

Università degli Studi di Padova
Dipartimento di Biologia
Corso di Laurea in Biologia Molecolare



Elaborato di Laurea

**OTTIMIZZAZIONE DEI PROTOCOLLI
DI ESTRAZIONE ED AMPLIFICAZIONE DEL DNA
OTTENUTO DA RESTI ARCHEOBOTANICI
NON CARBONIZZATI**

Tutor: Prof. Alessandro Vezzi
Dipartimento di Biologia

Co-tutor: Dott. Riccardo Schiavon

Laureando: Giorgio Licciardello

Anno Accademico 2016/2017

Indice

Abstract	p.	5
1. Introduzione	»	7
1.1. La difficoltà di estrarre il DNA dai resti archeobotanici non carbonizzati	»	7
1.2. Il sequenziamento high-throughput e la concentrazione del DNA estratto	»	7
1.3. La scelta della DNA polimerasi	»	8
2. Materiali e metodi	»	9
2.1. La scelta dei criteri di valutazione dei metodi di estrazione	»	9
2.2. I metodi di estrazione e la loro comparazione	»	10
2.3. La comparazione delle DNA polimerasi	»	13
3. Risultati	»	15
3.1. Comparazione dei metodi di estrazione	»	15
3.2. Comparazione delle DNA polimerasi	»	19
4. Conclusioni	»	22
5. Bibliografia	»	23
Appendice	»	25

Abstract

La maggior parte degli studi che riguardano l'aDNA sono stati condotti su resti archeologici animali, perché l'estrazione di DNA da questi campioni ha tendenzialmente una resa più alta. Questo studio sottolinea invece l'importanza di estendere gli studi all'aDNA vegetale, per sfruttarne il ricchissimo potenziale. Essi possono infatti fornire importanti evidenze nell'ambito della comprensione dei processi di domesticazione delle piante, delle strategie di sussistenza umana e dell'interazione umana con l'ambiente.

L'obiettivo del lavoro è quello di valutare in che modo è possibile massimizzare l'efficienza dell'estrazione di DNA da resti archeobotanici non carbonizzati, al fine di ottenere dei campioni quanto più possibili adatti ai metodi di sequenziamento di nuova generazione.

Vengono inoltre testate cinque DNA polimerasi per determinare la loro abilità nel superare le inibizioni dei contaminanti spesso presenti in estratti di aDNA vegetale. Ciò al fine di valutare le loro proprietà nelle reazioni di PCR, la loro compatibilità con l'aDNA altamente danneggiato e la loro fedeltà nella replicazione.

Lo studio intende incoraggiare i ricercatori all'adozione delle tecniche più efficienti, per assicurare che i preziosi resti archeobotanici siano utilizzati in nella maniera ottimale.

1. Introduzione

1.1. La difficoltà di estrarre il DNA dai resti archeobotanici non carbonizzati

Gli scopi del presente lavoro sono due: comparare otto protocolli di estrazione del DNA per valutare quali siano i più adatti all'estrazione da resti archeobotanici non carbonizzati; testare l'effetto dei contaminanti frequentemente presenti in estratti di DNA antico (aDNA) vegetale sull'inibizione della reazione a catena della polimerasi (PCR) [1].

Lo studio dell'aDNA ha fornito importanti risposte a quesiti relativi a diversi ambiti della storia dell'uomo, quali la sua evoluzione o le migrazioni che lo hanno interessato [2]. Più nello specifico, la maggior parte degli studi che riguardano l'aDNA sono stati condotti su resti archeologici animali, perché l'estrazione di DNA da questi campioni ha tendenzialmente una resa più alta e risulta essere meno laboriosa rispetto ai campioni di origine vegetale. Sarebbe dunque auspicabile riuscire ad estendere gli studi all'aDNA vegetale, in quanto ciò potrebbe essere utile a chiarire in che modo l'attività agricola dell'uomo abbia determinato l'attuale distribuzione delle specie vegetali e quindi la domesticazione delle medesime.

Lo studio dell'aDNA vegetale presenta però una serie di criticità di cui tener conto. Oltre alla criticità tipica di qualsiasi ricerca sull'aDNA – l'alto grado di frammentazione e degradazione dell'acido nucleico – il DNA estratto con i metodi classici dai campioni vegetali antichi fa registrare frequentemente la presenza di contaminanti polifenolici e polisaccaridici, pigmentati e non, che inibiscono le reazioni enzimatiche come quella di amplificazione mediante reazione a catena della polimerasi (PCR). Un'altra criticità di cui si deve tener conto nell'estrazione di DNA da resti archeobotanici è la presenza di acidi umici nei sedimenti associati ai semi conservati sotto terra, che costituiscono una ulteriore fonte di contaminanti inibitori delle PCR.

1.2. Il sequenziamento high-throughput e la concentrazione del DNA estratto

In linea teorica, per amplificare con successo il DNA mediante PCR potrebbe essere sufficiente estrarre una singola molecola di DNA stampo. Tuttavia, nell'era del sequenziamento high-throughput (HTS) diventa cruciale riuscire ad ottimizzare la resa in termini di concentrazione del DNA endogeno estratto, in quanto i metodi HTS, di nuova generazione, prevedono la creazione di librerie costituite da un enorme numero di frammenti di DNA, di cui è determinata la sequenza nucleotidica. Solo in questo modo sarà possibile sfruttare tutte le potenzialità offerte dal sequenziamento HTS.

Generalmente, la concentrazione di DNA estratto può essere determinata con metodi spettrofotometrici (Nanodrop), fluorimetrici (Qubit) e mediante PCR quantitativa (qPCR), metodologia che presuppone uno studio sulle DNA polimerasi in grado di amplificare il DNA danneggiato e il DNA in presenza di inibitori.

Fermiamoci un momento a considerare più in dettaglio la qPCR. La qPCR è una tecnica che consente di determinare la concentrazione iniziale di DNA presente

nel campione, mediante una amplificazione del DNA medesimo. La tecnica si basa su una curva di amplificazione ideale dalla quale si può dedurre detta concentrazione a partire dal numero minimo di cicli necessari per rilevare un aumento di segnale di fluorescenza. Più precisamente, il metodo valuta in tempo reale la quantità di DNA in doppia elica presente nella miscela di reazione dopo ciascun ciclo di amplificazione. Per far ciò si aggiunge all'interno della miscela un intercalante del DNA, fluorescente solo quando è intercalato alla doppia elica. A partire dalla fluorescenza misurata ad ogni ciclo è possibile tracciare una curva di amplificazione di andamento sigmoidale. Il tratto centrale della curva ha un andamento che con buona approssimazione può essere considerato lineare.

Il ciclo in cui avviene l'ingresso in fase lineare è definito "ciclo soglia" e questo in linea teorica dovrebbe essere raggiunto prima dai campioni aventi una concentrazione di DNA maggiore. Così, per esempio, se un estratto di DNA a concentrazione x risulta avere come ciclo soglia il ciclo n , un estratto di DNA a concentrazione $2x$ avrà come ciclo soglia il ciclo $n - 1$.

La curva reale può mostrare delle deviazioni da quella standard di amplificazione, essendo influenzata dalla presenza di additivi, di inibitori e/o di contaminanti. Pertanto, bisogna tenere conto di come queste variabili influiscano sulla determinazione del ciclo-soglia, al fine di poter quantificare correttamente la concentrazione del DNA nella soluzione iniziale, senza essere tratti in inganno dalle diverse cinetiche di reazione.

1.3. La scelta della DNA polimerasi

Per molte applicazioni, dopo l'estrazione di DNA può essere necessaria una amplificazione tramite PCR. Detta amplificazione prevede la scelta di un enzima replicativo tra DNA polimerasi con proprietà differenti, da scegliersi in funzione dell'esperimento progettato. Il DNA antico è infatti altamente degradato, non soltanto per la frammentazione in molecole di qualche centinaio di paia di basi (pb), ma anche per la deamminazione della citosina (C), fenomeno che può rappresentare un problema quando si decide di amplificare un DNA stampo così complesso.

La deamminazione della citosina, infatti, genera mutazioni all'interno del DNA. Ciò in quanto la citosina è una base azotata normalmente appaiata alla guanina (G) in un DNA a doppia elica, ma se deamminata si trasforma in una molecola di uracile (U). L'uracile a sua volta è una base azotata non canonica – ovvero normalmente non presente nel DNA – che ha la proprietà di appaiare un'adenina (A). Ciò significa che quando la reazione di polimerizzazione va avanti, l'adenina appaia in maniera del tutto canonica una timina (T), con il risultato finale che il processo di mutazione avrà generato una transizione C - T.

La frequenza delle transizioni C - T è funzione del tipo di DNA polimerasi usata. Ciò in quanto una polimerasi che incontra l'uracile sul filamento stampo può comportarsi in due diversi modi. Infatti, le polimerasi ad alta fedeltà si bloccano non consentendo l'ulteriore amplificazione di quel frammento e non introducendo alcuna transizione, mentre altre polimerasi prone all'errore sono in grado di proseguire la sintesi a partire dal filamento stampo, introducendo un errore.

La caratterizzazione della fedeltà di una DNA polimerasi diventa particolarmente rilevante prima di un sequenziamento HTS. Questo metodo prevede una fase di amplificazione in cui è opportuno ridurre al minimo la probabilità di introdurre nucleotidi errati nelle nuove molecole sintetizzate. Ciò in quanto sequenze in cui sono stati introdotti degli errori in fase di polimerizzazione potrebbero allinearsi in posizioni del genoma che differiscono da quelle corrette di pochi nucleotidi. Per questo motivo è opportuno l'utilizzo di una DNA polimerasi ad alta fedeltà.

Al contrario, se il DNA estratto è particolarmente degradato e deve essere semplicemente amplificato in PCR, basterà utilizzare una polimerasi prona all'errore, che presenta il vantaggio di amplificare il DNA con una maggiore probabilità di successo. Questo comunque determinerà un aumento di frequenza delle mutazioni puntiformi.

2. Materiali e metodi

Questa sezione è articolata in tre paragrafi. Nel primo vengono definiti alcuni criteri per stabilire l'efficienza dei metodi di estrazione. Nel secondo si descrivono gli otto protocolli di estrazione e le modalità con cui vengono messi a confronto. Nel terzo si comparano degli enzimi replicativi utilizzati nelle reazioni di PCR e qPCR, caratterizzati per le loro proprietà e la loro compatibilità con substrati altamente degradati e contaminati.

2.1. La scelta dei criteri di valutazione dei metodi di estrazione

Come detto nell'introduzione, uno degli obiettivi di questo studio è valutare l'efficienza di otto protocolli di estrazione. La suddetta efficienza viene comparata sulla base di tre parametri: la purezza del DNA estratto, la sua concentrazione e il successo nell'amplificarlo.

Prendiamo in considerazione il primo criterio: la purezza del DNA estratto. Lo spettrofotometro Nanodrop 1000 è uno strumento in grado di valutare la purezza della soluzione contenente DNA. Il DNA ha un picco di assorbanza (A) ad una lunghezza d'onda di 260 nm. Esistono dei contaminanti, come le proteine e i fenoli, che hanno un picco di assorbanza a 280 nm. I carboidrati complessi, invece, hanno un picco di assorbanza a 230 nm. Dunque è possibile determinare la natura molecolare dei contaminanti presenti in soluzione, misurando i tre valori di assorbanza. In generale, una soluzione di DNA si considera pura se:

$$\frac{A_{260}}{A_{280}} > 1.8 \quad \text{e} \quad \frac{A_{260}}{A_{230}} > 2$$

Passiamo adesso al secondo criterio: la concentrazione di DNA. Il Qubit 1.0 (Thermo Fisher) è uno strumento che consente di misurare la concentrazione di DNA sulla base di un colorante fluorescente, che si intercala nella doppia elica

quando essa è presente – escludendo tanto i nucleotidi liberi in soluzione quanto il DNA in singola elica. Tuttavia, la lettura del Qubit può essere alterata dalla presenza di altri contaminanti, per cui in alcuni casi è utile far seguire la misurazione della concentrazione mediante Qubit con una qPCR.

L'ultimo criterio considerato è il successo di amplificazione. La scelta di amplificare il locus *rbcL* non è casuale, visto che questo gene mappa sul genoma di plastidi, presenti in copia multipla all'interno del citoplasma delle cellule vegetali. Ciò è particolarmente significativo in quanto, come detto, l'aDNA è molto degradato e la presenza di un numero elevato di copie garantisce una maggiore probabilità di successo dell'amplificazione. Inoltre, anche se la resa dell'estrazione del DNA dovesse risultare bassa in termini di concentrazione, è comunque probabile che quel dato locus sia particolarmente rappresentato all'interno dell'estratto. Il locus *rbcL* codifica la subunità maggiore della ribulosio-bisfosfato carbossilasi ossidasi (RuBisCO) ed è sottoposto a forte pressione evolutiva, caratteristica che lo rende un marcatore universale per le piante [3].

2.2. I metodi di estrazione e la loro comparazione

Avendo stabilito nel paragrafo 2.1. i criteri utili per comparare i metodi di estrazione, vediamo in dettaglio come si procede per determinare il più performante.

Prima fase: metodi raffrontati e punti di debolezza dell'approccio

Si comparano fra loro cinque metodi, promuovendo alla fase successiva il più efficiente. I metodi di cui trattasi sono i seguenti:

Metodo Epicentre – Si utilizza un kit commerciale di estrazione. I semi vengono puliti e polverizzati e sottoposti a due fasi di incubazione: una a 65 °C per 6 minuti, l'altra a 98 °C per 2 minuti. Si tratta quindi di un metodo veloce per estrarre il DNA delle piante moderne, richiedendo meno di 10 minuti.

Metodo Finnzymes – I semi vengono puliti e polverizzati. L'estrazione avviene grazie al kit commerciale Phire Plant Direct PCR (Finnzymes). Anche questo metodo prevede dei passaggi di incubazione della durata massima di una decina di minuti, consentendo l'estrazione e l'amplificazione di DNA da materiale vegetale moderno.

Metodo Gilbert *et al.* – Prevede una estrazione con tampone Tris-HCl a pH 8.0 e l'utilizzo di detergenti ionici (SDS), chelanti dei cationi divalenti (EDTA) e proteasi (proteinasasi K) per disgregare le membrane e rompere le pareti cellulari da utilizzare su materiale archeobotanico pulito e polverizzato. È un metodo più lungo e laborioso rispetto ai precedenti, perché prevede una incubazione *overnight* a 55 °C ed una estrazione fenolo-cloroformio seguita da una cromatografia su colonne MinElute PCR purification kit (Qiagen) [4].

Metodo Japelaghi *et al.* – Metodo lungo, messo a punto per le piante moderne ricche di tannini (come la vite), che prevede la preparazione di un tampone d'estrazione con Tris-HCl a pH 8.0, un chelante dei cationi divalenti (EDTA), 2-mercaptoetanolo, polivinilpirrolidone (PVP), bromuro di cetil-trimetilammonio (CTAB). Fatto ciò viene effettuata una doppia estrazione fenolo-cloroformio ed una cromatografia su colonna MinElute PCR purification KIT (Qiagen) [5].

Metodo MO BIO – Metodo rapido e commerciale messo a punto per purificare gli estratti di materiale proveniente da suoli ricchi di acidi umici. Prevede la polverizzazione dei semi e il lavaggio in sei diverse soluzioni MO BIO, seguiti da una eluizione finale del DNA da colonna. Fra le sei soluzioni, la C2 e la C3 hanno la funzione di purificare gli estratti dalle contaminazioni di acidi umici. Per questo motivo verranno utilizzate nella seconda fase per testare il loro effetto sulla resa dell'estrazione, come reagenti aggiuntivi non indispensabili di altre metodiche di estrazione.

L'estrazione del DNA con i cinque metodi sopra descritti viene condotta su semi di *Vitis vinifera* – prelevati da contesti diversi (Areni, Argenta, Ferrara, Ferrara, Lugo, Parma e San Marino) – così denominati: ARE-A, ARG, CPR-A, VAD-A, LUG, PAR-A, SAM. Per i semi delle singole località si sarebbe voluto testare ciascun metodo in duplicato, conducendo due ripetizioni dell'estrazione (ripetizione A e ripetizione B). Tuttavia, a causa del basso numero di semi disponibili, ciò non è stato possibile. Pertanto, i protocolli Gilbert *et al.*, Japelaghi *et al.* e MO BIO sono stati ripetuti due volte su semi diversi, mentre per i protocolli Epicentre e Finnzymes si è prima utilizzato l'estratto puro (ripetizione A), poi l'estratto diluito al 10% in volume (ripetizione B) di DNA dello stesso seme.

