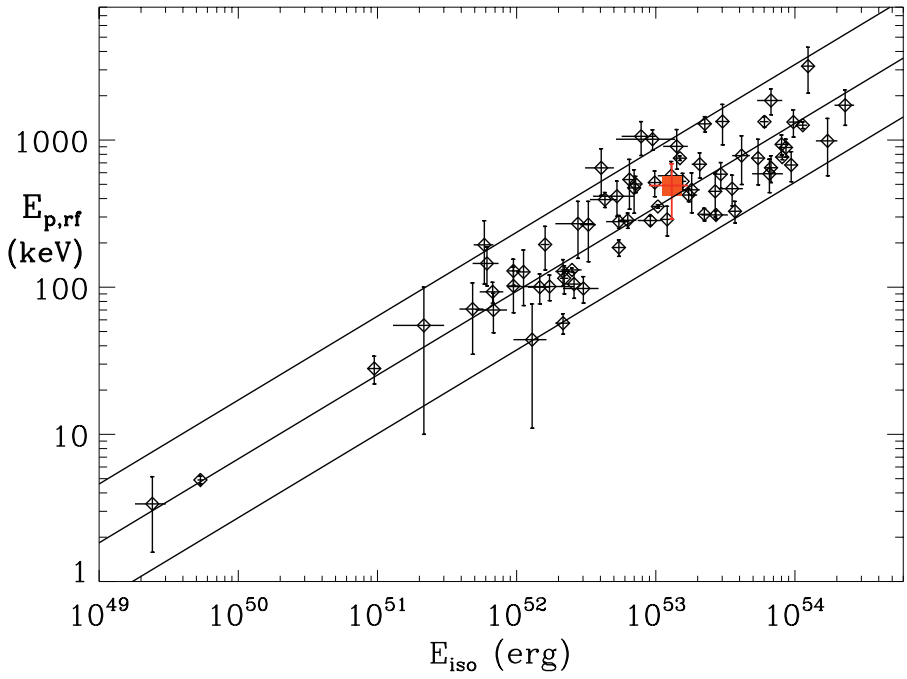
45. Preece R. D. *et al.* The BATSE Gamma-Ray Burst Spectral Catalog. I. High Time Resolution Spectroscopy of Bright Bursts Using High Energy Resolution Data. *Astrophys. J. Suppl. Ser.* **136**, 19-36 (2000).

46. Yonetoku D. *et al.* Gamma-Ray Burst Formation Rate Inferred from the Spectral Peak Energy-Peak Luminosity Relation. *Astrophys. J.* **609**, 935-951 (2004).

