



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL INSTITUTO
POLITÉCNICO NACIONAL

LABORATORIO NACIONAL DE GENÓMICA PARA LA BIODIVERSIDAD

Unidad Irapuato

*Análisis de perfiles de metilación en el paleogenoma de maíz: implicaciones
para el estudio de su adaptación a cambios ambientales.*

Tesis que presenta

I.Q. Sara Edith Garcia Morales

Para obtener el grado de

Maestra en Ciencias

En la especialidad de

Biología Integrativa

Directores de Tesis:

Dr. Jean Philippe Vielle Calzada

Dr. Rafael Montiel Duarte

Asesores:

Dra. Gertrud Lund

Dr. Alexander Correa-Metrio

El presente trabajo fue realizado en el grupo de Interacción Núcleo Mitocondrial y Paleogenómica y el grupo de Desarrollo Reproductivo y Apomixis del Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO) del Centro de Investigación y de Estudios Avanzados del Politécnico Nacional, CINVESTAV-IPN Unidad Irapuato, bajo la codirección del Dr. Rafael Montiel Duarte y del Dr. Jean Philippe Vielle Calzada así como la asesoría de la Dra. Gertrud Lund y el Dr. Alexander Correa-Metrio. Además se contó con el apoyo de CONACYT siendo el becario 326348.

AGRADECIMIENTOS

Agradezco a mi familia por su apoyo incondicional durante mi formación personal y profesional. Sin ellos, este trabajo no habría sido posible.

Agradezco a mis directores de tesis Dr. Jean-Philippe Vielle Calzada y Dr. Rafael Montiel Duarte, al permitirme formar parte de sus grupos de trabajo y por enseñarme con paciencia durante toda mi trayectoria en la maestría, además de su calidad humana y de compartirme su pasión por la ciencia.

A los miembros del comité tutorial Dra. Gertrud Lund y Dr. Alexander Correa-Metrio por compartir sus conocimientos conmigo, sus comentarios y permitir una discusión abierta para realizar este proyecto integrativo con componentes de áreas sumamente distintas.

A todos aquellos que forjaron las técnicas y protocolos paleogenómicos y quienes excavaron las muestras utilizadas en colaboración y bajo la guía del profesor Ángel García Cook ya que este proyecto tiene como cimientos todo ese trabajo previo.

A CINVESTAV Irapuato y el programa de Biología Integrativa por otorgar la flexibilidad de un programa de maestría sumamente enriquecedor y a CONACYT por la beca otorgada.

A todos los miembros de los laboratorios 1 y 13 por su apoyo incondicional, consejos, amistad y críticas constructivas que siempre tuvieron a bien compartirme. Agradezco el apoyo brindado por el técnico del laboratorio 13 de LANGEBIO, Christian Martínez González durante los procesos administrativos y su apoyo en programación. A la técnico del laboratorio 4 (epigenética) de CINVESTAV, BQ Dalia Rodríguez por su apoyo durante la elaboración de la técnica de bisulfito y a Miguel Vallebuena por discutir mi proyecto de manera exhaustiva a cada paso de su realización.

A todos mis amigos que me apoyaron durante esta trayectoria en CINVESTAV y LANGEBIO y a mis amigos de la infancia pero sobre todo, a Juan Carlos Robles y a Daniel Abel Madrigal por demostrarme que a pesar de las adversidades, siempre debemos de luchar por nuestros ideales.

ÍNDICE DE CONTENIDO

ÍNDICE DE TABLAS	1
ÍNDICE DE FIGURAS	1
RESUMEN	4
ABSTRACT	4
CAPÍTULO I. ANTECEDENTES	5
A) LAS RESPUESTAS Y ADAPTACIÓN DE LAS PLANTAS AL AMBIENTE.	5
A.1 Definición del ambiente	
A.1.1 Componente abiótico: sequía	
B) ADAPTACIÓN DE PLANTAS AL AMBIENTE	6
B.1 La respuesta fenotípica en plantas bajo estrés	
B.1.1 Respuesta fisiológica	
B.1.2 Respuesta genética	7
B.1.3 Respuesta epigenética	
B.1.3.1 Respuesta epigenética en plantas expuestas a estrés abiótico.	9
C) EL CLIMA: EL PRINCIPAL COMPONENTE ABIÓTICO DEL AMBIENTE.	10
C.1 El paleoclima: Inferencias a partir de evidencias indirectas	
C.1.1 Cambio climático durante el Neolítico	12
C.1.1.1 Forzamientos externos del Neolítico.	13
D) PRESENTE Y PASADO DEL AMBIENTE EN TEHUACÁN.	16
D.1 El clima en México	
D.1.1 El clima en Tehuacán	17
D.2 La agricultura durante el Neolítico.	20
D.2.1 La historia de la agricultura de Tehuacán.	
D.2.2 La historia del maíz en el Valle de Tehuacán.	21
E) EL GENOMA Y SU RESPUESTA AL AMBIENTE.	21
E.1 Particularidades del genoma de maíz.	
E.2 Los procesos de respuesta genómica y la importancia de la metilación.	22
E.3 La importancia del paleogenoma para entender la respuesta al ambiente.	
E.3.1 Paleoepigenética.	24
F) ESTRATEGIAS PARA EL ESTUDIO DE LA ADAPTACIÓN DE PLANTAS ANTIGUAS AL AMBIENTE.	
F.1 Estrategia geológica para inferir la historia climática en Tehuacán.	
F.2 Estrategia paleogenómica para estudiar la estructura genética del maíz antiguo.	26
F.2.1 Paleoepigenética para estudiar la respuesta ambiental del maíz.	
CAPÍTULO II. HIPÓTESIS Y OBJETIVOS	27
2.1 HIPÓTESIS DE TRABAJO	
2.2 OBJETIVOS	

CAPÍTULO III. MATERIALES Y MÉTODOS	27
3.1 REVISIÓN SISTEMÁTICA DE BIBLIOGRAFÍA	28
3.1.1 Criterios de clasificación de los estudios paleoclimáticos.	
i) <i>Proxys</i> estudiados.	29
ii) Tecnología para analizar datos.	30
iii) Análisis estadístico.	
iv) Objetividad en la discusión.	31
v) Datos continuos en la temporalidad de interés.	
3.2 SELECCIÓN DE MUESTRAS ARQUEOBOTÁNICAS.	32
3.2.1 Excavación 2012.	
3.2.2 Métodos de datación.	33
3.2.2.1 Datación por radioisótopos	
3.2.2.2 Principios de datación por ¹⁴ C	
3.2.2.3 Procedimientos de medición, materiales y problemáticas.	
3.2.2.4 Exactitud en las mediciones de radiocarbono.	34
3.2.2.5 Datación en las muestras de maíz.	35
3.3 EXTRACCIÓN ADN DE MUESTRAS PALEOBOTÁNICAS.	
3.3.1 Laboratorio de ADN antiguo.	
3.3.2 Extracción y purificación de ADN de maíz antiguo.	36
3.3.3 Amplificación diagnóstica para el gen Rubisco.	
3.4 ELABORACIÓN DE BIBLIOTECAS GENÓMICAS.	
3.4.1 Tratamiento con Bisulfito.	38
3.5 SECUENCIACIÓN MASIVA.	39
3.6 ANÁLISIS BIOINFORMÁTICO.	
CAPÍTULO IV. RESULTADOS	44
4.1 REVISIÓN SISTEMÁTICA DE BIBLIOGRAFÍA.	
4.2 MUESTRAS DE TEHUACÁN Y SU DATACIÓN.	47
4.3 EXTRACCIÓN DE ADN DE MUESTRAS PALEOBOTÁNICAS.	50
4.3.1 Amplificación diagnóstica para el gen Rubisco.	
4.4 ELABORACIÓN DE BIBLIOTECAS GENÓMICAS.	51
4.4.1 Tratamiento con Bisulfito.	53
4.5 SECUENCIACIÓN MASIVA.	54
4.6 ANÁLISIS BIOINFORMÁTICO.	55
4.6.1 Filtrado de calidad de secuencias.	56
4.6.2 Análisis de calidad de bibliotecas filtradas.	57
4.6.3 Mapeo genómico	61
4.6.3.1 Establecimiento de las condiciones óptimas de mapeo.	
4.6.4 Mapeo de <i>reads</i> contra genoma de referencia.	64

4.6.5 Estimación del daño molecular.	65
4.6.6 Búsqueda de sitios metilados.	65
CAPÍTULO V. DISCUSIÓN	69
5.1 PALEOCLIMA EN EL VALLE DE TEHUACÁN	
5.2 ADN ANTIGUO DE MUESTRAS PALEOBOTÁNICAS	70
5.3 BIBLIOTECAS GENÓMICAS	71
5.4 SECUENCIACIÓN MASIVA	72
5.4.1 Problemática durante el análisis bioinformático.	
CAPÍTULO VI. CONCLUSIONES	73
CAPÍTULO VII. BIBLIOGRAFÍA	74
7.1 BIBLIOGRAFÍA PARA REVISIÓN PALEOCLIMÁTICA	77
CAPÍTULO VII. ANEXOS	84

ÍNDICE DE TABLAS

Tabla 1. Principales metiltransferasas implicadas en la metilación de ADN de *Arabidopsis thaliana* y *Zea mays*.

Tabla 2. Fuentes principales de *proxys* con datos para reconstrucciones paleoclimáticas.

Tabla 3. Características de los archivos naturales en donde se encuentran *proxys*.

Tabla 4. Número de estudios paleoclimáticos categorizados mediante criterios de análisis.

Tabla 5. Vida media de radioisótopos usados en datación.

Tabla 6. Adaptadores para estrategias de *pooling* en kits TruSeq de ADN.

Tabla 7. Relación entre el Phred score (valor de calidad) y la exactitud de la secuenciación por sitio nucleotídico.

Tabla 8. Metilomas comparativos en líneas de maíz moderno.

Tabla 9. Caracterización de las extracciones de ADN en muestras paleobotánicas.

Tabla 10. Número de *reads* obtenidos de la secuenciación y *reads* sin duplicados de PCR.

Tabla 11. Caracterización de las bibliotecas genómicas.

Tabla 12. Caracterización de los fragmentos tratados con bisulfito.

Tabla 13.

ÍNDICE DE FIGURAS

Figura 1. Comparativa de fenotipos de peloria con respecto a *Linaria vulgaris*.

Figura 2. Esquema de metilación simétrica como resultado de la interacción con las distintas metiltransferasas.

Figura 3. Esquema general de los tres forzamientos orbitales.

Figura 4. La oblicuidad de la elíptica modula la insolación en distintas latitudes terrestres.

Figura 5. Impacto de los distintos forzamientos orbitales a lo largo del último millón de años antes del presente.

Figura 6. Distintos movimientos atmosféricos afectan el clima en México dependiendo de la estacionalidad.

Figura 7. Climatograma de Tehuacán, Puebla.

Figura 8. Mapa del Valle de Tehuacán.

Figura 9. Esquema de un fragmento de ADN en doble hélice con las cuatro bases nucleotídicas y los sitios de posible daño.

Figura 10. Mapa de la zona que comprende Mesoamérica.

Figura 11. Regiones de Mesoamérica de interés.

Figura 12. Diagrama de la conversión de bisulfito.

Figura 13. Esquema de la conversión mediante bisulfito de citosinas a uracilos.

- Figura 14. Esquema metodológico para el análisis comparativo de patrones de metilación en ADN antiguo.
- Figura 15. Mapa de estudios del Holoceno en México.
- Figura 16. Plano de planta del cuarto limpio en LANGEBIO-CINVESTAV Irapuato.
- Figura 17. Esquema metodológico modificado para la elaboración de bibliotecas genómicas
- Figura 18. Esquema de las alícuotas utilizadas durante la elaboración del tratamiento con bisulfito.
- Figura 19. Esquema de lectura de un segmento *paired-end*.
- Figura 20. Esquema general del procesamiento informático.
- Figura 21. Diagrama de la estrategia seguida para la identificación de alineamientos únicos.
- Figura 22. Diagrama representativo del discernimiento entre daño molecular con respecto a tratamiento de bisulfito.
- Figura 23. Mapa con la ubicación de los *proxys* analizados en los estudios paleoclimáticos selectos.
- Figura 24. Gráficas que muestran el estándar de crecimiento en los anillos de Ahuehuate estandarizados con PDSI con respecto a los años de la era común.
- Figura 25. Gráfica de los valores de isótopos ^{18}O en Aljojuca.
- Figura 26. Gráfica con los valores de isótopos ^{18}O en Juxtlahuaca.
- Figura 27. Mapa de México marcando las zonas con reconstrucción paleoclimática.
- Figura 28. Curva de calibración IntCal09 con edades corregidas de ^{14}C (‰) con una desviación estándar.
- Figura 29. Curva de calibración de radiocarbono a años calendario calibrados para SM1.
- Figura 30. Curva de calibración de radiocarbono a años calendario calibrados para Cox17.
- Figura 31. Muestras datadas de cada estrato temporal excavado en el Valle de Tehuacán.
- Figura 32. Gel de electroforesis de los extractos paleobotánicos con marcador de peso molecular de una escalera de 100 pb a los costados.
- Figura 33. Análisis cualitativo de electroforesis para los extractos de Cox17 y SM1.
- Figura 34. Resultados de los análisis cualitativos de electroforesis para seleccionar fragmentos antiguos.
- Figura 35. Resultados de la elaboración de bibliotecas mediante análisis cualitativos de electroforesis.
- Figura 36. Esquemas con calidad inicial de las bibliotecas secuenciadas.
- Figura 37. Esquema de calidad global de los *reads* de la biblioteca Cox17 en Fastqc.
- Figura 38. Esquema de la calidad global de los *reads* sin adaptadores ni duplicados de PCR.
- Figura 39. Número de secuencias con el porcentaje de GC en las bibliotecas.
- Figura 40. Gráfica con porcentaje de ácidos nucleicos en relación con la posición global en los reads de ambos clústeres.

- Figura 41 . Selección de reads con porcentaje menor a 45% de la biblioteca con tratamiento de bisulfito.
- Figura 42. Esquema de la calidad global de los reads con porcentaje menor a 45% GC de BisCox17.
- Figura 43. Número de secuencias con el porcentaje de GC de la muestra menor a 45% GC de BisCox17.
- Figura 44. Gráfica con porcentaje de ácidos nucleicos en relación con la posición en los reads de la muestra menor a 45% GC Biscox17.
- Figura 45. Gráficas con los valores de mapeo obtenidos con las dos modalidades de BWA mem.
- Figura 46. Esquema de la problemática del daño en muestras antiguas tratadas con bisulfito.
- Figura 47. Variación del parámetro A que equivale a la igualdad o *match* de un *read* con respecto a la referencia.
- Figura 48. Variación en el parámetro B que equivale a la penalización por *mismatch* de un *read* con respecto a la referencia.
- Figura 49. Gráfica con estimador de daño molecular.
- Figura 50. Gráficas con las frecuencias de bases nucleotídicas para regiones antiguas de Cox17 en yuxtaposición con regiones modernas.
- Figura 51. Esquema del visor IGV con un segmento del cromosoma 1 con cambios en bases nucleotídicas en reads con tratamiento de bisulfito.
- Figura 52. Visor IGV con un sitio del cromosoma 1 con tratamiento de bisulfito.
- Figura 53. Visor IGV con un sitio del cromosoma 8 con tratamiento de bisulfito.
- Figura 54. Visor IGV con un segmento del cromosoma 8 con tratamiento de bisulfito.
- Figura 55. Mapa de las zonas lacustres aledañas al Valle de Tehuacán.

RESUMEN

Varios estudios sugieren que el maíz fue domesticado a partir del teosinte (*Zea mays* spp. *parviglumis*) en la región del Balsas a través de eventos de domesticación que iniciaron en el centro de México alrededor de 9 000 años antes del presente. Aunque los mecanismos de desarrollo que dieron lugar a estas transformaciones fenotípicas han estudiado desde una perspectiva antropocéntrica, se conoce muy poco sobre los factores ambientales que influenciaron este proceso evolutivo.

En este proyecto, se obtuvo una selección de muestras paleobotánicas para realizar el análisis genético de temporalidades correspondientes a periodos de sequía reportados con base en información geológica. A partir del genoma antiguo de una muestra de maíz se utilizó la metodología de bisulfito para su posterior secuenciación masiva. Al explorar los perfiles de metilación identificados se concluye que es posible identificar citosinas metiladas en función de la preservación del ADN endógeno.

ABSTRACT

Several studies suggest that maize was domesticated from Balsas teosinte (*Zea mays* spp. *parviglumis*) through domestication events that initiated in Central Mexico circa 9,000 years before present (BP). Although the developmental mechanisms that gave rise to this phenotypic transformation are studied from a perspective centered in human selection, little is known about the environmental factors that influenced this evolutionary process. In this project, the genomic methylation patterns of ancient genomes are explored using a conventional bisulfite technique and massive sequencing. Few studies of this type have been conducted in ancient crops, as the rate of DNA deamination must be taken into consideration to avoid artifacts resulting from diagenetic processes.

In this project, we aim to better understand the influence of environmental factors in the genetic and epigenetic constitution of ancient maize. A paleoclimatic reconstruction from the Holocene was made based on geological reports. We then selected paleobotanical maize samples that presumably grew in drought conditions 1500 years BP. Selected paleobotanical maize samples were bisulfite-treated and further sequenced to identify methylated cytosines. We concluded that methylation profiles can be identified in ancient maize samples if endogenous ancient DNA is preserved. In the future, correlating methylation profiles of paleo-botanical maize samples with the environment may allow the identification of adaptive effects that could have environmentally influenced the domestication of maize during the Holocene in Mesoamerica.

CAPÍTULO I. ANTECEDENTES

A. Las respuestas y adaptación de las plantas al ambiente.

A.1 Definición del ambiente

Las plantas han evolucionado un amplio rango de mecanismos que les permiten lidiar con diversos factores ambientales, que pueden ser categorizados tanto como componentes bióticos como abióticos. Se define como componente biótico todos aquellos organismos simbioses, parásitos, patógenos, herbívoros y plantas con las que compitan por recursos para un óptimo desarrollo. El componente abiótico se define como todos aquellos recursos que son determinantes para el crecimiento de las plantas: luz, humedad, temperatura, agua, nutrientes minerales y dióxido de carbono además de viento, radiación o contaminación (Koyro *et al.* 2002). Dentro de esta categoría podríamos mencionar algunos ejemplos tales como temperatura (desde calor hasta el congelamiento), radiación muy alta o muy baja, estancamiento de agua o falta de ella, falta de nutrientes en el sustrato de crecimiento, salinidad excesiva o sequía. Aunque no siempre se trata de condiciones extremas a las que se someten las plantas, las fluctuaciones ambientales son importantes para su correcto desarrollo y el déficit o exceso de cualquier factor es crucial.

El efecto de cada factor ambiental depende de la cantidad o intensidad recibida en las plantas por lo que se requiere un equilibrio óptimo para su adecuado crecimiento, desarrollo o productividad (Bray *et al.* 2000). Al déficit de cualquier factor ambiental para el adecuado desarrollo de las plantas se le denomina estrés, y al ser percibido por las plantas, comienzan las señales que dan lugar a diversas respuestas de aclimatación.

A.1.1 Componente abiótico: sequía

El estrés ambiental es la principal razón de pérdida de cultivos a nivel mundial. Los principales factores que reducen la producción de cultivos son la sequía y el déficit de nutrientes pues conducen a la desertificación de los campos (Bray *et al.* 2000; Wang *et al.* 2003; Evans, 2005; Vinocur & Altman, 2005).

El concepto de sequía implica cambio de temperatura y estrés hídrico. El aumento en la temperatura es un factor ambiental abiótico que afecta los procesos biológicos de los organismos. Independientemente de tratarse de una exposición de corto o largo plazo, hay modificación de propiedades de las membranas, niveles de actividad enzimática, reacciones químicas y difusión, viscosidad de citoplasma, floema y xilema en plantas (Sung *et al.* 2003).

El estrés hídrico causa cambios en componentes celulares de las plantas (el correcto funcionamiento de proteínas y membranas dependen de una adecuada hidratación) lo que genera estrés osmótico el cual, a nivel bioquímico se traduce como un mal funcionamiento en la síntesis de aminoácidos, metabolismo de proteínas, fotosíntesis o respiración celular (Larcher 2001; Schulze *et al.* 2002).

Otro factor importante a tomar en cuenta que causa la desertificación del campo es el crecimiento de la población humana. En las últimas décadas se ha generado presión en el uso de campos cultivados, explotando diversos recursos para su mantenimiento pero al sumarle el efecto producido por el calentamiento global, se genera tal estrés ambiental que termina impactando directamente en los cultivos para consumo humano en todo el planeta (Ericson *et al.* 1999; Gleick 2000).

B. Adaptación de plantas al ambiente

La adaptación es la evolución de una población dadas condiciones ambientales alteradas, mientras que la aclimatación es cualquier respuesta fenotípica de individuos a un cambio ambiental no favorable o estrés, como la sequía, alterando su aptitud biológica. Las respuestas de aclimatación de los individuos permiten esa aptitud biológica que resultan en una posterior adaptación en una población a lo largo de varias generaciones (Angilleta, 2009; Pörtner, 2010).

B.1 La respuesta fenotípica en plantas bajo estrés

Las plantas estresadas muestran un amplio rango de respuestas y manifiestan mecanismos que pueden discernirse en toda la planta a simple vista o a nivel celular y molecular. Con un enfoque a sequía, se describen a continuación las respuestas de manera general en las plantas.

B.1.1 Respuesta fisiológica

El desarrollo vegetal está determinado por una amplia variedad de factores ambientales que tienen efectos combinados, muy rara vez se presentan de manera independiente. En el caso de estrés por sequía, las altas temperaturas podrían estar implicadas al igual que el estrés por carencia de agua y en consecuencia, estrés oxidativo (Fitter & Hay 2002).

La viabilidad de plantas en hábitat seco depende en parte de su habilidad para lidiar con el déficit de agua debido al bajo potencial del suelo, la restricción de absorción de dióxido de carbono y déficit de nutrientes. El déficit de agua causa cambios en los componentes celulares, ya que la conformación biológicamente activa y correcto funcionamiento de las células vegetales depende de una correcta hidratación (Santarius 1973; Berry & Bjorkman 1980). En caso de existir desequilibrio hídrico la planta entra en estrés osmótico. Si llega a ser severo, puede dar pie a una síntesis truncada de aminoácidos y alteración del metabolismo de proteínas, la fotosíntesis o la respiración celular (Larcher, 2001; Schulze *et al.* 2002).

Por otro lado, las altas temperaturas son determinantes para la biología de las plantas, al modificar propiedades de las membranas, niveles de actividad enzimática, reacciones químicas, difusión, viscosidad dentro de vacuolas así como en el citoplasma, floema y xilema, interfiriendo con la fotosíntesis, respiración, aumento en el índice de evapotranspiración, reducción del potencial osmótico de las hojas e incremento de fluorescencia en clorofila (Sung *et al.* 2003; Tsukaguchi *et al.* 2003; Huve *et al.* 2005; Taiz & Zeiger 2006).

El concepto de alta temperatura es una función compleja dada por la intensidad, duración y velocidad en el incremento de temperatura. Para las plantas en general, temperaturas que aumenten valores de 10-15 °C sobre el óptimo es considerado estrés por temperatura (Larkindale *et al.* 2005). Entiéndase como tolerancia al calor la habilidad de las plantas para crecer y producir rendimiento a altas temperaturas (Fitter & Hay 2002). Para lograr la tolerancia a temperaturas altas, las plantas tienen un par de mecanismos adaptativos fisiológicos: la habilidad de prevenir aumento en la temperatura excesiva en sus tejidos al disipar sus efectos (ejemplos de ello son la reducción de la absorción de radiación solar ya sea por la presencia de tricomas y el paraheliotropismo, que es el crecimiento en paralelo a los rayos solares) y la tolerancia de calor en el protoplasma (ejemplo de lo anterior es la síntesis de proteínas resistentes al calor o HSP por sus siglas en inglés) (Kota *et al.* 2007; Al-Whaibi 2010).

B.1.2 Respuesta genética

El estrés por calor también genera otro tipo de defensas a nivel molecular, a través de la regulación transcripcional (Singh *et al.* 2002). Esto es debido a la inducción de numerosos genes que codifican factores transcripcionales involucrados en la respuesta al estrés por calor y su tolerancia (Chen & Zhu 2004; Kotak *et al.* 2007). Diversos estudios revelan que algunos genes son regulados por estrés abiótico ambiental (Kawasaki *et al.* 2001; Provar *et al.* 2003; Nogueira *et al.* 2003).

La tolerancia al calor tiene un carácter multigenético ya que el shock térmico induce la transcripción de muchos genes, los cuales se atribuyen a elementos de shock térmico (HSE, por sus siglas en inglés) (Dong *et al.* 2003). Algunos de los genes expresados incluyen genes de proteínas de shock térmico (hsp), dehidrinas (dhn), genes asociados con senescencia (sag) y genes *staygreen* (sgr) (Wahid *et al.* 2007; Vilardell *et al.* 1990; Godoy *et al.* 1994; Smart *et al.* 1995; Rajcan & Tollenar 1999).

B.1.3 Respuesta epigenética

Se denomina una respuesta epigenética a cualquier cambio heredable que no depende de modificaciones en la secuencia de ADN (Berger *et al.* 2009). Las interacciones moleculares a nivel del genoma con relación al ambiente son importantes para comprender cómo afectan los factores ambientales a las plantas. Dichas interacciones promueven, en algunos casos, cambios en el ADN o mutaciones que permitan a las generaciones sucesivas la capacidad de aclimatarse al estrés ambiental aumentando el potencial de adaptarse.

La epigenética modifica la expresión de los genes sin necesidad de mutar la secuencia del ADN (Holliday, 1994; Esteller 2006). La metilación ocurre tanto en animales como en plantas. En particular para plantas, se ha reportado que tiene un papel importante en diversos procesos, los cuales incluyen: paramutaciones, en las que se describen cambios en la expresión de un alelo con respecto a la expresión diferencial de otro alelo en la misma secuencia; impronta genética, la cual se caracteriza por la expresión selectiva de genes heredados ya sea por línea materna o paterna; modificaciones de histonas; y de particular interés para este proyecto, la metilación de ADN, que es una adición covalente de un grupo metilo (CH₃) a la posición 5' de las citosinas del genoma nuclear. (Gehring *et al.* 2009; Su *et al.* 2011; Waters *et al.* 2011). En algunos casos, las modificaciones en la metilación de ADN permiten observar efectos fenotípicos, como por ejemplo, la floración radial de *Linaria vulgaris* (Figura 1).