Il punto di debolezza di questa metodologia di indagine è l'utilizzo di un solo seme per ciascuna estrazione. Ciò potrebbe costituire un elemento di grande variabilità per il singolo esperimento, in quanto non è affatto detto che il grado di conservazione dei reperti archeologici sia lo stesso in tutti i casi. Per questo motivo, le estrazioni delle fasi successive vengono condotte su più collezioni di materiale archeobotanico polverizzato.

Seconda fase: metodi raffrontati e utilizzo delle soluzioni C2 e C3 MO BIO

Nella seconda fase, il metodo risultato più performante tra questi cinque – il metodo Gilbert – viene comparato con due nuovi metodi: il metodo Palmer e il metodo Rohland e Hofreiter. A seguito della nuova comparazione, i due metodi più performanti vengono promossi all'ultima fase. Qui, in fase due, oltre all'amplificazione del locus *rbcL* si procede anche al sequenziamento del suo amplicone con il metodo Sanger.

In sintesi, ecco i capisaldi nuovi dei metodi di estrazione:

Metodo Palmer – Utilizza un tampone d'estrazione contenente Tris-HCl pH 8.0, chelanti dei cationi divalenti (EDTA) e CTAB. Seguono due estrazioni fenolo-cloroformio-isoammilalcol. Concentrazione e purificazione avvengono attraverso

un Ultra filter (30 kDa) Amicon (Millipore), a seguito di ripetute centrifugazioni. Il volume raccolto viene ulteriormente purificato col kit DNeasy (Qiagen) [6].

Metodo Rohland e Hofreiter – Prevede la realizzazione di matrici di silice, utilizzando diossido di silicio, acqua e HCl 37%. I campioni, puliti e polverizzati, sono messi in un tampone di digestione contenente Tris-HCl a pH 8.0, un detergente ionico come il sodio dodecilsolfato al 2% p/V, un chelante dei cationi divalenti (EDTA), ditioneitol (DTT) e proteinasi K. Il giorno seguente i campioni vengono centrifugati a 600 g per 5 minuti ed il surnatante viene trasferito. Ad ogni campione vengono aggiunti 50 microlitri del pellet di silice, 20 mL di buffer di binding Guanidinio tiocianato, Tris, EDTA, N-lauril sarcosina ed HCl. Il surnatante viene scartato e il pellet è risospeso in una soluzione creata mescolando 50 mL di tampone di binding con 250 microlitri di HCl 37%. I campioni vengono centrifugati a 4000 g ed il pellet viene risospeso in 1 mL di etanolo e nuovamente centrifugato a 4000 g. Il surnatante viene rimosso ed il pellet essiccato. Vengono aggiunti 60 microlitri di tampone EB (Qiagen), si mescola tramite Vortex e si incuba a 37 °C per 15 minuti. Si centrifuga nuovamente, dopodiché il pellet viene scartato ed il surnatante, contenente il DNA trasferito in un nuovo tubino [7].

Come già anticipato, in questa fase dell'esperimento si estrae il DNA da collezioni di materiale polverizzato ottenuto da più tessuti vegetali, conservati in condizioni e contesti omogenei (Areni, Modena, Modena e Ferrara). I quattro reperti archeologici utilizzati sono: ARE-B (cotone), CAS (corniolo), SAF (vite), VAD-B (olivo). I metodi di Gilbert, Palmer e Rohland vengono testati tre volte per verificare gli effetti dell'utilizzo delle soluzioni C2 e C3 MO BIO. Più esattamente, la prima ripetizione non prevede l'aggiunta di queste soluzioni; la seconda prevede l'aggiunta di C2 e C3 MO BIO dopo l'incubazione *overnight* nel tampone di digestione; la terza ripetizione prevede l'utilizzo delle soluzioni C2 e C3 MO BIO dopo l'estrazione del DNA.

Entrando ancor più nello specifico, le soluzioni C2 e C3 MO BIO sono dei reagenti progettati allo scopo di far precipitare gli acidi umici ed aumentare la purezza del campione. Ma se è vero che l'esperimento intende verificare l'effetto dell'utilizzo delle soluzioni C2 e C3 MO BIO come additivi dei metodi presi in considerazione in fase due per migliorare la purezza del DNA estratto, è pur vero che un incremento della purezza potrebbe determinare una riduzione significativa della concentrazione del DNA. Pertanto, sarà anche necessario valutare questo potenziale aspetto negativo nell'utilizzo di tali soluzioni.

Terza fase: metodi raffrontati e utilizzo della BSA

Nella terza fase, infine, i due metodi promossi dalla fase precedente vengono comparati con un metodo di estrazione in via di messa a punto. Questo metodo – ideato da Andersen per l'estrazione di aDNA da sedimenti ricchi in acidi umici – potrebbe non essere ancora pienamente ottimizzato. Anche per questo metodo occorre in sintesi i capisaldi:

Metodo Andersen – Metodo sviluppato per estrarre aDNA da sedimenti. Alcune modifiche sono ancora in fase di messa a punto e potrebbero portare ad una maggiore efficienza in termini di purezza, concentrazione e successo di amplificazione. Prima dell'estrazione viene preparato un tampone di lisi contenente Sodio N-lauril sarcosina, Tris-HCl pH 8.0, ed un chelante dei cationi divalenti (EDTA). Prima dell'estrazione vengono aggiunti DTT e 2-mercaptoetanolio. I resti vengono polverizzati e posizionati in una glass bead tube PowerLyzer MO BIO, con tampone di lisi e proteinasi K. Il tampone viene mescolato tramite Vortex per un minuto, per polverizzare ulteriormente i resti. Dopodiché viene aggiunta altra proteinasi K. Segue un'incubazione *overnight* a 55 °C. Il giorno seguente i campioni vengono portati a temperatura ambiente e centrifugati a 10000 g per 3 minuti. Il surnatante viene aggiunto alla soluzione C2 MO BIO. Segue una incubazione per 5 minuti a 4 °C ed una centrifugazione a 10000 g per altri 5 minuti. La procedura è ripetuta utilizzando la soluzione C3 MO BIO. Viene poi aggiunto un ugual volume di fenolo. I campioni vengono centrifugati delicatamente per 5 minuti, poi segue una fase di centrifugazione a massima velocità della durata di altri 5 minuti. La fase acquosa viene rimossa accuratamente, senza toccare la fase ricca in proteine. Viene nuovamente aggiunto un ugual volume di fenolo e si ripetono le stesse operazioni per migliorare la resa. La fase acquosa viene rimossa ed aggiunta ad un ugual volume di cloroformio. Una centrifugazione consente di separare le due fasi; il surnatante viene trasferito su una colonnina Amicon Ultra filter (30 kDa) (Millipore) per la concentrazione e la purificazione. La soluzione è lavata con tampone EB, passaggio che è ripetuto due volte. Il DNA viene eluito con acqua.

L'amplificazione del DNA estratto con ciascun metodo viene condotta due volte: una con l'albumina di siero bovino (BSA) in concentrazione finale pari a 0.8 mg/mL, una senza. Sempre in questa fase viene effettuato un ulteriore controllo per determinare l'esatta quantità di DNA estratto. Ciò in quanto, come accennato, la lettura al Qubit potrebbe essere alterata da alcuni contaminanti presenti in soluzione. Ed è sempre per questo motivo che si effettua anche una qPCR per il locus *rbcl*.

2.3. La comparazione delle DNA polimerasi

L'amplificazione del DNA estratto avviene tramite PCR, reazione di amplificazione che prevede l'utilizzo della DNA polimerasi. Esistono diversi tipi di DNA polimerasi, ingegnerizzate per obiettivi diversi. Alcune di queste hanno la caratteristica di possedere un'alta fedeltà, ovverosia hanno una probabilità molto bassa di introdurre una mutazione; altre, prone all'errore, riescono ad amplificare stampi più degradati, con lo svantaggio di introdurre mutazioni.

Le polimerasi sono scelte sulla base delle loro caratteristiche e di precedenti utilizzi su aDNA. Di seguito, in Tabella 1, una lista dei cinque enzimi utilizzati e delle rispettive caratteristiche:

DNA polimerasi	Marca	Caratteristica
AmpliTaq Gold	Applied Biosystems	Precedentemente utilizzata in ricerca sull'aDNA
Omni Klentaq	DNA Polymerase Technology	Progettata per superare molte fonti di inibizione
Pfu Turbo C_x Hotstart	Agilent Technologies	Si arresta sugli uracili mantenendo l'alta fedeltà
Phire Hot Start II	Finnzymes (Thermo Scientific)	Progettata per superare le inibizioni ed alta processività
Phusion Hot Start	Finnzymes (Thermo Scientific)	Progettata per alta fedeltà ed alta processività

Tabella 1 - DNA polimerasi testate

La fedeltà delle polimerasi e la loro compatibilità col DNA altamente danneggiato viene studiata tramite il *deep sequencing*. Utilizzando un marcatore endogeno *rbcL*, si effettuano delle amplificazioni di tre campioni: LUG, VAD-A e VAD-B. Dopodiché si procede con una fase di sequenziamento degli ampliconi tramite HTS, utilizzando il sequenziatore 454 (Roche). Poi le reads, ovverosia le sequenze ottenute, vengono allineate e sovrapposte ad una sequenza attesa tramite il software bioinformatico Geneious Pro, per valutare gli errori di incorporazione dei nucleotidi e la presenza di eventuali *indels* (*insertions/deletions*) e la loro estensione.

Per testare l'effetto inibitorio dei contaminanti degli eluati ottenuti dall'estrazione dell'aDNA, si scelgono invece due campioni con estratti pigmentati, contenenti in soluzione del DNA endogeno e dei contaminanti. Quindi verranno effettuate delle diluizioni scalari di questa soluzione, ciascuna con una concentrazione di inibitore variabile dallo 0 al 40% del volume di reazione. Per eliminare la variabilità legata all'efficienza di estrazione da ciascun campione, l'amplificazione non sarà condotta sul DNA endogeno delle piante, ma si aggiungerà alla soluzione del DNA esogeno di tigre (*Panthera tigris*), a concentrazioni note. In parallelo si testeranno anche degli additivi della PCR, reagenti non fondamentali in una reazione classica di PCR, ma che vengono utilizzati generalmente in contesti in cui la PCR potrebbe avere un esito incerto. Anche in questo caso ciascun esperimento è ripetuto due volte: con e senza BSA, in concentrazione finale pari a 0.8 mg/mL.

Le polimerasi AmpliTaq Gold, Omni Klentaq e Pfu Turbo C_x Hotstart sono state testate anche per l'utilizzo in qPCR. Stavolta l'obiettivo è stato quello di determinare l'effetto degli eventuali inibitori presenti negli estratti di DNA sulle curve di amplificazione di diverse tipologie di DNA polimerasi, nonché quello di determinare gli effetti dell'additivo BSA sulle medesime curve. Anche in questo caso si sono utilizzate concentrazioni crescenti di inibitore.

3. Risultati

3.1. Comparazione dei metodi di estrazione

I cinque metodi della prima fase performano in maniera piuttosto diversa per quanto riguarda il successo di amplificazione in PCR. Il DNA è stato fatto correre in gel elettroforesi al 2% di agarosio, per consentirne la visualizzazione. Di seguito, in Tabella 2, viene riportato l'esito dell'amplificazione in PCR sul locus plastidiale *rbcL*:

Campione	Replica	Finnzymes	Epicentre	Gilbert	Japelaghi	MO BIO
ARE-A	A	-	-	-	-	-
	B	-	-	-	-	-
ARG	A	-	-	(+)	-	-
	B	-	-	+	-	-
CPR-A	A	+	(+)	+	-	+
	B	-	-	+	(+)	-
LUG	A	-	+	+	+	+
	B	-	+	+	+	+
PAR-A	A	-	-	+	+	-
	B	-	(+)	+	-	-
SAM	A	-	-	-	(+)	-
	B	-	-	-	-	-
VAD-A	A	(+)	+	(+)	+	(+)
	B	-	(+)	+	+	-
Successi		2/14	6/14	10/14	7/14	4/14

Dove: + : presenza di una banda netta, distinta.
 (+) : presenza di una banda tenue.
 - : assenza di bande.

Tabella 2 - Esito dell'amplificazione in PCR

Il campione ARE-A non mostra alcun successo di amplificazione. Ciò potrebbe essere dovuto alla cattiva riuscita di tutti i metodi di estrazione oppure, più probabilmente, all'elevato grado di frammentazione del DNA. Il campione, infatti, potrebbe essere stato conservato in condizioni tali da favorire la degradazione del DNA in frammenti più piccoli delle 138 pb coperte dal locus *rbcL*.

Confrontando fra loro i metodi, si rileva una differenza statisticamente significativa in termini di successo di amplificazione del DNA estratto fra il metodo Gilbert e i metodi MO BIO e Finnzymes. Le differenze fra i campioni Gilbert, Japelaghi e Epicentre, invece, non sono statisticamente significative. Tuttavia, per il successo di amplificazione il metodo ritenuto più performante è il metodo Gilbert, perché le bande della gel elettroforesi sono più distinte, nitide ed intense. Si osserva altresì che la frequenza di successo è molto bassa nel caso di metodi rapidi, basati sull'estrazione del DNA con soluzioni commerciali (come MO BIO e Finnzymes). Tuttavia, il metodo Epicentre mostra una buona probabilità di successo nonostante la sua rapidità di esecuzione.

La Figura 1 mostra la concentrazione del DNA estratto (a sinistra) e la sua purezza (a destra):

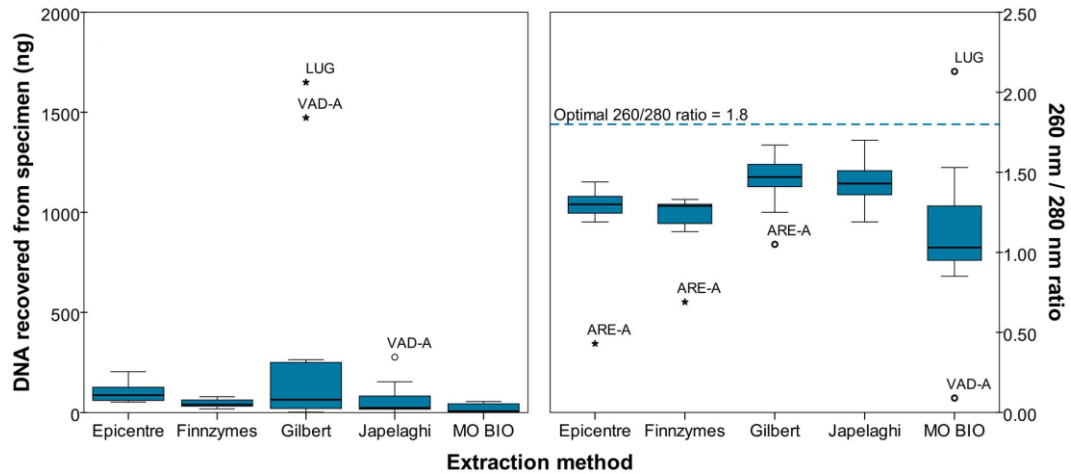


Figura 1 - Purezza e concentrazione del DNA estratto nella prima fase

Gli eluati ottenuti con i metodi Epicentre e Finnzymes fanno registrare una resa piuttosto bassa. Ma visto che le soluzioni ottenute con questi metodi sono pigmentate, si può supporre che le molecole di pigmento possano in qualche modo interferire col legame della sostanza fluorescente che intercala la doppia elica, rilevando una concentrazione minore di quella reale. Come mostrato in Figura 1, la quantità di DNA estratto con il metodo Gilbert è quasi tripla rispetto al metodo Epicentre, classificatosi al secondo posto. Le differenze di concentrazione del DNA estratto sono risultate statisticamente significative sia in termini di comparazione fra metodi, sia in termini di comparazione fra campioni.

Per quanto riguarda la purezza, infine, nessuno dei metodi testati ha raggiunto il livello soglia del rapporto $\frac{A_{260}}{A_{280}} > 1.8$, tale da poter considerare il campione sufficientemente puro. Il metodo che più si è avvicinato a tale soglia è ancora una volta il metodo Gilbert, che però non mostra differenze statisticamente significative rispetto ai metodi MO BIO e Japelaghi.

Per le considerazioni fatte, il metodo promosso alla seconda fase è stato il metodo Gilbert. Si ricordi che in questa nuova fase l'obiettivo era quello di paragonare questo metodo ai metodi Palmer e Rohland, verificando se l'aggiunta delle soluzioni C2 e C3 MO BIO potesse aumentare significativamente la resa, senza ridurre eccessivamente la probabilità di successo dell'amplificazione o la concentrazione del DNA.

