Electronic supplementary information

Stabilizing synthetic DNA for long-term data storage with earth alkaline salts

A. Xavier Kohll,[a] Philipp L. Antkowiak,[a] Weida D. Chen,[a] Bichlien H. Nguyen,[b] Wendelin J. Stark,[a] Luis Ceze,[c] Karin Strauss[b] and Robert N. Grass*,[a]

[a] Institute for Chemical and Bioengineering, ETH Zurich, Vladimir-Prelog-Weg 1, 8093 Zürich

[b] Microsoft, One Microsoft Way, Redmond, WA, 98052, USA

[c] Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA

Corresponding author:

Email: robert.grass@chem.ethz.ch

DNA and primer sequences

148 BP MS DNA: Seq4235_tr_wpr_re

AGACGTGTGCTCTTCCGATCTCTTGCTGACTTCTATGTAGCGCTATGGTAGTCGCG
AGCTAACTATATCGCGTGAGCACCGTTGAAGACGAGACTGCGAGGTGGAGATCT
ATCTAGTAGATGAGAGTCAGATCGG AAGAGCGTCGTGT

148 BP MS DNA: Seq4235_tr_wpr_fr

ACACGACGCTCTTCCGATCTGACTCTCATCTACTAGATAGATCTCCACCTCGCAGT
CTCGTCTTCAACGGTGCTCACGCGATATAGTTAGCTCGCGACTACCATAGCGCTA
CATAGAAGTCAGCAAGAGATCGGAAGAGCACACGTCT

Primer 0R: AGA CGT GTG CTC TTC CGA TCT

Primer 0F: ACA CGA CGC TCT TCC GAT CT

Sample preparation protocols

This protocol describes a method for encapsulating and de-encapsulating DNA into calcium phosphate. After the de-encapsulation of DNA from calcium phosphate, the nucleic acids were analyzed by qPCR. All other DNA salt samples were prepared following the same protocol by changing the used salt solutions.

Preparation of solutions:

1 M solutions of chemicals were prepared (Calcium chloride dihydrate (Ca solution), Potassium Phosphate Monobasic, and Di-Potassium Hydrogen Orthophosphate Trihydrate) in mQ water. KP solutions were prepared by mixing the two phosphate solutions 1:1.. Ca solution and KP solutions were diluted with mQ water to a concentration of to 0.1 mM.

DNA preparation

Prior to use DNA was desalted via drop dialysis and diluted to a concentration to 15 ng/µl with mQ water.

Sample preparation

To 2 µl of DNA solution (15 ng/µl) in an Eppendorf tube were added 5 µl of Ca solution (0.1 M) prior to adding 5 µl KP solution. The solutions were dried in a vacuum centrifuge to remove the water for at least 2 hours.

De-encapsulation

For retrival of the DNA from the dry material, 100 µl of a 1mM EDTA solution were added to the tube and vortexed. Prior to qPCR the sample was additionally diluted 1:100 to prevent interference of the salts in the amplification.

Accelerated aging experiments

All accelerated aging experiments were performed at 60°C or 70°C at 50% relative humidity in a closed desiccator, filled with a saturated NaBr solution. The desiccators were placed inside an oven at the respective temperature.

PCR amplification protocol

MS DNA and DNA pool encoding 115kB of data were amplified with a Roche Lightcycler 96. Each sample well was filled with a total volume of 20 µl containing: 5 µl sample volume, 10 µl of Kapa Sybr Fast qPCR Master Mix, 3 µL mQ water, 1 µl (10 µM) forward primer and 1 µl (10 µM) reverse primer. The qPCR for MS DNA consisted of a 3-step amplification protocol (95°C for 15 s, 56°C for 15 s and 72°C for 10 s), and each sample was analyzed in duplicates. The qPCR for the DNA pool also consisted of a 3-step amplification protocol (98°C for 20 s, 60°C for 15 s and 72°C for 20 s), and each sample was analyzed in duplicates.

Experiments with genomic DNA

For the measurements of the decay of genomic DNA with $MgCl_2$, Sodium salt of salmon testes DNA (1 mg) was dissolved in deionized water (1 mL), desalted over a membrane filter (0.025 µm pore size) and diluted to a concentration of 15 ng/µL with deionized water or TE buffer (DNA stock solutions). Concentration measurements were conducted with a ThermoFisher Nanodrop 2000 device. Samples were prepared by adding 5 µL of $MgCl_2$ solution (0.1 mM) to 2 µL of DNA stock solution in 2 mL eppendorf tubes. As controls, 2 µL of each, DNA-TE buffer stock solution and DNA-water stock solution without salt-encapsulation were aliquoted. The samples were placed in an oven for 0, 1, 2 days at 70 °C and 50 % relative humidity. De-encapsulation was performed by adding 100 µL of EDTA solution (1 mM) and subsequent dilution in deionized water (1:10). The samples were measured in duplicates with qPCR (Roche LightCycler 480 II). Each sample well was filled with a total volume of 20 µl containing: 5 µl sample volume, 10 µl of Kapa SYBR Fast qPCR Master Mix, 3 µL mQ water, 1 µl (1 µM) OnKeta-spDL-F3 forward primer and 1 µl (1 µM) OnKeta-spDL-R3 reverse primer. The protocol consisted of a 2 min activation step at 50 °C, 10 min

denaturation at 95 °C and a 3-step amplification (50 cycles, 95°C for 15 s, 60°C for 30 s and 72°C for 20 s).