En la naturaleza, la metilación permite regular la expresión genética en varios procesos biológicos a nivel molecular (Razin & Riggs, 1980; Fang *et al.* 2007). La variación de metilación en el ADN (además de las mutaciones genéticas) crea diversidad genética sobre la cual actúa la selección natural (Eichten *et al.* 2012; Li *et al.* 2014; Kawakatsu *et al.* 2016).



Figura 1. Comparativa de fenotipos de peloria con respecto a *Linaria vulgaris*.

La formación de epialelos da como resultado variación fenotípica en el reino vegetal como en el caso de peloria con respecto a *Linaria vulgaris* en donde peloria se encuentra metilada mientras que *Linaria vulgaris* está demetilada.

En plantas, la metilación ocurre en dos contextos simétricos: CG y CHG (donde H representa adenina, timina o citosina) y un contexto asimétrico CHH (Figura 2). Diversas DNA metiltransferasas actúan para mantener la metilación y metilar de novo, En el caso de la metilación de novo como su nombre lo indica, se metilan citosinas previamente no metiladas para formar nuevos patrones de metilación. En contraste, la metilación de mantenimiento, los patrones de metilación presentes son mantenidos después de la replicación (Pavlopoulou & Kossida, 2007; Matzke & Mosher, 2014).

En el caso de los contextos simétricos, la metilación *de novo* ocurre una sola vez mientras que en el contexto asimétrico ocurre continuamente (Cao *et al.* 2000). El proceso que involucra la metilación *de novo* involucra RNA pequeños de interferencia y a la vía de RdDM (RNA-directed DNA methylation) (Zhong *et al.* 2014). En el mantenimiento de metilación, en el contexto CG las metiltransferasas mantienen citosinas metiladas durante la replicación del contexto simétrico CG, en tanto que las cromometilasas (CMT3) las mantienen en el contexto simétrico CHG (Pavlopoulou & Kossida, 2007; Stroud *et al.* 2014).

La familia de las enzimas glicosilasas del ADN que incluyen DEMETER y el factor de represión de silenciamiento 1 (ROS1 por sus siglas en inglés) están implicadas en el proceso de desmetilación, encargándose de eliminar el grupo metilo en citosinas de en *Arabidopsis* (Zhu, 2009; Zhang & Zhu, 2012).

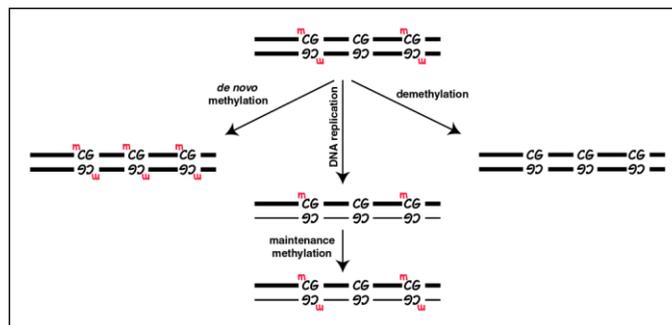


Figura 2. Esquema de metilación simétrica como resultado de la interacción con las distintas metiltransferasas.

Metiltransferasas y cromometiltransferasas involucradas en el mantenimiento de metilación; la ruta RdDM para el mantenimiento de metilación *de novo* y finalmente la pérdida de metilación por las ADN demetilasas. (Law & Jacobsen, 2011; Zhu 2009; Matzke & Mosher, 2014).

Aunque las enzimas encargadas de estos procesos de metilación y desmetilación han sido ampliamente estudiadas principalmente en *Arabidopsis*, también se han descrito homólogos en maíz (*Zea mays*) y la gran mayoría de estas rutas están conservadas y cumplen funciones similares, por ejemplo, CMT3 en *Arabidopsis* y sus homólogos Zmet2 y Zmet5 en maíz (Li *et al.* 2014) (Tabla 2).

<i>Arabidopsis thaliana</i>	Función	Homólogo en <i>Zea mays</i>
MET	Mantenimiento de metilación en contexto simétrico CG	Zmet1
CMT3	Mantenimiento de metilación en contexto simétrico CHG	Zmet2 & Zmet5
DRM1/DRM2	Metilación <i>de novo</i> en contexto asimétrico CHH	Zmet 3 & Zmet7
DRM3	Metilación <i>de novo</i> en contexto asimétrico CHH	Zmet6

Tabla 1. Principales metiltransferasas implicadas en la metilación de ADN de *Arabidopsis thaliana* y *Zea mays*. (Li *et al.* 2014)

En maíz, se tiene evidencia de genes altamente expresados que contienen altos niveles de metilación bajo contexto CHH, aún encontrándose a cientos de pares de bases de distancia del inicio del sitio de transcripción. En contraste, niveles altos de metilación en contexto CHG cercanos a los sitios de transcripción se asocian con reducción en la expresión genética (Gent *et al.* 2013). En resumen, la metilación del ADN por sí sola está asociada con activar, silenciar o regular la expresión de un gen dependiendo de los contextos metilados (Costello & Plass 2011; Duthie, 2011).

B.1.3.1 Respuesta epigenética en plantas expuestas a estrés abiótico.

Dada la naturaleza sésil de las plantas, los mecanismos para aclimatarse a diversas condiciones ambientales tienen una respuesta inmediata, gracias a los mecanismos epigenéticos. Diversos estudios en plantas sometidas a estrés ambiental han aportado evidencia de cambios consistentes en la metilación de ADN como respuesta a estrés biótico (Downen *et al.* 2012; Le *et al.* 2014).

En el caso de estrés abiótico, los estudios realizados toman en cuenta condiciones de estrés muy variables (por ej. salinidad y sequía) observando variación de metilación de ADN (Stroud *et al.* 2013; Stelpflug *et al.* 2014). Evidencia de que la variación en los niveles de metilación de ADN provee mecanismos para la adaptación de los organismos al estrés abiótico, fue estudiando dos accesiones de *A. thaliana* al crecer en dos temperatura distintas. Se observó un incremento de metilación en contexto CHH mientras que en el contexto CG se encontró altamente metilado para la accesión originaria de regiones frías. Se relacionaron estos patrones diferenciales de metilación con loci reguladores en trans, los cuales mostraban evidencia de adaptación local (Dubin *et al.* 2015). Adicionalmente, se ha reportado en una gran variedad de especies vegetales indicios de que la variación de la metilación responde a estrés abiótico (Richards, 2006; Finnegan, 2002; Steward *et al.* 2000; Kalisz & Purugganan, 2004; Lukens & Zhan, 2007; Richards, 2011).

Los metilomas de maíz actualmente reportados en la literatura científica han sido caracterizados mediante una amplia gama de técnicas para identificar metilación así como la interpretación de la misma. En algunos casos se ha usado la amplificación de polimorfismos sensible a metilación (MSAP, por sus siglas en inglés) para identificar cambios de metilación en distintos ambientes (Zheng *et al.* 2013; Jones, 2013; Wang *et al.* 2011). En otros estudios, se han reportado los perfiles a nivel genómico de la metilación del ADN de accesiones (B73 y Mo17) que crecieron bajo estrés de calor, frío, radiación UV, salinidad y condiciones de laboratorio (Tan, 2010; Stroud *et al.* 2013; Stelpflug *et al.* 2014; Eichten & Springer, 2015).

Gracias a la evidencia de respuesta a estrés abiótico, el estudio en los cambios de los perfiles de metilación permitirán estudiar la historia adaptativa de las plantas al ambiente (Feng *et al.* 2012; Sahu *et al.* 2013).

C. El clima: el principal componente abiótico del ambiente.

El clima es considerado como el promedio de las condiciones meteorológicas anuales en una zona geográfica durante los últimos treinta años. Diversos componentes y sus interacciones determinan el clima para una región: atmósfera, océanos, biosfera, superficie terrestre y criósfera. Dentro de este sistema climático se pueden dar cambios caracterizados por las diferencias de las condiciones medias entre un periodo y otro, y pueden darse en una región a diferente escala y de distintas maneras pero siempre con una tendencia que aumenta o disminuye de la media central (Hare, 1979).

El sistema climático depende de una multitud de interacciones entre diferentes factores: la temperatura, humedad, presión atmosférica, latitud y altitud, entre otros. Al ser un sistema interconectado, cualquier cambio de los factores afecta a todo el sistema y se considera un forzamiento interno. Estos cambios pueden ser estocásticos y las consecuencias climáticas no son predecibles más allá de la escala que inició el proceso (Mitchell, 1976).

En este complejo sistema climático hay cambios que lo alteran pero no modifican el factor inicial y se denominan forzamientos externos (Robock, 1978). Son predecibles ya que varían en forma conocida y su repercusión dentro del sistema climático sigue patrones similares; ejemplos de ello son las variaciones orbitales o cambios en la insolación (Berger and Loutre, 1991; Berger *et al.* 1991).

C.1 El paleoclima: inferencias a partir de evidencias indirectas.

Para entender cómo se dieron las distintas fluctuaciones climáticas a lo largo de las eras geológicas hasta llegar al sistema climático moderno, se necesitan conocer las interacciones entre procesos físicos y biológicos con el clima para estudiar evidencia que pueda ser indicativa de las condiciones climáticas del pasado. Estas evidencias indirectas también conocidas en inglés como *proxys* pueden ser de distintas índoles y permiten ver fluctuaciones climáticas, con ciertas limitaciones, por lo que a la hora de inferir condiciones climáticas, hay que ser cautelosos.

Los *proxys* permiten reconstruir climas del pasado o paleoclimas. Ejemplos de ellos son núcleos de hielo, núcleos sedimentarios, corales, espeleotemas, polen y dendrocronología, entre otros (Bradley & Eddy, 1991; Alverson *et al.* 1999, 2003; Bradley, 2008). Mientras que en tiempos históricos, las mediciones del tiempo se han hecho con instrumentos desde el año 1850 (Brohan *et al.* 2006). Aunque se tiene evidencia de mediciones instrumentales climáticas desde 1654 en Florencia, Italia y en 1659 en el centro de Inglaterra que son consecutivas y con poca resolución (Manley, 1953).

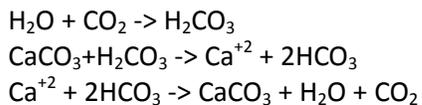
Uno de los *proxys* que funge como herramienta para establecer la línea temporal de reconstrucción ambiental son los isótopos ambientales. Se distribuyen a través de procesos biogeoquímicos en la naturaleza, representando uno de los instrumentos que evidencian variaciones climáticas. En este sentido, la datación isotópica se basa en la desintegración de isótopos radiactivos o acumulación de isótopos hijos permitiendo determinar la edad de aguas subterráneas, hielo, rocas y espeleotemas, entre otros.

Diversos métodos de datación isotópica se pueden aplicar en paleoclimatología para fijar la escala de tiempo correcta de sucesos climáticos anteriores; el intervalo de edad y la naturaleza del sistema son factores a tomar en cuenta para elegir entre radioisótopos naturales (cosmogénicos o in situ) o antropogénicos (Bradley, 2015). El radioisótopo más común es el carbono 14, el cual tiene un periodo de desintegración de 5730 años permitiendo determinar edades de hasta 400 000 años, lo cual, para aplicaciones biológicas es de gran utilidad.

Además del uso de isótopos radiactivos, los isótopos estables en compuestos naturales también son informativos, pues varían debido a su comportamiento físicoquímico, derivado de su diferente masa. La temperatura es el parámetro más importante dentro de esos procesos que tiene como efecto el fraccionamiento isotópico y de manera indirecta se obtiene información sobre las condiciones ambientales en el momento de ocurrir el proceso de composición isotópica del agua, hielo y minerales.

Los isótopos estables que más suelen aplicarse en investigaciones climatológicas son los de deuterio, oxígeno-18 y carbono-13; los dos primeros por indicar cambios inducidos en el ciclo hidrológico y el último por variar directamente con el clima.

En el caso de cuerpos lacustres, los isótopos de oxígeno se derivan de fuentes biogénicas (por ej. ostrácodos, moluscos) o fuentes endógenas (autigénicas) como la calcita, la cual precipita directamente de la columna de agua una vez que se satura el carbonato de calcio de acuerdo a las siguientes reacciones:



En cuerpos lacustres cerrados, esta precipitación puede darse debido a la evaporación ya que aumenta la concentración de minerales disueltos o por la fotosíntesis de algas y macrofitos al remover el dióxido de carbono del lago, causando de vuelta una saturación de carbonato de calcio. Al evaporarse el agua, hay remoción preferencial del isótopo más ligero (^{16}O) lo que deja agua enriquecida con ^{18}O . Para estandarizarlo, se compara con el estándar global δD conocido como la línea de agua global meteórica (GMWL, por sus siglas en inglés) (Ito, 2001; Leng, 2003).

$$\delta\text{D} = 8\delta^{18}\text{O} + 10$$

Como resultado, la evaporación enriquece el agua con deuterio y ^{18}O con respecto a la GMWL. Esto es constante para todos los lagos y es de esta manera que el balance precipitación - evaporación queda representado de manera particular en cada lago y difiere de manera constante con respecto al estándar global.

La evidencia empírica permite hacer aproximaciones de la temperatura. Bajo condiciones de equilibrio isotópico el $\delta^{18}\text{O}$ disminuye $\sim 0.23\text{‰} \text{ } ^\circ\text{C}^{-1}$ al aumentar la temperatura, con lo cual algunos estudios empíricos aproximan la temperatura con la siguiente fórmula (Leng & Marshall, 2004):

$$T (\text{ } ^\circ\text{C}) = 13.8 - 4.58 (\delta\text{c} - \delta\text{w}) + 0.08 (\delta\text{c} - \delta\text{w})^2$$

Donde los valores δc y δw son los valores de la calcita autógena y del agua, respectivamente. Aunque hay que tomar en cuenta carbonatos biogénicos los cuales pueden formar calcita en ciertas épocas del

año dependiendo de la especie. Aparte de este ruido biológico, se debe analizar si el valor del agua se mantiene constante, lo cual representa un escenario ideal más no real. También se puede hacer aproximaciones de temperatura a partir de diatomeas con el $\delta^{18}\text{O}$ en donde la disminución es de $\sim 0.2\text{‰}$ $^{\circ}\text{C}^{-1}$ (Moschen *et al.* 2005) pero en esos casos se usa la siguiente fórmula:

$$T (^{\circ}\text{C}) = 11.03 - 2.03(\delta^{18}\text{O}_{\text{Si}} - \delta^{18}\text{O}_{\text{w}} - 40)$$

Donde $\delta^{18}\text{O}_{\text{Si}}$ representa el contenido isotópico de sílice de la diatomea. Para usar cualquiera de estas fórmulas empíricas todas las condiciones deben ser cuidadosamente estudiadas y propiamente estandarizadas, lo cual requiere un estudio profundo de las condiciones del cuerpo lacustre (Shemesh *et al.* 2001).

En el caso de que se tengan registros naturales como los espeleotemas, es decir, formaciones minerales conocidas como estalagmitas o estalactitas, su propia formación se debe a la acumulación de carbonato de calcio. La formación de espeleotemas se debe a que precipitó en algún momento en la superficie y fue percolándose por los distintos sustratos terrestres (entrando en contacto con materia orgánica en descomposición) hasta llegar a la cueva en donde hubo acumulación de carbonato debido a evaporación del agua (Richards & Dorale, 2003).

Las variaciones estacionales de precipitación permiten identificar capas anuales o estacionales en algunas estalagmitas. Las estalagmitas son comúnmente datadas mediante métodos de series de uranio. Los isótopos de uranio que se encuentran en las cuevas precipitan al mismo tiempo que la calcita de los espeleotemas formando carbonato de uranilo o rutherfordina. Al decaer ^{234}U , su producto es ^{230}Th con lo cual se obtienen la edad del espeleotema. Es preferible este método pues permite un rango temporal de $\sim 500\ 000$ a 100 años antes del presente y requiere una muestra pequeña (<100 mg de calcita) (Richards & Dorale, 2003).

En las cuevas, a diferencia de los cuerpos lacustres, el aire y movimiento del agua suele ser muy lento, alcanzando fácilmente el equilibrio térmico entre la temperatura de las paredes en la cueva y el aire, por lo que al depositarse la calcita del agua que gotea, la fracción de los isótopos de oxígeno ocurre a un ritmo constante y el contenido de isótopos depende completamente de la temperatura ambiental externa. En caso de no tener equilibrio isotópico en la cueva sería imposible hacer una reconstrucción paleoclimática (O'Neil *et al.* 1969).

C.1.1 Cambio climático durante el Neolítico

El quinto reporte del Panel Intergubernamental de Cambio Climático define como cambio climático abrupto un cambio a gran escala que tiene lugar a nivel de décadas, o menos, pero que es suficiente para causar alteraciones en los sistemas naturales o humanos (IPCC, 2013).

Es en el contexto antropológico, que se dio uno de los más grandes cambios en la historia humana a nivel social, demográfico y económico, a partir de la transición cazador-recolector a sedentarismo en las poblaciones humanas y se definió como revolución neolítica (Childe, 1936; Crosby, 1986). Sin embargo, los procesos que causaron sincronía en diferentes continentes (ejemplo de ello son las civilizaciones Andinas, el creciente fértil o media luna fértil, el Oriente Asiático y Mesoamérica) son evidencia de que se trató de un proceso global que comenzó alrededor de 12 500 años antes del presente.

El cambio de la dieta en poblaciones neolíticas tuvo que ver con el uso intensivo de plantas y animales domesticados, resultando en un cambio de interacciones sociales. Este cambio global es estudiado a través de la contribución de disciplinas como arqueología y antropología, que no permiten explicarlo por sí solas, requiriendo la integración de otras disciplinas para entenderlo. La paleoclimatología permite estudiar la revolución neolítica desde otra perspectiva: la transición del Pleistoceno al Holoceno la cual ocurre en la misma temporalidad que la revolución neolítica (11 600 años antes del presente) y propone que posiblemente fue ese cambio climático el que alentó esta revolución global.

C.1.1.1 Forzamientos externos del Neolítico.

Durante el periodo Cuaternario (hace alrededor de 2.58 millones de años hasta el presente) se han dado ciclos glaciales-interglaciares en el último millón de años (Figura 5). La teoría más aceptada que explica estos ciclos se relaciona con fluctuaciones en la energía solar recibida por la tierra. James Croll, en 1867, desarrolló la hipótesis en la que atribuía dichas fluctuaciones a forzamientos orbitales. Posteriormente, esta hipótesis fue complementada por Milankovitch en 1941 (Berger, 1988; Laksar, 1993) transformándola en una teoría crucial de la paleoclimatología: La Teoría de Milankovitch (Imbrie & Imbrie 1979). Gracias a esta teoría, la comprensión de la relación entre la variación de la energía solar disponible en la atmósfera terrestre y los cambios orbitales permite comprender la transición del Pleistoceno al Holoceno.

Siguiendo la Teoría de Milankovitch, la transición al Pleistoceno-Holoceno ocurrió dada la combinación de tres parámetros orbitales que afectan la intensidad de la energía solar recibida por la tierra. Estos tres parámetros orbitales tienen periodos y frecuencias temporales distintas y son categorizadas como forzamientos externos (Ruddiman, 2006; Jouzel *et al.* 2007; DeConto *et al.* 2008; Bretagnon *et al.* 1982; Loutre *et al.* 1992 Berger & Loutre, 1991).

El primer parámetro se denomina precesión orbital y describe la variación del eje de rotación de la Tierra en una posición regular de solsticios y equinoccios (Figura 3). A este efecto hay que sumarle el hecho de que la órbita terrestre gira progresivamente alrededor del sol, por lo que la posición de la Tierra con respecto al sol varía con el paso del tiempo. El ciclo de la precesión orbital es de 19 000 a 24 000 años (Berger, 1977; Hays *et al.* 1976).

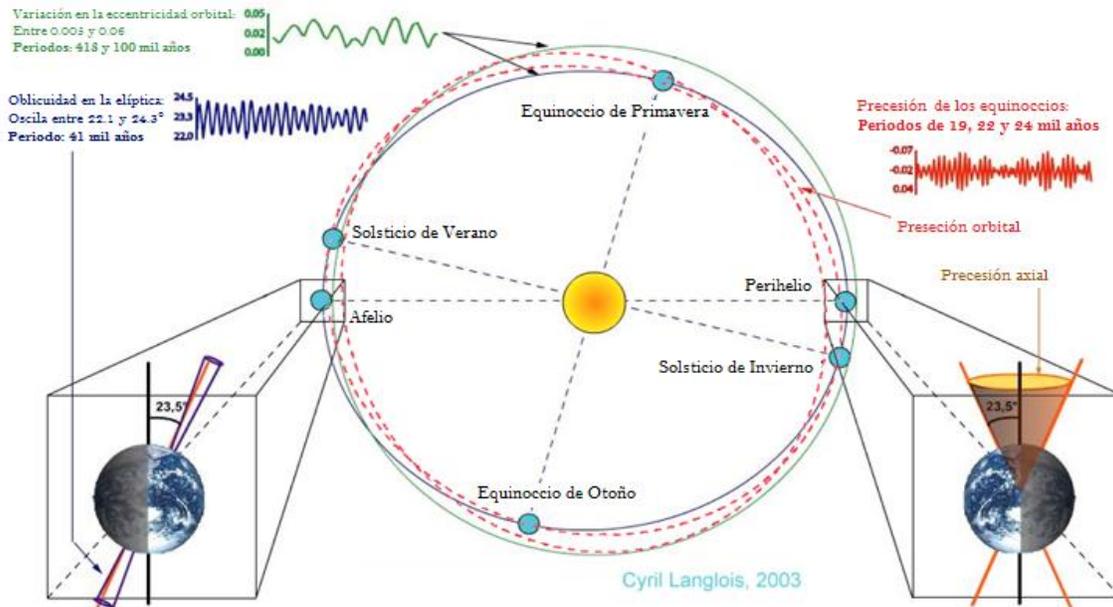


Figura 3. Esquema general de los tres forzamientos orbitales.

Los cambios en los forzamientos orbitales (precesión orbital, oblicuidad de la elíptica y excentricidad orbital) permitieron las fluctuaciones ambientales durante el Holoceno que posiblemente dieron lugar al desarrollo de la revolución neolítica (Langlois, 2003).

El segundo parámetro es la oblicuidad de la elíptica: se refiere al ángulo entre el eje de rotación y la perpendicular del orbital medio de la Tierra. Dicho ángulo se mueve a través del tiempo gracias a perturbaciones planetarias que hacen que la oscilación se modifique. Su efecto aumenta gracias a la irregularidad en la esfericidad de la tierra afectando las fuerzas gravitacionales del Sol y de la Luna. La periodicidad principal de dichas oscilaciones es de 41 000 años.

Debido a este fenómeno existe la estacionalidad en la tierra, así como la modulación de la cantidad de insolación recibida en distintas latitudes (Figura 4). Ambos efectos permiten una oscilación de la órbita Terrestre, que actualmente es de aproximadamente 23.5° (Berger, 1988; Tricot & Berger, 1988).

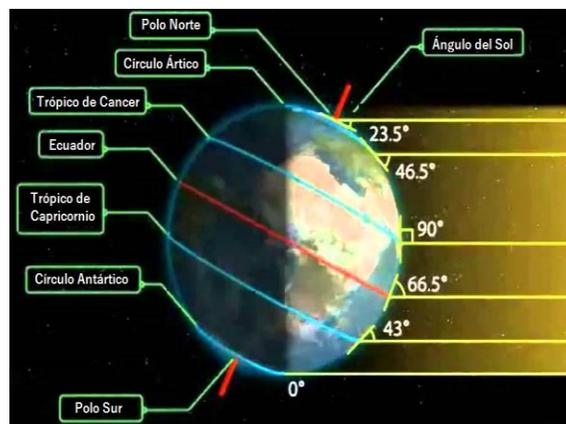


Figura 4. La oblicuidad de la elíptica modula la insolación en distintas latitudes terrestres.

Los cambios climáticos son fácilmente identificables en latitudes altas debido a que cualquier variación en la oblicuidad generará un cambio importante, a diferencia de aquellas zonas que se encuentran cercanas al ecuador y que reciben una mayor cantidad de insolación.

(Imagen obtenida de: <https://www.tes.com/lessons/G2PrAJfm2aJkFA/insolation>).

Finalmente, la excentricidad orbital es el grado de desplazamiento de la órbita de la Tierra con respecto al Sol. La frecuencia es de 100 000 años y se relaciona directamente con los forzamientos solares (Imbrie *et al.* 1993).

De acuerdo con la Teoría de Milankovitch, los tres parámetros orbitales descritos se relacionan directamente con las frecuencias de forzamiento solar en el último millón de años (Figura 5). De esta manera se comprende la relación de los cambios orbitales y cómo afectaron el sistema climático de la Tierra, obteniendo estadios glaciario-interglaciario durante el último millón de años.

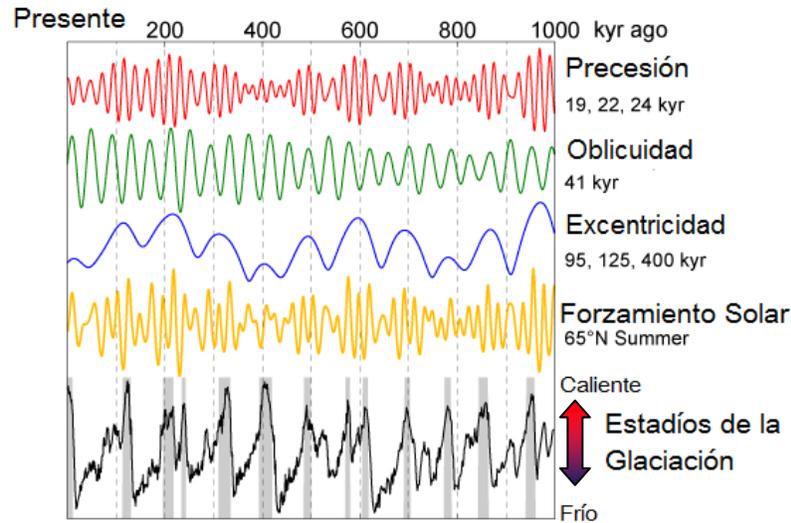


Figura 5. Impacto de los distintos forzamientos orbitales a lo largo del último millón de años antes del presente.