La Tabella 3 mostra il successo di amplificazione dei tre metodi testati in questa fase ed il rapporto fra le assorbanze a 260 nm e a 280 nm:

Campione	Aggiunta di C2/C3	Gilbert		Palmer		Rohland	
		Amp	A ₂₆₀ /A ₂₈₀	Amp	A ₂₆₀ /A ₂₈₀	Amp	A ₂₆₀ /A ₂₈₀
ARE-B	Prima	-	1.40	-	1.25	-	2.05
	Nessuna	-	1.37	-	1.33	-	1.99
	Dopo	-	1.35	-	1.22	-	2.07
CAS	Prima	-	1.43	-	1.39	(+)	2.67
	Nessuna	-	1.26	-	1.39	(+)	4.87
	Dopo	-	1.17	-	1.23	(+)	1.43
SAF	Prima	+	1.44	-	1	+	2.55
	Nessuna	+	1.50	+	1.57	+	2.74
	Dopo	-	1.25	-	1.06	(+)	3.22
VAD-B	Prima	+	1.72	-	1.17	+	1.47
	Nessuna	+	1.73	+	1.77	+	1.1
	Dopo	+	1.45	-	1.18	-	1.13

Dove: + : presenza di una banda netta, distinta.
 (+) : presenza di una banda tenue.
 - : assenza di bande.

Tabella 3 - Successo di amplificazione e rapporto fra assorbanze

L'amplificazione del DNA non ha avuto successo nel caso del campione ARE-B, indipendentemente dal metodo utilizzato per estrarlo. Tutti i metodi hanno invece amplificato i campioni SAF e VAD-B, mentre il campione CAS mostra una banda tenue solo quando amplificato con il metodo Rohland. Dall'analisi di sequenziamento dell'amplicone, risulta che i campioni SAF e VAD-B hanno una differenza inferiore di 2 pb rispetto alla sequenza attesa, tasso d'errore compatibile con l'amplificazione di aDNA. Il campione CAS, invece, differisce per molte posizioni con la sequenza attesa: per questo motivo si suppone che il DNA presente sia un contaminante. Come mostrato in Figura 2, l'aggiunta delle soluzioni C2 e C3 MO BIO riduce quasi sempre le concentrazioni di DNA degli eluati; ma d'altra parte, l'effetto vantaggioso in termini di aumento della purezza del campione non è di entità tale da mostrare un vantaggio effettivo nell'utilizzo delle due soluzioni:

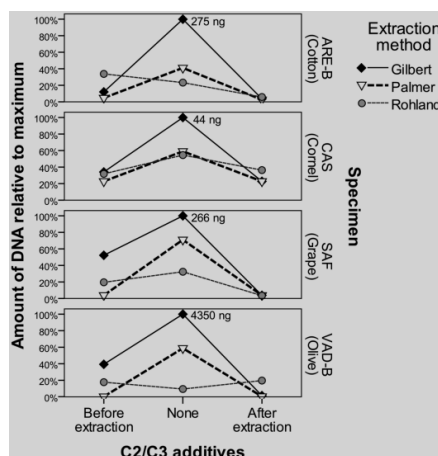


Figura 2 - Effetti degli additivi C2 e C3 sulla concentrazione di DNA estratto

Per questo motivo ci si è concentrati sui metodi non modificati, Gilbert e Palmer, promossi in terza fase. Il protocollo Gilbert, in particolare, è risultato essere più performante del metodo Rohland e Hofreiter, sia in termini di concentrazione del DNA, sia in termini di purezza dell'estratto; ma non ha presentato alcuna differenza significativa con il metodo Palmer.

Nella terza fase si sono paragonati i metodi Gilbert e Palmer al metodo Andersen, il cui protocollo è in via di messa a punto ed ha già fornito dei risultati molto promettenti. Qui viene testato anche l'effetto dell'additivo BSA sul successo di amplificazione del DNA, conducendo le PCR con e senza BSA. Questa differenza è risultata cruciale, dal momento che l'assenza di BSA ha determinato il fallimento di quasi tutte le reazioni. Al contrario, con l'aggiunta di BSA i metodi hanno performato in maniera simile. Più in dettaglio, i metodi di Andersen e Palmer sono stati in grado di amplificare sei campioni, il metodo Gilbert ne ha amplificati sette. Il numero di copie di *rbcL* estratto è variato in maniera considerevole sia al variare dei metodi, sia al variare dei campioni. I dati relativi alla concentrazione del DNA estratto ed al numero di copie di *rbcL* sono riassunti in Figura 3:

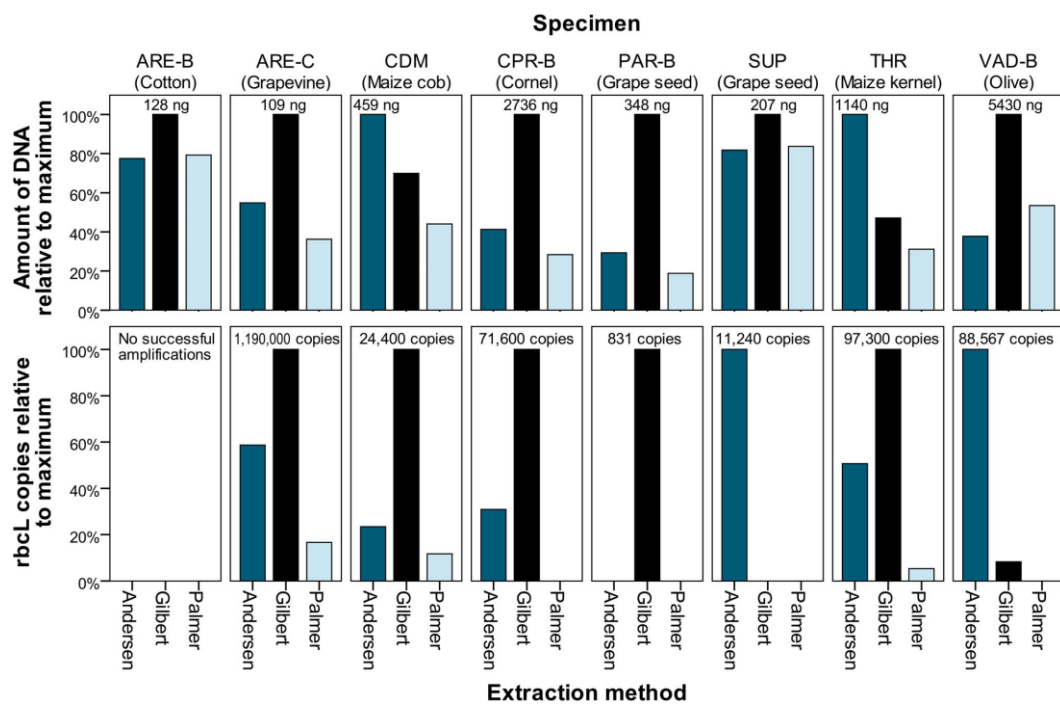


Figura 3 - Concentrazione del dna estratto e numero di copie di *rbcL*

Come detto, la concentrazione di DNA è quantificabile tanto utilizzando il Qubit, quanto utilizzando la qPCR. Poiché i risultati sono stati lievemente discordanti, si è ipotizzata la presenza di contaminanti che interferiscono con la suddetta misurazione. I livelli di purezza del DNA sono statisticamente identici per ciascun metodo, mentre in termini di concentrazione del DNA i metodi Andersen e Gilbert – pur non mostrando alcuna differenza statisticamente significativa tra loro – hanno fatto registrare una resa maggiore rispetto al metodo Palmer.

Per concludere, al termine di questo studio i metodi Andersen e Gilbert possono essere considerati come i metodi più efficienti nell'estrazione di aDNA da resti archeobotanici.

3.2. Comparazione delle DNA polimerasi

Le cinque polimerasi sono state classificate in base alla loro abilità nel superare le inibizioni, alla loro compatibilità con la qPCR ed alla loro fedeltà nella amplificazione.

Il superamento degli effetti degli inibitori della PCR di cui si è discusso al paragrafo 2.3. è stato testato estraendo DNA da due campioni: SPC e ARE-B. L'estrazione di DNA da questi campioni – prelevati rispettivamente a Modena ed Areni – ha prodotto delle soluzioni contenenti un inibitore pigmentato. Sono state quindi allestite delle diluizioni scalari di tale soluzione, ottenendo così delle nuove soluzioni contenenti percentuali diverse di inibitore. Ad ognuna di queste diluizioni si è aggiunto del DNA esogeno a concentrazione nota e costante, amplificato con una PCR. La Tabella 4 mostra il successo di amplificazione di tale DNA in concentrazioni crescenti di inibitore, in una semplice reazione di PCR senza additivi ed in una reazione con BSA 0.8 mg/mL:

Polimerasi							
	Soluzione inibente	% inibitore	AmpliTaq Gold	Omni Klentaq	PfuTurbo C _x	Phire	Phusion
No BSA	SPC	0%	+	+	+	+	+
		0.1%	-	+	-	+	-
		≥ 1%	-	-	-	-	-
	ARE-B	0%	+	+	+	+	+
		0.1%	-	+	-	+	-
		1%	-	(+)	-	(+)	-
BSA	SPC	≥ 2.5%	-	-	-	-	-
		0%	+	+	+	+	+
		0.1%	+	+	+	+	+
		1%	+	+	+	+	+
		2.5%	+	+	-	+	-
		5%	(+)	+	-	(+)	-
	ARE-B	10%	-	+	-	-	-
		≥ 20 %	-	-	-	-	-
		0%	+	+	+	+	+
		0.1%	+	+	+	+	+
		1%	+	+	+	+	+
		2.5%	+	+	(+)	+	(+)
		5%	-	+	-	+	-
		10%	-	+	-	-	-
		≥ 20 %	-	-	-	-	-

Dove: + : presenza di una banda netta, distinta.
 (+) : presenza di una banda tenue.
 - : assenza di bande.

Tabella 4 - Successo della PCR in presenza/assenza di BSA e concentrazioni crescenti di inibitore

Senza la BSA, soltanto le polimerasi Omni KlenTaq e Phire Hot Start II sono state in grado di amplificare il DNA esogeno in presenza di inibitori. In questi esperimenti, l'aggiunta di BSA sembra quindi essere stata cruciale perché ha consentito di migliorare la resa di ogni polimerasi, estendendo il successo di amplificazione almeno alle reazioni contenenti l'1% di inibitori. Sotto questo aspetto, le polimerasi più performanti sono dunque AmpliTaq Gold, Phire e Omni KlenTaq. Quest'ultima, in particolare, ha funzionato anche in esperimenti contenenti il 10% di sostanze inibitorie.

La fedeltà della polimerasi è stata investigata tramite sequenziamento HTS degli ampliconi *rbcL*. La maggior parte delle sequenze ottenute è risultata compatibile con quelle attese. Alcuni ampliconi sono risultati essere chimerici, quindi non allineabili con le sequenze attese. La polimerasi che ha avuto la maggiore percentuale di mutazione è proprio la Omni KlenTaq, risultata la migliore anche in termini di successo di amplificazione. Talvolta le elettroforesi su gel d'agarosio di ampliconi ottenuti con Omni KlenTaq non hanno mostrato bande nette ma piuttosto degli smear, indice di errori di replicazione. I test statistici hanno mostrato che il tasso di errore della polimerasi Phusion sono significativamente più bassi rispetto a tutte le altre polimerasi. Tutte le altre differenze non sono risultate significative. I dati sulle percentuali di mutazione sono riassunti in Figura 4:

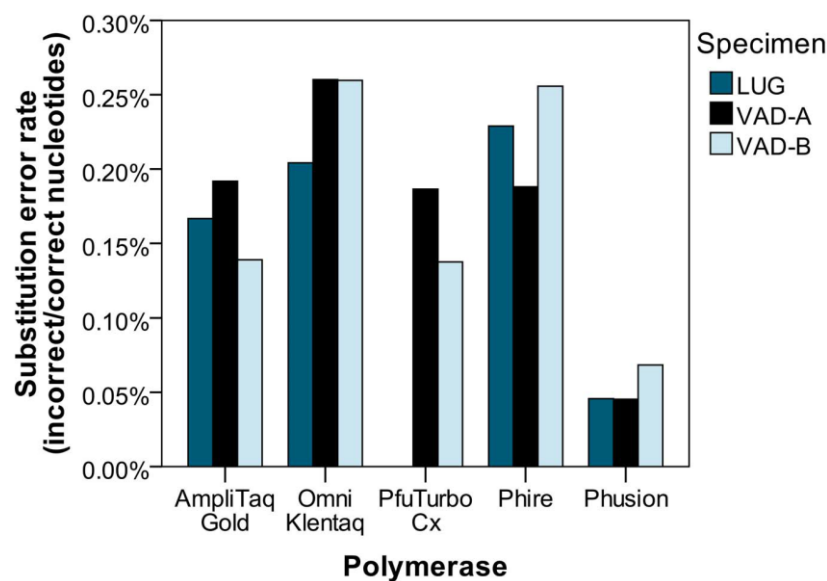


Figura 4 - Tasso di sostituzioni nucleotidiche delle DNA polimerasi

I due aspetti di fedeltà della polimerasi e di successo della reazione di amplificazione possono essere ricompresi nel concetto di compatibilità con l'aDNA. Come infatti già visto, i campioni di DNA antico sono soggetti a degradazioni quali la deaminazione della citosina a uracile. La polimerasi Phusion – risultata ottimale nelle sue caratteristiche di alta fedeltà – è pubblicizzata dal suo produttore come “non-compatibile” con l'uracile, al contrario della Pfu Turbo Cx, che introduce una transizione C - T sul filamento stampo e una G - A sul filamento complementare. Il problema dell'utilizzo di una DNA polimerasi ad alta fedeltà nell'amplificazione di aDNA è che questa, essendo incompatibile con l'uracile, andrà in

stallo se e quando incontrerà una citosina deaminata. Per questo motivo amplificherà più efficacemente una eventuale fonte di contaminante esogeno rispetto al DNA endogeno che si intende caratterizzare. Per i campioni più preziosi e difficili è stato messo a punto un protocollo [8] volto a ridurre la frequenza di mutazioni. Esso prevede due fasi: l'utilizzo di una DNA polimerasi prona all'errore come AmpliTaq Gold per i primi dieci cicli di PCR, sostituita da una DNA polimerasi ad alta fedeltà come Phusion per i cicli successivi. Esistono anche delle altre strategie [9] per trattare l'uracile, ma non riguardano l'utilizzo di DNA polimerasi per cui non sono state oggetto di questi studi.

Tre polimerasi sono state testate anche per il loro utilizzo in qPCR. L'additivo BSA ha avuto un impatto negativo sulla curva di amplificazione con AmpliTaq Gold, ma non sulle altre due polimerasi, come mostra la Figura 5:

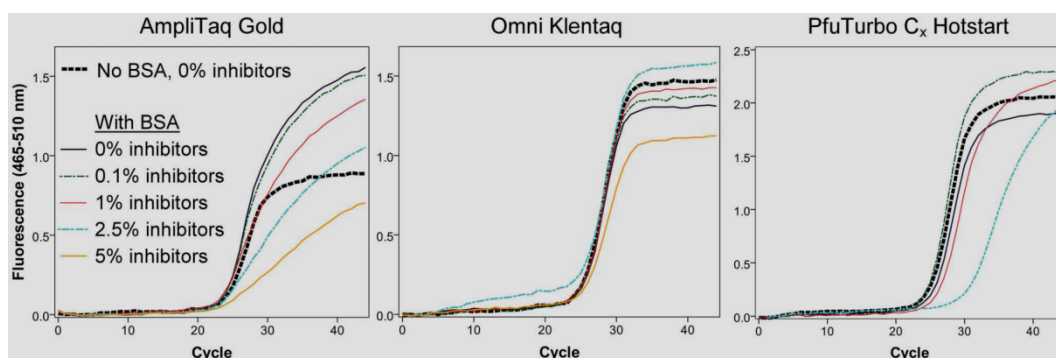


Figura 5 - Curva di amplificazione delle tre DNA polimerasi in presenza/assenza di BSA, a concentrazione crescente di inibitore

Le alte concentrazioni di inibitore hanno ridotto la pendenza della curva sia nell'amplificazione con AmpliTaq Gold, sia con PfuTurbo C_x Hotstart, ma non in quella con Omni KlenTaq.

4. Conclusioni

Gli esperimenti condotti hanno fornito una nuova prospettiva sull'estrazione e sull'amplificazione del DNA endogeno da resti archeobotanici non carbonizzati. Per beneficiare al massimo delle opportunità offerte dall'HTS è necessario massimizzare la resa dell'estrazione, e per far ciò il miglior metodo è risultato essere quello di Gilbert *et al*, che fornisce eluati con una maggior concentrazione di DNA endogeno ed una maggiore purezza.