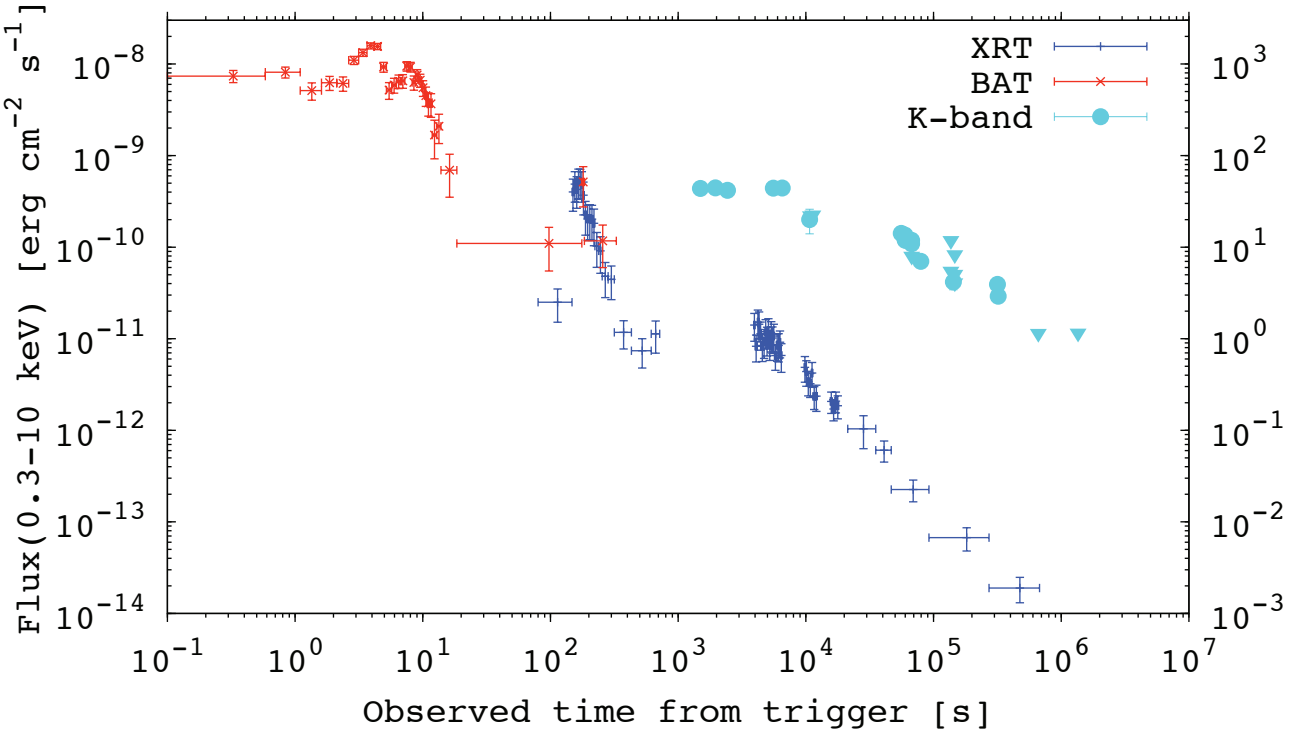
47. Porciani, C. & Madau, P. On the Association of Gamma-Ray Bursts with Massive Stars: Implications for Number Counts and Lensing Statistics. *Astrophys. J.* **548**, 522-531 (2001).



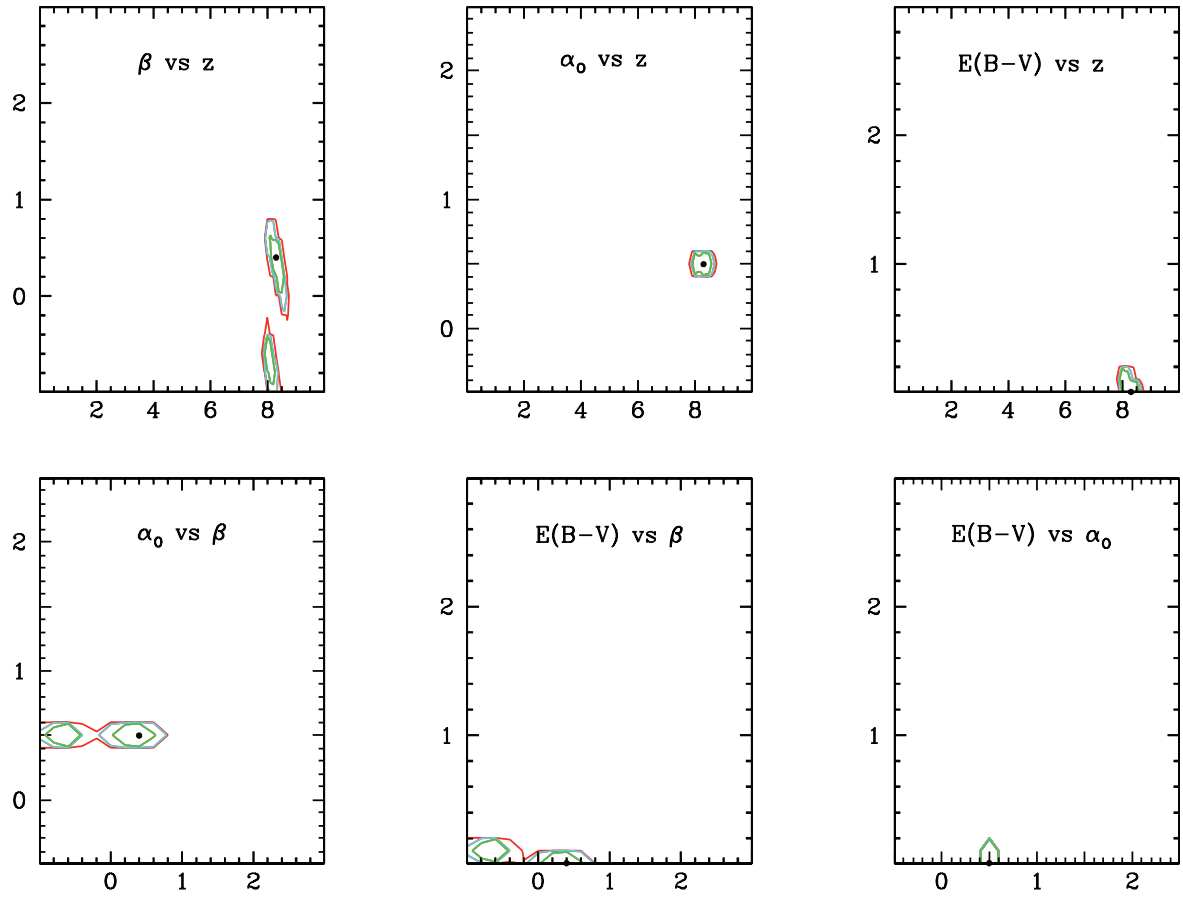
Supplementary Figure 1. BAT mask-weighted light curve. Four channels and combined 0.512 s mask-weighted light curve. The light curve of the 100-150 keV energy channel shows a weak signal, because of the soft spectrum; the corresponding integration time is 2.048 s. Errors are at 1σ level.