Salmon DNA forward primer: OnKeta-spDL-F3

5'-CCC GCA CAT TTG TAA ATG C-3'

Salmon DNA reverse primer: OnKeta-spDL-R3

5'-TGA TGT ATG AGG GGT TAA AAT AAG-3'

For STEM/EDX and PXRD measurements, DNA (sodium salt from salmon testes, Sigma-Aldrich) was dissolved in deionized water (Merck Milli-Q, ~ 1 g/L) and desalted by applying to a membrane filter (Merck MF-Millipore, 0.025 µm pore size) for 30 min. Samples were prepared by precipitating K_2HPO_4 /KH_2PO_4 (1:1 (v/v), 500 µL, 0.1 M) and $CaCl_2$ (500 µL, 0.1 M) in the presence of DNA (1.5 mg). The resulting CaP-DNA solution was dried under vacuum for 5 hours at room temperature before measuring.

Scanning transmission electron micrographs and EDX spectra were recorded with a FEI, NovaNanoSEM 450 and a beam voltage of 30.0 kV. For this, the dried CaP-DNA was dispersed in ethanol and ultrasonicated.

Powder X-ray diffraction was performed with a PANalytical X'Pert PRO-MPD diffractometer using Cu-Kα (λ = 0.154 nm) radiation and a step size of 0.033°.

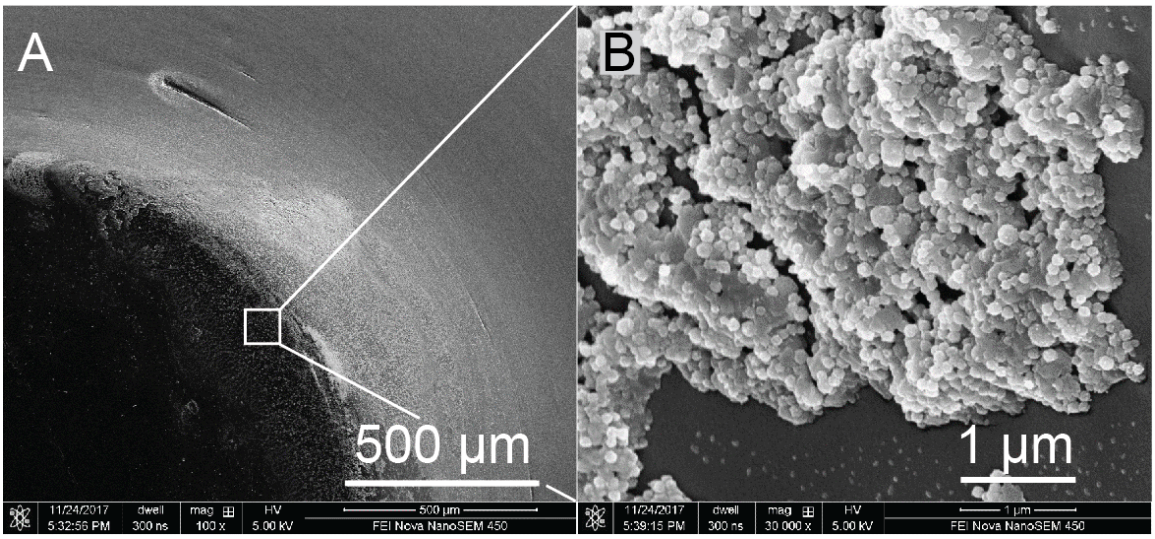

Figure S1. *SEM pictures (**A**) and (**B**) at a different magnification of a dried sample of DNA with CaP (18 wt%) in a 2 ml Eppendorf tube.*

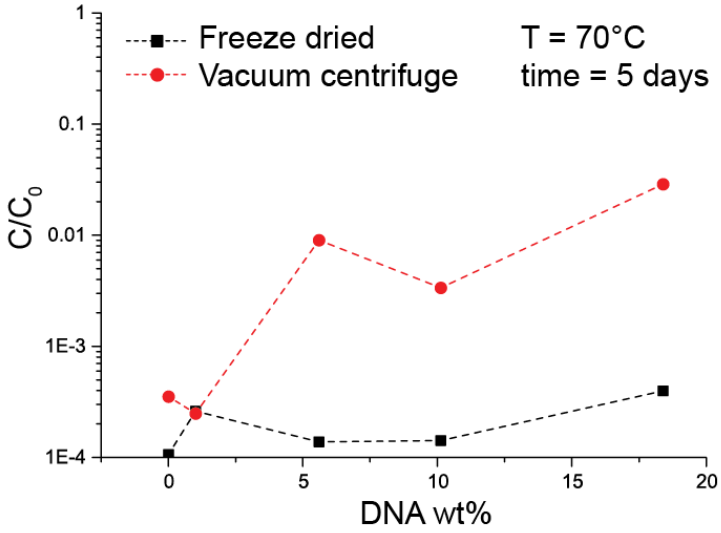

Figure S2. *The relative concentration of DNA stored with different amounts of calcium phosphate after accelerated aging at 70°C and 50% RH for 5 days. Samples were dried via freeze drying or a vacuum centrifuge. The slow solidification of the CaP/DNA sample during vacuum drying resulted in a significantly increased DNA stability, if compared to CaP/DNA solidified via freezing in liquid nitrogen.*