Tuttavia, dalla terza fase di comparazione risulta evidente che per i campioni più preziosi – che andranno inevitabilmente distrutti al fine di estrarne il DNA – l'ideale è testare più di un metodo. Ed in effetti, anche il metodo di Andersen ha fornito degli ottimi risultati. In accordo con quanto ci si poteva anche aspettare, i metodi commerciali rapidi per estrarre il DNA dalle piante moderne (l'ordine di idee è quella dei dieci minuti) non sono adatti per l'estrazione del DNA da resti archeobotanici. Altra evidenza è che i metodi più performanti sono stati quelli che non richiedono l'uso dell'additivo CTAB, un reagente che serve a rimuovere i contaminanti polisaccaridici dalle piante moderne.

Il test di comparazione fra le DNA polimerasi ha fornito nuove evidenze. In primo luogo, nessuna polimerasi è universalmente valida per qualsiasi tipo di applicazione sperimentale. Bisogna quindi sceglierla tenendo in considerazione gli obiettivi da raggiungere ed i metodi impiegati. Si suggerisce l'uso della polimerasi AmpliTaq Gold con l'aggiunta di BSA per superare quantità moderate di inibitore, per le applicazioni di biologia molecolare che non prevedono un sequenziamento HTS. Una valida alternativa è la polimerasi Omni KlenTaq, che riesce a superare quantità ancora più elevate di sostanze inibitorie. Questa polimerasi è anche consigliata nel caso di esperimenti di qPCR in cui sono presenti dei contaminanti. Ciò in quanto, a differenza di AmpliTaq Gold, la suddetta polimerasi mostra una curva di amplificazione standard anche in presenza di BSA ed inibitori. Al contrario, essa non è indicata in esperimenti che prevedono il sequenziamento dell'amplicone generato, presentando il maggior tasso d'errore fra tutte le cinque polimerasi testate. Si consiglia infine l'utilizzo dell'additivo BSA in tutte le reazioni di amplificazione di DNA da resti archeobotanici, specialmente quelli particolarmente contaminati.

Per concludere: come previsto da Palmer *et al*. [10] il futuro della ricerca sull'aDNA vegetale sembra essere promettente, ma per sfruttare al meglio tutte le opportunità che possono derivare da metodi di sequenziamento di nuova generazione occorre ottimizzare i protocolli che consentono un recupero efficiente del materiale da sequenziare. Questo studio ha inteso incoraggiare i ricercatori all'adozione delle tecniche più performanti. Questo genere di ricerca costituirà dunque una linea guida per gli studiosi di archeogenetica. L'obiettivo è quello di rendere la ricerca più fruttuosa, al fine di comprendere sempre meglio argomenti quali l'evoluzione delle piante, la loro domesticazione, la loro interazione con l'uomo.

5. Bibliografia

1. Wales N., Andersen K., Cappellini E., Ávila-Arcos M. C., Gilbert M. T. P., Optimization of DNA Recovery and Amplification from Non-Carbonized Archaeobotanical Remains, PLoS ONE, 2014, 9: e86827.
2. Rasmussen M., Li Y., Lindgreen S., Pedersen J. S., Albrechtsen A., Moltke I. *et al.*, Ancient human genome sequence of an extinct Palaeo-Eskimo, Nature, 2010, 463: 757–762.
3. Poinar H. N., Hofreiter M., Spaulding W. G., Martin P. S., Stankiewicz B. A., Bland H. *et al.*, Molecular Coproscopy: Dung and Diet of the Extinct Ground Sloth *Nothrotheriops shastensis*, Science, 1998, 281: 402-406.
4. Gilbert M. T. P., Wilson A. S., Bunce M., Hansen A. J., Willerslev E., Shapiro B. *et al.*, Ancient mitochondrial DNA from hair, Curr Biol, 2004, 14: R463-R464.
5. Japelaghi R., Haddad R., Garoosi G., Rapid and Efficient Isolation of High Quality Nucleic Acids from Plant Tissues Rich in Polyphenols and Polysaccharides, Mol Biotechnol, 2011, 49: 1-9.
6. Palmer S. A., Moore J. D., Clapham A. J., Rose P., Allaby R. G., Archaeogenetic Evidence of Ancient Nubian Barley Evolution from Six to Two-Row Indicates Local Adaptation, PLoS ONE, 2009, 4: e6301.
7. Rohland N., Hofreiter M., Comparison and optimization of ancient DNA extraction, BioTechniques, 2010, 42: 343-352.
8. Green R. E., Krause J., Briggs A. W., Maricic T., Stenzel U., Kircher M. *et al.*, A Draft Sequence of the Neanderthal Genome, Science, 2010, 328: 710-722.
9. Briggs A. W., Stenzel U., Meyer M., Krause J., Kircher M. Pääbo S. *et al.*, Removal of deaminated cytosines and detection of *in vivo* methylation in ancient DNA, Nucleic Acids Res, 2010, 38: e87.
10. Palmer S. A., Smith O., Allaby R. G., The blossoming of plant archaeogenetics, Ann Anat, 2012, 20: 146-156.

Appendice

Optimization of DNA Recovery and Amplification from Non-Carbonized Archaeobotanical Remains

Nathan Wales^{1,2*}, Kenneth Andersen², Enrico Cappellini², María C. Ávila-Arcos², M. Thomas P. Gilbert^{2,3}

¹ Department of Anthropology, University of Connecticut, Storrs, Connecticut, United States of America, ² Centre for GeoGenetics, University of Copenhagen, Copenhagen, Denmark, ³ Department of Environment and Agriculture, Curtin University, Perth, Western Australia, Australia

Abstract

Ancient DNA (aDNA) recovered from archaeobotanical remains can provide key insights into many prominent archaeological research questions, including processes of domestication, past subsistence strategies, and human interactions with the environment. However, it is often difficult to isolate aDNA from ancient plant materials, and furthermore, such DNA extracts frequently contain inhibitory substances that preclude successful PCR amplification. In the age of high-throughput sequencing, this problem is even more significant because each additional endogenous aDNA molecule improves analytical resolution. Therefore, in this paper, we compare a variety of DNA extraction techniques on primarily desiccated archaeobotanical remains and identify which method consistently yields the greatest amount of purified DNA. In addition, we test five DNA polymerases to determine how well they replicate DNA extracted from non-charred ancient plant remains. Based upon the criteria of resistance to enzymatic inhibition, behavior in quantitative real-time PCR, replication fidelity, and compatibility with aDNA damage, we conclude these polymerases have nuanced properties, requiring researchers to make educated decisions as to which one to use for a given task. The experimental findings should prove useful to the aDNA and archaeological communities by guiding future research methodologies and ensuring precious archaeobotanical remains are studied in optimal ways, and may thereby yield important new perspectives on the interactions between humans and past plant communities.

Citation: Wales N, Andersen K, Cappellini E, Ávila-Arcos MC, Gilbert MTP (2014) Optimization of DNA Recovery and Amplification from Non-Carbonized Archaeobotanical Remains. PLoS ONE 9(1): e86827. doi:10.1371/journal.pone.0086827

Editor: Carles Lalueza-Fox, Institut de Biologia Evolutiva - Universitat Pompeu Fabra, Spain

Received: October 7, 2013; **Accepted:** December 14, 2013; **Published:** January 27, 2014

Copyright: © 2014 Wales et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Funding is from: the Danish Council for Independent Research grant 10-081390 (NW, EC, and MTPG), the Danish National Research Foundation (DNRF94) (MCAA and KA), and a dissertation fellowship from the American-Scandinavian Foundation (NW). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: M. Thomas P. Gilbert (Tom Gilbert) is an Editorial Board member for PLOS ONE. However, that in no way should influence the reception of the manuscript. The authors confirm that his position does not alter adherence to the PLOS ONE policies on sharing data and materials.

* E-mail: nathan.wales@snm.ku.dk

Introduction

Ancient DNA (aDNA) studies have become an integral part of Quaternary research, providing invaluable anthropological and biological insights, on issues as diverse as human evolution [1], modern human migrations [2–4], plant and animal domestication [5,6], and paleoecology [7]. Research on plant aDNA from archaeological contexts is of particular interest because archaeobotanical remains can provide important data on subsistence patterns, human behavioral variability, domestication, and broader environmental issues [8–10]. Despite this rich potential, relatively few researchers have studied aDNA from plant materials [9,11]; the scarcity of this line of research can be partially attributed to the many methodological challenges posed by ancient plant materials.

In addition to the issues of contamination and biomolecular degradation faced by all aDNA research [12], ancient plant materials frequently contain compounds that impede DNA extraction and enzymatic reactions, including the polymerase chain reaction (PCR). In modern plant materials, polysaccharides and polyphenols, such as tannins, pose significant problems for the extraction of nucleic acids [13]; these compounds may still thwart geneticists millennia after the death of a plant. In addition, archaeological plant materials are often rich in humic acids, some

of which originate from associated sediments. These darkly-pigmented compounds are often inadvertently extracted together with DNA and inhibit many DNA polymerases which are required for genetic analyses [14]. Even when DNA eluates are visually transparent, inhibitors may still be present, leading to PCR failures.

In their systematic review of aDNA techniques, Rohland and Hofreiter [15] explore numerous protocols, the use of different binding salts, incubation modifications, PCR additives, and DNA polymerases. The results of the study have been influential in the aDNA community and have been adopted by a number of researchers, including for the prominent Neanderthal genome project [1]. Nevertheless, Rohland and Hofreiter's [15] investigation focused only upon aDNA from bones, and therefore the findings may not be applicable to other aDNA source materials, including ancient plant remains. In this article, we expand upon Rohland and Hofreiter's [15] work by examining the effectiveness of various extraction techniques on non-charred archaeobotanical remains and the relative capabilities of different polymerases to amplify aDNA. Given the growing importance of high-throughput sequencing (HTS) technologies in plant aDNA research [9], issues and goals related to HTS are given special attention.

Goals for aDNA Extractions

The fundamental aim of DNA extractions of archaeobotanical remains is to isolate as much endogenous DNA from a sample as possible. Ancient samples characteristically have few copies of endogenous DNA, and these molecules are usually fragmented into segments less than a few hundred base pairs (bp) in length [16]. Optimizing aDNA recovery has become even more important in the era of HTS [17]. For conventional PCR-based studies, it is only necessary for the locus of interest to be amplified, and amplification can theoretically initiate from a single template molecule. HTS, on the other hand, require a much larger “library” of DNA molecules (that is, DNA molecules from a sample with special nucleotide adapters attached to each end). HTS platforms require libraries to be amplified to a specified starting concentration, and if DNA extract concentrations are low, more amplification cycles are required, leading to PCR drift and clonality [18,19].

While it is important to extract as much DNA from an ancient sample as possible, the DNA must also be relatively pure: clear of other cellular components like proteins and lipids that might otherwise hinder downstream analyses. For archaeobotanical remains in particular, it is vital to remove substances which impair enzymatic reactions, including humic acids and polyphenols.

Goals for Polymerases in aDNA Amplification

Ideally all traces of inhibitory substances would be removed in the course of DNA extraction; however, in some instances these substances remain, often leaving DNA eluates pigmented [20]. Such recalcitrant samples presumably contain humic acids and DNA strands of the same molecular weight, and these molecules consequently coprecipitate in purifications due to their shared anionic properties [21]. Repeated purifications using silica and other methods have been investigated [22–24], but since every additional purification step can reduce DNA yield, and because PCR inhibitors may not manifest themselves as obvious pigmentation, it is advantageous to use polymerases that tolerate residual inhibitors.

Real-time quantitative PCR (qPCR) experiments have been designed to study DNA from archaeobotanical remains [25], but there has been little research into the compatibility of different polymerases and PCR additives in qPCR. Exploratory experimentation (N. Wales, unpublished data) suggested that some polymerases do not exhibit normal amplification curves when samples are pigmented or when certain PCR additives are included in the reaction. As departures from ideal amplification curves may lead to inaccurate DNA quantification, it is important to know which polymerases yield consistent qPCR results under a broad range of conditions.

The fidelity of polymerases is an important concern, especially when aDNA libraries are amplified for HTS. Ancient samples frequently yield low levels of coverage for all loci, making it challenging to identify which genetic motif is real and which is the result of polymerase copy errors.

The degraded and damaged nature of aDNA has a profound effect on the performance of polymerases. In particular, research has identified cytosine deamination, a hydrolysis reaction in which cytosine is converted to uracil, as the main source of the problem [26–28]. The presence of uracil in aDNA molecules has adverse effects in PCR because DNA polymerases cannot add the appropriate complementary nucleotide to the opposite DNA strand. Instead, polymerases either 1) stop replicating the DNA molecule, or 2) insert adenine which is complementary to uracil in RNA. The latter scenario leads to an apparent C-to-T transition in

the template molecule [28,29]. Depending on the research goals, either of the available options may be preferable. For example, if a polymerase does not copy damaged DNA molecules, bioinformatic analyses are simplified as it can be assumed that damage is not a factor in generating sequence variation. On the other hand, if nearly all molecules are damaged, the polymerase may fail to amplify anything, thus providing no data at all. Additionally, by using a polymerase which pairs uracil with adenine, one may argue for the authenticity of aDNA based upon damage patterns [30,31]. It is therefore important to be fully aware of how a given polymerase handles damage.

Materials and Methods

The authors thank the following researchers for permission for destructive analysis of archaeobotanical remains: Boris Gasparyan, Institute of Archaeology and Ethnology, National Academy of Sciences, Yerevan, Armenia; Giovanna Bosi and Anna Maria Mercuri, Museo Di Paleobiologia e dell’Orto Botanico, Università di Modena e Reggio Emilia, Modena, Italy; Girolamo Fiorentino, Dipartimento di Beni Culturali, University of Salento, Lecce, Italy; Mike Jacobs, Arizona State Museum; and José Luis Punzo-Díaz, Instituto Nacional de Antropología e Historia, Centro INAH Michoacán, Mexico.

Comparison of Extraction Methods

All extractions and PCR setups were performed in a dedicated clean laboratory at the University of Copenhagen, which conforms to the highest standards for the field [32]. Methodological experiments on plant aDNA are fundamentally complicated by limited numbers of suitable specimens and potentially variable DNA preservation among samples, however extractions were designed to minimize variability within a collection of samples. Over three rounds of experiments, sets of archaeobotanical remains were extracted using three to five different methods, and tested for DNA yield and purity. We refer to the methods according to the leading author of the first publication to describe the technique or the commercial name, as listed in Table 1. Appendix S1 provides detailed protocols for all methods, including any modifications from the authors’ or manufacturers’ specifications.

Archaeobotanical remains from a variety of contexts were extracted, listed in Table 2. When deemed sufficiently intact, seeds were cleaned in 0.5% bleach (NaClO) and rinsed in molecular grade water before being extracted; seeds with small cracks or other imperfections, indicated in Table 2, were instead wiped with a towel. The cleaning of other types of archaeobotanical remains, such as maize cobs and grape branches, was conducted by removing exterior surfaces with sterile tools. Most archaeobotanical remains were desiccated, although one set was waterlogged. No charred archaeobotanical remains were tested in these experiments because burned remains often contain little or no endogenous DNA that can be amplified by PCR [33–35]. This is an important consideration because macrobotanical remains are most frequently preserved at archaeological sites through charring or carbonization [36]. Desiccation and waterlogging are comparatively less common processes by which plant remains become preserved; nonetheless, desiccated and waterlogged macrobotanicals have been recovered from archaeological sites around the world and are much more likely to contain endogenous aDNA since they have not been exposed to high temperatures. Thus, these experiments are most pertinent to non-charred remains, although some findings may prove applicable to charred remains in subsequent analyses.

Table 1. Extraction techniques compared in this study.