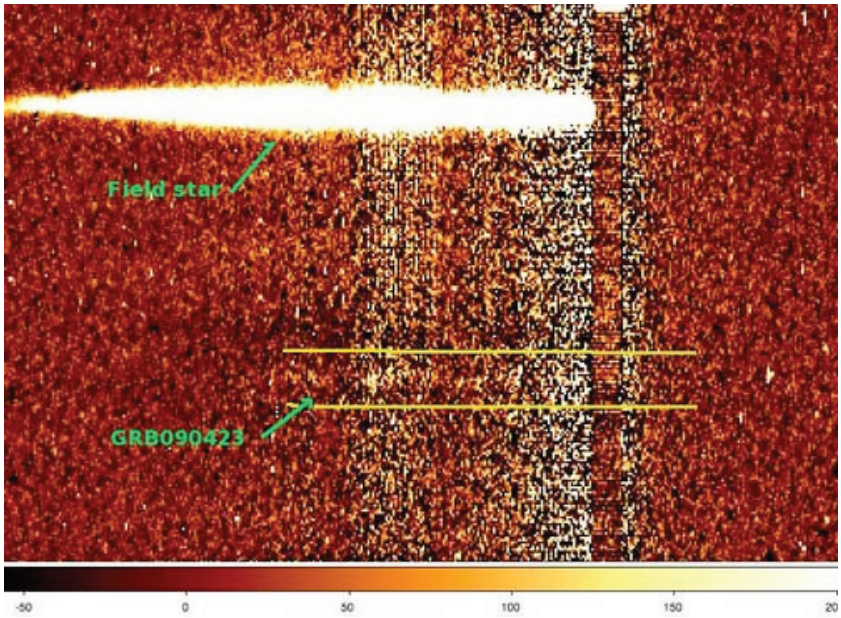
Supplementary Figure 2. Isotropic energy and peak energy correlation. Position of GRB 090423 in the $E_{\rm p,rf}$ – E_{iso} plane based on Swift/BAT[4] and Fermi/GBM[10] (fitted with the Band function) results. The lines show the best–fit power–law and the $\pm 2\sigma$ region of the correlation as derived by [13]. Also shown are the 70 GRBs included in the sample analyzed in that work (errors on individual bursts are at 1σ level). Given that short GRBs do not follow the correlation[13], this evidence supports the hypothesis that, despite its cosmological rest–frame duration of $\sim$ 1.3 s, GRB 090423 belongs to the long GRB class.