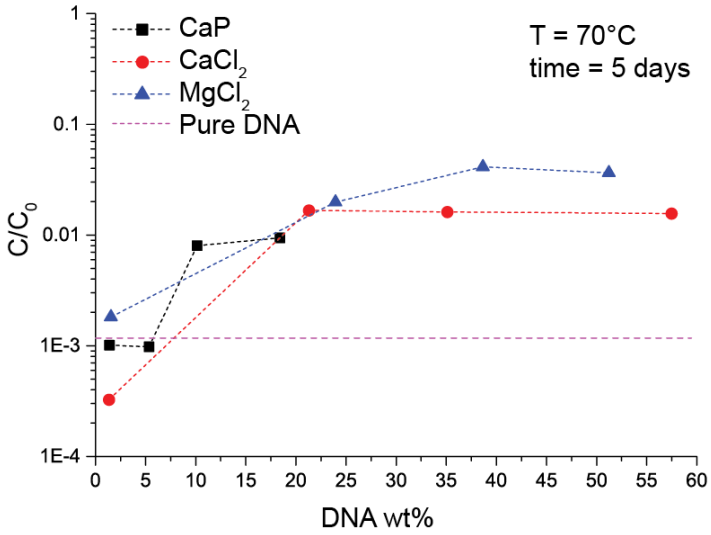

Figure S3. *The relative concentration of DNA pool consisting of 7'323 distinct oligonucleotide sequences (150 bp long) stored with different amounts of salts (CaP, $MgCl_2$, and $CaCl_2$) after accelerated aging at 70°C and 50% RH for 5 days. For CaP, higher loadings resulted in a reduced stability (data not shown).*

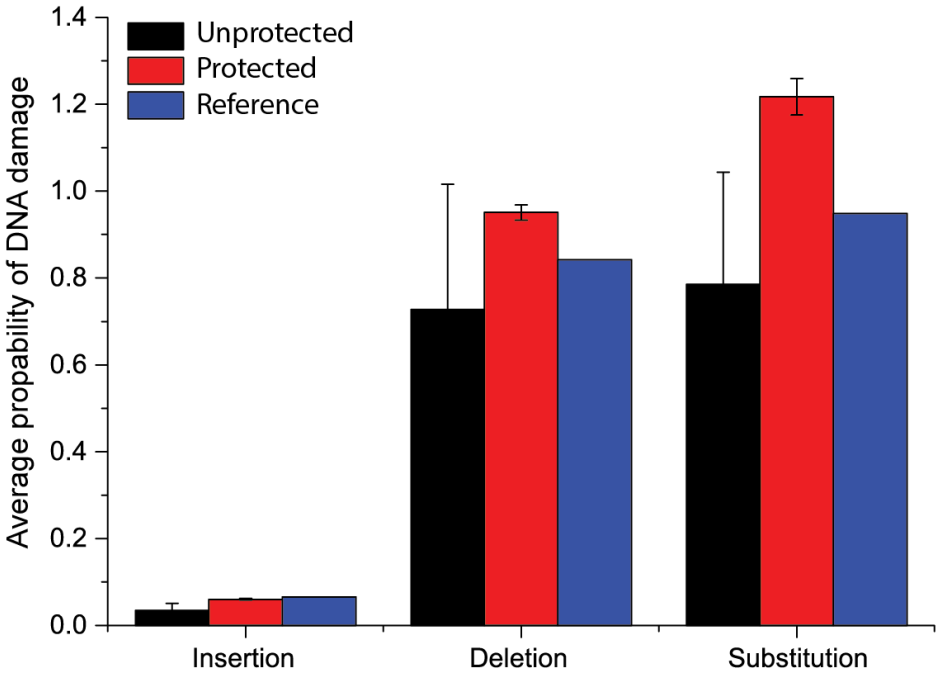

Figure S4. *The average probability of DNA damage by insertion, deletion, or substitution per 110 base pairs (payload region) of DNA protected by $MgCl_2$ (red, n = 5), unprotected DNA (black, n = 5) and a reference (=unaged) sample (blue), indicating that DNA damage during rapid aging does not involve the deletion, insertion or substitution of individual bases, but rather the loss of whole DNA strands due to strand break (see Fig 4 in main text).*

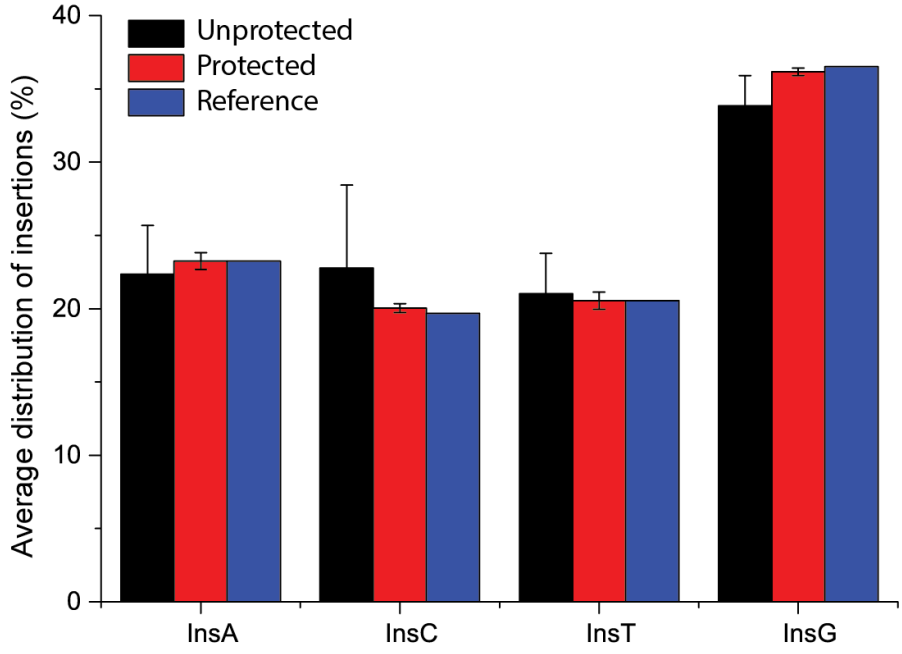

Figure S5. *The average distribution of bases inserted in DNA samples after storage protected by $MgCl_2$ (red, n = 5), unprotected DNA (black, n = 5), and a reference sample (blue).*