Evidencia encontrada en distintos núcleos de hielo demuestran que las oscilaciones de forzamientos orbitales coinciden con los estadios glaciario-interglaciarios durante el último millón de años. Al comprender el impacto de los forzamientos en el sistema climático terrestre y la relación de sus componentes en el presente, se podrían anticipar posibles cambios ambientales a futuro y a su vez, inferir interacciones del pasado. (Imagen obtenida de <https://es.wikipedia.org/wiki/Glaciación>)

El último interglaciario o Eemense tuvo lugar hace 124 000-119 000 años durante el Pleistoceno, y después de él la temperatura media global presentó una tendencia a la baja, hasta alcanzar mínimos globales durante el Último Máximo Glacial (23 000-19 000 años antes del presente, Mix *et al.* 2001). Alrededor de 12 000 años AP, se inició una tendencia hacia condiciones más cálidas, con aumentos de la concentración tanto de dióxido de carbono como de metano y el nivel del mar. Estos factores modificaron el sistema climático marcando el principio del Holoceno, hace 11 700 años (Monnin *et al.* 2001).

La modificación de las condiciones durante el Holoceno entre los distintos componentes, dio pie a eventos propios de esta época, el más conocido es El Niño-Oscilación Sur (ENOS). Se trata de un evento cálido que varía las temperaturas superficiales del mar en las costas del Ecuador y al norte de Perú, causando cambios convectivos en las regiones costeras del noroeste Sudamericano. Gracias a las evidencias geológicas se ha estimado la frecuencia a la que suceden los eventos ENOS la cual tiende a ser un proceso auto-regresivo de primer orden por lo que hay un pulsamiento gradual desde el Holoceno temprano en donde se considera que el ENOS era muy débil o inexistente hasta llegar a los 7000 años antes del presente teniendo como máxima entre 2000-1000 años antes del presente.

Posterior a esa fecha la frecuencia ha decaído. Otros cambios climáticos más prominentes registrados durante el Holoceno son el cálido medieval (800-1300 d.C.) y la pequeña edad de hielo (1300-1850 d.C.) (Mann *et al.* 2009; Conroy *et al.* 2008; Moy *et al.* 2002; Toth *et al.* 2015).

D. Presente y pasado del ambiente en Tehuacán

D.1 El clima en México

Se ha demostrado que la energía extraterrestre recibida por la tierra ejerce un fuerte control en el clima a escalas de milenios o sub-milenios (Lozano-García, Torres-Rodríguez, Ortega, Vázquez, & Caballero, 2013). En escalas menores, los patrones regionales de variabilidad climática que afectan la evolución de los ecosistemas son mayoritariamente modulados por las dinámicas internas del planeta.

En México, el clima se vuelve particular para cada zona a lo largo de la franja latitudinal de Mesoamérica. Depende de varios factores de los cuales, el movimiento atmosférico juega un papel fundamental en nuestro país ya que es el movimiento resultante de las diferencias de presiones y la densidad entre masas de aire. Dicho movimiento tiene como fuente de energía al sol, que gracias a la excentricidad orbital permite recibir mayor intensidad de energía solar en sitios cercanos al Ecuador (Balvanera *et al.* 2016).

Para comprender el papel que juega la temperatura en el movimiento atmosférico hay que tener claras las relaciones fisicoquímicas entre el vapor de agua y la atmósfera. La humedad relativa es el vapor de agua expresado en porcentaje presente con respecto a la capacidad de carga del aire. Al acercarse al 100%, ocurre condensación del vapor de agua. Al disminuir la presión o temperatura del aire, la habilidad de retener agua en forma gaseosa se verá reducida. Debido a que la elevación modifica la masa de aire, en general hay un cambio inverso en su densidad y temperatura: el cambio de temperatura en aire seco es de aproximadamente 1 °C cada 100 m de cambio en la altitud (MacNeish *et al.* 1967).

Además hay cambios convectivos involucrados en donde el aire tenderá a expandirse al calentarse y a elevarse si la orografía lo facilita (por ejemplo las faldas de una montaña) elevándose a zonas más altas, pero si se trata de una masa de aire fría, tenderá a bajar. Así, la masa caliente que inicialmente se expandió tiende a subir, se reduce su presión y se enfría adiabáticamente formando así condensación (nubes) o precipitación (lluvia). La masa de aire fría que tiende a bajar (en el otro costado de la montaña) se comprime, se calienta y pierde elevación; mientras va calentándose, su capacidad de retener vapor de agua aumenta y va secando el ambiente. Este fenómeno fisicoquímico explica el por qué se tiene sombra de lluvia en zonas montañosas y está presente en la parte oriental de México (MacNeish *et al.* 1967).

Es por ello que la estabilidad de una columna de aire depende de su gradiente de temperaturas y las zonas en donde hay un intercambio térmico, debido a que las masas de aire tienden a ser secas y usualmente sin lluvias. Cada año, el intercambio térmico varía su latitud, conformando una franja de baja presión y la humedad, conocida como la zona de convergencia intertropical (ZCIT). La posición de esta franja se determina en promedio por la latitud sobre la cual se presenta la mayor insolación, de manera que está regulada por el ciclo de insolación intra-anual. Así, durante el verano boreal, el ZCIT trae las lluvias para la zona de América Central y el sur de México mientras que, durante el invierno, ésta migra hacia el sur, provocando la estación seca. Los vientos que pasan por las aguas oceánicas tienden a calentarse en verano, acarreando más agua, permitiendo lluvias, aunque sus efectos sean poco notorios en las costas mexicanas (MacNeish *et al.* 1967).

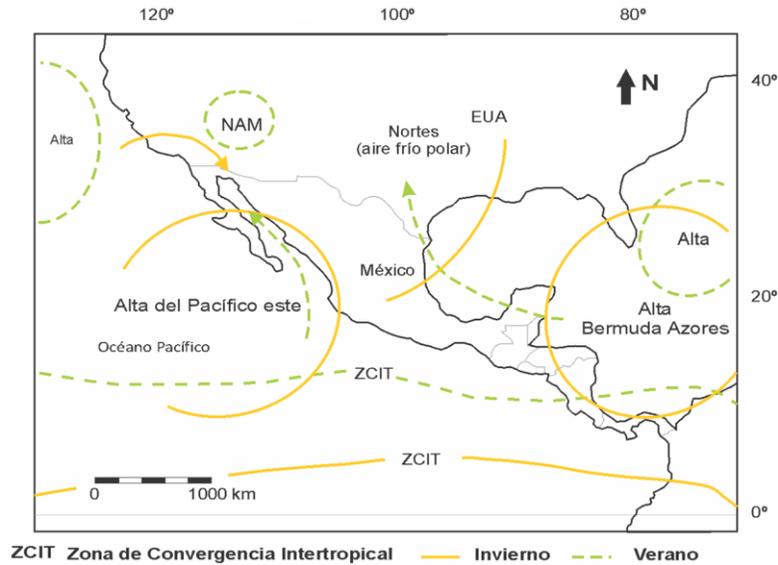


Figura 6. Distintos movimientos atmosféricos afectan el clima en México dependiendo de la estacionalidad.

NAM - Monzón Norte-Americano ZCIT- Zona de Convergencia Intertropical y Celdas atmosféricas: Alta del Pacífico Este y Alta Bermuda Azores. Las líneas amarillas muestran cómo afectan las celdas atmosféricas, el movimiento de las corrientes y la zona de convergencia intertropical durante la estación de invierno y verano a México (Imagen donada por la Dra. Esperanza Torres).

También hay que tomar en cuenta disturbios atmosféricos: el aire fluye en el sentido de las manecillas del reloj cuando proviene de los trópicos, mientras que los vientos polares provenientes del norte vienen en el mismo sentido pero hacia el sur, dando así una zona de convergencia turbulenta dado el choque térmico, logrando que el aire caliente suba y precipite como lluvia o nieve mientras que el aire polar baja y se mueve hacia el sur provocando así los “Nortes” que llegan hasta el Caribe. Las corrientes oceánicas también disipan calor de la zona ecuatorial hacia zonas más frías (Bradley, 2015).

D.1.1 El clima en Tehuacán

El clima en Tehuacán es producto de las características físicas de la atmósfera y la posición del valle con respecto a la atmósfera, afectado principalmente por las corrientes del hemisferio norte. En segundo lugar, las corrientes oceánicas provenientes del Golfo de México y el Océano Pacífico. Finalmente, el factor decisivo en términos climáticos es la topografía de la región.

Actualmente el valle de Tehuacán se caracteriza por un clima cálido y seco. Las lluvias ocurren en dos periodos con una duración 6 a 8 semanas en los meses de mayo a julio y agosto a septiembre. El promedio anual de lluvias en el centro del valle es de aproximadamente 500mm. Se encuentra a una elevación de 1676 metros sobre el nivel del mar (msnm). Durante el invierno hay heladas ocasionales y los días son cálidos. Ocasionalmente, durante el invierno ocurren “Nortes” que traen oleadas de viento frío causando una acumulación de nubes y lluvia sobre los picos cercanos al eje neo-volcánico en los que se forma una capa de nieve (MacNeish *et al.* 1967).

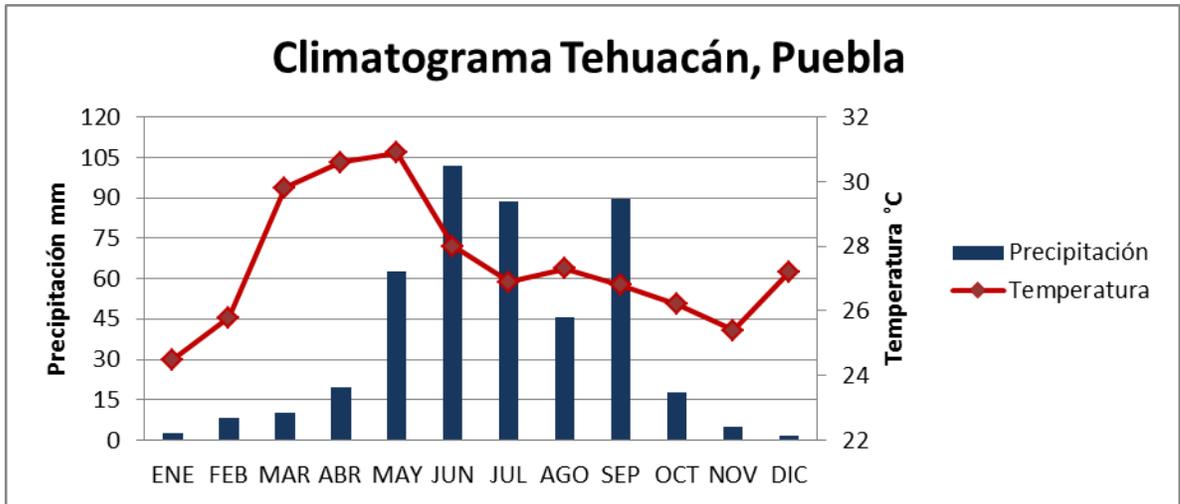


Figura 7. Climatograma de Tehuacán, Puebla.

Datos del promedio de precipitación y de temperatura de los últimos 30 años en Tehuacán, Puebla (Datos obtenidos del servicio meteorológico nacional).

Los disturbios atmosféricos como los “Nortes” no producen lluvias en el Valle de Tehuacán, sólo bajas temperaturas, pero sí hay efectos al presentarse tormentas tropicales de las cuales no se comprenden las causas pero tienen como consecuencia las lluvias conocidas como “veranillo”, que permiten un cultivo temprano (MacNeish *et al.* 1967).

La Sierra Madre de Oaxaca provee una barrera a las tormentas del este debido a sus 2600 msnm y provee sombra de lluvia al valle especialmente en el noroeste. La humedad proveniente del Golfo de México queda en la parte este de la Sierra por lo que al descender los vientos por la parte oeste, se calientan adiabáticamente y se transforman en vientos desecantes, lo que aumenta la temperatura en el valle. En la parte noreste del valle que presenta mayores elevaciones ocurren comúnmente temperaturas de congelamiento (MacNeish *et al.* 1967).

D.2 La agricultura durante el Neolítico.

Desde hace 11 000 años, en distintas partes del globo, los distintos grupos humanos comenzaron a seleccionar las plantas que tenían a su alrededor, particularmente a las gramíneas, en las que lograron fijar características como la pérdida de la dispersión y el aumento de productividad de la semilla (Purugganan & Fuller, 2009).

Se estima que más de 118 especies de plantas, pertenecientes a 70 géneros y 39 familias, han sido domesticadas en México (Hernández-Xolocotzi, 1993). Algunas como el maíz fueron dispersadas hacia todo el territorio, forzándolas a adaptarse a los diversos ambientes, lo que generó una gran cantidad de razas propias de cada región, obteniendo como resultado características muy variadas en cuanto a diversidad genética y morfología (Vigouroux *et al.* 2008).

La evidencia genética ha identificado el posible epicentro de domesticación del maíz en la región del Balsas, alrededor de 9 000 años antes del presente, en elevaciones de alrededor de 1 500 metros sobre el nivel del mar, teniendo como ancestro al teocintle (*Zea mays ssp. parviglumis*) (Matsuoka, 2002).

D.2.1 La historia de la agricultura de Tehuacán.

En el Valle de Tehuacán en Puebla, una excavación en los años 60 encabezada por el Arqueólogo Richard S. MacNeish se llevó a cabo con un acercamiento interdisciplinario, al integrar estudios de paleobotánica (destacan maíz *Zea mays*, frijol *Phaseolus vulgaris* y calabaza *Curcubita mixta*), palinología, geología, geografía, antropología física, etnografía y etnohistoria, además de arqueología. De esa excavación se obtuvieron alrededor de 750 000 especímenes, de los cuales alrededor de 24 000 eran restos de maíz. Gracias a la evidencia arqueológica se logró la reconstrucción del modo de vida de los habitantes antiguos, incluyendo sus hábitos alimenticios (MacNeish, 1964).

La reconstrucción del estilo de vida de los habitantes en el Valle de Tehuacán se divide en siete fases: Se determinó que en la primera fase, “Ajuerado”, que termina alrededor de 7200 A.C., los habitantes eran recolectores de plantas y animales. Cabe resaltar que durante esta fase ocurrió el final de la última glaciación, dando inicio al Holoceno, por lo que la fauna pudo ser distinta a la que cazaron en fases posteriores (MacNeish, 1964).

Durante la segunda fase, del “Riego”, que va de 7200 a 5200 A.C., se tiene registro de actividades nómadas, pero la recolección de plantas jugó un papel importante durante la temporada de lluvias. La evidencia incluye plantas que posteriormente pudieron ser cultivadas: calabaza (*Curcubita mixta*), chile (*Capsicum annuum*) y aguacate (*Persea americana*) (MacNeish, 1964).

La tercera fase es la conocida como “Coxcatlán”, datada alrededor de 5200 a 3400 A.C. y se caracterizó por la agricultura incipiente con el uso de maíz (*Zea mays*), chile (*Capsicum annuum*), aguacate (*Persea americana*) y guaje (*Leucaena leucocephala*). A mediados de dicha fase se encontraron restos de amaranto (*Amaranthus cruentus*), frijol (*Phaseolus vulgaris*), zapote amarillo (*Pouteria campechiana*) y calabaza (*Curcubita mixta*). Para el final de la fase ya se tenía zapote negro (*Diospyros digyna*) y blanco (*Casimiroa edulis*) (MacNeish, 1964).

La cuarta y quinta fases denominadas “Abejas” (3400-3200 A.C.) y “Purron” (2300-1500 A.C.) respectivamente, dan muestras de un estilo de vida sedentario, al contar con alimentos provenientes de plantas domesticadas. Sin embargo, el 70% de los alimentos provenía de plantas y animales salvajes.

Es en estas etapas en donde se cuenta con evidencia del proceso de domesticación del maíz (MacNeish, 1964).

En la sexta fase “Ajaltzan” 1500-900 A.C. la sociedad era de agricultores mientras que en la séptima fase, “Santa María” 900-200 A.C. se infiere que hubo conocimiento de sistemas de riego para los cultivos. Las últimas dos etapas, conocidas como “Palo Blanco” y “Venta Salada” van de 200 A.C. a 1540, en las que se encuentra evidencia de la existencia de un sistema religioso, político y social complejo, que ya contaba con sistemas intrincados de irrigación (MacNeish, 1964).

D.2.2 La historia del maíz en el Valle de Tehuacán.

En cinco cuevas del Valle de Tehuacán fueron encontrados restos de maíz durante la expedición de MacNeish: Coxcatlán, Purrón, San Marcos, Tecorral y El Riego. Entre estas muestras se encuentran mazorcas de 6500 años calibrados de antigüedad, conservadas en buen estado y disponibles para análisis botánicos (MacNeish, 1964). Morfológicamente son muy similares al teocintle debido a su tamaño uniforme, raquis frágil y glumas largas. Los especímenes recolectados revelaron una secuencia de maíz que había evolucionado a través de la domesticación desde hace 6500 años en la cueva de San Marcos.

Además de los restos arqueobotánicos de maíz, se tiene evidencia de un área de recolección de agua cercana a la cueva de San Marcos para el cultivo, posiblemente de maíz. Existe evidencia adicional de presas prehispánicas, como es el caso de Mequitongo; presas cercanas a San José Tilapa que llegan hasta los resguardos rocosos de Purrón y Abejas. Estas presas fueron construidas en piedras de fácil tallado. No sólo crearon las presas sino además estructuras que las pudiesen unir y permitir un correcto almacenamiento y distribución de agua (MacNeish *et al.* 1967).

Finalmente, la evidencia arqueológica y paleobotánica sugiere que a partir de la fase Abejas se practicaba la agricultura en la región, pero además se trataba ya de una civilización de agricultores establecidos (MacNeish, 1964).

E. El genoma y su respuesta al ambiente.

E.1 Particularidades del genoma de maíz

El maíz es una gramínea monocotiledónea cuyo genoma fue completamente secuenciado en 2009 (Vielle-Calzada *et al.* 2009; Schnable, P. S. *et al.* 2009). El tamaño de su genoma varía dependiendo de la raza; mientras que el tamaño de la variedad B73 es de aproximadamente 3.2 Gb, la raza palomero toluqueño tiene un genoma estimado en 2.9 Gb. Esta diversidad se traduce en materia agrícola como componentes de diversidad biológica que se vuelven esenciales para alimentar poblaciones humanas y mejorar su calidad de vida (FAOSTAT 2014).

A nivel genómico el maíz es sumamente interesante, dado que su genoma contiene alrededor de 50 000 genes y más de un millón de transposones y elementos repetitivos (Schnable *et al.* 2009). Esta variación permite estudiar la naturaleza propia de los genes que va desde silenciamiento de transposones, expresión de estado activo a inactivo (*cytling*), heredable por una o dos generaciones (*presetting*) hasta impronta genómica y paramutación, generando oportunidades abundantes para explicar características difíciles de entender únicamente con variación alélica. (Fedoroff, 1999; Gottlieb *et al.* 2002; Rabinowicz & Bennetzen, 2006; Schnable, 2009).

El maíz es utilizado gracias a sus cualidades alimenticias, producción de proteína animal y uso industrial. En el ciclo comercial de 2015/2016 la producción mundial de maíz se estima que fue de 972.6 millones de toneladas (Panorama agroalimentario, 2015) lo que vuelve a esta gramínea interesante desde el punto de vista comercial.

E.2 Los procesos de respuesta genómica y la importancia de la metilación.

El maíz, al igual que todas las plantas debe responder a estrés ambiental y en los últimos años se ha estudiado la relación no sólo de las respuesta a nivel genómico sino también epigenético al estrés (Finnegan, 2002; Richards, 2011). El análisis de la acumulación de mutación espontánea provee evidencia de que los cambios en la metilación del ADN en nucleótidos ocurren a un índice mayor que las mutaciones puntuales (de un solo nucleótido) (Becker et al. 2001). Sin embargo, los cambios en niveles regionales de metilación en regiones diferencialmente metiladas (DMR) ocurren a una frecuencia muy similar a los índices de mutación de nucleótidos únicos (Schmitz *et al.* 2011).

Al comparar niveles de metilación del ADN en distintos genotipos, la evidencia indica que una porción substancial de las DMR se asocian con el propio genotipo. Como ejemplo de lo anterior tenemos en plantas de maíz que la comparativa de metilación de ADN en dos genotipos (B73 y Mo17) revelaron que una parte de las citosinas metiladas se debía a un proceso meramente epigenético mientras que otras regiones metiladas eran resultado de la inserción por transposones (Regulski, 2013). Sin embargo, existen ejemplos de metilación diferencial del ADN en maíz donde regiones genómicas no muestran diferencias genéticas y no son influenciadas por otras regiones (Eichten, 2011).

En el caso de metilación de ADN causada específicamente por estrés abiótico, se han encontrado DMR en plantas crecidas en cultivo de tejido con cambios específicos en el metiloma (Stelpflug *et al.* 2014; Stroud *et al.* 2014) pero no de manera estadísticamente significativa en estudios donde se ha sembrado maíz (Eichten & Springer 2015), aunque si han mostrado un aumento en la frecuencia de regiones hipermetiladas en plantas estresadas. Por lo que es posible que la estocasticidad de la metilación pueda ser detectada de manera más fina en estudios con una mayor muestra poblacional (Eichten & Springer 2015; Jiang *et al.* 2014).

También existe evidencia de la herencia de cambios en metilación del ADN mediante mitosis y meiosis bajo estrés ambiental (Eichten *et al.* 2014), siendo posible que los cambios en la metilación del ADN puedan ser alterados y mantenidos mediante mitosis para regular características genómicas durante las distintas etapas de desarrollo de plantas (Law & Jacobsen, 2010).

Es necesario desentrañar las distintas capas en las que el genoma se estructura, siendo una de ellas el epigenoma. En particular, no se sabe nada a la fecha sobre los efectos que el epigenoma pudo jugar durante el proceso de adaptación local del maíz, aunque en otros modelos de plantas se comienzan a investigar estos efectos (Lasky *et al.* 2014).

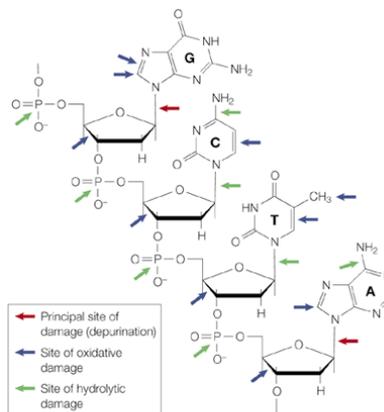
E.3 La importancia del paleogenoma para entender la respuesta al ambiente.

Se denomina paleogenómica al estudio de la secuencia, estructura y función de genomas extintos (Poinar *et al.* 2006); por lo que, desde hace una década, con la aplicación de nuevas técnicas de biología molecular y secuenciación masiva, se ha logrado recuperar e interpretar valiosa información codificada en el material genético preservado de distintos tipos de restos orgánicos antiguos (Willerslev & Cooper, 2005).

En organismos vivos, las moléculas de ADN están constantemente siendo mantenidas gracias a una maquinaria de reparación enzimática (Lindahl, 1993). Al morir el organismo, los mecanismos de reparación cesan sus funciones y el genoma queda expuesto a factores producidos por el decaimiento del organismo, desde químicos (ej. hidrolíticos, oxidativos) hasta térmicos. Dichos factores terminan por degradar al genoma de tres maneras: Fragmentación, lesiones que bloquean la replicación del ADN endógeno y daño químico causante de incorporaciones erróneas de nucleótidos cuando el ADN es replicado *in vitro*.

El primer estudio sistemático de las propiedades del ADN antiguo fue realizado por Pääbo hace casi 30 años, en donde demostró que el ADN extraído de muestras de 4 a 13 000 años estaba degradado en fragmentos de 40-500 pares de bases (Pääbo, 1989). Durante la última década, los métodos de secuenciación nos permitieron tener un mejor entendimiento de los procesos de fragmentación en ADN antiguo: las condiciones climáticas como la humedad, temperatura, salinidad, pH y factores bióticos como el crecimiento microbiano tendrían un fuerte efecto en la preservación del ADN (Briggs *et al.* 2007).

Este daño post-mortem limita el acceso a secuencias de ADN antiguo y presenta algunas complicaciones para su manipulación, aunque también permite autenticar el ADN, es decir, nos indica que se trata de ADN antiguo y no de contaminación moderna. Las incorporaciones erróneas son causadas principalmente por formas deaminadas de citosinas es decir, uracilos (Briggs *et al.* 2007).



Nature Reviews | Genetics

Figura 9. Esquema de un fragmento de ADN en doble hélice con las cuatro bases nucleotídicas y los posibles sitios de daño.

Los principales sitios de daño son indicados por flechas rojas. Los ataques hidrolíticos son indicados por flechas verdes y el daño oxidativo por flechas azules (Hofreiter *et al.* 2001).

Actualmente, los genomas más antiguos que se han logrado recuperar corresponden a un équido del Pleistoceno de 780 000 años de antigüedad (Orlando *et al.* 2013) y a una planta e insectos encontrados en núcleos de Groenlandia que van de 450 000 a 800 000 años de antigüedad (Willerslev *et al.* 2007). Los genomas que han generado mayor interés han sido los de los homínidos (Meyer *et al.* 2012; Prüfer *et al.* 2014). Ejemplos de otros organismos antiguos con secuenciación total o parcial de su genoma incluyen al oso de las cuevas (Dabney *et al.* 2013) y al mastodonte (Rohland *et al.* 2010), entre otros (Llamas *et al.* 2012; Miller *et al.* 2012; Cooper *et al.* 2001). Las condiciones ambientales reinantes en yacimientos de restos orgánicos inertes determinan en gran medida su conservación, a esto se le denomina procesos diagenéticos (Allaby *et al.* 2014). Dichas condiciones pueden ser a nivel de microclima, lo cual podría explicar niveles distintos de conservación de ADN en diferentes muestras del mismo sitio (Sawyer *et al.* 2012).

El ADN antiguo permite identificar cambios genómicos a través del tiempo al ofrecer una perspectiva histórica única, pues permite recuperar información de organismos extintos o dinámicas poblacionales del pasado, aunque no permite ver cómo ocurrieron esos cambios a nivel de adaptaciones, dejando un nicho de oportunidades abierto para reconstruir señales epigenéticas a partir de los paleogenomas. Es así como se pretende comprender las transiciones sufridas en las poblaciones a lo largo del tiempo (Gokhman, 2016).

E.3.1 Paleoepigenética

La primera medición directa de metilación antigua se hizo en un bisón del plesitoceno de aproximadamente 26 000 años, el cual fue conservado en permafrost, permitiendo la preservación de la metilación (Llamas *et al.* 2012). Trabajos más recientes en la investigación de la metilación de ADN se basaron en el estudio de 30 esqueletos de nativos Americanos con antigüedades de 230 a 4500 años antes del presente (Smith *et al.* 2015).