Experiment phase	Name	Method synopsis and relevant information	Reference
Phase 1	Epicentre	QuickExtract Plant DNA Extraction Solution. Designed to extract DNA from modern plant remains in 10 minutes.	Epicentre, Madison, WI
	Finzymes	Phire Plant Direct PCR kit. Sample incubated for 3 minutes in buffer and immediately amplified.	Thermo Fisher Scientific, Waltham, MA
	Gilbert	Digestion in SDS, DTT, and Proteinase K, followed by phenol and chloroform extraction. Previously used to extract DNA from ancient grapes [58].	Gilbert <i>et al.</i> [37]
	Japelaghi	Digestion in PVP, CTAB, and 2-mercaptoethanol followed by chloroform-isoamylalcohol extraction. Method designed for modern plant remains rich in tannins.	Japelaghi <i>et al.</i> [13]
	MO BIO	PowerLyzer PowerSoil DNA Isolation kit. Used to recover aDNA from humic-rich soils [43,59].	MO BIO Laboratories, Carlsbad, CA
Phase 2 ¹	Gilbert	See phase 1.	See phase 1.
	Palmer	Digestion in CTAB, followed by chloroform-isoamyl alcohol extraction, and purification in Qiagen MinElute column.	Modified from Palmer <i>et al.</i> [39]
	Rohland	Digestion in SDS, DTT, and Proteinase K, followed by DNA binding to silica pellet. Silica extraction previously found to be optimal for extracting aDNA from bones.	Modified from Rohland and Hofreiter [15]
Phase 3	Andersen	Digestion in 2-mercaptoethanol, DTT, and Proteinase K, followed by MO BIO inhibitor removal, phenol and chloroform extraction, and Millipore filter purification. Designed to recover aDNA from sediment.	Experimental method developed by Kenneth Andersen
	Gilbert	See phase 1.	See phase 1.
	Palmer	See phase 2, but with purification in Millipore filter and Qiagen DNeasy silica column. Exact method used to recover aDNA from ancient barley remains [39].	Palmer <i>et al.</i> [39]

¹Extraction methods in phase 2 were conducted in three ways: according to the specified directions, with MO BIO C2 and C3 solutions added before extraction, and with MO BIO C2 and C3 solutions used after extraction. See the text and Appendix S1 for further details.

doi:10.1371/journal.pone.0086827.t001

In extraction phase 1, seven sets of *Vitis vinifera* pips were extracted. Grapes were tested because they contain a number of PCR inhibitors and provide a challenge even for genetic studies of modern material [13]. In extraction phase 1, a single seed was extracted with a given method. Recognizing that DNA within samples may be differentially preserved, phases 2 and 3 were conducted on a homogenized collection of seeds from a given context, thereby standardizing the amount of aDNA, contaminant DNA, and inhibitory substances. In addition, a wider range of species and contexts were tested in later extraction phases: four sets of archaeobotanical remains were tested in phase 2 and eight sets in phase 3.

In phase 1, we compared five extraction techniques which have been designed for either ancient materials or modern plant remains. Samples were tested in duplicate for the Gilbert *et al.* [37], Japelaghi *et al.* [13], and MO BIO methods; however, due to a limited number of seeds from identical contexts, it was not possible to perform duplicate extractions for the Epicentre and Finzymes techniques. Extraction methods were compared on the basis of three criteria: DNA concentration measured on a Qubit 1.0 Fluorometer (Invitrogen, Carlsbad, CA), sample purity measured on a NanoDrop 1000 spectrophotometer (Thermo Scientific, Waltham, MA), and amplification success for the ribulose-bisphosphate carboxylase (*rbcL*) gene, a universal plant marker [38]. PCR conditions for the *rbcL* locus are listed in Appendix S1.

The most promising method was advanced to phase 2, where it was compared with Palmer *et al.*'s [39] extraction method (with minor modifications as listed in Table 1) for ancient plants and a silica pellet extraction, the top performing technique in Rohland and Hofreiter's [15] study on isolating aDNA from bones. In

addition to conducting the extractions according to the specified directions, the methods were modified with the addition of MO BIO 'C2' and 'C3' solutions (MO BIO Laboratories, Carlsbad, CA), reagents designed to precipitate humic acids and increase sample purity. This modification was conducted either after overnight incubation in digestion buffer or directly after DNA extraction. The same criteria were used to compare the methods as in phase 1, with the addition of sequencing *rbcL* products to determine if endogenous DNA was recovered.

For phase 3, the two top performing methods were further compared, along with an experimental technique developed by one of the authors of this paper (KA). This method, referred to as the Andersen method, is part of an ongoing project to extract aDNA from sediments, and therefore may not be fully optimized. Nonetheless, preliminary findings suggest the Andersen method readily handles humic-rich sediments, and it was hypothesized the technique may also effectively isolate DNA from archaeobotanical remains. In addition to the above previously used testing criteria, the three methods were compared using a qPCR assay for the *rbcL* generic marker to more precisely determine the amount of plant DNA recovered (for details, see "qPCR assay for quantifying DNA in extraction phase 3" in Appendix S1). This approach was deemed necessary because pigmentation in some extracts could lead to erroneous DNA concentration readings in the Qubit Fluorometer.

Comparison of DNA Polymerases

Enzymatic inhibition of five polymerases was tested by amplifying exogenous tiger (*Panthera tigris*) DNA "spiked" into pigmented plant eluates. As indicated in Table 2, heavily pigmented DNA extracts from two ancient plant samples were

Table 2. Archaeobotanical remains analyzed.

Sample information			Archaeological context		Provenience and age			Extraction phase			PCR tests	
Name	Species	Tissue ¹	Site	Geographic location	Repository ²	Provenience and age		1	2	3		
ARE-A	<i>Vitis vinifera</i>	Pips	Areni-1	Areni, Armenia	IAE	Trench 1, square P30/31, locality 2, spit 6. Medieval context.		X				
ARG	<i>Vitis vinifera</i>	Pips	Fossato	Argenta (FE), Italy	UMeRE	SU ³ 2.2. 1275–1325 A.D.		X				
CPR-A ⁴	<i>Vitis vinifera</i>	Pips	Corso Porto Reno – Via Vasporgolo	Ferrara, Italy	UMeRE	SU 1703. Medieval context.		X				
LUG	<i>Vitis vinifera</i>	Pips	Piazza Baracca	Lugo (RA), Italy	UMeRE	SU 144. 15 th –16 th c. A.D.		X				Polymerase fidelity
PAR-A	<i>Vitis vinifera</i>	Pips	Piazza Municipale	Parma, Italy	UMeRE	SU 320. 4 th –2 nd c. B.C.		X				
SAM ⁴	<i>Vitis vinifera</i>	Pips	Pozzo 1 Domagnano	San Marino, Republic of San Marino	UMeRE	SU 565. Late Roman–Gothic context.		X				
VAD-A	<i>Vitis vinifera</i>	Pips	Vasca Ducale Piazza Municipale	Ferrara, Italy	UMeRE	SU 1050. 2 nd half 15 th c. A.D.		X				Polymerase fidelity
CAS	<i>Cornus mas</i>	Seeds	Cassa di Risparmio	Modena, Italy	UMeRE	SU 31. Roman context.			X			
SAF	<i>Vitis vinifera</i>	Pips	Piazzale San Francesco	Modena, Italy	UMeRE	SU 16. 10 th –11 th c. A.D.			X			
ARE-B ⁵	<i>Gossypium</i> sp.	Seeds	Areni-1	Areni, Armenia	IAE	Trench 1, square K35, spit 4. Medieval context.		X	X			Inhibition
VAD-B	<i>Olea europaea</i>	Pits	Vasca Ducale Piazza Municipale	Ferrara, Italy	UMeRE	SU 1050. 2 nd half 15 th c. A.D.			X			Polymerase fidelity
ARE-C ⁵	<i>Vitis vinifera</i>	Branch	Areni-1	Areni, Armenia	IAE	Trench 1, square N16, locality 29. Medieval context.				X		
CDM ⁵	<i>Zea mays</i>	Cob	Cueva del Maguey 1	Pueblo Nuevo, Durango, Mexico	INAH	Specimen ID: 10189. 1410±25 ¹⁴ CYBP				X		
CPR-B	<i>Cornus mas</i>	Seeds	Corso Porta Remo-Via Vesporgolo	Ferrara, Italy	UMeRE	SU 2597. Medieval context.				X		
PAR-B	<i>Vitis vinifera</i>	Pips	Piazza Municipale	Parma, Italy	UMeRE	SU 165. Medieval context.				X		
SUP ⁴	<i>Vitis vinifera</i>	Waterlogged pips	Loc. Scorpo	Supersano (LE), Italy	US	Excavated from well, 7 th –8 th c. A.D.				X		
THR	<i>Zea mays</i>	Kernels	Turkey House Ruin	Navajo County, Arizona	ASM	Specimen ID: 935. 723±23 ¹⁴ CYBP				X		
SPC	<i>Vitis vinifera</i>	Pips	Via San Pietro	Modena, Italy	UMeRE	Excavated from composting feature. Medieval.						Inhibition

¹Tissues are desiccated except where noted.²Samples provided by archaeologists and curators, as listed in acknowledgements. ASM: Arizona State Museum; IAE: Institute of Archaeology and Ethnology, National Academy of Sciences, Yerevan, Armenia; INAH: Instituto Nacional de Antropología e Historia, Centro INAH Michoacán, Mexico; UMeRE: Università di Modena e Reggio Emilia, Modena, Italy; US: University of Salento, Lecce, Italy.³SU: Stratigraphic unit.⁴Seeds cleaned by wiping with a dry paper towel. All other seeds cleaned by washing in 0.5% bleach.⁵Samples cleaned by removing exterior (seed coat or bark) with sterile scalpel.

doi:10.1371/journal.pone.0086827.t002

used as inhibiting substances: medieval cotton (*Gossypium* sp.) seeds from the Areni-1 site in Areni, Armenia, and medieval grape (*Vitis vinifera*) pips from the Via San Pietro site in Modena, Italy. Varying amounts of inhibiting solutions were added to PCR reactions, with pigmented extracts representing up to 40% of the reaction volume. Polymerases were selected based upon either their ubiquity in aDNA research, advertised fidelity, or purported ability to overcome inhibition, as summarized in Table 3. PCR details for each polymerase are located in Table S1. As bovine serum albumin (BSA) has been shown to prevent inhibition and increase the likelihood of amplification success in ancient samples [15,40], reactions were conducted with and without 0.8 mg/mL BSA additive.

AmpliTaQ Gold, Omni KlenTaq, and PfuTurbo C_x Hotstart were further tested for potential use in qPCR assays by amplifying spiked DNA in varying concentrations of inhibitors. Experimentation suggested that BSA occasionally interfered with the detection of fluorescence with AmpliTaQ Gold; therefore, for each polymerase, reactions were conducted with and without 0.8 mg/mL BSA. The effects of BSA and inhibition were observed through changes in cycle threshold (C_t) and amplification curves. Experiment conditions are listed in Appendix S1 in the section “qPCR inhibition testing.”

Polymerase fidelity and compatibility with aDNA damage were investigated through “deep sequencing” of an endogenous DNA marker from ancient plant samples. This approach is commonly used to characterize biodiversity in environmental samples [41], including ancient ones [42]. In such studies, a universal genetic marker for a group of organisms, such as plants or animals, is amplified and sequenced on a HTS platform to identify all species present in the sample and their relative proportions [43]. Here, the aim of deep sequencing is to test thousands of copies of the plant *rbcL* marker amplified from a sample to infer how often polymerases make errors. PCR products from three ancient plant samples, listed in Table 2, were sequenced on a Roche/454 Genome Sequencer FLX platform (for further information, see “Deep sequencing of *rbcL* products” in Appendix S1). Reads were aligned to the expected sequence in Geneious Pro 5.5.7 [44] and nucleotide misincorporations, insertions, and deletions were analyzed.

Results

Extraction Comparisons

Phase 1. The five extraction methods yielded highly variable DNA concentrations, amplification success rates, and purity levels. The Epicentre and Finnzymes extraction methods frequently yielded DNA eluates that were darkly pigmented. This is significant because DNA concentrations, as measured on the fluorometer, could produce anomalous readings if the pigmented eluates prevent accurate assessment of DNA-binding dyes.

Therefore, the primary indication of success was taken to be the rate of successful amplification of genetic plant markers. Based on this criterion, the Gilbert method was the top performer, with successful amplification of the *rbcL* marker in 10 out of 14 specimens, as listed in Table 4. Japelaghi’s method scored the second most successes: 7 of 14.

Amplification successes were compared using the generalized estimating equations function in PASW Statistics 18.0 [45]. This approach accommodates the presence of replicates for a given method and controls for success rates within each set of samples, even with limited numbers of samples. The Wald test found the best performing technique, the Gilbert method, to have significantly higher odds of amplifying the *rbcL* marker than the Finnzymes and MO-BIO techniques ($p=0.001$ and 0.018 , respectfully). The difference between the Gilbert method and the other two methods was not statistically significant (Epicentre, $p=0.061$; Japelaghi, $p=0.273$); however, qualitatively, it yielded stronger, more distinct PCR bands than the others.

None of the methods yielded amplifiable DNA from ARE-A, but this could be due to degradation of the sample (i.e. the endogenous DNA was shorter than the 138 bp *rbcL* marker). Therefore, DNA concentrations and purity readings for this sample are still considered germane. The mean amount of DNA for the Gilbert method was 304.5 ng, nearly triple the second highest value, 102.2 ng by Epicentre. After omitting the outliers shown in the left side of Figure 1, values were compared using a univariate generalized linear model (mixed model ANOVA) to control for differences between specimens. The model determined the method [$F(4, 24)=6.771$, $p=0.001$], specimen [$F(6, 24)=5.566$, $p=0.001$], and interaction between method and specimen [$F(18, 24)=9.607$, $p<0.001$] to be statistically significant. Tukey’s HSD post-hoc test finds the Epicentre and Gilbert methods yield statistically significant greater amounts of DNA than other methods ($p<0.001$), but the difference between the two is not statistically significant ($p=0.885$).

The right side of Figure 1 depicts the ratio of light absorbance at 260 and 280 nm, where a ratio of 1.8 is commonly considered to represent pure DNA [46]. None of the five methods consistently reached a ratio of 1.8, perhaps due to the low amount of aDNA in specimens, but the Gilbert method was the closest. After omitting the five outliers, an ANOVA test found statistical differences in the ratio of 260/280 between methods [$F(4, 41)=10.862$, $p<0.001$], and Tukey’s HSD post-hoc test found the Gilbert method to have a statistically higher 260/280 ratio than the Finnzymes ($p=0.014$) and MO BIO ($p<0.001$), but not the Epicentre ($p=0.116$) or Japelaghi methods ($p=0.867$).

Phase 2. As the Gilbert method performed the most consistently in phase 1, with the highest rate of successful amplification, the most DNA, and the purest eluates, it was promoted to more testing in phase 2. In terms of amplification success, all methods without C2/C3 solutions yielded PCR bands

Table 3. Polymerases tested.

Polymerase	Vendor	Notable features
AmpliTaQ Gold	Applied Biosystems, Foster City, CA	Commonly used in aDNA research
Omni KlenTaq	DNA Polymerase Technology, St. Louis, MO	Engineered to overcome multiple sources of inhibition, including blood and soil
PfuTurbo C _x Hotstart	Agilent Technologies, La Jolla, CA	Purportedly reads through uracil while maintaining high fidelity
Phire Hot Start II	Finnzymes (Thermo Fisher Scientific), Waltham, MA	Designed to overcome inhibition and features rapid processivity
Phusion Hot Start	Finnzymes (Thermo Fisher Scientific), Waltham, MA	Engineered for high fidelity and rapid processivity

doi:10.1371/journal.pone.0086827.t003

Table 4. Amplification success for extractions, phase 1.