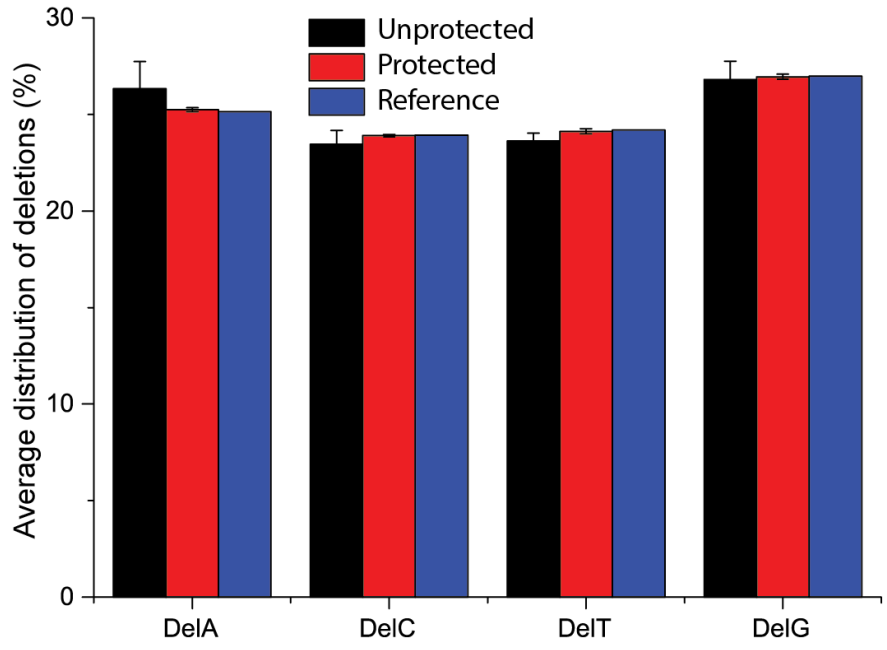

Figure S6. *The average distribution of bases deleted from the sequenced DNA samples protected by $MgCl_2$ (red, n = 5), unprotected DNA (black, n = 5), and a reference sample (blue).*

Protocol for sequencing preparation

In short, a random 25N region was added via PCR.[1] Following an adapted TruSeq® Nano Library Prep protocol (Illumina, June 2015, Rev. D) and a TruSeq® ChiP sample preparation protocol (Illumina, October 2013, Rev. B), sequencing adapters were ligated to the samples. After quality control of the samples via Gel electrophoresis, samples were cut out from the gel a released with QIAX Gel Extraction Kit. Next, a Qubit concentration measurement was done, and the samples were diluted and spiked with 2% of Illumina's PhiX control before being sequenced on the DNA sequencer iSeq 100 (Illumina).

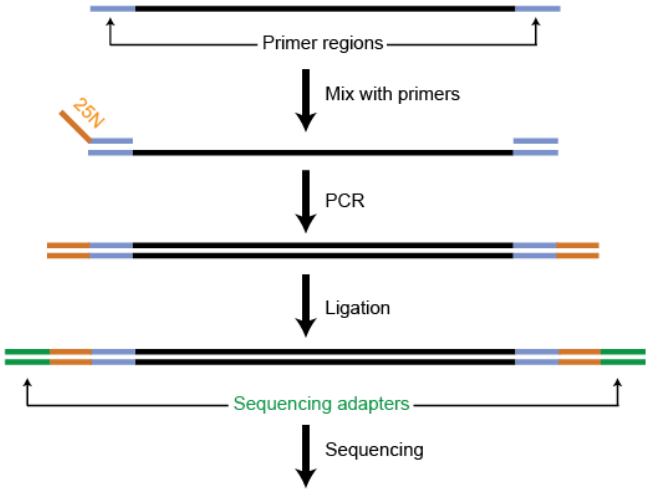

Figure S7.*Schematic overview of the main steps involved in DNA sequencing preparation*.

Detailed sequencing Preparation protocol:

Addition of 25N region by PCR amplification

- Amplify protected, and unprotected samples with two separate PCR runs.
- Dilute the samples 1:500 with mQ water.
- Prepare PCR mix:
- 5 µl DNA sample; 1 µl of 10 mM forward primer and 1 µl of 10 mM reverse primer, 10 µL of 2x Kapa cybr Fast mix, 3 µl of mQ water.
- Thermocycling protocol: (1) 95 °C for 3 min, (2) 98°C for 20 s, (3) 62 °C for 20 s, (4) 72 °C for 15 s, repeat steps 2-4 as needed.
- Stop the PCR when the signal reaches the plateau phase.

Conversion to blunt ends (adapted from page 14 of TruSeq® Nano Library Prep protocol)

- Add 20 µL ERP2 to each well with 10 µL of PCR product and 20 µL of mQ water.
- Pipette up and down to mix.

- Place on the thermal cycler and run the ERP program (30 min at 30 °C) then place it on ice. Each well contains 50 µl.

Sample purification with AMPure XP beads (adapted from page 12 of TruSeq® ChiP sample preparation protocol)

- Add 80 µl well-mixed AMPure XP beads to each well of the PCR plate containing 50 µl End Repair Mix.
- Gently pipette the entire volume up and down 10 times to mix thoroughly.
- Incubate the PCR plate at room temperature for 15 minutes.
- Place the PCR plate on the magnetic stand at room temperature for 15 minutes or until the liquid is clear.
- Using a 200 µl pipette set to 125 µl, remove and discard 125 µl of the supernatant from each well of the PCR plate.
- 5 µl are left in each well.