La única evidencia que se tiene hasta el momento del estudio de metiloma en cultivos agrícolas antiguos, pertenece a muestras de centeno provenientes de Egipto, con una antigüedad de hasta 3000 años antes del presente. Los patrones de metilación evidencian que una de las muestras fue infectada por el virus del mosaico (Llamas *et al.* 2012). Aunque la cantidad de patrones de metilación obtenida correlaciona con la cantidad de ADN que se obtuvo para cada una de las muestras, esto nos da indicios de preservación de metilación en muestras paleobotánicas (Allaby, 2015).

Los trabajos previamente citados demuestran que las señales epigenéticas pueden ser reconstruidas en genomas antiguos proveyendo información adicional que complementa nuestro entendimiento de las señales regulatorias en el genoma de diversos organismos. A este campo de estudio se le conoce como paleoepigenética (Gokham, 2016).

La preservación bajo condiciones ambientales adecuadas tales como baja temperatura y baja humedad permiten que el metiloma se preserve por miles de años, mientras que en condiciones cálidas, debido a la degradación de ADN, su vida media puede ser tan corta como 1500 años (Llamas *et al.* 2012, Allaby *et al.* 2015).

F) Estrategias para el estudio de la adaptación de plantas antiguas al ambiente.

Nuevas disciplinas tales como la genética y la paleoclimatología abren panoramas para entender no sólo procesos de domesticación a partir de muestras arqueobotánicas sino que, por medio de la epigenética, se pretende contribuir con un entendimiento integral de los cambios ambientales que tuvieron lugar durante el Holoceno para dilucidar la adaptación del maíz al clima Mesoamericano. Así, el estudio sobre procesos de aclimatación y posterior adaptación al cambio climático actual será de gran ayuda para el futuro cultivo de esta gramínea.

F.1 Estrategia paleoclimática para inferir la historia climática en Tehuacán.

Al estimar las fluctuaciones ambientales que ocurrieron durante el Holoceno en el Valle de Tehuacán se usó un enfoque interdisciplinario mediante el análisis de evidencia geológica para llevar a cabo la reconstrucción ambiental de Mesoamérica.



Figura 10. Mapa de la zona que comprende Mesoamérica.
(Imagen de *wikimedia commons*)

A pesar de existir una gran cantidad de estudios paleoclimáticos en toda la zona mesoamericana, hay una mejor resolución en un par de regiones específicas para esta temporalidad: Región central y Maya lo cual no es representativo del Valle de Tehuacán.

Al no contar con evidencia geológica del sitio de estudio para elaborar una reconstrucción paleoclimática, la estrategia se centra en conocer las tendencias climáticas generales de la latitud en donde se encuentra Puebla desde el este al oeste a lo ancho de México. Analizando las publicaciones científicas del Holoceno, los *proxys* utilizados, interpretación y contexto biológico, se podrán encontrar puntos de inflexión interesantes en cuanto a variación climática que podrían ser indicativos de las condiciones climáticas en la región del Valle de Tehuacán o sus cercanías.



Figura 11. Regiones de Mesoamérica de interés.
Dentro del recuadro negro se señala la región de estudios paleoclimáticos analizados.
(Imagen de *wikimedia commons*)

Debido a la naturaleza extensiva del análisis se le denominó revisión sistémica de la bibliografía pues toma en consideración todos los estudios sin importar el *proxy* ni la implicación biológica que pudiesen arrojar adicional a la interpretación paleoclimática.

F.2 Estrategia paleogenómica para estudiar la estructura genética del maíz antiguo.

F.2.1 Paleoepigenética para estudiar la respuesta ambiental del maíz.

Diversos trabajos han demostrado que los patrones de metilación del ADN pueden reconstruirse a partir de secuencias antiguas empleando protocolos usados regularmente en muestras modernas, por ejemplo: conversión de bisulfito o métodos de enriquecimiento en grupos metilos (Pomraning *et al.* 2009; Clark *et al.* 2012). Éstos métodos generan datos con resolución de un sólo nucleótido, aunque en el caso de la secuenciación por bisulfito se pierde la muestra, mientras que el de enriquecimiento presenta un sesgo hacia islas CG, así como a fragmentos largos de ADN (Seguin-Orlando *et al.* 2015).

La secuenciación utilizando el tratamiento de bisulfito involucra la conversión de citosinas no metiladas a uracilos, seguida de secuenciación para determinar las posiciones metiladas del genoma (Figura 12).

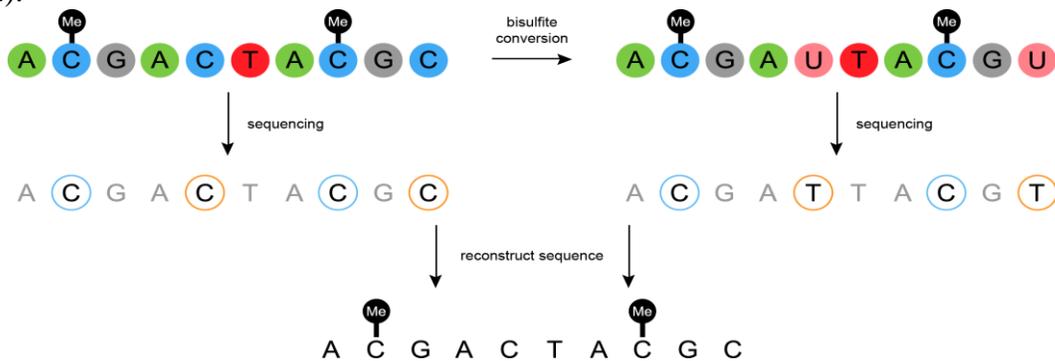


Figura 12. Diagrama de la conversión de bisulfito.

Al comparar las secuencias con y sin tratamiento nos da una reconstrucción de los sitios metilados. El principio de esta comparativa nos permite identificar las citosinas que no estaban metiladas en el genoma y como resultado, se infiere que aquellas que se mantienen como citosinas estaban metiladas originalmente. (Fuente: <http://www.slideshare.net/gattacaliao/workflow-of-analyzing-dna-methylation-data>)

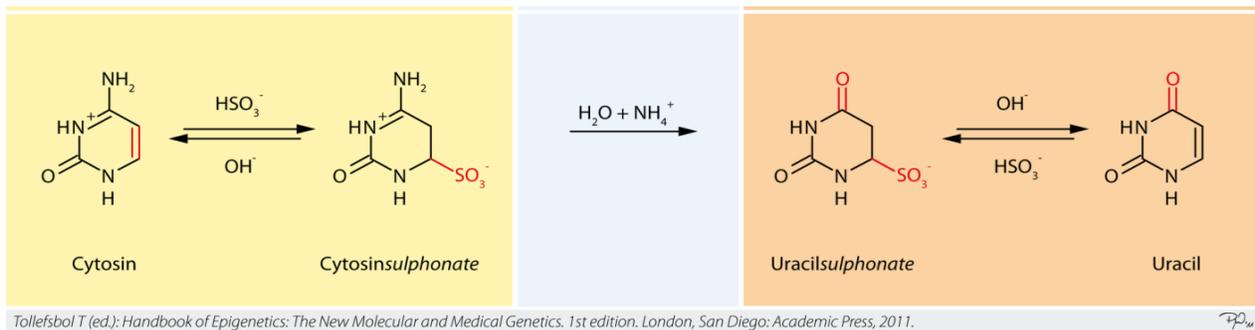


Figura 13. Esquema de la conversión mediante bisulfito de citosinas a uracilos.

Reacciones químicas que permiten la transformación de citosinas a uracilos en caso de no estar metiladas. De izquierda a derecha: Sulfonación, deaminación hidrolítica y desulfonación.

(Imagen obtenida de <http://core-genomics.blogspot.mx/2016/10/controlling-for-bisulfite-conversion.html>)

Cualquier cambio elemental que pudiese ocurrir en el estado epigenético de un locus es determinado por diversos factores ya mencionados: diferenciación en los tejidos dependiendo de su etapa de desarrollo, metilación en los loci controlados por alguna proteína que regule la metilación y factores ambientales tanto bióticos como abióticos, entre otros. Aunque los estudios paleoepigenéticos están confinados a ciertos tipos de tejidos que tienden a preservarse mejor dados los efectos diagenéticos (e.g. hoja, raíz, mazorca, semilla), varias de las regiones diferencialmente metiladas (DMR) que se detectan podrían ser funcionalmente relevantes para otros tejidos (Seguin-Orlando *et al.* 2015).

La comparativa de genomas debe tomar en cuenta el índice de mutaciones en el locus examinado; aquellos que evolucionan rápidamente (DNA mitocondrial y microsatélites) son adecuados para estudiar poblaciones y especies cercanas, mientras que los genes ribosomales ultraconservados son adecuados para comparar especies distantes. Consideraciones similares deben de tomarse en cuenta en epigenómica comparativa (Schmitz *et al.* 2013; Hernando-Herraez *et al.* 2013; Gokhman *et al.* 2014).

En este proyecto se pretende encontrar perfiles de metilación en muestras antiguas para dilucidar la magnitud de los efectos diagenéticos en el metiloma antiguo y entender cómo se comportaban las firmas epigenéticas durante épocas de sequía en maíz en pleno proceso de domesticación ya que para modelar el futuro debemos ser capaces de estimar el costo de las adaptaciones del pasado. En este proyecto interdisciplinario el contexto paleoclimático juega un papel fundamental para comprender el proceso de domesticación dependiente tanto de factores antropogénicos como de cambios climáticos y de paisaje.

CAPÍTULO II. HIPÓTESIS DE TRABAJO Y OBJETIVOS

2.1 HIPÓTESIS DE TRABAJO

Perfiles de metilación en muestras paleobotánicas de maíz nos permitirán estudiar cambios adaptativos al estrés ambiental.

2.2 OBJETIVOS

- 1) Estimar las fluctuaciones ambientales que ocurrieron 11 600 años antes del presente en el Valle de Tehuacán.
- 2) Determinar perfiles de metilación en el paleogenoma de al menos una muestra antigua de maíz (fecha de entre 2000 y 500 años antes del presente) del Valle de Tehuacán.
- 3) Buscar posibles relaciones ambientales en los perfiles de metilación en las muestras analizadas.

CAPÍTULO III. MATERIALES Y MÉTODOS

A continuación se muestra el esquema metodológico que se siguió durante este proyecto para estudiar posibles cambios adaptativos al estrés ambiental en muestras paleobotánicas mediante la implementación del tratamiento de bisulfito:



Figura 14. Esquema metodológico para el análisis comparativo de patrones de metilación en ADN antiguo

3.1 Revisión sistemática de bibliografía.

La revisión sistemática de bibliografía paleoclimática de la región geográfica de interés (Figura 11) se llevó a cabo en el Instituto de Geología de la Universidad Nacional Autónoma de México bajo la dirección del Dr. Alexander Correa-Metrio. Se comenzó con un acercamiento a las distintas evidencias indirectas o *proxys*, sus alcances y las interpretaciones de los expertos en cada región.

Comenzando por una rápida inspección al número de estudios realizado a lo largo del territorio mexicano, las zonas con mayor número de publicaciones se encuentran en la región central del país, la Zona Maya y el noroeste de la república Mexicana (Figura 15).

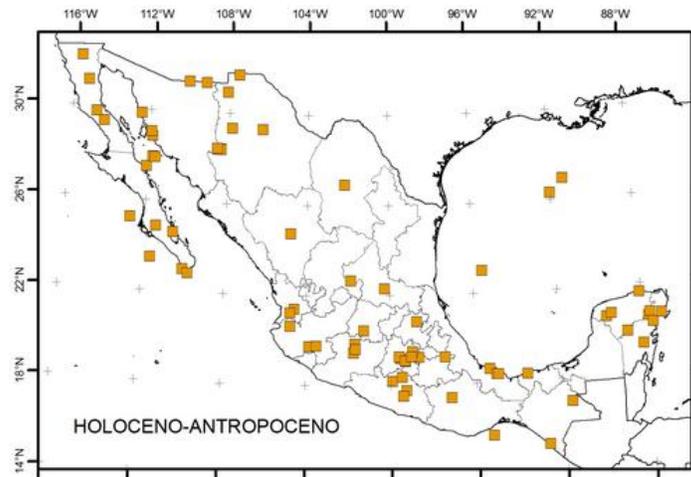


Figura 15. Mapa de estudios paleoclimáticos del Holoceno en México.

Los cuadros amarillos muestran los sitios con estudios paleoclimáticos datados en el Holoceno que se han realizado en México. (Lozano-García *et al.* 2016).

Debido a que la zona de interés se encuentra en Puebla, se eligieron todos aquellos estudios en donde se represente el clima holocénico en la franja latitudinal de 18-22° N a lo ancho del país. El número inicial de estudios científicos revisados fue de 115 (Apéndice). En dichos estudios se utilizaron diversos métodos de datación de acuerdo al tipo de material analizado y se estudió una amplia diversidad de *proxys* con lo cual los análisis estadísticos para cada uno es de distinta índole.

3.1.1 Criterios de clasificación de los estudios paleoclimáticos.

A pesar de contar con una amplia diversidad de datos para las distintas zonas de estudio, la continuidad en la cronología no siempre fue la más adecuada para representar fidedignamente las fluctuaciones ambientales de los distintos sitios. Algunos estudios presentan estimaciones paleoclimáticas basadas en evidencias insuficientes, proponiendo posibles fluctuaciones se contradicen con otros estudios de la misma zona o región. Así, en algunas regiones, diferentes *proxys* que representan la misma temporalidad condujeron a interpretaciones conflictuadas.

Mientras algunos estudios presentan aspectos puramente metodológicos en los que los autores se limitan a explicar los *proxys* estudiados, otros son de naturaleza subjetiva. Así, la generación de un marco paleoclimático regional se convirtió en un reto, no sólo a nivel técnico y metodológico sino también de interpretación. Para eliminar el sesgo de la subjetividad se analizaron alrededor de 115 estudios bajo los siguientes criterios:

i) *Proxys* estudiados.

En la paleoclimatología hay una amplia variedad de *proxys* dependiendo de las características geológicas de la zona de estudio. Las más comunes para la región de Mesoamérica son de varias índoles como se presentan en la Tabla 1.

Geológicas	Marinas (núcleos sedimentarios del océano): Sedimentos biogénicos (fósiles bénticos y planctónicos) Composición isotópica de oxígeno Abundancia de fauna y flora Elementos traza (ejemplos: Mg y Ca) Biomarcadores orgánicos Sedimentos inorgánicos (loess) Polvo terrestre, tamaño del grano Radios elementales (Pa y Th)
	Terrestres: Espeleotemas (isótopos estables y composición de elementos traza) Sedimentos lacustres y meteorización Depósitos eólicos Líneas lacustres
Biológicos	Anillos de árboles (espesor, densidad y composición isotópica) Polen (tipo, abundancia relativa y concentración absoluta) Diátomeas y ostrácodos en sedimentos lacustres (reconstrucciones, abundancia, geoquímica y biomarcadores). Insectos Corales Macrofósiles botánicos Distribución moderna de nichos ecológicos (refugios de poblaciones animales y de plantas)

Tabla 2. Fuentes principales de *proxys* con datos para reconstrucciones paleoclimáticas.

El cambio de Pleistoceno a Holoceno se dio hace 11 700 años antes del presente, por lo que aquellos estudios en donde los *proxys* tengan una resolución anual son extremadamente informativos. La duración de la temporalidad que se está describiendo depende del tipo de archivo natural obtenido, como se muestra en la siguiente tabla.

Archivo natural	Intervalo mínimo de muestreo	Orden temporal	Información potencial derivada
Anillos de árboles	Anual / Estacional	$\sim 10^4$	T, P, B, V, S
Sedimentos lacustres	Anual – 20 años	$\sim 10^4$ - 10^6	T, B, M, P, V, Cw
Corales	Anual	$\sim 10^4$	Cw, L, T, P
Polen	20 años	$\sim 10^6$	T, P, B
Espeleotemas	Anual	$\sim 5 \times 10^5$	Cw, T, P, V, B
Loess	100 años	$\sim 10^6$	P, B, M
Características geomórficas	100 años	$\sim 10^6$	T, P, V, L, P
Sedimentos marinos	100 años	$\sim 10^7$	T, Cw, B, M, L, P, S

Tabla 3. Características de los archivos naturales en donde se encuentran *proxys*.

T: temperatura; P: precipitación, humedad o balance de agua (precipitación- evaporación); C: composición química del aire (Ca) o agua (Cw); B: información de biomasa o patrones de vegetación; V: erupciones volcánicas; M: variaciones en campos geomagnéticos; L: nivel del mar; S: actividad solar. (Bradley & Eddy, 1991).

En un mismo sitio se puede tener más de un tipo de *proxy* y al ser las interpretaciones consistentes para las distintas evidencias, la reconstrucción paleoclimática en una temporalidad dada se va reforzando. Así, el primer criterio de selección de estudios de caso estuvo basado en la consistencia de la interpretación para toda evidencia indirecta obtenida dentro de un mismo sitio.

ii) Tecnología para analizar datos.

La tecnología en el área de la geología al igual que en otras disciplinas ha mejorado. Algunos de los estudios analizados más antiguos se hicieron para algunas regiones (Michoacán y el Valle de México) que presentaban *proxys* abundantes. Si bien la calibración de la datación por radiocarbono tiene como base la fecha de 1950 y está presente en la gran mayoría de los estudios, algunos estudios fueron datados utilizando series de uranio o de manera indirecta debido a una excelente identificación estratigráfica.

Aquellos que utilizaron tecnología que no permitía una resolución temporal tan fina, simplemente por el hecho de que en el momento de su estudio la tecnología presentaba ciertas limitantes, fueron omitidos de los análisis. Para esos mismos sitios, fue tomada en consideración la información que arrojaron trabajos modernos con una mejor resolución. En algunos casos, se revisaron *proxys* de núcleos extraídos en décadas pasadas pero se analizaron con tecnologías actuales, por lo que la incorporación de datos finos se volvió de suma importancia para obtener una adecuada interpretación por parte de los expertos.

iii) Análisis estadístico.

Las reconstrucciones paleoclimáticas siguen un número de etapas o niveles para su interpretación. En el primer nivel se recolectan datos en el campo y se continúa con mediciones y análisis iniciales de laboratorio (Hecht *et al.* 1979; Peterson *et al.* 1979). Para el segundo nivel, los datos extraídos son calibrados, convirtiéndolos en estimadores del paleoclima. La calibración puede ser totalmente cualitativa (condiciones tibias, húmedas o frías) o puede incluir un procedimiento que provee estimados cuantitativos de la variación climática a través del tiempo en un sitio geográfico específico.

Distintos procedimientos de calibración son, en potencia, diferentes reconstrucciones paleoclimáticas, por lo que la evaluación de los procedimientos es crítica para obtener una interpretación consistente. Para lograr tal consistencia, se utiliza un tercer nivel en donde se busca un patrón climático en una temporalidad específica (Mann *et al.* 2009). En algunos casos los arreglos del segundo nivel son transformados en estadísticos que los convierten en datos derivados del tercer nivel.

Aquellos estudios que cuentan con datos cuantitativos estadísticos (del tercer nivel), no solamente cualitativos o descriptivos (primero y segundo nivel), permiten obtener una probabilidad de certeza objetiva, a diferencia de aquellos que son únicamente interpretativos y dependen únicamente de la experiencia y agudeza del investigador que los conduce.

Tomemos en cuenta que puede existir riesgo de contaminación de los *proxys* debido a su manipulación, por lo que es fundamental realizar un análisis estadístico en donde se correlacionen *proxys* disponibles para eliminar falsos positivos. Como ejemplo tenemos estudios actuales donde las capas estratigráficas superficiales son descartadas dada la influencia humana en el sitio de estudio. Además puede existir contaminación aún sin la influencia humana. Ejemplo de ello es el crecimiento desmesurado de alguna raíz que al ser removida, modifica la estratificación reflejándose así un error en la interpretación.

Finalmente, la correcta datación (ver “Métodos de datación” en la página 36) nos permite tener un contexto temporal definido, con el cual se obtiene una interpretación paleoclimática consistente. Los estudios que contaban únicamente con datos del primer o segundo nivel, fueron descartados del meta-análisis.

iv) Objetividad en la discusión.

Una proporción considerable de estudios realizados en el Valle de México usan como principal *proxy* la palinología, es decir, el estudio del polen que se encuentra en sedimentos lacustres. El polen es un excelente indicador biológico del tipo de vegetación que creció en un sitio y dado el conocimiento actual de las condiciones en las que ciertas plantas crecen, es posible inferir cómo era el clima del sitio en su momento.

Desafortunadamente, si había influencia humana, el tipo de vegetación se modificaba, por lo que es importante complementar con otros *proxys* para darle un mayor peso a las interpretaciones. Algunos estudios al no presentar más de una evidencia indirecta sobre el paleoclima o elucubrar sobre las razones por las cuales se encuentran ciertos tipos de taxones, sin reforzar con otros *proxys*, fueron descartados, debido a que en algunos casos los autores proponían escenarios que posteriormente, al contrastarlos con las reconstrucciones de otros autores, no coincidían.

Dado que se trata meramente de un análisis bibliográfico, es imposible conocer la geografía, hidrología y orografía de cada sitio por lo que aquí la voz más importante es siempre la del autor que llevó a cabo el estudio. Al no compaginar las versiones de los distintos autores, los estudios menos objetivos, fueron descartados para no crear inconsistencias al interpretar el paleoclima holocénico.

v) Datos continuos en la temporalidad de interés.

Como criterio final, se valoró la continuidad de datos en la temporalidad de interés. Dada la naturaleza propia de los *proxys*, en ocasiones no guardan un registro continuo de las fluctuaciones ambientales. En el caso de sedimentos lacustres, esto se debe al movimiento propio de la tectónica de placas, vulcanismo, meteorización o deslave de tierras altas.

Además, la temporalidad holocénica, si bien es relativamente corta en comparación con otras épocas geológicas, presenta fluctuaciones variables en los registros naturales, obteniendo datos discretos. Es por ello que se requiere una mayor cantidad de estudios para una acertada reconstrucción paleoclimática.

Aquellos estudios que nos permiten una reconstrucción paleoclimática continua ya sea por sí solos o que aportan datos faltantes en otros estudios con la misma temporalidad, cumplen el último criterio necesario para este proyecto. Siguiendo estos criterios obtuvimos tres estudios paleoclimáticos que cumplían con todos ellos. Los demás estudios aunque interesantes, quedaron en las siguientes categorías de acuerdo a la clasificación antes mencionada:

Criterio a evaluar	No. de artículos que cumplen el criterio
Análisis estadístico	60
Tecnología para analizar datos	30
Proxys estudiados	18
Objetividad en la discusión	11
Datos continuos en la temporalidad de interés	3

Tabla 4. Número de estudios paleoclimáticos categorizados mediante criterios de análisis.

3.2 SELECCIÓN DE MUESTRAS ARQUEOBOTÁNICAS.

3.2.1 Excavación de 2012

En 2012 se formó un grupo interdisciplinario conformado por el Profesor Ángel García Cook (quien perteneció al grupo de trabajo de la expedición de MacNeish en 1964) de la Subdirección de Investigación y Conservación del Patrimonio Arqueológico, el Arqueólogo Javier Martínez González adscrito al centro INAH en la Subdirección de Salvamento Arqueológico y los Doctores Jean-Philippe Vielle-Calzada y Rafael Montiel Duarte, así como el entonces estudiante de Maestría Miguel Andrés Vallebuena Estrada, estos últimos provenientes del CINVESTAV Irapuato, para buscar evidencia paleobotánica en el Valle de Tehuacán.

En el pasado, las técnicas arqueológicas de colecta y preservación de restos antiguos se basaban en análisis morfológicos por lo que el manejo y reconstrucción de las piezas derivaba en contaminación de las muestras, como la aplicación de adhesivos o barnices en la superficie de las muestras arqueológicas.

La contaminación es una de las principales restricciones en ADN antiguo, no sólo del tratamiento posterior a la extracción de las muestras: también hay contaminación de ADN moderno desde el proceso de excavación hasta el análisis en el laboratorio, por lo que se implementó un protocolo especializado en el grupo de trabajo interdisciplinario para excavaciones arqueológicas en general sin comprometer la integridad del material, que cumpliera con estándares de calidad de extracción permitiendo así la preservación de material genético endógeno para posteriormente aplicar protocolos experimentales de análisis paleogenómicos.

Los resultados de esta excavación fueron fructíferos, se obtuvieron aproximadamente 30 restos de maíz (olotes en su mayoría), agave, frijol, aguacate, guaje y chile entre otros. Se realizó un registro fotográfico y morfológico de cada muestra para documentar el fenotipo en caso de ser procesadas posteriormente ya que la extracción de ADN es un método destructivo.

3.2.2 Métodos de datación

Es fundamental una correcta datación, eliminando trazas de contaminantes para asegurar que el contexto estratigráfico sea interpretado correctamente. De igual manera hay que comprender los alcances y limitaciones del procedimiento. En ocasiones es más importante conocer el margen de error que la fecha por sí misma (Walker, 2005). Los métodos de datación se dividen en cuatro categorías. La

primera son los radiosotópicos que están basados en el índice de desintegración en una muestra o sus cercanías. Los métodos de correlación paleomagnética se basan en cambios del campo magnético terrestre y sus posibles efectos en las muestras. Los métodos químicos pueden ser orgánicos o inorgánicos y están basados en cambios químicos dependientes del tiempo en una muestra o características químicas propias de las muestras. Finalmente los métodos biológicos los cuales dependen del crecimiento de un organismo para datar el sustrato en el que se encontró (Colman *et al.* 1987).

Dependiendo de la temporalidad de interés hay métodos que se vuelven más adecuados. No todos los métodos dan una edad certera, en muchos casos es relativa, permitiendo calibrar dicha edad relativa a métodos numéricos (caso de radioisótopos). Al igual que en los métodos de datación, los métodos de calibración tienen un amplio espectro de métodos: numéricos, calibrados, relativos y aquellos que involucran correlaciones estratigráficas (Colman *et al.* 1987). Para las muestras paleobotánicas el más adecuado es el de datación por radioisótopos como se discutirá a continuación.

3.2.2.1 Datación por radioisótopos

Los átomos están conformados por neutrones, protones y electrones. Para cualquier elemento el número de protones (número atómico) no es variable pero el de neutrones sí lo es, dando como resultado isótopos del mismo elemento. El carbono por ejemplo, tiene tres isótopos con masa atómica distinta: 12, 13 y 14 que se denominan ^{12}C , ^{13}C y ^{14}C . Generalmente cada elemento tiene más de un isótopo estable, mientras que aquel isótopo que es inestable pasa por decaimiento radioactivo espontáneo al perder partículas nucleares (partículas α ó β), transmutando en un nuevo elemento. Siguiendo con el ejemplo del carbono, éste decae a ^{14}N . El índice de decaimiento no varía por lo que podemos estimar el tiempo en el que decaerá a su isótopo hijo. Ésta es la base de datación por cualquier método isotópico.