Sample	Replicate ²	Generic <i>rbcl</i> plant marker amplified for given method ¹				
		Finnzymes	Epicentre	Gilbert	Japelaghi	MO BIO
ARE-A	A	—	—	—	—	—
	B	—	—	—	—	—
ARG	A	—	—	(+)	—	—
	B	—	—	+	—	—
CPR-A	A	+	(+)	+	—	+
	B	—	—	+	(+)	—
LUG	A	—	+	+	+	+
	B	—	+	+	+	+
PAR-A	A	—	—	+	+	—
	B	—	(+)	+	—	—
SAM	A	—	—	—	(+)	—
	B	—	—	—	—	—
VAD-A	A	(+)	+	(+)	+	(+)
	B	—	(+)	+	+	—
Amplification successes		2/14	6/14	10/14	7/14	4/14

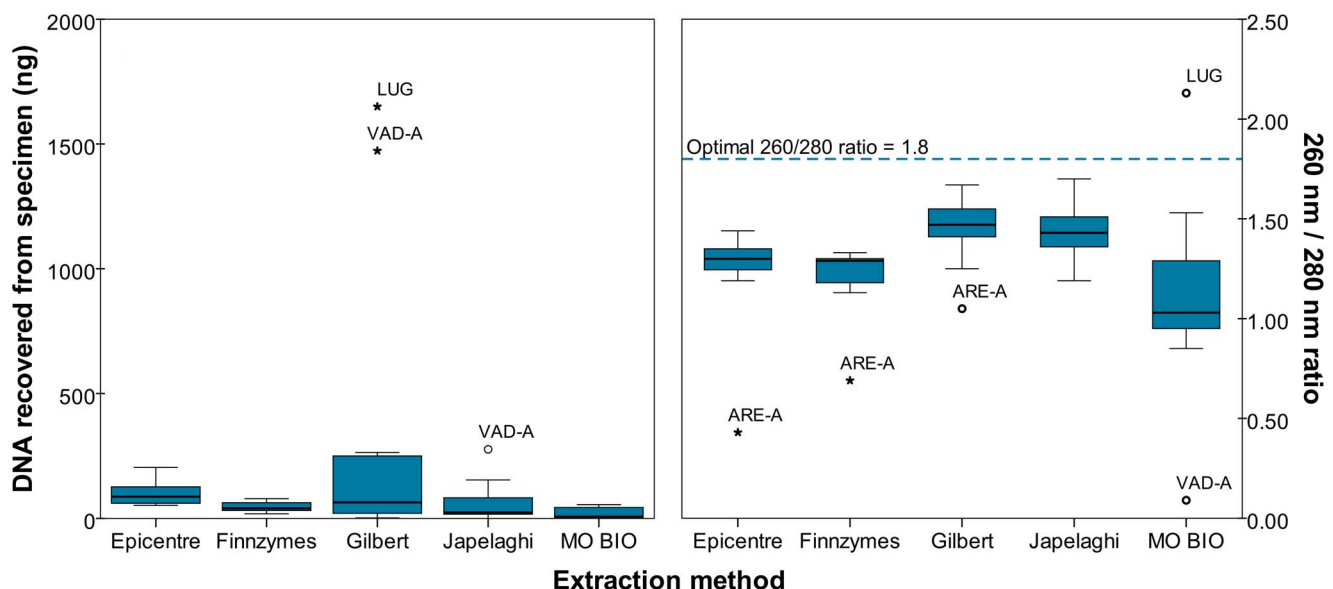
¹+ indicates a distinct band on 2% agarose gel, (+) indicates a faint band, and — indicates no band.

²Amplifications for Finnzymes and Epicentre were conducted at (A) full strength and (B) 10% dilutions to test for enzymatic inhibition rather than two separate extractions of different seeds.

doi:10.1371/journal.pone.0086827.t004

for SAF and VAD-B samples, and the Rohland method also produced a weak band for the CAS sample, as listed Table S2. Cloning and sequencing of the PCR bands showed that sequences for the SAF and VAD-B samples were identical to the expected sequence, or <2 bp different from the sequence, an error rate generally consistent with damaged DNA. None of the recovered sequences of CAS sample from the Rohland method were closer than 2 bp to the expected sequence and therefore likely represent contamination.

The unmodified Gilbert method yielded more DNA than the other methods, and the addition of C2/C3 nearly always decreased DNA yield, as seen in Figure 2. To control for major differences in DNA recovery between specimens, DNA yield values were compared after logarithmic transformation. Log values were tested in a univariate generalized linear model controlling for differences in specimens, and found to have significant effects of extraction method [$F(2, 28) = 3.563$, $p = 0.042$], C2/C3 additives [$F(2, 28) = 14.278$, $p < 0.001$], and specimen [$F(3, 28) = 13.239$, $p < 0.001$]. Tukey's HSD post-hoc

**Figure 1.** DNA yield and purity for extractions, phase 1.

doi:10.1371/journal.pone.0086827.g001

identifies the addition of C2/C3 solutions before ($p=0.010$) or after ($p<0.001$) an extraction to significantly reduce DNA recovery. An ANOVA test on the extractions not modified with C2/C3 solutions was significant for method [$F(2, 6)=10.109$, $p=0.012$] and specimen [$F(3, 6)=42.802$, $p<0.001$]. The Gilbert method was found to recover significantly more DNA than the Rohland method ($p=0.010$), but there was no significant difference between the Gilbert and the Palmer methods ($p=0.236$).

The unmodified Gilbert and Palmer methods have statistically identical mean 260/280 ratios: 1.465 and 1.515, respectively. The Rohland method yielded ratios ranging from 1.10 to 4.87, likely due to low DNA content or residual particles from the silica extraction. When modified by C2/C3, the 260/280 ratios were not consistently brought closer to the ideal value of 1.8, as can be seen in Table S2. In all, there was no compelling evidence that the C2/C3 additions improved DNA purity, however, they certainly reduced DNA content.

Phase 3. The Gilbert and Palmer techniques were further tested in the final extraction phase, along with the Andersen sediment-style extraction. In terms of amplification success, the methods performed similarly: the Andersen and Palmer methods amplified six samples, while the Gilbert method amplified the same six as well as PAR-B. PCR was also tested without BSA,

leading to the failure of nearly every reaction. The only samples amplifiable without BSA were THR (successful in all three methods) and VAD-B (a faint band in Palmer's method). This finding may have important implications for the use of BSA in PCR on aDNA from non-charred archaeobotanical remains, as discussed below.

DNA purities were statistically identical, with mean 260/280 ratios of 1.527 ($sd=0.188$), 1.558 ($sd=0.157$), and 1.524 ($sd=0.245$) for the Andersen, Gilbert, and Palmer methods, respectively. The amount of DNA recovered by the methods was more variable, as shown in top half of Figure 3. Mean DNA recovery was highest in the Gilbert method (1226.9 ng, $sd=1909.1$), followed by the Andersen (651.1 ng, $sd=722.2$) and Palmer (597.6 ng, $sd=968.6$) methods. Log transformed DNA yields were tested in a univariate generalized linear model controlling for differences in specimens (mixed model ANOVA), and were found to have significant effects for extraction method [$F(2, 14)=6.539$, $p=0.012$] and specimen [$F(3, 28)=13.239$, $p<0.001$]. Post-hoc testing with Tukey's HSD test found the Gilbert method recovered a statistically significantly greater amount of DNA than the Palmer method ($p=0.007$), but not the Andersen method ($p=0.184$).

The number of copies of the *rbcL* gene recovered by each method varies dramatically between methods and samples. As seen

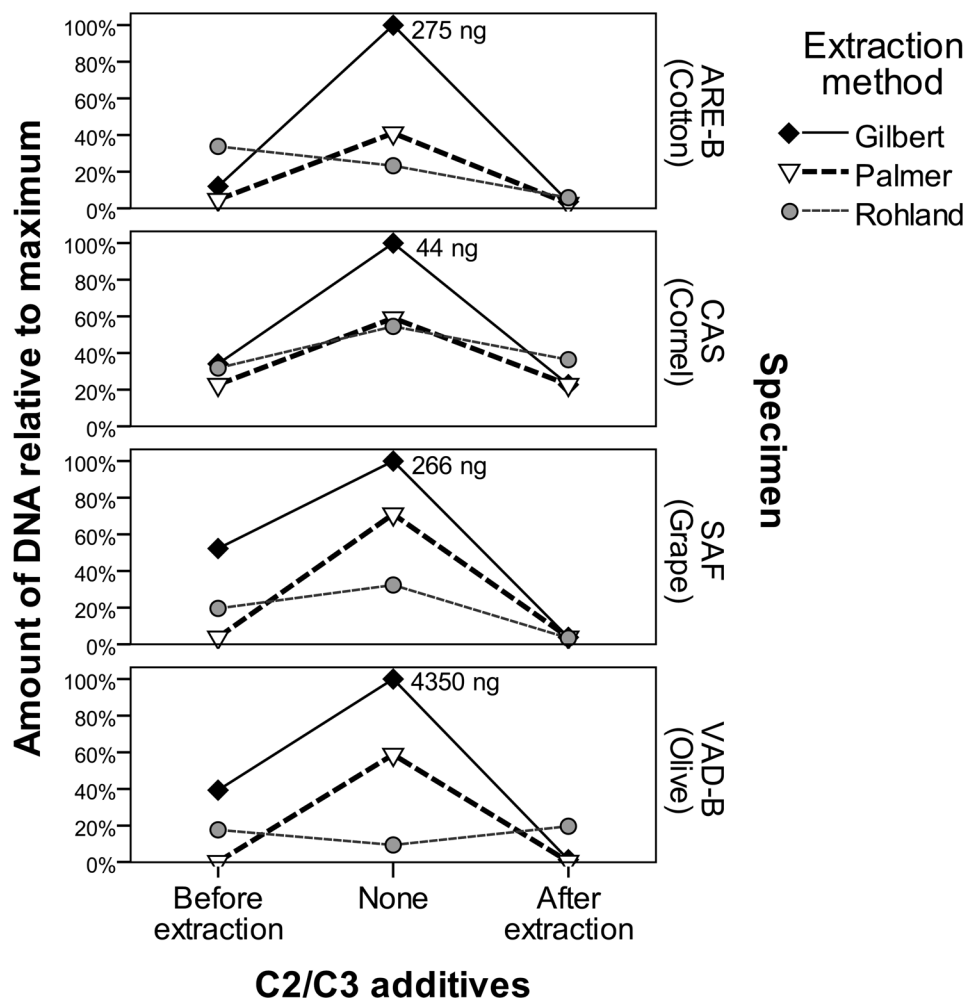


Figure 2. DNA yield from extractions, phase 2. Maximum amount of DNA recovered in each specimen listed by corresponding symbol. doi:10.1371/journal.pone.0086827.g002

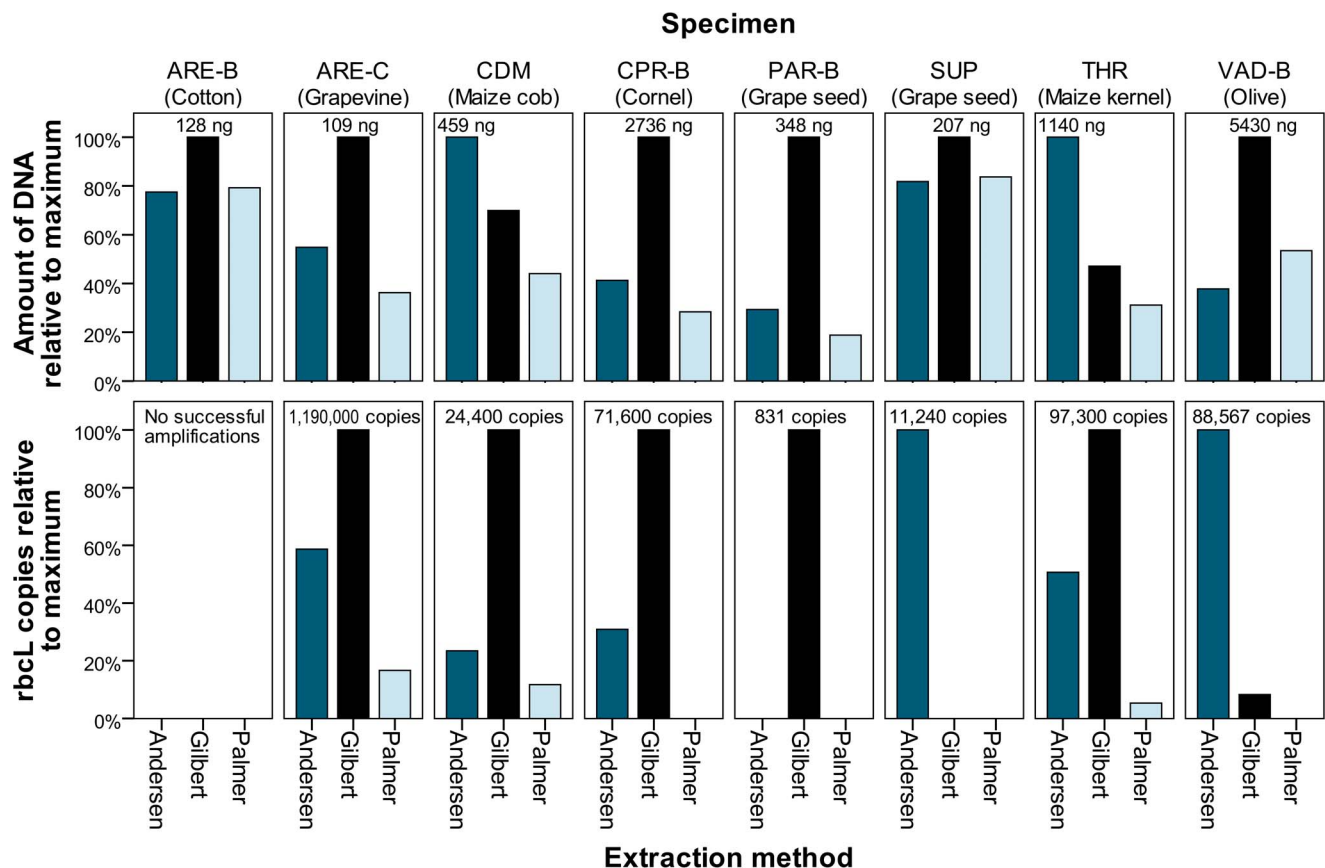


Figure 3. DNA yield and *rbcL* copies extracted during phase 3. DNA yield (top) calculated using a Qubit 1.0 Fluorometer and *rbcL* copies (bottom) determined by qPCR. Values are scaled to the maximum value of each sample, with the highest value listed above the corresponding bar. Missing bars in lower portion of figure indicates that a sample did not amplify in qPCR. doi:10.1371/journal.pone.0086827.g003

in Figure 3, the number of *rbcL* copies as determined by qPCR does not perfectly reflect the amount of DNA measured on the Qubit fluorometer. This may indicate less pure eluates occasionally yield errant values. It could also be possible the methods differ in their ability to extract endogenous and exogenous DNA. To control for the wide large range of values, a logarithmic transformation was done, using $\log(x+1)$ to incorporate zero values. A mixed model ANOVA found the method [$F(2, 14) = 4.707$, $p = 0.027$] and specimen [$F(7, 14) = 5.646$, $p = 0.003$] to be significant factors in the number of recovered *rbcL* copies. Tukey's HSD post-hoc test determined the Andersen method recovers significantly more *rbcL* copies than the Palmer method ($p = 0.043$), but there is not statistical difference between the Andersen and Gilbert methods ($p = 0.995$). Results provided by the Gilbert method are also found to differ from those provided by the Palmer method, but the differences are just beyond the threshold of statistical significance ($p = 0.051$).

DNA Polymerases

Ability to overcome inhibition. The five polymerases demonstrated great variability in overcoming inhibition from substances found in ancient plant materials, as shown in Table 5. Without BSA additives, only Omni KlenTaq and Phire Hot Start II were successful amplifying spiked tiger DNA in the presence of inhibitors, yielding PCR bands in reactions containing up to 1% of the ARE-B eluate. The addition of BSA enabled all polymerases to be functional in reactions containing at least 1% inhibiting

substances. With BSA, AmpliTaq Gold, Omni KlenTaq, and Phire overcame inhibition in at least one sample with 5% inhibitors. Omni KlenTaq particularly excelled when BSA was added, successfully amplifying reactions containing 10% inhibiting solutions.

Compatibility with qPCR. The three polymerases tested in qPCR behaved very differently when amplifying spiked DNA in the presence of BSA and inhibitors, as can be observed in Figure 4 and Table S3. The addition of BSA had a negative impact on the amplification curve in AmpliTaq Gold, but not the other polymerases. Increasing concentrations of inhibitors further reduced the slope of the amplification phase of AmpliTaq Gold reactions, and also affected PfuTurbo C_x Hotstart when inhibitors reached 2.5%. Conversely, Omni KlenTaq was remarkably resilient to amplification inefficiencies due to inhibition.

Fidelity. The PfuTurbo C_x Hotstart polymerase was unable to amplify plant DNA in the LUG sample; therefore, 14 of the 15 possible combinations of specimens and polymerases were analyzed. Deep sequencing of the *rbcL* plant marker showed the vast majority of recovered sequences were consistent with the expected endogenous sequence, listed in Table S4. The entire dataset of sequencing reads is available online in Data S1–S14. All reads differing from the expected sequence by more than 3 bp were excluded from analyses, leaving 99.2%–99.9% of the original data for each case.

Three polymerases yielded a small number of sequences that could not be aligned to *rbcL* markers, shown in Table S4. Some of

Table 5. Amplification of spiked DNA in the presence of inhibiting substances.