NOTE Leave the PCR plate on the magnetic stand while performing the following 80% EtOH wash steps.

- With the PCR plate on the magnetic stand, add 100 µl freshly prepared 80% EtOH to each well without disturbing the beads. (*Pipette to the side of the wall)*
- Incubate the PCR plate at room temperature for 30 seconds, and then remove and discard all of the supernatants from each well. Take care not to disturb the beads.
- Repeat steps g and h one time for a total of two 80% EtOH washes.
- Let the PCR plate stand at room temperature for 15 minutes to dry, and then remove the plate from the magnetic stand.
- Resuspend the dried pellet in each well with 17 µl Resuspension Buffer (RSB). Gently pipette the entire volume up and down 10 times to mix thoroughly.
- Incubate the PCR plate at room temperature for 2 minutes.
- Place the PCR plate on the magnetic stand at room temperature for 5 minutes or until the liquid is clear.
- Transfer 15 µl of the clear supernatant from each well of the PCR plate to the corresponding well of a new 96-well 0.2 ml PCR plate.

Ligation of adapters

- Add 15 µl of ATL to the 15 µl of supernatant from the previous step.
- Briefly centrifuge (3 min at 600 rpm).
- Place on the thermal cycler and run the ATAIL70 program. Each well contains 30 µl
- Choose the preheat lid option and set it to 100°C.
- 37°C for 30 minutes
- 70°C for 5 minutes
- 4°C for 5 minutes

Addition of the adapters: add the in the following order:

- RSB (2.5 μl)
- LIG2 (2.5 μl)
- DNA adapter (2.5 μl)

- Pipette up and down, centrifuge briefly (3 min at 600 rpm).
- Run **lig program**
- Choose the preheat lid option and set to 100°C
- 30°C for 10 minutes
- Hold at 4°C
- Add 5 µl STL to each well, and then mix well.

NOTE: Can hold this mixture at -20°C overnight.

- Vortex SPB until well dispersed.
- Add 42.5 µl of SPB (AMPure XP beads) to each well, and then mix thoroughly as follows. Pipette up and down.
- Incubate at room temperature for 5 minutes.
- Place on a magnetic stand and wait until the liquid is clear (2–5 minutes).
- Remove and discard all supernatant from each well.
- Wash 2 times as follows.
- Add 200 µl freshly prepared 80% EtOH to each well.
- Incubate on the magnetic stand for 30 seconds.
- Remove and discard all supernatant from each well.
- Use a 20 µl pipette to remove residual EtOH from each well.
- Air-dry on the magnetic stand for 5 minutes.
- Add 52.5 µl RSB to each well.

- Remove from the magnetic stand, and then mix thoroughly as follows. Pipette up and down.
- Incubate at room temperature for 2 minutes.
- Place on a magnetic stand and wait until the liquid is clear (2–5 minutes).
- Transfer 25 µl supernatant to the corresponding well of the PCR plate.

SAFE STOPPING POINT: Ligation is done. If you are stopping, seal the plate and store at -25°C to -15°C for up to 7 days.

Enrichment for sequencing:

Sample	Concentration	Volume (µL)
DNA mix		3
PPC (PCR Primer Cocktail)	10X	3
EPM (Enhanced PCR Mix)	2.5X	12
H_2O		12
Total		30

- Choose the preheat lid option and set to 100°C
- 95°C for 3 minutes
- 10 total cycles of:
- 98°C for 20 seconds
- 60°C for 15 seconds
- 72°C for 30 seconds
- 72°C for 3 min
- Hold at 4°C

Quality control and dilution for sequencing:

- Run a gel electrophoresis experiment (E-gel®EX, SYBR Gold II, 2% Agarose) with ligated samples and control samples with 25N region.
- Cut out DNA band of ligated samples and release the DNA with QIAX Gel Extraction Kit (Qiagen).
- Measure DNA concentration with dsDNA assay for Qubit Fluorometer.
- Dilute each sample to 1 nmol/L.
- Add 10 µl of each the samples to one tube.
- Dilute the DNA solution to a total volume of 100 µl and 50 pmol/L and add 2 µl of 50 pmol/L PhiX. For details, see Illumina iSeq 100 protocol.

Comparison of DNA storage media

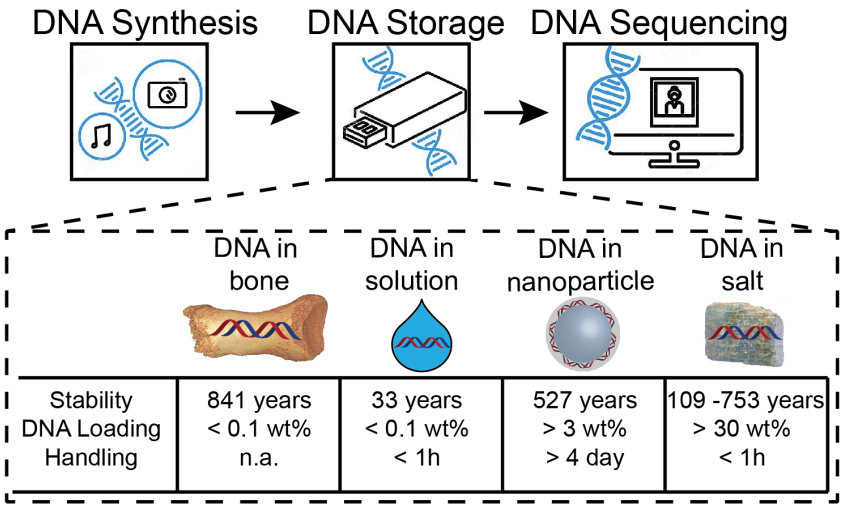

Figure S8. Comparison of different DNA storage media. Estimates for stability (half-life), DNA loading and synthesis speed of the storage media are shown. Calculation of values given below.[2–8]

DNA in bone

Stability:

Half-life ($t_{1/2}$) of 242 bp mtDNA at 13.1 °C is 521 years.[2] $t_{1/2}$ scaled to 150 bp equals 841 years.