Si el reloj isotópico comienza cercano a la fecha de interés en un estrato, la medición actual de la concentración isotópica indicará la cantidad de tiempo que ha pasado desde que la muestra quedó en el sitio de descanso. La cantidad de tiempo que tarda el material radioactivo en decaer la mitad de su concentración inicial es conocida como vida media. Como ejemplo de lo anterior, si una planta murió 5730 años atrás con un 50% de contenido de ^{14}C , en 5730 años quedará con el 25%. En 1955 Libby explicó los principios de la datación por radiocarbono y calculó una vida media de 5568 años. En 1960 se demostró que había un error del 3% (Godwin, 1962). Para evitar confusión, se utiliza la vida media de Libby pero al comparar los años de radiocarbono con el calendario actual hay que ajustar esa diferencia.

Átomo	Vida media
^{14}C	5.730×10^3 años
^{238}U	4.468×10^9 años
^{235}U	0.704×10^9 años
^{40}K	1.310×10^9 años

Tabla 5. Vida media de radioisótopos usados en datación.

Aunque la vida media del ^{14}C es de 5730 ± 40 años, por convención "Libby" se usan 5568 ± 30 años.

Para que un radioisótopo sea útil para datación debe contar con algunos atributos:

- i. El isótopo o sus productos deben ser cuantificables y diferenciables de otros isótopos o que su decaimiento pueda ser medible.
- ii. Su vida media debe estar en el periodo de interés apropiado.

- iii. La concentración inicial debe ser conocida
- iv. Debe existir alguna correlación entre el evento que se pretende datar y el inicio del decaimiento radioactivo.

3.2.2.2 Principios de datación por ^{14}C

En general, para las fluctuaciones climáticas del cuaternario, la datación por radiocarbono es la más útil. Dada su distribución en todo el mundo se puede encontrar en madera, hueso, conchas, sedimentos marinos y lacustres. Se produce en la atmósfera por el bombardeo de neutrones a átomos de Nitrógeno:



Los neutrones se producen por la radiación que entra a la tierra. La difusión de átomos de ^{14}C en la atmósfera baja permite una homogénea distribución en la biosfera y una vez ahí se oxida a $^{14}\text{CO}_2$ y se mezcla con el resto del dióxido de carbono presente. Existe un equilibrio entre el índice de nuevo ^{14}C y el decaimiento en las reservas existentes de carbono lo que permite una adecuada aplicación del método aunque existen ciertas excepciones a la regla.

3.2.2.3 Procedimientos de medición, materiales y problemáticas

A principios de los años 80, las técnicas de datación se basaba en contadores proporcionales de gas y contadores de centelleo líquido (Povinec *et al.* 2009). En el contador proporcional de gas, el carbón se convertía en gas (metano, dióxido de carbono o acetileno) y posteriormente se medía la variación de los pulsos de voltaje al ser proporcionales a las emisiones de partículas- β en un contador proporcional. En los contadores de centelleo líquido, el carbono se convertía en benceno o cualquier otro líquido inorgánico y se colocaba en cualquier instrumento que producía centelleos de luz cuando las partículas- β interactuaban con fósforo que había sido añadido previamente a la solución orgánica. En ambos métodos son necesarias medidas estrictas para proteger a las muestras de radioactividad creada por componentes instrumentales. Pocos laboratorios usan estas técnicas actualmente por ser lentas y requerir cantidades que se cuentan en gramos de carbono por muestra.

Existió una revolución a finales de 1970 y principios de 1980 al desarrollarse un método de datación utilizando poca muestra (~1mg de carbono) y un acelerador acoplado a un espectrómetro de masas (AMS, por sus siglas en inglés) (Muller, 1977; Nelson *et al.* 1977; Litherland & Beukens, 1995). En esta técnica, en vez de medir la cantidad de ^{14}C de manera indirecta por la emisión de partículas β , las concentraciones de iones individuales (^{12}C , ^{13}C y ^{14}C) son cuantificadas directamente. Se pretende que esta tecnología permita desarrollos tecnológicos en datación por AMS permitiendo datar muestras de carbono tan pequeñas como ~1 μg (Povinec *et al.* 2009).

3.2.2.4 Exactitud en las muestra de radiocarbono.

Las fechas de radiocarbono tienen un rango de probabilidad, al igual que todas las medidas radiométricas. La desintegración radioactiva varía aleatoriamente alrededor de un valor medio. Es esta incertidumbre estadística en cada muestra, además de la calibración del comienzo del decaimiento y cualquier ruido externo de radiación, que crea ruido inherente en la datación de ^{14}C . Para dar certeza a las mediciones de ^{14}C , se puede tomar como referencia el isótopo ^{13}C debido a que el fraccionamiento del radioisótopo es el doble del ^{13}C (Craig, 1953). El contenido de ^{13}C es expresado a partir del estándar PeeDee Belemnite:

$$\delta^{14}\text{C} = 2\delta^{13}\text{C} = \frac{\left(\frac{^{13}\text{C}}{^{12}\text{C}}\right)_{\text{sample}} - \left(\frac{^{13}\text{C}}{^{12}\text{C}}\right)_{\text{PDB}}}{\left(\frac{^{13}\text{C}}{^{12}\text{C}}\right)_{\text{PDB}}} \times 10^3\text{‰}$$

Al adoptar este valor de referencia, se permite hacer una comparativa entre todos los valores de radioisótopo sin importar su origen. La datación reportada es el punto medio de la distribución de probabilidad de Poisson junto con su desviación estándar lo que define un nivel conocido de probabilidad para una muestra (Craig, 1953).

3.2.2.5 Datación en las muestras de maíz

Las muestras paleobotánicas de maíz pasaron por un proceso de selección, teniendo como resultado un conjunto de 29 muestras que presentaban buena preservación morfológica. Para cada muestra se extrajo un fragmento de 20mg y mediante la técnica de espectrometría por aceleración de masas se procedieron a datar en un servicio de datación externo ofrecido por la empresa Beta Analytic.

3.3 Extracción de ADN en muestras paleobotánicas

3.3.1 Laboratorio de ADN antiguo

Dentro de las instalaciones del Laboratorio Nacional de Genómica para la Biodiversidad (Langebio), se tienen instalaciones especializadas denominadas cuarto limpio para la extracción de ADN de muestras antiguas, diseñadas para minimizar las posibilidades de que las muestras extraídas de excavaciones se contaminen, ya sea por contaminación cruzada entre muestras o de forma exógena al manipularlas.

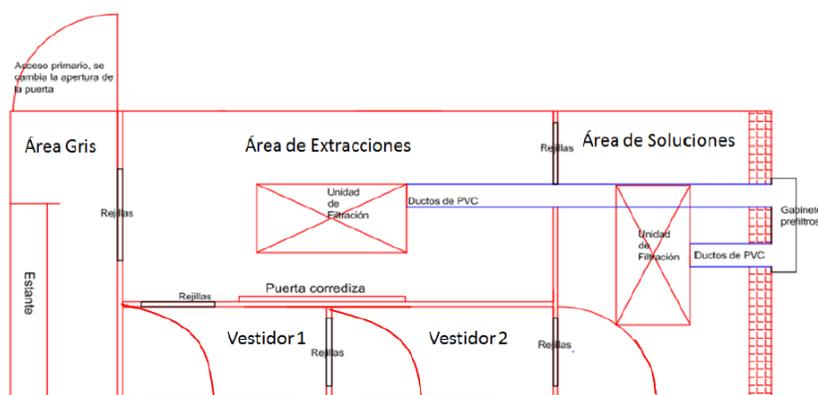


Figura 16. Plano de planta del cuarto limpio en LANGEBIO-CINVESTAV Irapuato.

El laboratorio de ADN antiguo está acondicionado siguiendo especificaciones recomendadas para cuartos limpios de uso militar (Engineers, 2004). Está conformado por paneles de cristales que lo aíslan del ambiente exterior simulando una campana de flujo de aire con presión positiva. El aire dentro del cuarto es filtrado mediante filtros HEPA no permitiendo la entrada al 99.97% de partículas mayores a 0.03 micras. Antes de entrar al área de trabajo el aire es irradiado con lámparas de luz UV, eliminando así bacterias e inactivando el posible ADN contaminante, al propiciar la generación de anillos de ciclobutilo en timinas adyacentes (Brown, 2002).

Existe un protocolo especializado para entrar al cuarto limpio (Anexo 1), que comienza con el ingreso al laboratorio a través del área gris en donde se colocan guantes, botas quirúrgicas, cofia y cubre bocas. Posteriormente se entra al área de vestidores que cuentan con equipo especial para entrar al área de soluciones o de extracciones respectivamente.

El área de soluciones tiene la función de preparar soluciones químicas necesarias para los procesos de extracción, PCR y preparación de bibliotecas genómicas. Una vez preparadas, las soluciones se llevan al área de extracciones donde se procesan las muestras antiguas procedentes de contextos arqueológicos. El procesamiento de las muestras se realiza en una campana de seguridad biológica que impide la entrada o salida de aire, disminuyendo el riesgo de contaminación por partículas y aerosoles durante la manipulación de las muestras y soluciones durante la extracción del ADN.

3.3.2 Extracción y purificación de ADN de maíz antiguo

Se comienza preparando las soluciones en el área de soluciones siguiendo el protocolo modificado de Capellini y col. (2010) (Anexo 2). Posteriormente, en el área de extracciones se hace una digestión bioquímica logrando lisis celular para liberar el contenido de la célula. Para purificar y obtener únicamente el ADN se utilizan solventes orgánicos como fenol y cloroformo. Se separan los solventes del ADN usando centrifugación, al separar gracias a las diferentes densidades en los componentes orgánicos y obtener finalmente el ADN antiguo purificado.

3.3.3 Amplificación diagnóstica para el gen Rubisco

Posterior a la extracción del ADN se realizó una reacción en cadena de la polimerasa (PCR) para identificar muestras candidatas a ser secuenciadas masivamente. Para lograrlo se eligió un gen que estuviese presente en todas las plantas y que fuese esencial. Se eligió el de rubisco (*rbcL*) presente en cloroplasto y del que existen muchas copias en una sola célula, lo que aumenta las probabilidades de amplificación en los candidatos.

Se utilizaron iniciadores previamente reportados para plantas antiguas (Poinar *et al.* 1998; Bærholm *et al.* 2010; Wales *et al.* 2012) con temperatura y ciclos previamente reportados en trabajos previos en el laboratorio (Vallebuena Estrada, 2014). Las reacciones fueron reveladas en geles de agarosa al 3% y aquellas que presentaron banda de amplificación de 140 pares de bases fueron consideradas candidatas para análisis posteriores.

3.4 Elaboración de bibliotecas genómicas.

Una vez confirmada la presencia de ADN de planta en las muestras de maíz se procedió a elaborar las bibliotecas genómicas para enviarlas a secuenciar. La elaboración de bibliotecas consiste en una serie de pasos a nivel molecular permitiendo fragmentar, reparar, etiquetar y purificar el ADN proveniente de una alícuota de ADN extraída de cualquier muestra biológica.

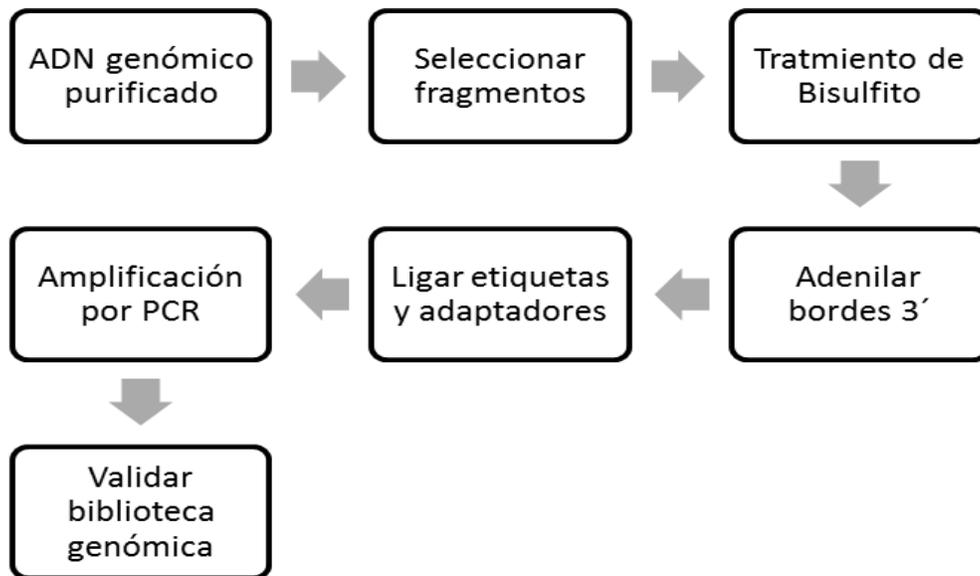


Figura 17. Esquema metodológico modificado para la elaboración de bibliotecas genómicas.

La preparación de las soluciones para el tratamiento se realizó en el área de soluciones. El tratamiento se realizó en el área de extracciones partiendo de una fracción de ADN antiguo (aADN) purificado mientras que la otra alícuota quedó intacta (Figura 18).



Figura 18. Esquema de las alícuotas utilizadas durante la elaboración del tratamiento de bisulfito.

Tanto de la muestra de SM1 como Cox17 se tomaron dos alícuotas de las cuales una fue tratada con bisulfito. Las alícuotas dentro de los cuadros rojos se usaron para seguir con la elaboración de las bibliotecas.

El kit usado para la elaboración de bibliotecas genómicas fue el *TruSeq Nano DNA* con un par de modificaciones como se muestran en el esquema de la Figura 17.

- Paso adicional para realizar el tratamiento de bisulfito en una alícuota del extracto de ADN.
- Omisión de la fragmentación genómica del ADN. La única ventaja técnica que ofrece el ADN antiguo es su naturaleza fragmentada gracias a los procesos diagenéticos.

A excepción del tratamiento de bisulfito, todos los demás pasos de la elaboración de la biblioteca genómica se realizaron para ambas alícuotas siguiendo las instrucciones del protocolo, de manera que quedaran dos bibliotecas con dos adaptadores distintos para cada muestra como se observa en la figura 18.

Gracias al trabajo previo llevado a cabo en el laboratorio, los niveles de fragmentación son conocidos para las muestras de maíz del Valle de Tehuacán, aunque hay que considerar que los procesos

diagnéticos pueden afectar de distintas maneras en las muestras de un mismo sitio, por lo que es necesario descartar una alteración en el tamaño de los fragmentos.

Se confirmó que la fragmentación fuera consistente en todas las muestras mediante un análisis cualitativo de electroforesis. Este análisis fue pedido en el área de servicios genómicos de LANGEBIO y se realizó en un Bioanalyzer con un chip de alta sensibilidad de ADN. Es posible cuantificar la cantidad de ADN a partir de este mismo ensayo pero se eligió un análisis fluorimétrico para proceder con la cuantificación por ser más preciso. En este trabajo se considera al análisis de electroforesis como cualitativo. Se procedió a elegir los fragmentos con una degradación ya conocida que pudiesen pertenecer al ADN endógeno del maíz y no a contaminación moderna. Se seleccionaron los fragmentos deseados usando perlas AMPURE (ver anexo 4) para proceder con el tratamiento de bisulfito.

3.4.1 Tratamiento de Bisulfito

El tratamiento de bisulfito se llevó a cabo con el kit de zymo research “EZ DNA Methylation-Direct”. Se eligió este kit gracias a su eficiencia al completar la conversión de bisulfito con bajas concentraciones de ADN como 10 células o 50 pg (Anexo 3). Se puede resumir en tres pasos:

- Reacción de conversión
- Lavado de desulfonación
- Elución.

El tratamiento de bisulfito fue elaborado bajo la supervisión de Dalia Rodríguez, técnico de laboratorio de la Dra. Gertrud Lund, responsable del laboratorio de epigenética del Cinvestav Irapuato. Durante la última fase de elución, Dalia sugirió el uso de una elución adicional, dada su experiencia con las columnas usadas en los kits de zymo research. Este paso adicional en el protocolo promovió una extracción mayor del aADN tratado.

Ambas bibliotecas (con y sin tratamiento) serían secuenciadas en una misma línea por lo que los adaptadores debían ser compatibles entre ellos para mezclarlos (*pooling*). Al contar con adaptadores tipo A en nuestro kit, los candidatos para las bibliotecas son los que se muestran en la tabla a continuación:

Plexity	Option	Set A Only	Set B Only
2	1	AD006 and AD012	Not recommended
	2	AD005 and AD019	
3	1	AD002 and AD007 and AD019	AD001 and AD010 and AD020
	2	AD005 and AD006 and AD015	AD003 and AD009 and AD025
	3	2-plex options with any other adapter	AD008 and AD011 and AD022
4	1	AD005 and AD006 and AD012 and AD019	AD001 and AD008 and AD010 and AD011
	2	AD002 and AD004 and AD007 and AD016	AD003 and AD009 and AD022 and AD027
	3	3-plex options with any other adapter	3-plex options with any other adapter

Tabla 6. Adaptadores para estrategias de *pooling* en kits TruSeq de ADN

Por la naturaleza del tratamiento en nuestras muestras, se tendrían dos bibliotecas en el pool y a nivel técnico se secuenciarían dos genomas simultáneamente, proceso conocido en inglés como *multiplexing*. Siguiendo la tabla, las opciones 1 y 2 del set A con un nivel 2 de *multiplexing* fueron las adecuadas para cualquier maíz candidato.

Finalmente, para la validación de las dos bibliotecas (con y sin tratamiento) se volvió a realizar un análisis de electroforesis y un análisis cuantitativo fluorimétrico. El primero se realizó en LANGEBIO con el equipo antes descrito y el último se realizó en la Universidad de California en Davis con un equipo Qubit (Life Technologies).

3.5 Secuenciación masiva

La tecnología de secuenciación utilizada en este proyecto fue Illumina *Next Generation Sequencing* (NGS). El principio de secuenciación consiste en identificar bases de ADN y de manera simultánea procesarlas en una cadena de ácidos nucleicos. Cada tipo de base emite una señal fluorescente única conforme se une a la cadena creciente, con lo que se determina el orden de las secuencias del ADN y se va guardando en bases de datos computacionales. La secuenciación se hizo de tipo *paired-end* (figura 19), es decir, el mismo fragmento fue secuenciado a partir de los dos extremos para generar datos en forma de lecturas (*reads*) que facilitarían su alineamiento y así aumentar la calidad, sin importar si hay elementos repetitivos en el genoma de referencia además de detectar inserciones, deleciones e inversiones (Heather & Chain, 2015).



Figura 19. Esquema de lectura de un segmento *paired-end*.

Se muestran cómo se detectan los *reads* (azul y rojo) en un mismo fragmento (escala de morado a amarillo) de una cadena de ADN.

Las dos bibliotecas de cada muestra (con y sin tratamiento de bisulfito) fueron secuenciadas de manera simultánea en un mismo pool y se generaron *reads* para ambas bibliotecas que posteriormente se separaron gracias a la identificación de los adaptadores utilizados durante la elaboración de las bibliotecas genómicas.

El trabajo de secuenciación fue realizado en la Universidad de California con un equipo Illumina HiSeq 3000, en el centro de genomas con la ayuda del equipo de Lutz Froenicke quien además apoyó durante el envío de las muestras.

3.6 Análisis Bioinformático

A partir de la secuenciación se generan bases de datos en donde se encuentran las secuencias de nucleótidos que además cuentan con características como identificador para cada una de las secuencias o *reads*, calidad de secuenciación a nivel de sitio nucleotídico, información sobre la corrida del secuenciador y el sentido de lectura del *read*. Esto es lo que se denominan datos crudos en lo sucesivo.

Para obtener respuestas biológicas hay que procesar los datos crudos. Para cada set de datos crudos en general se sigue la misma forma de procesamiento informático (*pipeline*) (Figura 20). Sin embargo, hay ligeras modificaciones de parámetros dependiendo de la naturaleza técnica de los datos y sobre todo, del tipo de preguntas que pretendamos resolver.

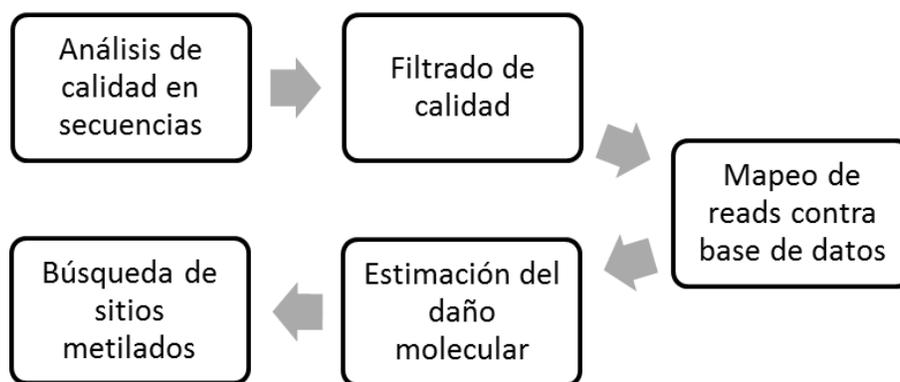


Figura 20. Esquema general del procesamiento informático

Para las secuencias de maíz paleobotánico, el análisis de calidad inicial en las secuencias crudas se realizó con el programa Fastqc (Andrews, 2010) permitiendo ver la tasa de errores en cada *read* mediante el *Phred score*: este parámetro guarda una relación con la probabilidad de error en cada base leída por el secuenciador como se muestra a continuación:

Q	P	Exactitud
10	1 en 10	90%
14	1 en 41	94%
17	1 en 50	98%
20	1 en 100	99%
30	1 en 1,000	99.90%
40	1 en 10,000	99.99%

Tabla 7. Relación entre el Phred score (valor de calidad) y la exactitud de la secuenciación por sitio nucleotídico.

En donde: $Q = -10 \log_{10} P$

Q= Calidad en Phred score, P= Probabilidad de error

Una vez analizados los valores del Phred score se procede con el filtrado de calidad utilizando el programa Cutadapt (Marcel, 2011). Posteriormente se “rasuran” los adaptadores, se eliminan *reads* con baja calidad y finalmente, se eliminan aquellos *reads* que tuvieran un tamaño menor de 20pb. Este valor es conocido a priori gracias al análisis cualitativo de electroforesis durante la elaboración de las bibliotecas.

Una vez que se cuenta con las secuencias sin adaptadores procedemos a mapear. Para ello se toma un genoma de referencia contra el cual se van a comparar cada par de *reads* para identificar el segmento secuenciado con la referencia, conocer su posición, calidad y descartar aquellos *reads* no reconocidos como maíz. Este paso se realizó con el programa BWA con la versión mem (Li & Durbin 2009); dicha modalidad permite mapear fragmentos pequeños.

El genoma de referencia utilizado para este proyecto fue la versión 3 del genoma de la accesión B73 de *Zea mays*. Adicionalmente se usó la versión enmascarada, la cual carece de las regiones repetitivas del genoma, al utilizar un algoritmo que permite omitir regiones que aparecen en más de un sitio al reemplazarlas con “N”. De esta manera, aquellos *reads* mapeados contra regiones únicas no representarán duplicados biológicos en esta fase. Los conceptos de profundidad y cobertura en los estudios genómicos se refieren a la cantidad del genoma representado. Idealmente se querría que todos los nucleótidos del genoma se representaran al menos una vez independientemente del método de secuenciación utilizado, a eso se le llama cobertura.

Los *reads* obtenidos por cualquier método de secuenciación son menores a 250 nucleótidos, al mapearlos la representación del genoma de referencia es muy variable, en muchos casos no está cubierto en su totalidad. Además puede contener errores de secuenciación, lo cual no permite distinguir entre error y variación en la secuencia. Para resolver ese problema se incrementan el número de *reads* secuenciados y al obtener varias representaciones de la misma posición del nucleótido (profundidad) es posible discernir entre un error ya sea por la secuenciación o daño molecular de un sitio metilado en el genoma antiguo.

Otra consideración muy importante en el mapeo son los parámetros, pudiendo ser más flexibles o estrictos según la semejanza que se busque entre los *reads* y la referencia. Para identificar la estrategia de mapeo idónea para cada biblioteca, se tomaron secuencias ya conocidas de maíz antiguo que fueron sometidas al tratamiento de bisulfito “in silico” y después fueron mapeadas, ajustando los parámetros mediante la comparación de estos resultados con los del mapeo de los mismos *reads* sin el tratamiento “in silico”. Esto nos ayudó a determinar el punto intermedio idóneo de los parámetros a utilizar para ambas bibliotecas. Los *reads* de maíz antiguo utilizados provenían de una muestra paleobotánica de San Marcos, identificada como SM10, con una edad aproximada de 5100 años antes del presente. Los resultados de esta comparativa permitieron ajustar los parámetros de mapeo adecuados para proceder con la estimación del daño molecular como se muestra en la sección de resultados.

Trabajos previos dentro del área de ADN antiguo permiten estimar el nivel de fragmentación y daño molecular de muestras, independientemente de la cantidad de ADN endógeno en la muestra causado por los procesos diagenéticos. A pesar de que en pasos previos se estimó el tamaño de fragmento mediante el análisis cualitativo de electroforesis, falta cuantificar incorporaciones erróneas promedio para los segmentos secuenciados. Este daño consiste en patrones ya conocidos (Citosinas > Uracilo) que nos permiten autenticar secuencias antiguas distinguiéndolas de las modernas; sin embargo, al secuenciar usando plataformas NGS, los uracilos son traducidos como Timinas.

Dependiendo de la plataforma utilizada para secuenciar, los patrones encontrados de deaminación pueden variar debido a la forma de construcción de la biblioteca o su amplificación (Meyer *et al.* 2012; Ginolhac *et al.* 2011). Si se sigue el protocolo de Illumina, es posible encontrar un número mayor de citosinas deaminadas en los extremos 5', un incremento de sustitución de C > T al inicio de las secuencias, además de un incremento complementario en G > A hacia el final de las secuencias (Meyer & Kircher, 2010; Briggs *et al.* 2007). Al compararlo con el protocolo en el que se utilizan plantillas de una sola hebra de ADN, las sustituciones de C > T se elevan en ambos extremos del fragmento (Meyer *et al.* 2012).