BSA additive	Inhibiting solution	Amount of inhibitor in reaction	Polymerase ¹				
			AmpliTaq Gold	Omni KlenTaq	PfuTurbo C _x	Phire	Phusion
No BSA	SPC	0%	+	+	+	+	+
		0.1%	—	+	—	+	—
		≥1%	—	—	—	—	—
	ARE-B	0%	+	+	+	+	+
		0.1%	—	+	—	+	—
		1%	—	(+)	—	(+)	—
		≥2.5%	—	—	—	—	—
BSA added	SPC	0%	+	+	+	+	+
		0.1%	+	+	+	+	+
		1%	+	+	+	+	+
		2.5%	+	+	—	+	—
		5%	(+)	+	—	(+)	—
		10%	—	+	—	—	—
		≥20%	—	—	—	—	—
	ARE-B	0%	+	+	+	+	+
		0.1%	+	+	+	+	+
		1%	+	+	+	+	+
		2.5%	+	+	(+)	+	(+)
		5%	—	+	—	+	—
		10%	—	+	—	—	—
		≥20%	—	—	—	—	—

¹+ indicates a distinct band on 2% agarose gel, (+) indicates a faint band, and — indicates no band.
doi:10.1371/journal.pone.0086827.t005

these were determined to be chimeras of amplicons. Notably, Omni KlenTaq had a relatively high percentage of non-aligning reads. Additionally, Omni KlenTaq was observed to occasionally yield DNA smears on agarose gels, a characteristic consistent with replication errors.

Nucleotide substitution rates were calculated as the number of incorrect nucleotides divided by the number of correct nucleotides

[47], listed in Table S5. Sequencing errors and DNA damage undoubtedly contribute to the overall error rate, but they are expected to be relatively constant across samples. As seen in Figure 5, Phusion polymerase had a consistently lower error rate than the other polymerases. A one-way ANOVA test found statistical differences in the error rates between polymerases [$F(4, 9) = 20.022$, $p < 0.001$] and Tukey's HSD post-hoc test found

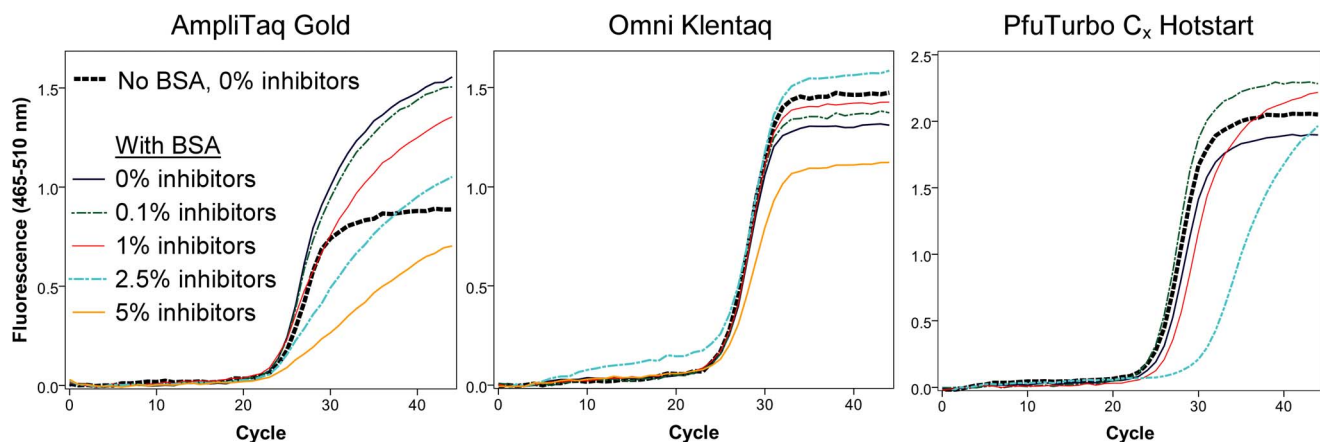


Figure 4. Compatibility of polymerases with qPCR. Inhibitory substances extracted from the SPC sample prevented amplification of spiked DNA in all reactions not including BSA, except for Omni KlenTaq in 0.1% inhibitors (not shown). Unsuccessful amplifications, including PfuTurbo C_x Hotstart in 5% inhibitors, are not included in figure.
doi:10.1371/journal.pone.0086827.g004

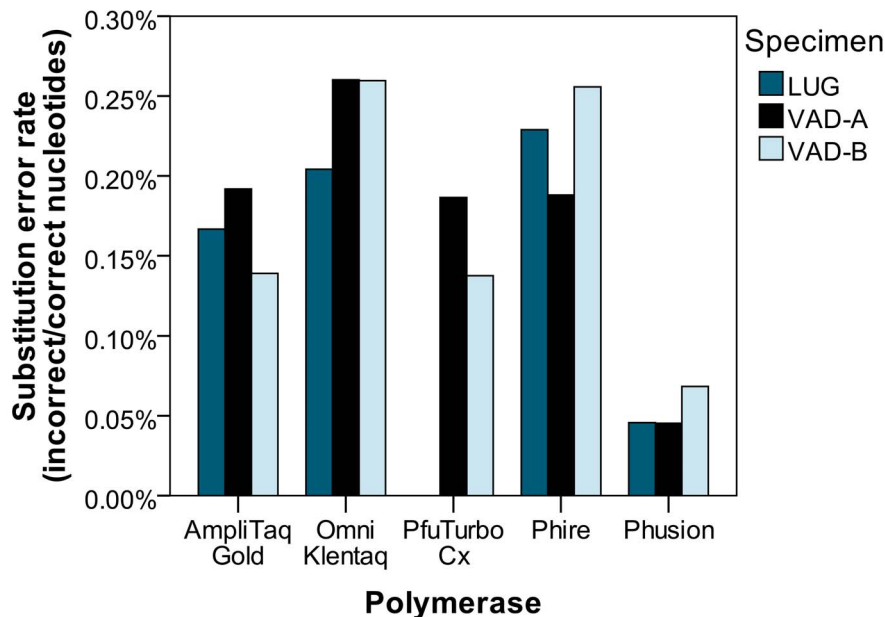


Figure 5. Overall substitution error rates on endogenous aDNA. Shorter bars represent fewer nucleotide misincorporations (higher polymerase fidelity). Sequencing reads that differed from the expected *rbcl* sequence by >3 nucleotide substitutions were omitted prior to tallying nucleotide calls and errors. As stated, the PfuTurbo C_x polymerase did not amplify the LUG sample.
doi:10.1371/journal.pone.0086827.g005

Phusion's error rate to be significantly lower than the other polymerases (versus AmpliTaq Gold: $p=0.006$, PfuTurbo C_x: $p=0.015$, and Omni Klentaq and Phire: $p<0.001$). Differences among the other polymerases were not statistically significant.

Phusion was also found to have the lowest error rates for nucleotide insertions and deletions, but several other polymerases had similar rates, as seen in Table S5. A one-way ANOVA test found statistically significant differences among the samples in nucleotide deletion rates [$F(4, 9) = 3.976$, $p=0.040$], but not insertion rates [$F(4, 9) = 2.031$, $p=0.173$]. Tukey's HSD post-hoc test found the deletion rate in Phusion to be statistically different from that of AmpliTaq Gold ($p=0.025$).

Compatibility with damaged DNA. According to the manufacturer, Phusion polymerase is incompatible with uracil, causing DNA replication to stall. Conversely, PfuTurbo C_x is advertised as able to read uracil, resulting in an apparent C-to-T transition on the template strand and G-to-A transition on the complementary strand. Nucleotide substitutions rates in the other three polymerases were compared to those of Phusion and PfuTurbo C_x to determine if they follow similar patterns. As seen in Figure 6, Phusion has lower error rates in C-to-T and G-to-A transitions than the other polymerases. An ANOVA test on the error rates for individual samples found statistically significant differences in error rates for C-to-T [$F(4, 9) = 30.846$, $p<0.001$] and G-to-A [$F(4, 9) = 7.045$, $p=0.007$] transitions. Tukey's HSD post-hoc test on the C-to-T transitions found Phusion to have a statistically different error rate than the other polymerases ($p\leq 0.002$ for each pairwise comparison). Tukey's HSD post-hoc test on the G-to-A transitions found Phusion to have a statistically different error rate than AmpliTaq Gold ($p=0.028$), Omni Klentaq ($p=0.005$), and PfuTurbo C_x ($p=0.034$), but not Phire ($p=0.103$). Overall, none of the polymerases tested have a pattern consistent with Phusion, indicating they pair uracil with adenine rather than stalling. The expanded dataset with error rates for all substitution types is available in Table S6.

Discussion

These experiments provide a new perspective on how to extract and amplify endogenous DNA from non-charred archaeobotanical remains. Now that researchers are incorporating HTS technologies into the study of aDNA from ancient plant remains [19,35,48], these findings should prove especially useful, and may aid future research on critical issues surrounding plant evolution, domestication, and cultivation.

In order to fully profit from HTS of ancient remains, steps should be taken to optimize aDNA recovery. For archaeobotanical remains, these concerns are not trivial, because samples are often small and suboptimal approaches yield insufficient quantities of DNA, potentially leading to the destruction of samples for little or no gain. In the extraction experiments conducted here, the method that consistently performed the best is that described by Gilbert *et al.* [37]. While this method was developed by one of the authors, it was tested impartially, and found to recover more DNA with fewer co-extracted inhibiting substances than other techniques, even across a wide range of species and plant tissues. For previously untested archaeobotanical remains, it logically follows the Gilbert method provides the greatest chance for successful aDNA recovery. That being said, in the final round of testing, an extraction method developed for humic-rich sediments recovered more DNA from a few specimens, suggesting that it may be necessary to test a couple of methods for the most precious of samples. Of course, the insights garnered during this testing are limited to the set of extraction techniques used in the experiments. However, most methods commonly employed on ancient plant remains combine elements of the already tested approaches, so we do not anticipate such techniques to perform drastically differently.

It is interesting to consider how the best extraction methods compare to some others used in the field. For instance, the top two performing methods do not include cetyltrimethylammonium bromide (CTAB), a reagent used in many extraction methods on ancient plant remains, both charred [35,49] and non-charred

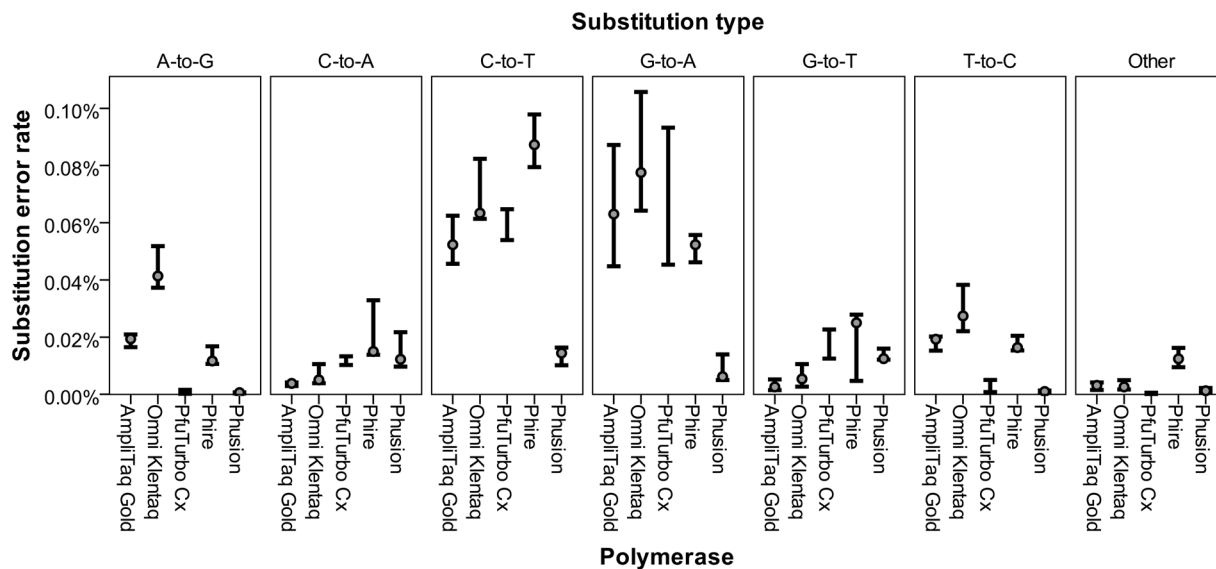


Figure 6. Error rates of most frequent substitution types. High-low chart depicts the maximum and minimum error rates within the three tested samples. Median values, represented by circles, are not included for PfuTurbo C_x because only two samples were amplified. doi:10.1371/journal.pone.0086827.g006

[48,50]. CTAB is used to remove polysaccharides in modern plants [51], but contrary to conventional wisdom, it may not be necessary for non-charred archaeobotanical remains. Likewise, silica pellet extractions have been shown to excel at isolating aDNA from bones [15], but they did not perform as well on ancient plant samples in our testing. Unsurprisingly, commercial DNA extraction kits designed for use on freshly sampled modern plants were found to perform very poorly on ancient samples. Therefore, we would generally discourage aDNA researchers from using such kits on archaeobotanical remains, although similar kits have successfully yielded plant aDNA in some instances [52,53].

Comparative testing of polymerases also yielded a number of important insights. One of the key findings is that no polymerase excels in all categories; rather, they have nuanced properties and should be selected with care, according to the goals and methods in a given research project, as outlined below. Some of the findings about particular polymerases have been reported [20,54], but the results of these experiments can help select which polymerase to use in different circumstances.

One of the most commonly used polymerases in aDNA research, AmpliTaq Gold, was found to perform well in many categories, making it a good all-around polymerase. When used in conjunction with BSA, it can overcome moderate amounts of inhibition. Furthermore, it handles the most common form of nucleotide damage, cytosine deamination. Therefore, AmpliTaq Gold is well suited to amplify markers of interest in aDNA libraries, albeit with some reservation due to its replication error rate.

Phusion, a polymerase designed to have very high fidelity, was indeed found to have a much lower error rate than the other polymerases. However, Phusion is incompatible with uracil and stalls on damaged DNA templates. This is a critical concern for amplification of genetic markers or aDNA libraries, because Phusion will preferentially amplify non-damaged molecules, precisely those originating from modern contaminants. Therefore, some aDNA researchers, such as Green *et al.* [1], have devised a two-step amplification approach to retain damaged DNA but keep replication errors to minimum. First, a uracil-friendly polymerase, such as AmpliTaq Gold, is used to amplify over damaged

nucleotides in a genetic marker or DNA library with a limited number of PCR cycles (10 cycles, for example). Then, in a second reaction, a high-fidelity polymerase, such as Phusion, is used to copy DNA with minimal errors, and reach the required number of DNA copies. Note that other strategies to deal with uracil in aDNA exist [55], but they are not based on polymerases and are therefore outside the realm of this article.

One of the most striking findings of the polymerase tests was the ability of Omni Klentaq to overcome inhibitory substances, consistent with findings on archaeological fish bone samples [20]. Even in high levels of inhibitory substances derived from non-charred ancient plant materials, like humic acids, Omni Klentaq successfully amplified spiked DNA when used with BSA. Without BSA, Omni Klentaq could still amplify DNA in the presence of low levels of inhibitors, a feat not matched by AmpliTaq Gold or Phusion. The significance of this property should not be overlooked, because enzymatic inhibition is not always recognized in the laboratory. For example, some DNA extracts in these studies contain inhibiting substances even though they lacked pigmentation. Omni Klentaq is also reliable in qPCR experiments where enzymatic inhibition may be encountered. Unlike AmpliTaq Gold, Omni Klentaq exhibits an exemplar qPCR amplification curve in the presence of BSA and inhibitors. Conversely, Omni Klentaq may have slightly lower fidelity than AmpliTaq Gold, and occasionally yields chimera amplicons, something not observed in other polymerases. Therefore, it is not an ideal polymerase to amplify libraries or other templates which will be sequenced. Nevertheless, it is an excellent choice for amplifying genetic markers in reticent samples and qPCR assays as it provides a safeguard against undetected enzymatic inhibition.

Another key discovery was that nearly all polymerases fail in the presence of inhibiting substances from non-charred archaeobotanical remains, unless BSA is added. In reactions without BSA, only Omni Klentaq and Phire could amplify spiked DNA, and even then, only the smallest concentrations of inhibitors could be overcome. When BSA was added to reactions containing small amounts of inhibiting substances, all polymerases were successful. This finding is even more important given the amplification tests from the third phase of extractions: irrespective of extraction

method, AmpliTaq Gold nearly always failed to amplify endogenous plant markers unless BSA was added. While it might be assumed that plant-specific extraction protocols, such as those using CTAB, adequately purify DNA, they failed at virtually the same rate as other methods. Thus, we encourage adding BSA in PCR on non-charred archaeobotanical remains, contrary to the approach in most plant aDNA studies [39,50,56].

As we have not extracted charred archaeobotanical remains in these studies, we cannot directly test Giles and Brown's [57] argument that BSA has no benefit for PCR on charred archaeobotanical remains and may reduce amplification success because DNA molecules become bound to BSA along with contaminants. However, it should be noted their study was based on artificially charred seeds and may not reflect the complexity of some archaeobotanical remains. For example, sediments adhering to charred cereals may contain humic acids that could inhibit PCR. Other things being equal, we suggest it is worth conducting PCR with BSA to ensure enzymatic inhibition does not lead to false negative results.