DNA loading:

$$DNA\ loading\ in\ bones = 0.05 * 0.01\ wt\% = 0.05\ wt\%$$

Calcified bones contain approximately up to 5 wt% cells.[3] More than 80% of the inorganic bone material is calcium and phosphate.[9] DNA mass in a human cell is about 10 pg/cell.[10] The cell volume is around 1000 µm^3 (≈ 1000 pg). Therefore, a cell is approximately composed of 1 wt% DNA.

DNA in solution

Stability:

DNA in solution has a reported activation energy of 129 kJ/mol.[8] Extrapolation of the rate constant of depurination per nucleotide to 10°C allows the calculation of the half-life, which is equal to 33 years (scaled to DNA strands of 150 bp). All data was taken

from Lindahl et al.[8] K_0 = 5.45 E12; Scaling to 150 bp strands: $t_{1/2}^{150nt} = t_{1/2}^{1\,nt}/75$ (on average, every second nucleotide is a purine).

Oxford Gene Technology suggests storage of DNA samples at 4°C for only a limited number of weeks to ensure high-quality microarray results.[4]

DNA loading:

$$DNA\ loading\ in\ solution = 500\ \text{ng}/\mu\text{L} = 0.05\ wt\%$$

The typical concentration of dsDNA from Microsynth AG is approximately 500 ng/µL.

Handling and synthesis time: DNA annealing time plus sample preparation time.

DNA in nanoparticle

Stability: Half-life of 500 years at 10°C was calculated with data from Grass et al. for 158 bp long DNA.[5] $t_{1/2}$ scaled to 150 bp equals 527 years.

DNA loading is reported by Chen et al.[6] Handling time is reported by Paunescu et al.[11]

DNA in salt

Stability:

DNA pool in the presence of $MgCl_2$ (38 wt%) was aged for three weeks and compared to data presented by Chen *et al.* (see **Fig. S2.**). The experimentally determined DNA decay rate constant at 70°C ($k_{70°C} = 2.95*10^{-6}\ s^{-1}$) and two different activation energies found in the literature (129 kJ/mol and 155 kJ/mol) were used to calculate the rate constant at different temperatures using Arrhenius equation. Next the half-life ($t_{1/2} = \ln(2) / k$) of this storage method was calculated for different temperatures (see **Fig. S10.**). [5,8,12]

Handling and synthesis time: Sample preparation time and drying time.

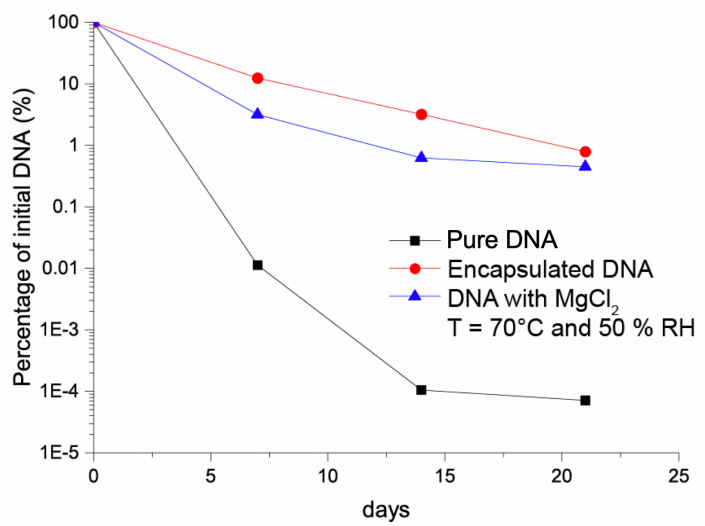

Figure S9. *The decay of the DNA pool with $MgCl_2$ (38 wt%) by exposure to accelerated aging conditions. Pure DNA and encapsulated DNA decay data presented by Chen et al. were compared to this storage method.*[6]

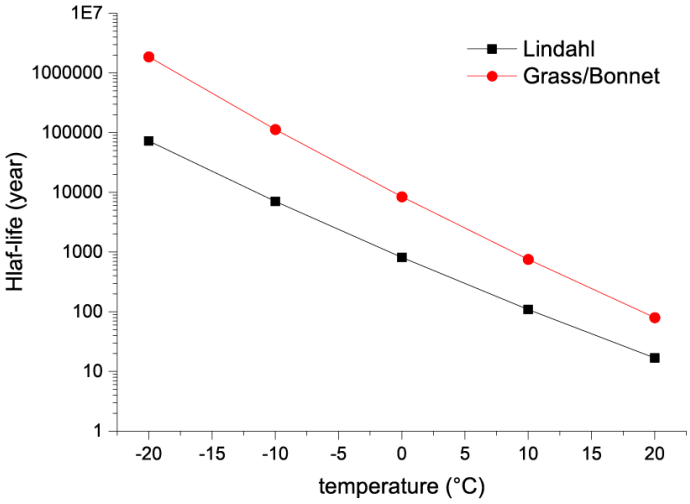

Figure S10. *Half-lives for DNA stored with $MgCl_2$ (38 wt%) extrapolated from experimental rate constant and activation energies from literature. Estimated half-lives at 10°C vary between 109 and 753 years with different activation energies taken from Lindahl et al. (129 kJ/mol), Grass et al. (155 kJ/mol) and Bonnet et al. (155 kJ/mol).*[5,8,12]