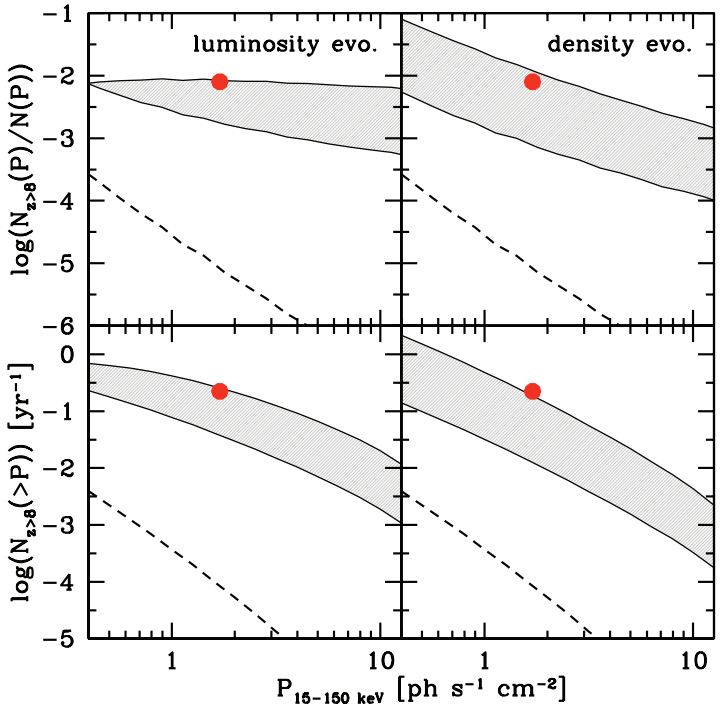
Supplementary Figure 3. Observed light curve. Light curve of GRB 090423 as observed by Swift/BAT (red crosses), Swift/XRT (blue plus) and in the NIR (cyan points). Errors on fluxes are at 1σ level and horizontal bars refer to the integration time interval. The XRT 0.3–10 keV light-curve, starting at 73 s after the burst, shows a prominent flare at $t \sim 170$ s (also detected by BAT), and a flat phase ($\alpha_{X,1} = 0.13 \pm 0.11$) followed by a rather typical decay (starting at $t = 4513 \pm 491$ s) with power-law index $\alpha_{X,2} = 1.3 \pm 0.1$. Available photometric data are plotted in the K band (AB magnitude) by transforming the fluxes, when the observations have been taken in a different filter, using a power law with $\beta = 0.4$, as estimated from the NIR spectral energy distribution. A small displacement in time for contemporary data in different bands is applied in order to increase the visibility. The NIR light curve is consistent with a plateau phase ($t \sim 10^2 - 10^3$ s) followed by a decay with $\alpha_O \sim 0.5$ ($t \sim 10^3 - 10^5$ s). This decay phase is shallower than the X-ray decay in the same time interval. Triangles at $t \sim 10^5$ s report NIR upper limits as obtained by our second epoch TNG observation with the NICS camera in the Y and J band and by GROND in the JHK band. These limits are consistent with the temporal decay observed by XRT.



Supplementary Figure 4. Multi-parameter analysis of the photometric data. Analysis of available photometric data for GRB090423 in the interval 70 min $< t <$ 1100 min. The code fits a model function with temporal index α_O and spectral index β, dust extinction $E(B-V)$, and redshift z. The different panels show the projection of the four-dimensional confidence intervals on the different two-dimensional planes of interest. The best-fit is marked by the black dot, with the red, cyan, and green contours defining respectively the 68%, 95%, and 99.5% confidence areas. The apparently bimodal distribution in the β direction is an artifact of the parameter space discretization.



Supplementary Figure 5. TNG 2-dimensional spectrum. The spectrum has been taken by $\sim$ 14 hrs from the trigger. The spectrum of the nearby reference star is also shown.



Supplementary Figure 6. Probability of the occurance of GRB 090423. Top panels: probability for a GRB with peak photon flux P to be detected by Swift at $z \geq 8$. Luminosity evolution models are shown in the left panel, where shaded area refers to a typical burst luminosity increasing as $L_{\rm cut} \propto (1+z)^{\delta}$ with $\delta = 1.5 - 3$. Density evolution models are shown in the right panel, where shaded area refers to a metallicity threshold for GRB formation $Z_{th} = 0.02 - 0.2\ Z_{\odot}$ (the lower bound refers to the higher Z_{th}). In both panels, the dashed line shows the no evolution case. The red point marks the position of GRB 090423. We note that the point represents a lower limit on the number of detection at $z > 8$ since a few very high-z bursts may be hidden among those bursts that lack of an optical detection. Bottom panels: cumulative number of GRBs at $z > 8$ to be detected by Swift with photon flux larger than P in one year of Swift observations.