Some of the experimental methodology developed and refined over the course of this study could also provide guidance for future aDNA comparative experiments. For example, spectrophotometric detection of DNA in pigmented eluates was found to be occasionally misleading, so quantification of endogenous aDNA can be more reliably measured with qPCR and sequencing of PCR products. Testing of newly engineered polymerases will continue to be invaluable, and as demonstrated here, comparisons of fidelity and compatibility with damaged nucleotides can be successfully explored via HTS. Considering little is known about the inhibitory effects on polymerases and other enzymes used in the construction of DNA libraries, a similar set of experiments could be undertaken to optimize this fundamental step of HTS research.

Conclusions

As foreseen by Palmer *et al.* [9], the future of plant aDNA research is very bright indeed. The introduction of high-throughput sequencing technologies allows geneticists to delve into ancient genomes in new and exciting ways. In fact, these technologies have already been tested on aDNA extracted from archaeobotanical remains [19,35,48]. However, in order for such studies to become more widespread and for the discipline to reach its full potential, it is critical the best available methods are used to extract, amplify, and analyze DNA from ancient specimens. For desiccated and waterlogged plant remains, this study is a step in that direction, and to that end, we strongly encourage fellow researchers to adopt the best performing extraction techniques, or at a minimum, conduct head-to-head comparisons with more familiar methods. Such experimentation will help advance plant archaeogenetics into a more fruitful discipline, yielding unprecedented understandings of plant evolution, domestication, and human-plant interactions.

Supporting Information

Table S1 PCR conditions for polymerases.
(DOCX)

Table S2 Extraction phase 2 data.
(DOCX)

Table S3 qPCR C_t values for polymerase inhibition testing.
(DOCX)

Table S4 Deep sequencing of *rbcl* markers to investigate polymerase fidelity.
(DOCX)

Table S5 Polymerase error rates. Sequencing reads that differed from the expected *rbcl* sequence by >3 nucleotide substitutions were omitted prior to tallying nucleotide calls and errors.
(DOCX)

Table S6 Specific substitution frequencies and corresponding error rate. Sequencing reads that differed from the expected *rbcl* sequence by >3 nucleotide substitutions were omitted prior to tallying nucleotide calls and errors.
(DOCX)

Appendix S1 Text with detailed extraction protocols, PCR information, DNA sequencing, and expanded results.
(DOCX)

Data S1 Deep-sequencing amplicon data of AmpliTaq Gold on LUG sample in FASTQ format.
(FASTQ)

Data S2 Deep-sequencing amplicon data of AmpliTaq Gold on VAD-A sample in FASTQ format.
(FASTQ)

Data S3 Deep-sequencing amplicon data of AmpliTaq Gold on VAD-B sample in FASTQ format.
(FASTQ)

Data S4 Deep-sequencing amplicon data of Omni KlenTaq on LUG sample in FASTQ format.
(FASTQ)

Data S5 Deep-sequencing amplicon data of Omni KlenTaq on VAD-A sample in FASTQ format.
(FASTQ)

Data S6 Deep-sequencing amplicon data of Omni KlenTaq on VAD-B sample in FASTQ format.
(FASTQ)

Data S7 Deep-sequencing amplicon data of PfuTurbo C_x on VAD-A sample in FASTQ format.
(FASTQ)

Data S8 Deep-sequencing amplicon data of PfuTurbo C_x on VAD-B sample in FASTQ format.
(FASTQ)

Data S9 Deep-sequencing amplicon data of Phire on LUG sample in FASTQ format.
(FASTQ)

Data S10 Deep-sequencing amplicon data of Phire on VAD-A sample in FASTQ format.
(FASTQ)

Data S11 Deep-sequencing amplicon data of Phire on VAD-B sample in FASTQ format.
(FASTQ)

Data S12 Deep-sequencing amplicon data of Phusion on LUG sample in FASTQ format.
(FASTQ)

Data S13 Deep-sequencing amplicon data of Phusion on VAD-A sample in FASTQ format.
(FASTQ)

Data S14 Deep-sequencing amplicon data of Phusion on VAD-B sample in FASTQ format. (FASTQ)

Acknowledgments

The authors thank the following researchers for permission for destructive analysis of archaeobotanical remains: Boris Gasparyan, Institute of Archaeology and Ethnology, National Academy of Sciences, Yerevan, Armenia; Giovanna Bosi and Anna Maria Mercuri, Museo Di Paleobiologia e dell'Orto Botanico, Università di Modena e Reggio Emilia, Modena, Italy; Girolamo Fiorentino, Dipartimento di Beni Culturali, University of Salento, Lecce, Italy; Mike Jacobs, Arizona State Museum; and José Luis Punzo-Díaz, Instituto Nacional de Antropología e Historia, Centro INAH Michoacán, Mexico. From the Centre for GeoGenetics, we

thank Aurélien Ginolhac for additional programming advice and Ludovic Orlando for recommendations on silica pellet extractions. We acknowledge our appreciation to Morten Rasmussen, Kim Magnussen, and the Danish National High-throughput Sequencing Centre for assistance in generating the Roche/454 GS FLX data. The authors also thank two anonymous reviewers who helped improve the manuscript. Special thanks goes to Randi Garcia, University of Connecticut, for thoughtful considerations about statistical tests.

Author Contributions

Conceived and designed the experiments: NW KA EC TG. Performed the experiments: NW KA EC. Analyzed the data: NW MCAA. Contributed reagents/materials/analysis tools: TG. Wrote the paper: NW KA EC MCAA TG.

References

1. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, et al. (2010) A Draft Sequence of the Neandertal Genome. *Science* 328(5979): 710–722.
2. Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, et al. (2010) Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities. *PLoS Biol* 8(11): e1000536.
3. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, et al. (2010) Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* 463(7282): 757–762.
4. Rasmussen M, Guo X, Wang Y, Lohmueller KE, Rasmussen S, et al. (2011) An Aboriginal Australian Genome Reveals Separate Human Dispersals into Asia. *Science* 334(6052): 94–98.
5. Jaenicke-Després V, Buckler ES, Smith BD, Gilbert MTP, Cooper A, et al. (2003) Early Allelic Selection in Maize as Revealed by Ancient DNA. *Science* 302(5648): 1206–1208.
6. Larson G, Liu R, Zhao X, Yuan J, Fuller D, et al. (2010) Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proc Natl Acad Sci USA* 107(17): 7686–7691.
7. Anderson-Carpenter L, McLachlan J, Jackson S, Kuch M, Lumibao C, et al. (2011) Ancient DNA from lake sediments: Bridging the gap between paleoecology and genetics. *BMC Evol Biol* 11(1): 30.
8. Schlumbaum A, Tensen M, Jaenicke-Després V (2008) Ancient plant DNA in archaeobotany. *Veg Hist Archaeobot* 17(2): 233–244.
9. Palmer SA, Smith O, Allaby RG (2012) The blossoming of plant archaeogenetics. *Ann Anat* 20: 146–156.
10. Wales N, Allaby RG, Willerslev E, Gilbert MTP (2013) Ancient Plant DNA. In: Elias SA, editor. *Encyclopedia of Quaternary Science*, vol. 2. Oxford: Elsevier. 705–715.
11. Gugerli F, Parducci L, Petit RJ (2005) Ancient plant DNA: review and prospects. *New Phytol* 166(2): 409–418.
12. Willerslev E, Cooper A (2005) Ancient DNA. *Proc Biol Sci* 272(1558): 3–16.
13. Japelaghi R, Haddad R, Garoosi G (2011) Rapid and Efficient Isolation of High Quality Nucleic Acids from Plant Tissues Rich in Polyphenols and Polysaccharides. *Mol Biotechnol* 49: 1–9.
14. Matheson CD, Gurney C, Esau N, Lehto R (2010) Assessing PCR Inhibition from Humic Substances. *Open Enzym Inhib J* 3: 38–45.
15. Rohland N, Hofreiter M (2007) Comparison and optimization of ancient DNA extraction. *BioTechniques* 42(3): 343–352.
16. Pääbo S, Poinar H, Serre D, Jaenicke-Després V, Hebler J, et al. (2004) Genetic Analyses from Ancient DNA. *Annu Rev Genet* 38(1): 645–679.
17. Knapp M, Hofreiter M (2010) Next Generation Sequencing of Ancient DNA: Requirements, Strategies and Perspectives. *Genes* 1(2): 227–243.
18. Polz MF, Cavanaugh CM (1998) Bias in Template-to-Product Ratios in Multitemplate PCR. *Appl Environ Microbiol* 64(10): 3724–3730.
19. Ávila-Arcos M, Cappellini E, Romero-Navarro JA, Wales N, Moreno-Mayar JV, et al. (2011) Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. *Sci Rep* 1: 74.
20. Monroe C, Grier C, Kemp BM (2013) Evaluating the efficacy of various thermostable polymerases against co-extracted PCR inhibitors in ancient DNA samples. *Forensic Sci Int* 228(1–3): 142–153.
21. Técher D, Martinez-Chois C, D'Innocenzo M, Laval-Gilly P, Bennisroune A, et al. (2010) Novel perspectives to purify genomic DNA from high humic acid content and contaminated soils. *Sep Purif Technol* 75(1): 81–86.
22. Hänni C, Brousseau T, Laudet V, Stehelin D (1995) Isopropanol precipitation removes PCR inhibitors from ancient bone extracts. *Nucleic Acids Res* 23(5): 881–882.
23. Kemp BM, Monroe C, Smith DG (2006) Repeat silica extraction: a simple technique for the removal of PCR inhibitors from DNA extracts. *J Archaeol Sci* 33(12): 1680–1689.
24. Pandey RK, Singh DP, Sudhakar G, Rao VR (2011) Ethanol re-precipitation removes PCR inhibitors from ancient DNA extract. *Antrocom Online Journal of Anthropology* 7(2): 173–179.
25. Wales N, Romero-Navarro JA, Cappellini E, Gilbert MTP (2012) Choosing the Best Plant for the Job: A Cost-Effective Assay to Prescreen Ancient Plant Remains Destined for Shotgun Sequencing. *PLoS ONE* 7(9): e45644.
26. Pääbo S (1989) Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. *Proc Natl Acad Sci USA* 86(6): 1939–1943.
27. Hansen AJ, Willerslev E, Wiuf C, Mourier T, Arcander P (2001) Statistical Evidence for Mismatching Lesions in Ancient DNA Templates. *Mol Biol Evol* 18(2): 262–265.
28. Hofreiter M, Jaenicke V, Serre D, von Haeseler A, Pääbo S (2001) DNA sequences from multiple amplifications reveal artifacts induced by cytosine deamination in ancient DNA. *Nucleic Acids Res* 29(23): 4793–4799.
29. Gilbert MTP, Hansen AJ, Willerslev E, Rudbeck L, Barnes I, et al. (2003) Characterization of Genetic Mismatching Lesions Caused by Postmortem Damage. *Am J Hum Genet* 72(1): 48–61.
30. Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S (2001) Ancient DNA. *Nat Rev Genet* 2(5): 353–359.
31. Ginolhac A, Rasmussen M, Gilbert MTP, Willerslev E, Orlando L (2011) mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* 27(15): 2153–2155.
32. Cooper A, Poinar HN (2000) Ancient DNA: Do It Right or Not at All. *Science* 289(5482): 1139.
33. Allaby RG, O'Donoghue K, Sallares R, Jones MK, Brown TA (1997) Evidence for the survival of ancient DNA in charred wheat seeds from European archaeological sites. *Anc Biomol* 1(2): 119–129.
34. Brown TA, Allaby RG, Sallares R, Jones G (1998) Ancient DNA in charred wheats: Taxonomic identification of mixed and single grains. *Anc Biomol* 2(2): 185–193.
35. Bunning SL, Jones G, Brown TA (2012) Next generation sequencing of DNA in 3300-year-old charred cereal grains. *J Archaeol Sci* 39(8): 2780–2784.
36. Zohary D, Hopf M, Weiss E (2012) Domestication of Plants in the Old World: The origin and spread of domesticated plants in south-west Asia, Europe, and the Mediterranean Basin. Oxford: Oxford University Press. 316 p.
37. Gilbert MTP, Wilson AS, Bunce M, Hansen AJ, Willerslev E, et al. (2004) Ancient mitochondrial DNA from hair. *Curr Biol* 14(12): R463–R464.
38. Poinar HN, Hofreiter M, Spaulding WG, Martin PS, Stankiewicz BA, et al. (1998) Molecular Coproscopy: Dung and Diet of the Extinct Ground Sloth *Nototheriopsis shastensis*. *Science* 281(5375): 402–406.
39. Palmer SA, Moore JD, Clapham AJ, Rose P, Allaby RG (2009) Archaeogenetic Evidence of Ancient Nubian Barley Evolution from Six to Two-Row Indicates Local Adaptation. *PLoS ONE* 4(7): e6301.
40. Pääbo S, Gifford JA, Wilson AC (1988) Mitochondrial DNA sequences from a 7000-year old brain. *Nucleic Acids Res* 16(20): 9775–9787.
41. Mardis ER (2008) Next-Generation DNA Sequencing Methods. *Annu Rev Genomics Hum Genet* 9(1): 387–402.
42. Rasmussen M, Cummings LS, Gilbert MTP, Bryant V, Smith C, et al. (2009) Response to Comment by Goldberg et al. on “DNA from Pre-Clovis Human Coprolites in Oregon, North America”. *Science* 325(5937): 148–d.
43. Andersen K, Bird KL, Rasmussen M, Haile J, Breuning-Madsen H, et al. (2012) Meta-barcoding of ‘dirt’ DNA from soil reflects vertebrate biodiversity. *Mol Ecol* 21(8): 1966–1979.
44. Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, et al. (2011) Geneious. Version 5.5.7.
45. SPSS. (2009) PASW Statistics. Version 18.0.
46. Desjardins P, Conklin D (2010) NanoDrop Microvolume Quantitation of Nucleic Acids. *J Vis Exp* 45(45): e2565.
47. Bertram JG, Oertel K, Petruska J, Goodman MF (2010) DNA Polymerase Fidelity: Comparing Direct Competition of Right and Wrong dNTP Substrates with Steady State and Pre-Steady State Kinetics. *Biochemistry* 49(1): 20–28.
48. Palmer SA, Clapham AJ, Rose P, Freitas FO, Owen BD, et al. (2012) Archaeogenomic Evidence of Punctuated Genome Evolution in *Gossypium*. *Mol Biol Evol* 29: 2031–2038.

49. Banerjee M, Brown TA (2002) Preservation of Nuclear but not Chloroplast DNA in Archaeological Assemblages of Charred Wheat Grains. *Anc Biomol* 4(2): 59–63.
50. Schlumbaum A, van Glabeke S, Roldan-Ruiz I (2012) Towards the onset of fruit tree growing north of the Alps: Ancient DNA from waterlogged apple (*Malus* sp.) seed fragments. *Ann Anat* 194(1): 157–162.
51. Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Mol Biol* 5(2): 69–76.
52. Mukherjee A, Roy S, De Bera S, Jiang H, Li X, et al. (2008) Results of molecular analysis of an archaeological hemp (*Cannabis sativa* L.) DNA sample from North West China. *Genet Resour Crop Evol* 55(4): 481–485.
53. Li C, Lister DL, Li H, Xu Y, Cui Y, et al. (2011) Ancient DNA analysis of desiccated wheat grains excavated from a Bronze Age cemetery in Xinjiang. *J Archaeol Sci* 38: 115–119.
54. Dabney J, Meyer M (2012) Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *BioTechniques* 52(2): 87–94.
55. Briggs AW, Stenzel U, Meyer M, Krause J, Kircher M, et al. (2010) Removal of deaminated cytosines and detection of *in vivo* methylation in ancient DNA. *Nucleic Acids Res* 38(6): e87.
56. Elbaum R, Melamed-Bessudo C, Boaretto E, Galili E, Lev-Yadun S, et al. (2006) Ancient olive DNA in pits: preservation, amplification and sequence analysis. *J Archaeol Sci* 33(1): 77–88.
57. Giles RJ, Brown TA (2008) Improved methodology for extraction and amplification of DNA from single grains of charred wheat. *J Archaeol Sci* 35(9): 2585–2588.
58. Cappellini E, Gilbert MTP, Geuna F, Fiorentino G, Hall A, et al. (2010) A multidisciplinary study of archaeological grape seeds. *Naturwissenschaften* 97(2): 205–217.
59. Haile J, Froese DG, MacPhee RDE, Roberts RG, Arnold LJ, et al. (2009) Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proc Natl Acad Sci USA* 106: 22363–22368.