Encapsulation of genomic DNA

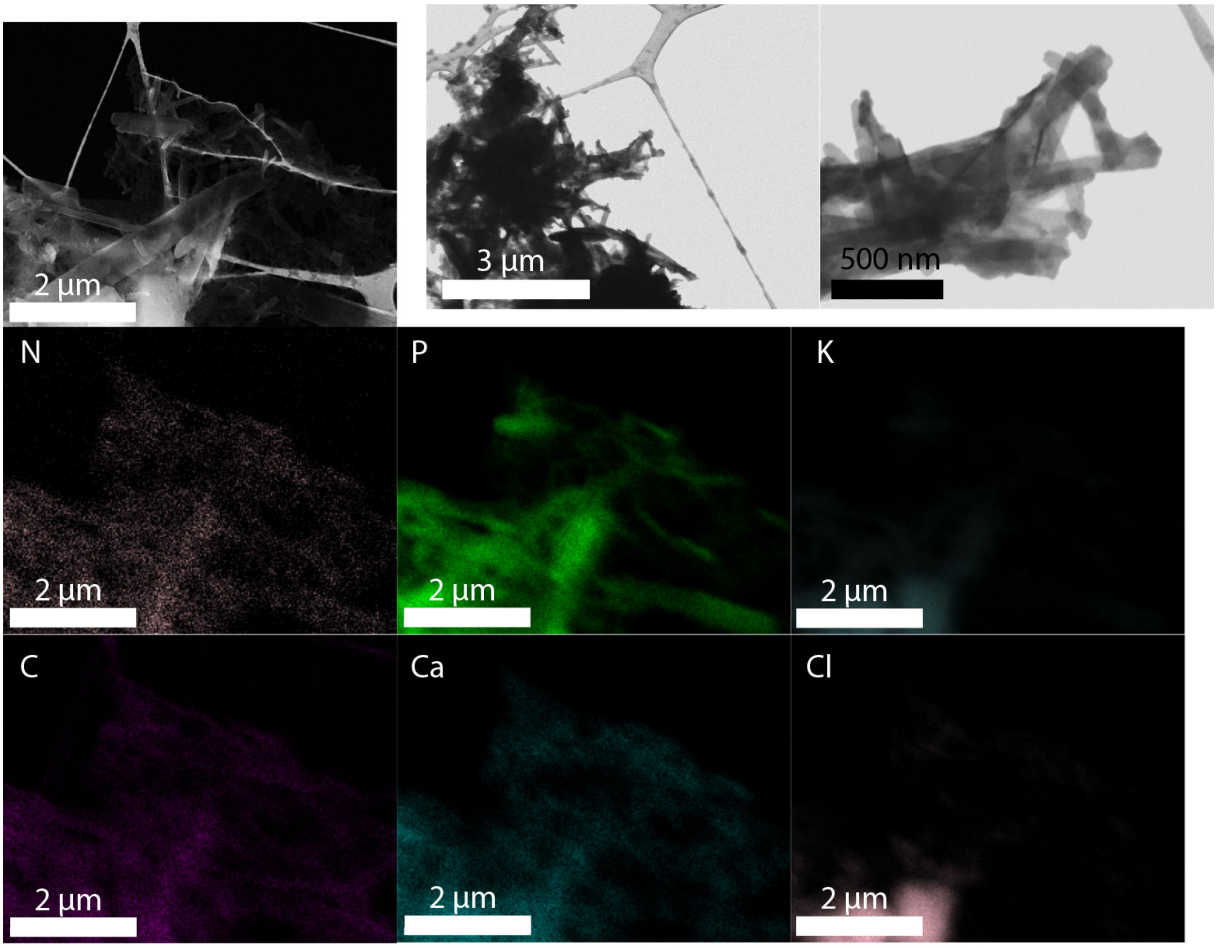

Figure S11 *STEM images and EDX maps of genomic DNA co-precipitated with calcium phosphate revealing needle-like structures. Based on the EDX maps, the needle like structures contain elements corresponding to calcium phosphate (Ca, P), as well as DNA (C, N). In the lower left corner the STEM image also shows some cubic structures, which according to the EDX maps consist of KCl. These images show that the products of co-precipitating DNA with solutions of potassium phosphates and calcium chloride results in DNA/CaP structures and KCl crystals as byproducts.*