Se utilizó el programa mapDamage 2.0 (Jónsson *et al.* 2013) que integra un modelo estadístico bayesiano del daño de ADN en secuencias NGS tomando como factores no sólo la deaminación sino además el rompimiento de hebras de ADN. Utiliza los alineamientos de los *reads* para referenciar genomas estimando cuantitativamente parámetros de daño molecular. Esta aproximación permite comparar niveles de daño en ADN a través de gradientes temporales y ambientales al comparar diversas muestras de un mismo sitio.

Una vez estimado el daño molecular para nuestro genoma antiguo, en las secuencias se procedió a buscar sitios metilados en los que el cambio de C > T no fuera estocástico. Para lograrlo, se mapearon con base en genomas en los cuales se enmascararon las citosinas a timinas y se corroboró con otro mapeo, esta vez enmascarando las guaninas a adeninas, esto se hizo con el fin de identificar sitios únicos que fueron efectivos bajo el tratamiento de bisulfito. Lo anterior queda descrito en el siguiente esquema y la finalidad es evitar encontrar sitios erróneamente mapeados debido a al alto contenido de timinas en un read o por estocasticidad.

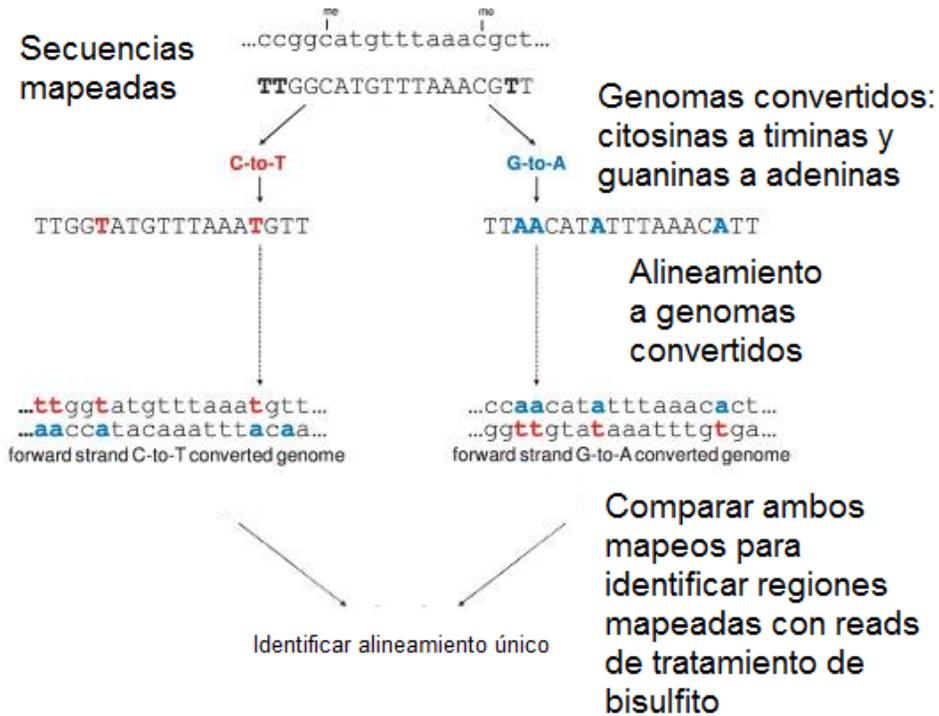


Figura 21. Diagrama de la estrategia seguida para la identificación de alineamientos únicos.

A partir del genoma de maíz se convirtió en otros dos genomas de referencia, uno con todas las citosinas convertidas a timinas y el segundo con las guaninas a adenina y se procedió a mapear los *reads* con tratamiento de bisulfito (secuencias mapeadas) a los genomas anteriormente mencionados. Una vez obtenidos los mapeos, se compararon las secuencias identificadas en ambos mapeos para determinar que fueran correctamente alineadas. (Diagrama modificado de Krueger, 2011).

Posteriormente, las secuencias encontradas fueron filtradas por la profundidad en una posición nucleotídica que tenga como mínimo tres *reads* (3x). En la parte superior de la figura 22 se muestra el ADN de maíz con secuencias únicas que ha sido tomado como referencia, mientras que las secuencias subsecuentes representan ese mismo fragmento cubierto en un determinado porcentaje y con una profundidad de 4x de la biblioteca de aADN en donde se encuentran citosinas deaminadas por procesos diagenéticos y además con el tratamiento de bisulfito.

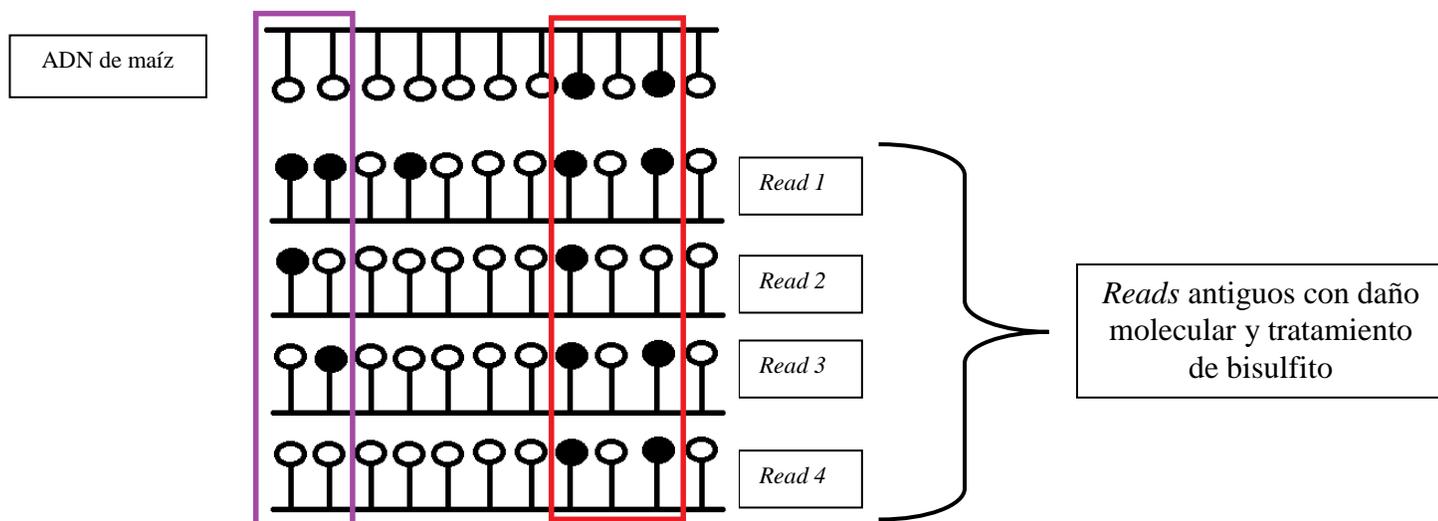


Figura 22. Diagrama representativo del discernimiento entre daño molecular con respecto a tratamiento de bisulfito.

En donde los reads en con una citosina metilada $\overset{\bullet}{\text{C}}$ > T [Esta timina sufrió deaminación ya sea por procesos diagenéticos, por error de secuenciación o por tratamiento de bisulfito en el caso de los reads antiguos]; $\overset{\circ}{\text{C}}$ representa una citosina no metilada. Las demás bases [A,G] no son representadas pero están implícitas en la línea base. Las comparativas dentro del recuadro morado son ejemplos de daño molecular causado por daño molecular y/o errores en la secuenciación (Meyer & Kircher, 2010). En el recuadro rojo se tienen ejemplos de lo que consideraríamos citosinas metiladas, dada la alta frecuencia representada en un mismo sitio aún si se desconoce el estado inicial de la secuencia.

De esta manera se puede discernir entre errores de secuenciación o daño molecular y sitios metilados con cierta probabilidad de certeza.

Finalmente, para inspeccionar los sitios previamente identificados como posibles blancos de metilación, procedimos a utilizar el programa IGV (Robinson *et al.* 2011) para visualizar los niveles de hiper o hipometilación en dos accesiones de maíz moderno y comparar en los sitios compartidos por ambas bibliotecas antiguas. Se utilizaron los siguientes metilomas reportados para compararlos con el metiloma antiguo de Cox17:

Accesión	Tejido	Número de individuos	Condiciones de crecimiento
B73	Embrión	3	Invernadero
	Hoja	2	Invernadero, Frío, Calor, UV
	Inflorescencia masculina	1	Invernadero
Mo17	Embrión	1	Invernadero
	Hoja	1	Invernadero
	Inflorescencia masculina	1	Invernadero

Tabla 8. Metilomas comparativos en líneas de maíz moderno.

CAPÍTULO IV. RESULTADOS

4.1 REVISIÓN SISTEMÁTICA DE BIBLIOGRAFÍA

Los resultados obtenidos de la revisión sistemática de bibliografía a nivel geológico representan una revisión correspondiente a una temporalidad geológica relativamente corta: 11 700 años antes del presente; sin embargo, a nivel biológico es notorio que los cambios demográficos de distintos organismos se han dado gracias a las fluctuaciones ambientales holocénicas (Cheng *et al.* 2009).

A nivel global, se registraron de manera muy marcada dos cambios climáticos que si bien no se sabe cuál fue su causa, la evidencia apunta a un forzamiento externo, causado por variabilidad en la actividad solar:

- Cálido medieval (800-1300 d.C.)
- Pequeña edad de hielo (1300-1850 d.C.)

En México se reportan dichas fluctuaciones climáticas (Mann *et al.* 2009). Específicamente, en la región de Mesoamérica se tienen evidencias de estos eventos (Cooke, 2005; Piperno, 2006; Lozano García *et al.* 2007 PNAS), aunque el efecto climático de cada zona se vuelve particular dadas sus características geológicas. Es por ello que el análisis de toda la franja latitudinal de Mesoamérica es importante, para conocer cómo los cambios climáticos holocénicos influyeron a lo largo del proceso de adaptación del maíz.

Los tres artículos científicos informativos resultado de la revisión sistemática que cumplieran con los criterios de selección (*proxys* estudiados, tecnología para analizar datos paleoclimáticos, análisis estadísticos, objetividad en la discusión y datos continuos en la temporalidad de interés) usaron como *proxys* principales los isótopos estables de oxígeno y dendrocronología (figura 22).



Figura 23. Mapa con la ubicación de los *proxys* analizados en los estudios paleoclimáticos selectos.

Se señala la ubicación de los *proxys* encontrados en los tres artículos científicos. La estrella dentro del círculo señala la ubicación del lugar de muestreo en Tehuacán, Puebla.

El primer sitio de estudio se encuentra en la Barranca de Amealco, Querétaro. Se desarrolló una cronología de 1 238 años que correlacionó con precipitación y migraciones antropológicas pues el sitio de estudio se sitúa a 60 Km de Tula y a 90 Km de Tenochtitlan. Se extrajeron placas de árboles de Ahuehuete (*Taxodium mucronatum*), se procedió a datarlos y posteriormente se analizó el crecimiento de los anillos, obteniendo 74 núcleos procedentes de 30 árboles (Stahle *et al.* 2011).

Una vez datados, se relacionó la cronología de los anillos con el Índice de Severidad de Sequía de Palmer (PDSI, por sus siglas en inglés). Este índice consiste en el balance de humedad del suelo basado

en datos de precipitación y temperatura que se tiene en distintos puntos de Norteamérica en los últimos 50 años. De esta manera provee una aproximación razonable para el crecimiento de los árboles dadas las condiciones paleoclimáticas (Cook *et al.* 1999).

Este estudio provee el primer estimado del paleoclima en Mesoamérica basado en anillos de árboles y, aunque los otros estudios también proveen información consistente en la misma franja temporal usando otros *proxys*, la precisión de la datación es mayor con la dendrocronología. La desventaja que este *proxy* presenta se debe a la naturaleza biológica del árbol: El crecimiento es estacional por lo que en algunas franjas de sequía no hay crecimiento, lo que podría entorpecer la datación.

Adicionalmente se compararon las condiciones climáticas encontradas en los anillos con los periodos de sequías en México y se encontró lo siguiente:

- La sequía del clásico terminal se extendió a lo largo del altiplano mexicano de 897 a 922 AD y ha sido la de mayor intensidad.
- Una sequía durante el clásico tardío de 810 a 860 AD había sido reportada en el caribe pero no en México.
- Periodo de sequía de 19 años de 1149-1167 AD, siendo ésta la primera evidencia en México de la sequía que azotó el Norte de América.
- Sequía posterior de 1378-1404 DC, que coincide con la sequía de “Uno Conejo” de los Aztecas.
- Posterior a la conquista de México se presentan 26 años de sequía: 1514-1539 DC.

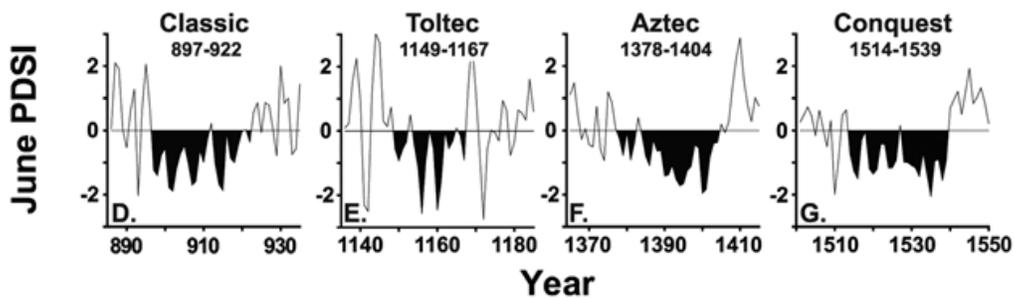


Figura 24. Gráficas que muestran el estándar de crecimiento en los anillos de Ahuehuate estandarizados con PDSI con respecto a los años de la era común.

Valores positivos indican crecimiento y por lo tanto, precipitación. Valores negativos indican sequía (Stahle, 2012).

En los dos artículos restantes (Bhattacharya *et al.* 2015; Lachniet *et al.* 2013) el *proxy* que permite reconstruir los niveles de precipitación es el isótopo estable de oxígeno (Leng, 2006; Leng & Henderson, 2013). En ambos artículos científicos, la reconstrucción paleoclimática fue realizada en el lago Aljojuca en Puebla (Bhattacharya *et al.* 2014) y en un espeleotema proveniente de la cueva de Juxtlahuaca en Guerrero (Lachniet *et al.* 2013).

En el caso del lago Aljojuca, se tomaron núcleos sedimentarios con una profundidad de 12 m que cubrían los últimos 6 200 años antes del presente. La sección de los últimos 3 800 años del núcleo sedimentario muestra evidencia robusta de una fase de sequía muy marcada en los ~800 - 1450 años antes del presente (Figura 25).

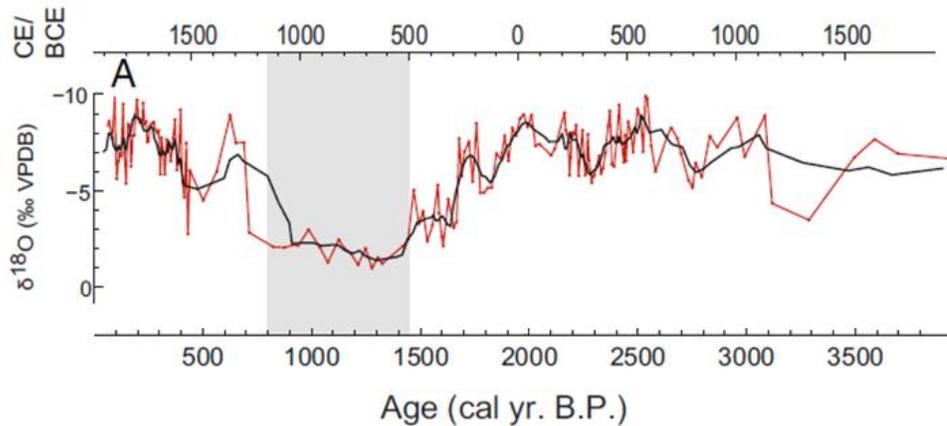


Figura 25. Gráfica de los valores de isótopos ^{18}O en Aljojuca.

El valor medio corresponde a -8‰ . Las líneas rojas representan los puntos muestreados y la línea negra una suavización. La franja gris señala la época de sequía más representativa con respecto a años antes del presente (eje Y inferior) o era común (eje Y superior). Cuando la tendencia es a 0, indica sequía en esa temporalidad (Bhattacharya *et al.* 2014).

Se encuentra otra depresión de la gráfica de ~ 400 a 650 años, consistente con sequías reportadas en otras regiones al norte de México (Roy *et al.* 2014). Sin embargo, en la región de Puebla, la evidencia muestra mayor sequía en la región delimitada por la zona gris en la gráfica anterior.

El espeleotema de la cueva de Juxtlahuaca, Guerrero, permitió estudiar los isótopos de oxígeno y aunque a simple vista no tiene una tendencia tan clara, hay un par de periodos que sobresalen de la media ($\delta^{18}\text{O} -7.6$ para el sitio) y van de 1100 - 1500 años antes del presente y otra de 1900 - 2400 años antes del presente, siendo esto consistente con datos de sequía reportados en Jalisco.

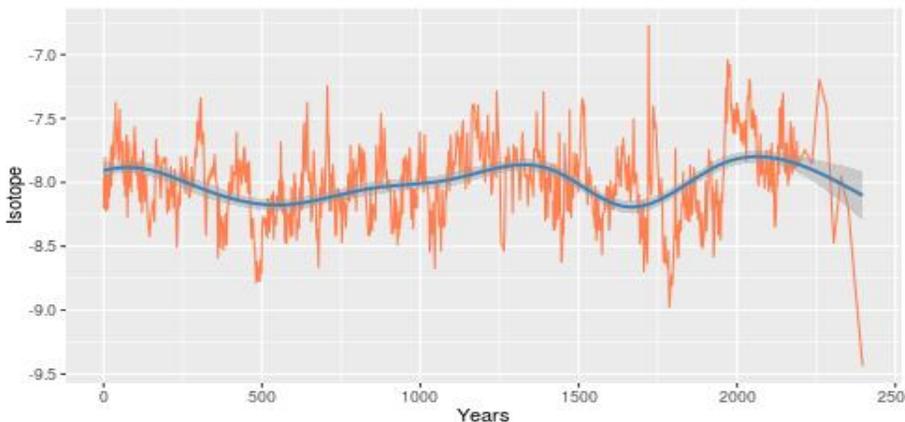


Figura 26. Gráfica con los valores de isótopos ^{18}O en Juxtlahuaca.

El valor medio para este sitio corresponde a -7.6‰ y se grafica con respecto a años antes del presente. La línea naranja representa los datos muestreados. En este caso, dada la tendencia a 0 se indica sequía y está determinada por los picos de la línea azul suavizada. (Datos obtenidos de Lachniet *et al.* 2013)

La evidencia de los espeleotemas sugiere la prevalencia de condiciones secas y cálidas entre ~ 950 a 1400 CE (Lachniet *et al.*, 2004) al igual que en la península de Yucatán de ~ 950 a 1250 CE (Correa-Metrio *et al.* 2011) y otros reportes tanto regionales como continentales (Roy *et al.* 2014; Peterson *et al.* 1991). Esta época representa el más grande periodo de sequía de los últimos dos milenios (Hodell *et al.* 2001).



Figura 27. Mapa de México marcando las zonas con reconstrucción paleoclimática.

La estrella indica el sitio de muestreo arqueobotánico en el Valle de Tehuacán. En la temporalidad de 1500 años antes del presente hay evidencia de sequía en los sitios de Aljojuca (Bhattacharya *et al.* 2014) y Juxtlahuaca (Lachniet *et al.* 2013).

En estos tres estudios (Bhattacharya *et al.* 2014; Lachniet *et al.* 2013; Stahle *et al.* 2011) hay consistencia con una temporalidad de sequía, aunque la evidencia más cercana a la zona de estudio es la de Aljojuca, Puebla, por lo que nos centramos en la franja temporal de ~800 a 1450 años antes del presente para seleccionar muestras arqueobotánicas correspondientes.

4.2 MUESTRAS DE TEHUACÁN Y SU DATACIÓN

Algunas de las muestras paleobotánicas de la excavación de 2012 en el Valle de Tehuacán fueron datadas por medio de radiocarbono con la técnica de espectrometría de acelerador de masas. Una vez obtenida la datación de las distintas muestras, quedaron en el acervo del grupo de paleogenómica, dirigido por el Dr. Rafael Montiel.

Se eligieron aquellas muestras de maíz que estuvieran dentro de la temporalidad de sequía antes identificada para Mesoamérica, para proceder con la extracción de ADN. Todas las muestras consistían en olotes de maíz que fueron calibrados siguiendo el estándar IntCal09 (Figura 28). Dicho estándar es ajustado a las muestras paleobotánicas de maíz de acuerdo a la concentración de radiocarbono y desviación estándar obtenidas en cada una de ellas.

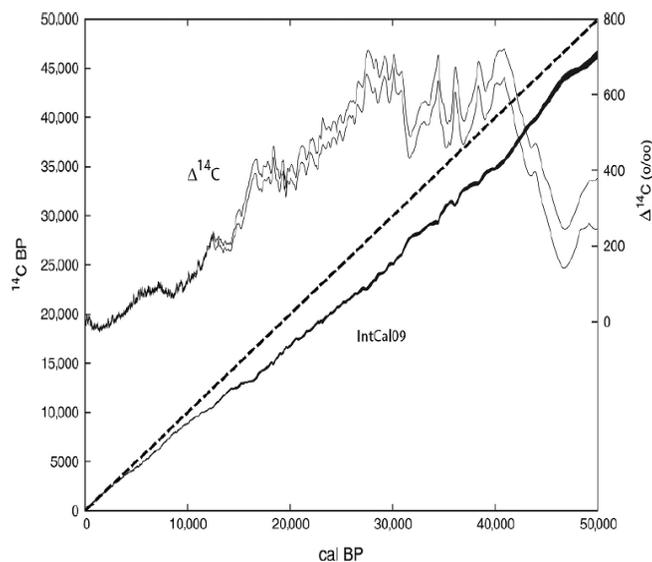


Figura 28. Curva de calibración IntCal09 con edades corregidas de $\Delta^{14}\text{C}$ (‰) con una desviación estándar.

La línea punteada muestra la correspondencia 1:1 si las edades ^{14}C y los años calendarios fuesen idénticos (Reimer *et al.* 2009). La calibración se logra al correlacionar la concentración de radiocarbono en la muestra (eje Y izquierdo) con la edad corregida con respecto al estándar (eje Y derecho) y la curva de calibración IntCal09 (eje X) obteniendo la edad calibrada antes del presente.

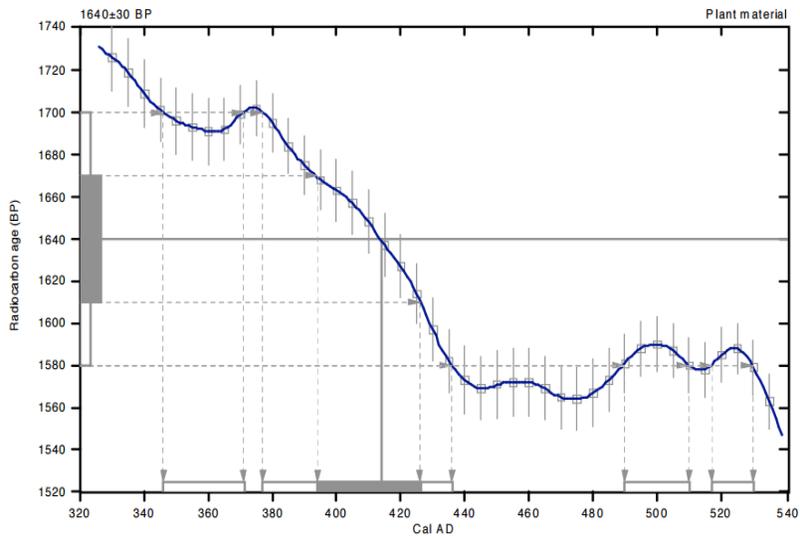


Figura 30. Curva de calibración de radiocarbono a años calendario calibrados para Cox17.

La línea azul representa las mediciones de radiocarbono en la muestra de maíz. En el eje Y se tiene la edad que representa la concentración de radiocarbono en la muestra de acuerdo a una distribución de Poisson; aquel intervalo en donde se encuentra la edad más probable (98%) está representado por una barra gris con bordes blancos, de la cual se tiene como valor medio la edad más probable, representada por la línea gris horizontal. Al intersectar la media de la concentración de radiocarbono con las mediciones de la muestra, se correlaciona con su edad calibrada en el eje de las X.

Cinco de los olotes datados estaban dentro del rango temporal de interés. Al analizar sus curvas de calibración, sólo cuatro muestras tenían un rango de distribución con un intervalo del 95% de confianza en la temporalidad buscada que está dado por la desviación estándar por lo que fueron seleccionadas para su posterior extracción de ADN (Figura 31).

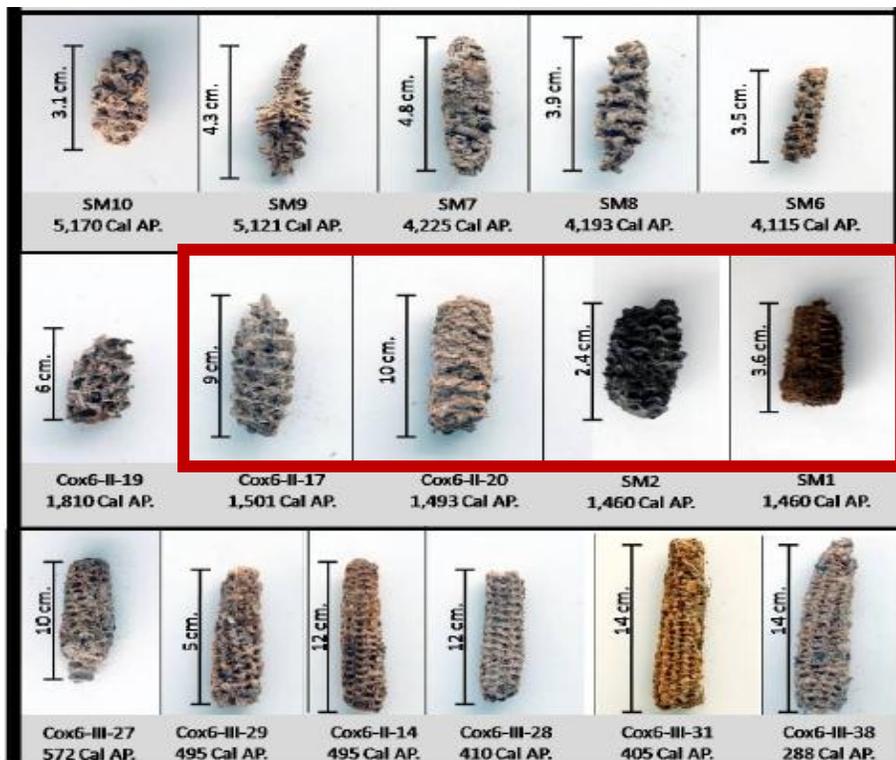


Figura 31. Muestras datadas de cada estrato temporal excavado en el Valle de Tehuacán.

La morfología de los olotes más antiguos muestra similitudes con el teocintle mientras que las modernas tienen fenotipo de maíz moderno. Las muestras dentro del recuadro rojo fueron seleccionadas para la extracción de ADN y posterior análisis de metilación.

4.3 EXTRACCIÓN DE ADN DE MUESTRAS PALEOBOTÁNICAS

Las cuatro muestras seleccionadas provenientes de las cuevas de San Marcos y Coxcatlán pertenecen a la temporalidad en la que se presume hubo sequía. Se procedió a preparar las soluciones en el cuarto de ADN antiguo siguiendo los protocolos antes indicados (Anexos 1 y 2) y posteriormente se introdujeron las muestras para comenzar con la extracción de ADN.

La muestra de San Marcos 2 (SM2) estaba carbonizada por lo que fue descartada para la extracción de ADN. El estado de preservación de las tres muestras restantes fue adecuada para su extracción de ADN: SM1, Cox6-II-17 (Cox17) y Cox6-II-20 (Cox20), obteniendo como resultado el siguiente volumen de extracto para cada una:

Muestra	Años antes del presente	Volumen extraído	Observaciones
Cox17	1501.25	200µl	
Cox20	1493	200µl	
SM1	1460	200µl	Olote hueco

Tabla 9. Caracterización de las extracciones de ADN en muestras paleobotánicas.

4.3.1 Amplificación diagnóstica para el gen Rubisco

Después de la extracción de ADN se procedió a verificar que se trata de ADN endógeno de maíz. Las muestras fueron extraídas bajo estándares estrictos para evitar la contaminación durante la excavación. Desafortunadamente, las muestras estuvieron expuestas a otros factores bióticos y abióticos en la intemperie, por lo que durante la extracción, el ADN recuperado podría no pertenecer a maíz. Para asegurarnos que se tratara de ADN endógeno de planta, se usó la amplificación diagnóstica del gen Rubisco, el cual es abundante en cloroplasto de plantas, por lo que, al observar una banda con un peso molecular de 140 pares de bases (pb) indicaría la presencia de ADN endógeno en las distintas muestras.

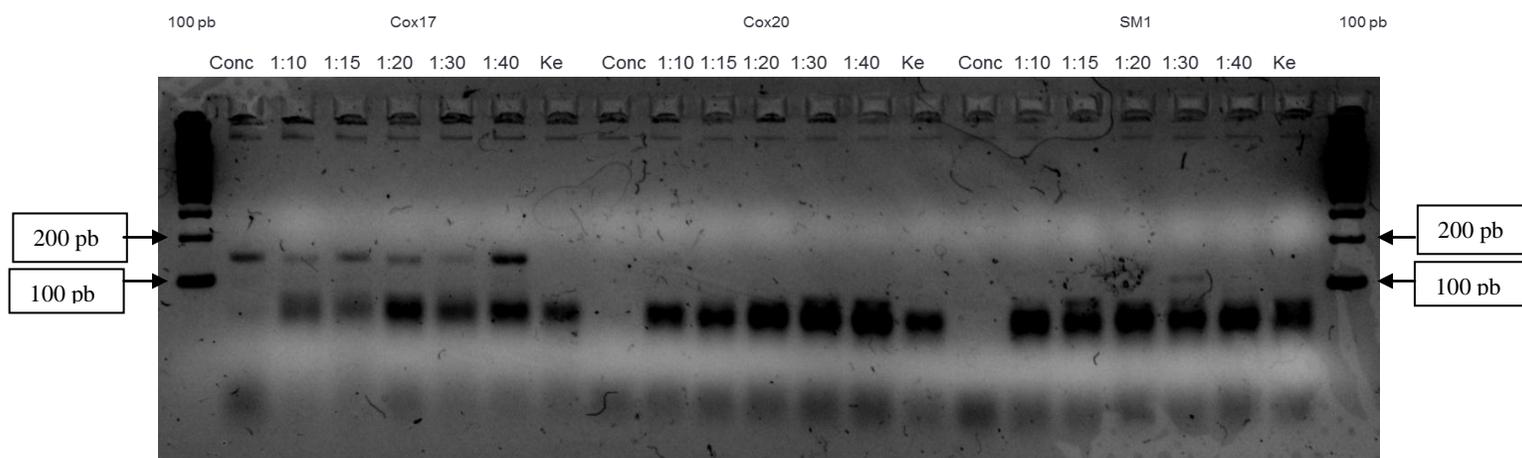


Figura 32. Gel de electroforesis de los extractos paleobotánicos con marcador de peso molecular de una escalera de 100 pb a los costados.

Los primeros carriles se cargaron con 7µl de muestra concentrada. Los siguientes se cargaron con diluciones del extracto para eliminar ruido producido por contaminantes que pudiesen inhibir la reacción de PCR (Reacción en cadena de la polimerasa por sus siglas en inglés). Ke se refiere al control negativo. Se obtuvieron bandas amplificadas de 140pb en Cox17 y SM1.

En la figura 32 se observa a los costados del gel de electroforesis los marcadores moleculares para determinar el tamaño de las bandas. La primera banda en la parte inferior indica un peso molecular de 100pb por lo que se esperarí encontrar en las muestras una banda en la zona intermedia entre esa

banda y la segunda hacia arriba, que corresponde a 200pb. La primera muestra que se cargó en el gel fue Cox17, seguida de Cox20 y finalmente, SM1. En los tres casos se cargaron 7µl de producto de PCR a partir de la alícuota extraída de ADN para cada muestra. En los pozos siguientes se cargó la misma cantidad de producto de PCR pero en esta ocasión se hicieron previas diluciones del extracto de ADN con la finalidad de diluir trazas de inhibidores que no permitieran una correcta amplificación durante la PCR.

Como resultado, la muestra de Cox17 mostró una banda de alrededor de 140pb en todos los pozos con ADN antiguo a distintas diluciones. Al diluir el extracto inicial, el efecto observado en la variación del fragmento, es causado por la dilución de inhibidores en cada pozo. La presencia de bandas de 140pb indica que Cox17 contiene ADN endógeno de planta.

En la muestra Cox20 no hubo banda visible en ninguno de los pozos mientras que en SM1 se encuentra una banda cercana a 100pb en la dilución 1:30, que se debe posiblemente a una amplificación de un fragmento indefinido y una banda de 140pb en la dilución 1:20. Basándonos en la presencia de bandas de 140pb, siendo indicativas de presencia de ADN endógeno de planta, se procedió a trabajar únicamente con las diluciones en las cuales hubo amplificación para Cox17 y SM1.

4.4 Elaboración de bibliotecas genómicas.

Una vez determinada la existencia de ADN de planta, muy posiblemente endógeno de maíz, en la muestra paleobotánica elegida, se procedió a analizar la calidad de los fragmentos de ADN de ambas muestras, enviando para ello 2 µl del extracto inicial al servicio genómico de LANGE BIO para realizar el análisis cualitativo de electroforesis utilizando un bioanalyzer (Figura 33).

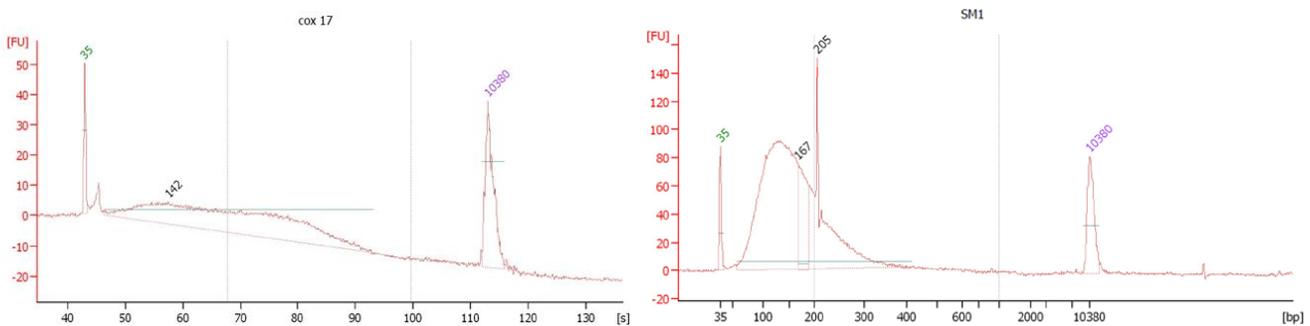


Figura 33. Análisis cualitativo de electroforesis para los extractos de Cox17 y SM1.

En el eje de las X se mide la longitud en pares de bases de los fragmentos de ADN contenidos en el extracto (pb) o el tiempo en el que se detectaron fragmentos de ADN dada la migración electroforética (s). En el eje de la Y las unidades de fluorescencia (FU), proporcionales a la cantidad de ADN en la muestra. Se pidió al servicio de genómica que se reportara en pares de bases pero para la muestra Cox17 se reportó en segundos de migración electroforética. Los picos a 35 pb (43s) y 10 380 pb (110s) corresponden a los marcadores de peso molecular contenidos en el chip del análisis de electroforesis. Las curvas entre esos dos picos representan la calidad del ADN.

En las gráficas de la figura 33, se espera ver que la mayoría de los fragmentos fueran menores a 400pb u 80s debido a la fragmentación ocurrida en el ADN antiguo. En caso de contar con una curva con fragmentos de mayor tamaño, la probabilidad de que sea ADN exógeno o contaminante aumenta. En estudios previos realizados en el laboratorio de ADN antiguo se ha trabajado exitosamente con fragmentos antiguos que van de 30 a 90pb.

La elección del tamaño del fragmento depende en gran medida de la calidad propia de cada muestra por lo que tener un rango estimado para seleccionar fragmentos de ADN antiguo brinda flexibilidad a la hora de seleccionar muestras que sean adecuadas para la elaboración de las bibliotecas genómicas.

Las unidades de fluorescencia son muy bajas en la muestra cox17 (10 FU) y la curva se vuelve más pronunciada alrededor de los 100-110 segundos por lo que es altamente probable que la cantidad de ADN endógeno de maíz sea muy pobre para esta muestra. En contraste, la calidad en la muestra SM1 es superior a la de Cox17 debido a que se tienen arriba de 100 unidades de fluorescencia en la curva que representa los fragmentos de 150pb.

Se eligió seleccionar fragmentos menores a 80pb utilizando perlas Ampure (ver anexo 4) para seleccionar los fragmentos de interés. Lo anterior es debido a que aquellas secuencias que sufrieron daño y presentan una mayor fragmentación son secuencias antiguas. Los resultados de la selección de fragmentos se muestran en la figura 33, en donde, la columna izquierda muestra las calidades de los fragmentos menores a 100pb para ambas muestras. En la columna derecha se muestran los fragmentos mayores a 150pb para ambas muestras.

A pesar de carecer de interés en fragmentos mayores a 150pb, se hizo el análisis cualitativo de esos fragmentos como controles negativos, pues en caso de tener una distribución menor a 150pb, sería posible que se debiera a errores técnicos durante la selección de fragmentos y los fragmentos de interés no se hubiesen seleccionado correctamente.

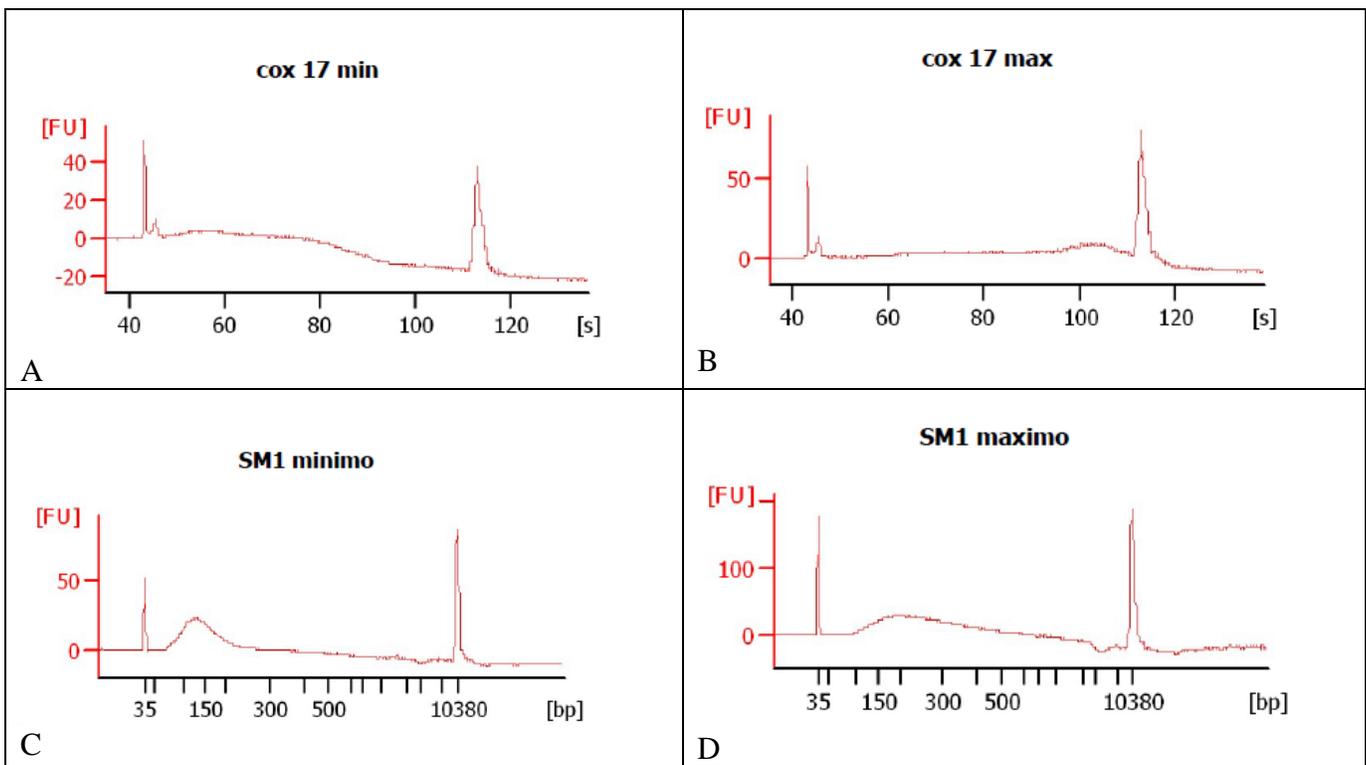


Figura 34. Resultados de los análisis cualitativos de electroforesis para seleccionar fragmentos antiguos.

A-Selección de fragmentos menores a 100 pb para la muestra Cox17. B-Selección de fragmentos mayores a 100pb para la muestra Cox17. C-Selección de fragmentos menores a 100 pb para la muestra SM1. D-Selección de fragmentos menores a 100 pb para la muestra SM1.

El resultado mostrado en la figura 33 del cuadro A corresponde a los fragmentos menores a 100pb provenientes de la muestra de Cox17 y se puede observar que las unidades de fluorescencia para esta muestra son, hasta los 60 segundos. A partir de los 80 segundos, no se encuentran unidades de fluorescencia en la muestra. En el cuadro B, se muestran los fragmentos mayores a 100pb y la curva de distribución corresponde con la curva de distribución de la figura 32, es decir, del extracto inicial con lo

que queda validado que la separación de fragmentos fue exitosa para esta muestra.

En el caso de la muestra de SM1, en el cuadro C de la figura 33, se muestra la distribución de fragmentos pequeños con una distribución de alrededor de 150pb. A comparación de la distribución de la alícuota original en la figura 32, las unidades de fluorescencia disminuyeron considerablemente, pasaron de 100FU a menos de 50FU. Para la distribución con los fragmentos de mayor tamaño en el recuadro D, las unidades de fluorescencia también disminuyeron considerablemente y el tamaño de fragmentos esperado va desde los 200pb. Al contrastar con la muestra original en la figura 32, se ve un decaimiento pronunciado en los fragmentos mayores de 500pb, por lo que en ambas alícuotas de SM1 hubo una disminución de fragmentos dado el proceso de selección con perlas Ampure. Se decidió proceder con la elaboración de las bibliotecas utilizando las muestras que fueron seleccionadas con fragmentos mínimos tanto de Cox17 como de SM1.

4.4.1 Tratamiento de bisulfito

A partir de las muestras seleccionadas con los fragmentos menores a 100pb, se tomaron alícuotas de las muestras antiguas (figura 18) y se procedió con el tratamiento de bisulfito según el protocolo descrito en el anexo 3. Posteriormente se continuó con la elaboración de las bibliotecas, siguiendo los pasos del protocolo True NanoSeq. Durante la amplificación por PCR se realizaron 10 ciclos para amplificar los fragmentos antiguos de ADN. Finalmente, para validar las bibliotecas genómicas se volvió a analizar cualitativamente por electroforesis obteniendo los siguientes resultados:

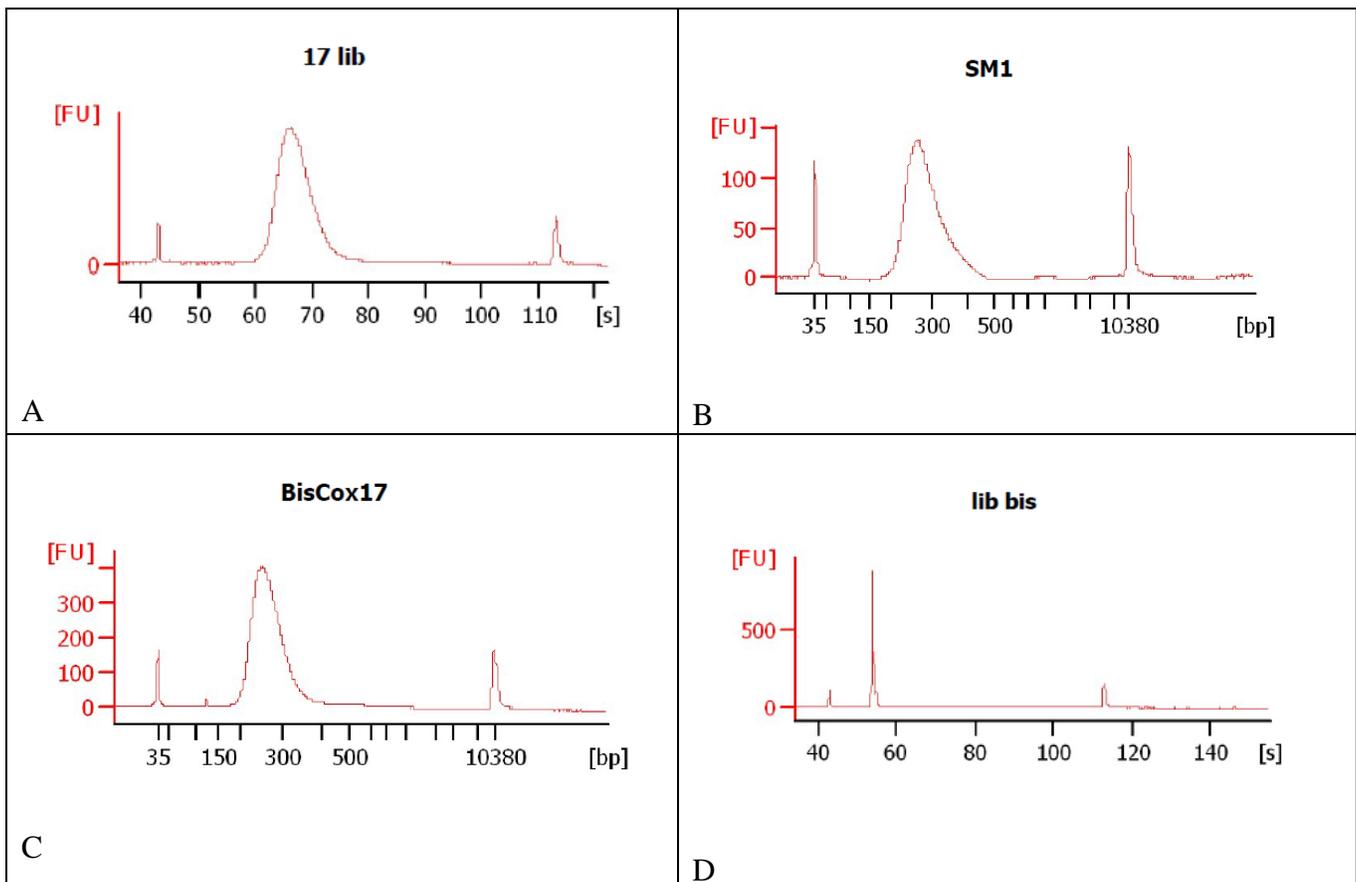


Figura 35. Resultados de la elaboración de bibliotecas mediante análisis cualitativos de electroforesis.

A. Biblioteca Cox17. B. Biblioteca SM1. C. Biblioteca Cox17 con tratamiento de bisulfito: BisCox17. D. La biblioteca SM1 no contaba con las unidades de fluorescencia necesarias para continuar con la secuenciación por lo que fue descartada después de 3 intentos (Ver Anexo 5).

En la figura 34-A se tiene la calidad de la biblioteca sin tratamiento de la muestra Cox17, 65 segundos corresponden a 250pb por lo que la distribución de la curva es la esperada para los segmentos menores de 100 pares de bases con adaptadores y barcode incluidos después de realizar la biblioteca genómica. Aunque no se muestra la escala de unidades de fluorescencia, se corroboró directamente con el personal de servicios genómicos y tiene 200FU por lo que la biblioteca tiene calidad suficiente para su posterior secuenciación.

En la figura 34-B se encuentra representada la calidad de la biblioteca sin tratamiento de la muestra SM1, en este caso la escala representa pares de bases y al igual que en el recuadro A, la distribución corresponde a 250pb como se esperaba. Se reportan 150FU para esta muestra.

En el caso de la figura 34-C, se trata de la muestra Cox17 con tratamiento de bisulfito (BisCox17) y la distribución de los *reads* es al igual que en la biblioteca sin tratamiento, de 250 pb y con 400FU es decir, la calidad es adecuada para continuar con la secuenciación. En este punto se dan por terminadas las bibliotecas de Cox17 y se procede a prepararlas para el envío al servicio de secuenciación.

Para la alícuota de SM1 que fue tratada con bisulfito, se elaboró la biblioteca BisSM1 aunque los resultados no fueron alentadores (anexo 5) debido a que no se mostraba que hubiese fragmentos en la biblioteca. Se repitió el procedimiento un par de veces y con la ayuda del personal de servicios genómicos de LANGEBIO se verificó que durante los pasos de la elaboración de la biblioteca no hubiese pérdidas de ADN. Finalmente en la figura 34-D se muestra la biblioteca que cualitativamente resultó ser la mejor en comparación con intentos anteriores. Las unidades de fluorescencia son mayores a 500 pero el tamaño de los fragmentos no corresponde al tiempo de migración, que debería de ser de 65s. De esta manera, al no poder realizar exitosamente una biblioteca con tratamiento de bisulfito para la muestra SM1, se descartó su secuenciación.

Finalmente, se realizó un análisis cuantitativo fluorimétrico de la muestra seleccionada para secuenciar (Cox17) utilizando Qubit:

- ζ Cox17- 3.18 ng/μl
- ζ BisCox17- 0.74 ng/μl (primer envío); 0.12 ng/μl (segundo envío)

Dado que la concentración en BisCox17 era muy baja la primera vez que se envió a secuenciar, se preparó una segunda biblioteca tratada con bisulfito y se envió por segunda vez, pero la concentración seguía siendo muy baja, por lo que se procedió a amplificar la biblioteca por PCR usando *primers* universales P5 y P7 para contar con una concentración suficiente para proceder a secuenciar.

4.5 Secuenciación masiva

Se enviaron las alícuotas congeladas y empaquetadas dentro de una caja con bolsas “*blue ice packs*” para que la muestra no sufriera daños térmicos durante el viaje. La secuenciación dio como resultado seis archivos de 58Gb que fueron guardados en la computadora del laboratorio 13 denominada “Vader”, como sigue:

- Undetermined_S0_L005_R2_001.fastq.gz [928Mb]
- Undetermined_S0_L005_R1_001.fastq.gz [832Mb]
- Cox17_S1_L005_R1_001.fastq.gz [16Gb]

- Cox17_S1_L005_R2_001.fastq.gz [17Gb]
- BisCox17_S2_L005_R1_001.fastq.gz [12Gb]
- BisCox17_S2_L005_R2_001.fastq.gz [13Gb]

Se copiaron al servidor del laboratorio 1 conocido como “Apomonster” y a partir de ahí se comenzaron los análisis bioinformáticos con los archivos catalogados como Cox17 y BisCox17.

4.6 Análisis Bioinformático

Siguiendo el esquema general de procesamiento informático (Figura 19), se filtró por calidad utilizando el programa Fastqc para los clusters identificados como Cox17 y BisCox17. Este primer análisis no sólo permitió conocer la calidad de los *reads* de acuerdo al phred score para cada secuencia, sino también conocer el porcentaje de GC en ambas bibliotecas.