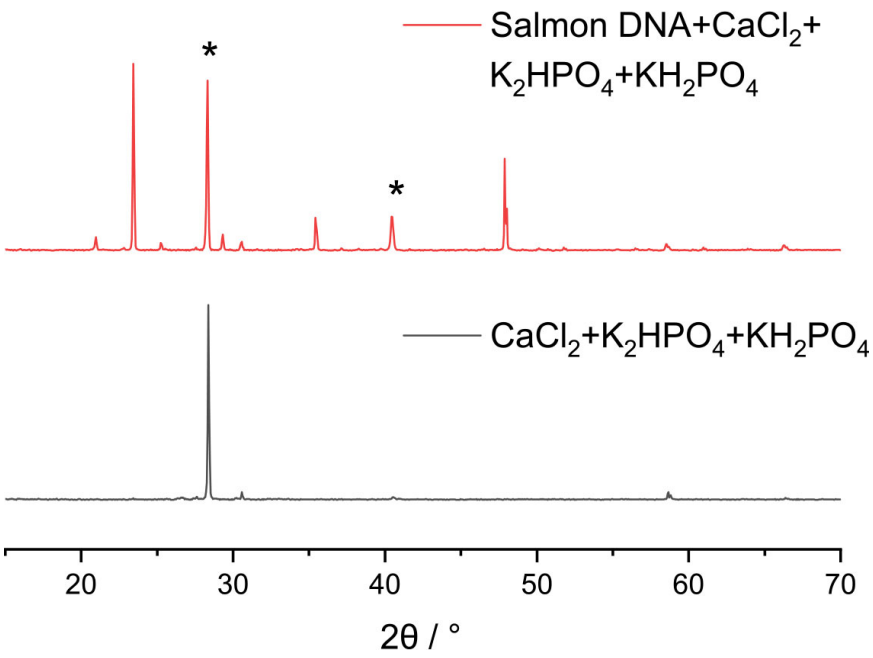

Figure S12. *Bottom: XRD pattern from the co-precipitation and crystallization by evaporation of calcium chloride and potassium phosphates showing the formation of KCl. No crystalline features of calcium phosphates are visible, which indicates the formation of amorphous calcium phosphate. Top: Adding genomic DNA (30wt%) to the same salt solution and precipitating/drying results in additional crystal phases, which could not be attributed to common calcium phosphates. *denotes reference crystal pattern of sylvite*

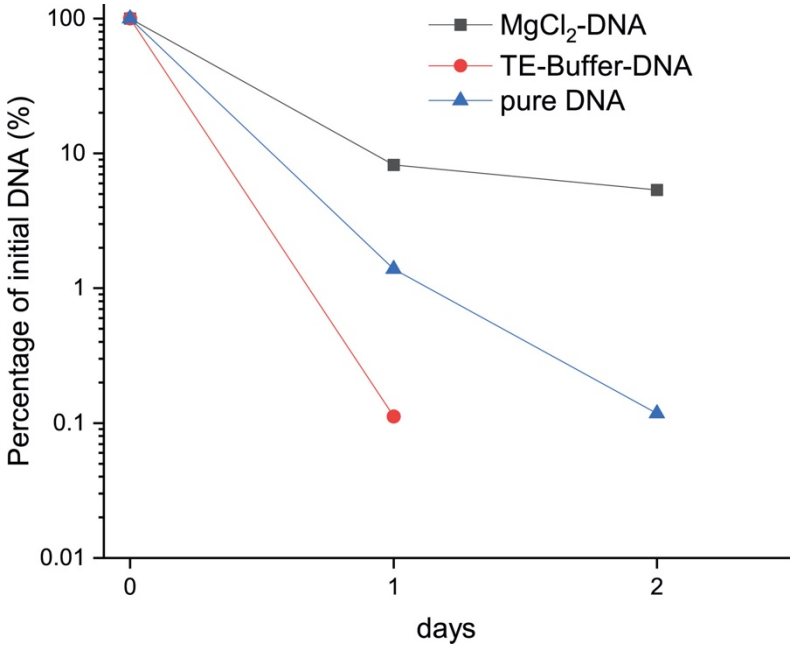

Figure S13. *DNA stability of genomic salmon DNA measured by qPCR stored in the absence of salts, dried from TE buffer and in the presence of $MgCl_2$. Similarly, to the data on the stability of short DNA fragments (see Fig S9), the presence of $MgCl_2$ significantly reduces DNA decay.*

1 L. Organick, S. D. Ang, Y.-J. Chen, R. Lopez, S. Yekhanin, K. Makarychev, M. Z. Racz, G. Kamath, P. Gopalan, B. Nguyen, C. N. Takahashi, S. Newman, H.-Y. Parker, C. Rashtchian, K. Stewart, G. Gupta, R. Carlson, J. Mulligan, D. Carmean, G. Seelig, L. Ceze and K. Strauss, *Nat. Biotechnol.*, 2018, **36**, 242–249.

2 M. E. Allentoft, M. Collins, D. Harker, J. Haile, C. L. Oskam, M. L. Hale, P. F. Campos, J. A. Samaniego, M. Thomas, P. Gilbert, E. Willerslev, G. Zhang, R. P. Scofield, R. N. Holdaway and M. Bunce, *Proc. R. Soc. B*, 2012, **279**, 4724–4733.

3 Introduction to Bone Biology: All About our Bones | International Osteoporosis Foundation, https://www.iofbonehealth.org/introduction-bone-biology-all-about-our-bones, (accessed 19 August 2019).

4 DNA Storage and Quality - Oxford Gene Technology, https://www.ogt.com/resources/literature/403_dna_storage_and_quality, (accessed 19 August 2019).

5 R. N. Grass, R. Heckel, M. Puddu, D. Paunescu and W. J. Stark, *Angew. Chemie - Int. Ed.*, 2015, **54**, 2552–2555.

6 W. D. Chen, A. X. Kohll, B. H. Nguyen, J. Koch, R. Heckel, W. J. Stark, L. Ceze, K. Strauss and R. N. Grass, *Adv. Funct. Mater.*, 2019, 1901672.

7 D. Paunescu, R. Fuhrer and R. N. Grass, *Angew. Chemie - Int. Ed.*, 2013, **52**, 4269–4272.

8 T. Lindahl and B. Nyberg, *Biochemistry*, 1972, **11**, 3610–3618.

9 K. M. Iyer and W. S. Khan, in *General Principles of Orthopedics and Trauma*, Springer International Publishing, 2019, pp. 1–779.

10 J. F. Gillooly, A. Hein and R. Damiani, *Cold Spring Harb. Perspect. Biol.*

11 D. Paunescu, M. Puddu, J. O. B Soellner, P. R. Stoessel and R. N. Grass, *Nat. Protoc.*, 2013, **8**, 2440–2448.

12 J. Bonnet, M. Colotte, D. Coudy, V. Couallier, J. Portier, B. Morin and S. Tuffet, *Nucleic Acids Res.*, 2009, **38**, 1531–1546.