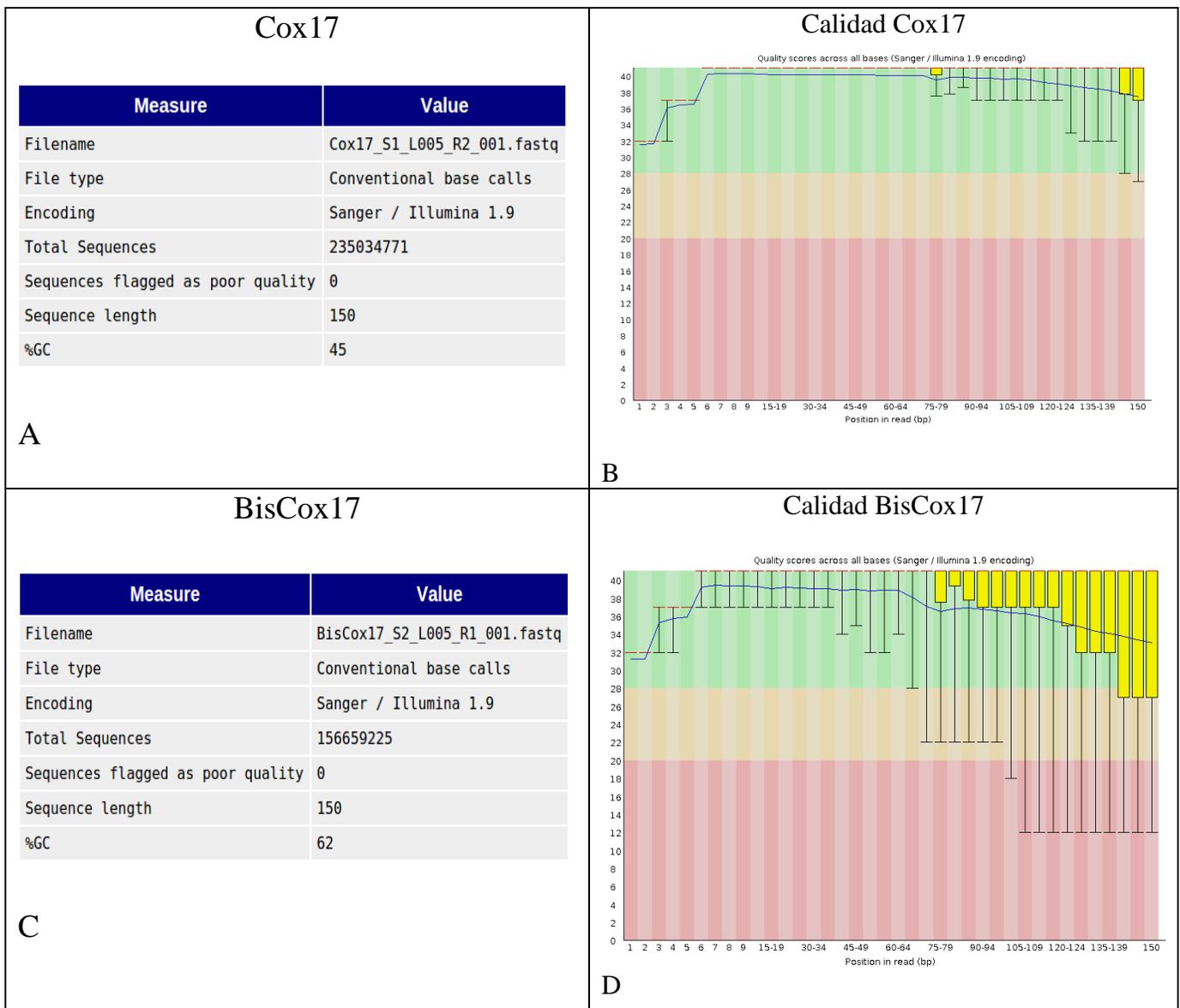


Figura 36. Esquemas con calidad inicial de las bibliotecas secuenciadas.

A- Resumen del clúster sin tratamiento de bisulfito. B- Calidad global de los *reads* sin tratamiento de bisulfito. C- Resumen del clúster con tratamiento de bisulfito. D- Calidad global de los *reads* con tratamiento de bisulfito.

A continuación se explica la anatomía de las gráficas obtenidas, tomando como ejemplo la biblioteca de Cox17:

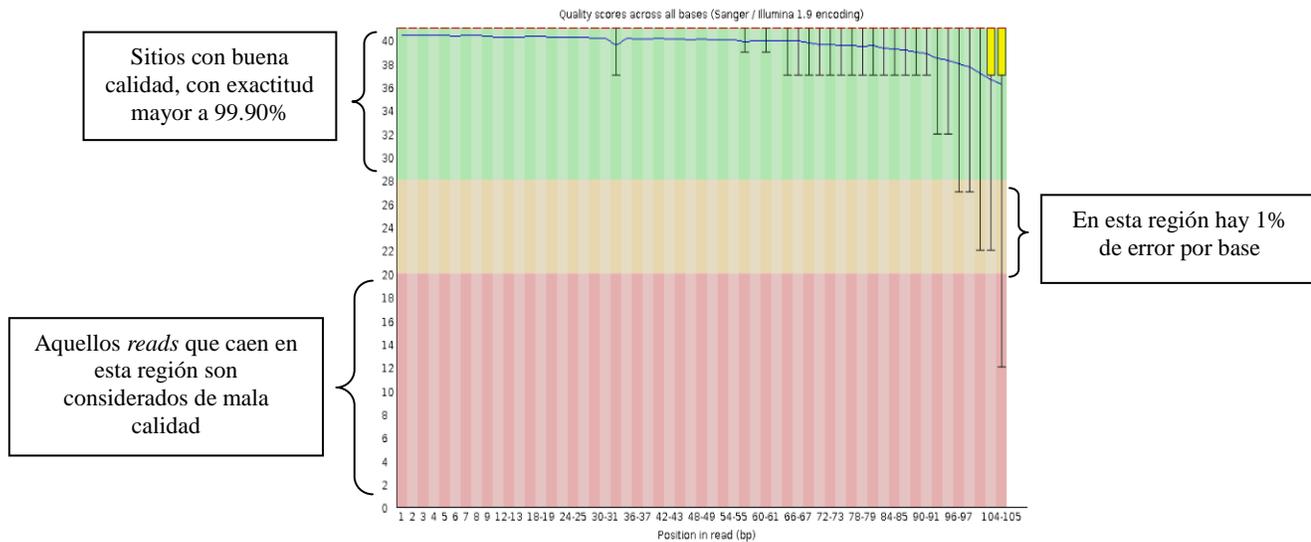


Figura 37. Esquema de calidad global de los reads de la biblioteca Cox17 en Fastqc.

En el eje y se representa el phred score mientras que en el eje x se tiene la posición de los nucleótidos en los reads. Las barras amarillas representan los cuantiles (25% y 75%), las líneas rojas son la media para cada posición, la línea azul es la calidad media.

La calidad de la secuenciación fue excepcionalmente buena para ambas bibliotecas pues en ambos clústeres la calidad media global que es representada por la línea azul se encuentra en el intervalo de phred score con exactitud mayor a 99.90%.

Teóricamente, el valor de la biblioteca con tratamiento de bisulfito debería de ser menor que la que no tiene el tratamiento debido a que la cantidad de citosinas disminuye al no estar metiladas. Sin embargo, la calidad en la biblioteca con tratamiento de bisulfito es mayor que aquella que no sufrió tratamiento.

Se procedió a filtrar las secuencias por calidad y a eliminar los adaptadores para eliminar el posible sesgo en el porcentaje en las bibliotecas.

4.6.1 Filtrado de calidad de secuencias

Al eliminar los adaptadores que se agregaron a las secuencias antiguas durante la elaboración de la biblioteca genómica, se obtienen fragmentos antiguos con una longitud de aproximadamente 70pb. Se usó el programa Cutadapt de manera que las secuencias quedaran con las siguientes características:

- Calidad mayor a 28 en phred score.
- Eliminar adaptadores universales y adaptadores con índices.
- Eliminar cualquier posible cola poli-A en las secuencias.
- Eliminar bases no identificadas (N).

Recordando que mediante la secuenciación por illumina es posible encontrar un número mayor de citosinas deaminadas en los extremos 5', un incremento de substitución de C > T al inicio de las secuencias y un incremento complementario en G > A hacia el final de las secuencias (Meyer & Kircher, 2010; Briggs *et al.* 2007) es imprescindible recortar las secuencias eliminando en los extremos regiones que presenten baja calidad. Adicionalmente se filtró el tamaño mínimo de las secuencias, siendo descartadas si tienen un tamaño de 20 pares de bases o menos.

Posteriormente se filtraron aquellos *reads* clonales, es decir, que fueron resultado de las múltiples amplificaciones de PCR. Para cada muestra fue distinto el porcentaje de *reads* clonados debido a que el número de amplificaciones en cada biblioteca fue distinta dada su concentración inicial. Recordando que hubo dos pasos de amplificación para la biblioteca con tratamiento de bisulfito (uno durante la elaboración de la biblioteca y otro después de la cuantificación fluorimétrica) era de suma importancia eliminar duplicados técnicos en este punto para finalmente hacer la cuantificación de *reads* a partir de los cuales se trabajaría en el resto del pipeline bioinformático.

Bibliotecas	Cox17	BisCox17
Total de secuencias	235 034 771	156 659 225
Secuencias sin duplicados PCR	157 299 472	41 324 236

Tabla 10. Número de *reads* obtenidos de la secuenciación y *reads* sin duplicados de PCR.

4.6.2 Análisis de calidad de bibliotecas filtradas.

Una vez obtenidas las secuencias sin adaptadores ni duplicados, se procedió a analizar la calidad de las secuencias antiguas:

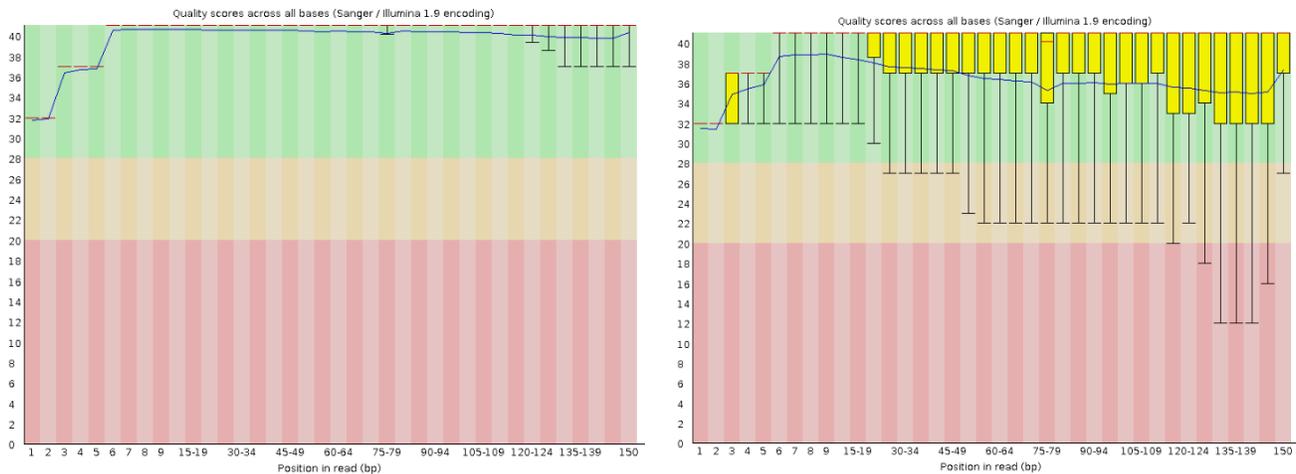


Figura 38. Esquema de la calidad global de los *reads* sin adaptadores ni duplicados de PCR.

Gráfica de calidad global en los *reads* de Cox17 (izquierda) y BisCox17 (Derecha) sin adaptadores ni duplicados de PCR. La calidad global en ambas bibliotecas tiene una calidad mayor a 99.90%. En el caso de BisCox17 los cuantiles presentan una mayor variación (barras amarillas) sin adaptadores.

De manera global, la calidad se mantuvo con una exactitud mayor a 99.90%. El clúster de *reads* con el tratamiento de bisulfito, tienen una calidad media (línea azul) con tendencia a disminuir pero de manera global se mantiene arriba de 28 en la escala de phred score.

Al verificar los porcentajes de GC en la figura 37 podemos observar que los valores no varían de sobremanera con respecto a las curvas teóricas, a pesar de la eliminación de *reads* con baja calidad y los adaptadores universales.

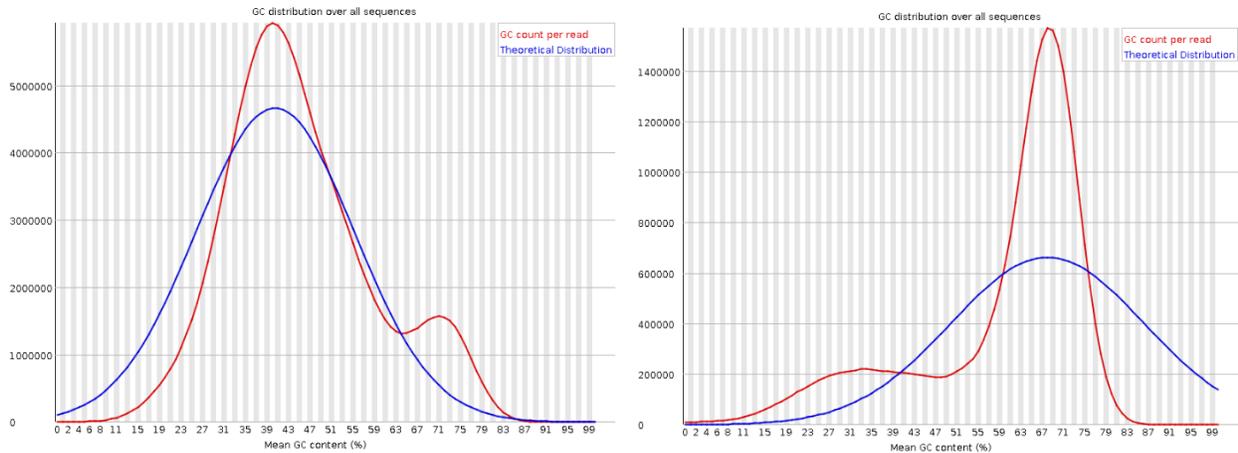


Figura 39. Número de secuencias con el porcentaje de GC en las bibliotecas.

La gráfica muestra el número de *reads* contra la media porcentual de GC. La línea azul es la distribución teórica. La línea roja es el conteo por *read* de Cox17. El valor medio para esta biblioteca sin tratamiento (izquierda) es de 44%. El valor medio para BisCox17 (derecha) es de 60%.

El porcentaje de Cox17 es de 44%, lo cual corresponde con el valor teórico esperado en maíz, el cual, dependiendo de la accesión va de 40-49%. El porcentaje de GC para la biblioteca con tratamiento de bisulfito fue de 60%. Este aumento porcentual es un primer indicio de que el tratamiento no fue exitoso pues los valores son distintos a los teóricamente esperados. Al comparar ambas figuras podemos observar que el número de *reads* con porcentaje promedio de GC de 70% es cercano a 1 500 000 en ambos casos, indicando que son éstas secuencias las que resistieron al tratamiento de bisulfito.

Se procedió a analizar el porcentaje de ácidos nucleicos en relación con su posición en los *reads*:

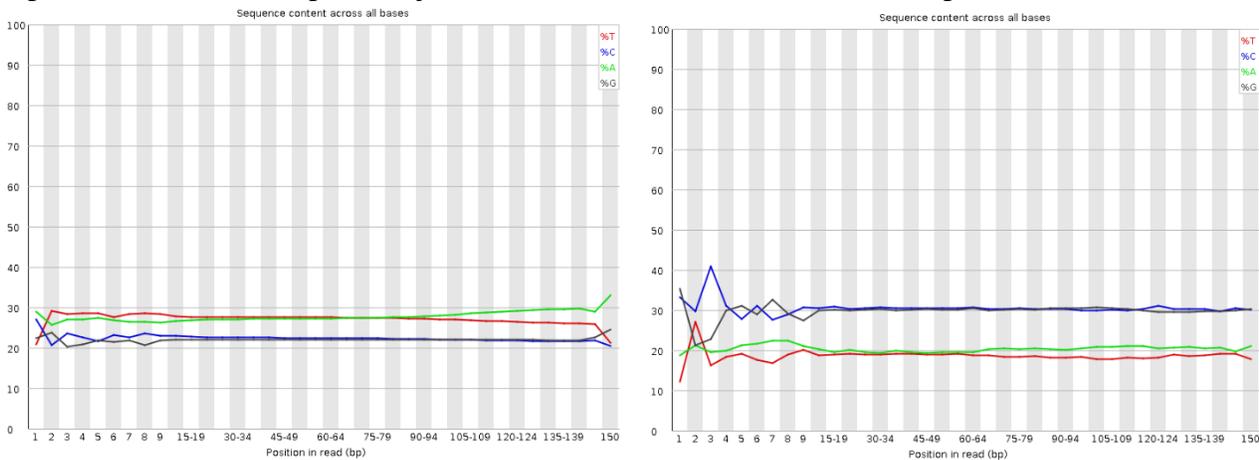


Figura 40. Gráfica con porcentaje de ácidos nucleicos en relación con la posición global en los *reads* de ambos clústeres.

Se representa el porcentaje contra la posición global en los *reads* para cada uno de los ácidos nucleicos en Cox17 (izquierda) y BisCox17 (Derecha). El porcentaje de timina (línea roja) y citosina (línea azul) presenta variación en la porción 3' del *read*.

En la figura 38 podemos observar el contenido de ácidos nucleicos en relación con la posición global de los *reads* y en general, se observa que el porcentaje de timinas y de adeninas es de alrededor del 30%. La proporción global de ácidos nucleicos en BisCox17 no corresponde con lo esperado para una biblioteca con tratamiento de bisulfito, ya que en vez de disminuir el número de citosinas, aumentó a 30% y el número de timinas y adeninas disminuyó hasta 20%.

En estudios realizados con genomas con alto contenido de GC tratados con bisulfito, se tienen resultados similares, ya que puede existir formación de estructuras secundarias, las cuales son difíciles de desnaturalizar y por ende, no hay una reacción eficiente (Ji et al. 2014). Por otro lado, al realizar varios pasos de amplificación por PCR posteriores a la metilación de las secuencias, es decir, durante la realización de la biblioteca, hay un sesgo en la amplificación de secuencias con alto contenido de GC.

Enfatizando el hecho de que en la figura 37 se encuentra una curva con alrededor de 200 000 *reads* que corresponden a valores porcentuales bajos de GC, se realizó un corte en los *reads* con tratamiento de bisulfito de manera que fuera posible discernir entre *reads* con un porcentaje alto de GC y *reads* que sufrieron el tratamiento de bisulfito pero que al ser una muestra muy pequeña, no fueran tan evidentes en la curva porcentual global.

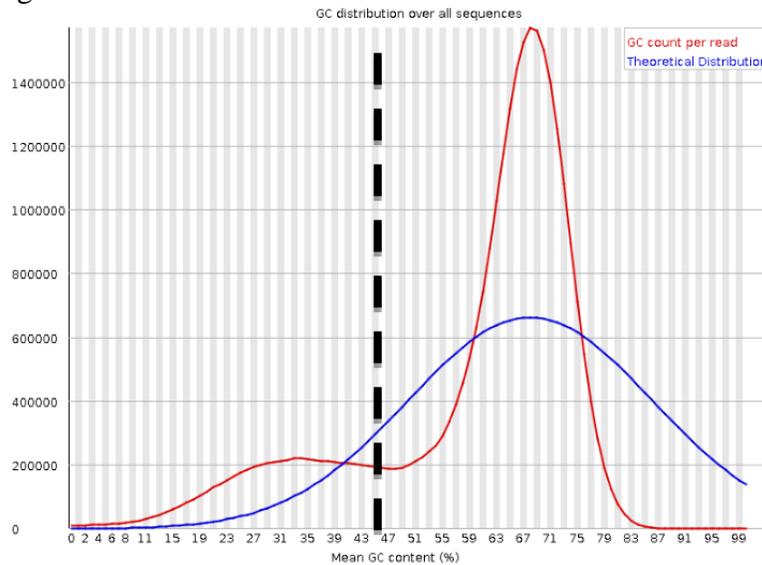


Figura 41. Selección de *reads* con porcentaje menor a 45% de la biblioteca con tratamiento de bisulfito.

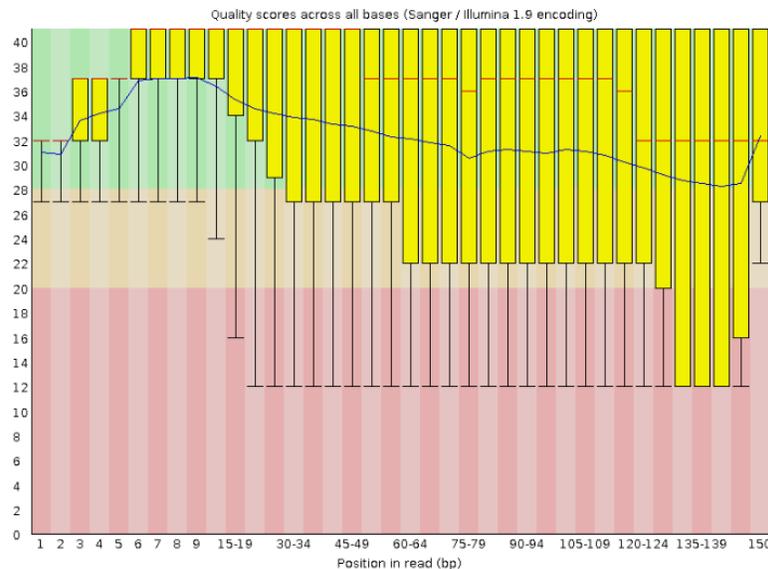


Figura 42. Esquema de la calidad global de los *reads* con porcentaje menor a 45% GC de BisCox17.

Gráfica de calidad global en los *reads* de la selección de BisCox17. La calidad global para esta biblioteca tiene su media con exactitud mayor a 99.90%, pero los cuantiles presentan una mayor variación (barras amarillas) que tiene una exactitud menor a 90% en el extremo 5' de los *reads*.

Al seleccionar *reads* con valores porcentuales bajos de GC, es decir, menores a 45%, se revisó la calidad de ese clúster seleccionado y hay una disminución en la calidad global, aunque la calidad media se mantiene arriba del 99.90% como se observa en la figura 36, por lo que se procedió a revisar el valor porcentual de GC para esta selección de fragmentos.

En la figura 43, se puede observar que el porcentaje de GC disminuyó un 10% en comparación con el porcentaje de la biblioteca sin tratamiento de bisulfito, esto es alentador debido a que es posible que el tratamiento no haya sido exitoso en toda la biblioteca pero una fracción del extracto sí sufriera el tratamiento. Si bien la distribución de campana para esta fracción de *reads* no se encuentra tan definida como la distribución teórica, el contenido porcentual global se encuentra alrededor de 35%.

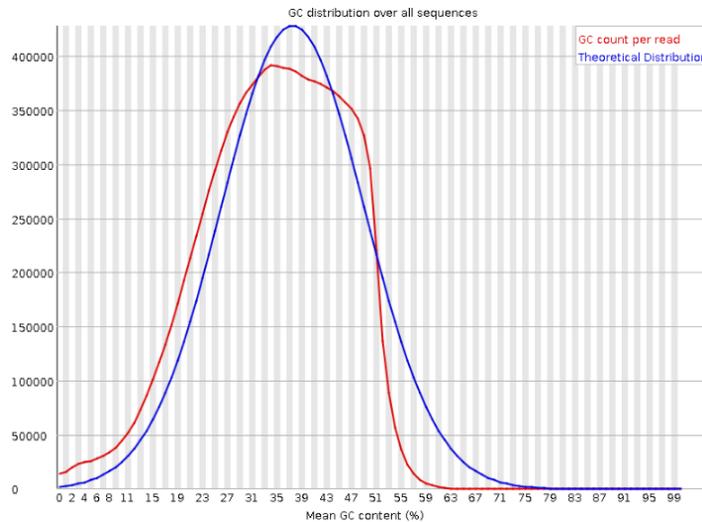


Figura 43. Número de secuencias con el porcentaje de GC de la muestra menor a 45% GC de BisCox17.

La gráfica muestra el número de *reads* contra la media porcentual de GC. La línea azul es la distribución teórica de BisCox17. El valor medio para esta biblioteca es de 35%

Finalmente, el porcentaje de timinas para esta selección de fragmentos es mayor al 30% original que estaba reportado en la biblioteca sin tratamiento mientras que el porcentaje de adeninas se mantiene en alrededor de 30%, al igual que en la biblioteca sin tratamiento.

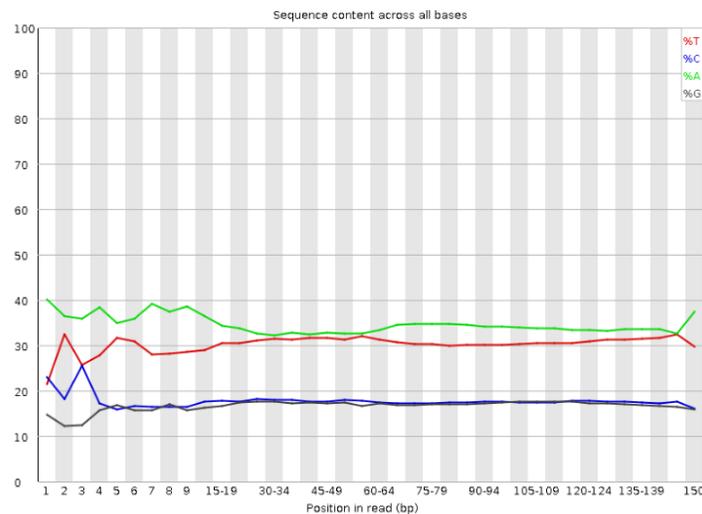


Figura 44. Gráfica con porcentaje de ácidos nucleicos en relación con la posición en los *reads* de la muestra menor a 45% GC BisCox17.

El porcentaje de timina (línea roja) y citosina (línea azul) presenta variación en la porción 3' del *read*.

4.6.3 Mapeo genómico

Se procedió a elaborar el mapeo genómico de las bibliotecas sin tratamiento y de los fragmentos seleccionados con un menor porcentaje de GC de la biblioteca con tratamiento para identificar sitios metilados. Además se realizaron diversas pruebas con la biblioteca con tratamiento para verificar parámetros de mapeo como se describe a continuación.

4.6.3.1 Establecimiento de las condiciones óptimas de mapeo.

En trabajos anteriores del laboratorio de paleogenómica, dentro del ámbito de genomas de maíz antiguo realizados por Miguel Vallebuena, se optimizaron las condiciones de mapeo para aADN probando diversas pruebas de mapeo para eficientar el alineamiento de secuencias con respecto a un genoma de referencia. Si bien se determinó que para un eficiente mapeo de ADN endógeno debían de flexibilizarse los parámetros de mapeo usando un algoritmo especializado en *reads* cortos de aproximadamente 80pb, las simulaciones se hicieron con bibliotecas de maíz antiguo contando con una calidad mínima de 10 phred score.

En este proyecto se decidió retomar la estrategia exitosa para *reads* con un tamaño menor a 100pb, posteriormente se llevó a cabo una prueba de mapeo con el programa BWA usando el algoritmo mem con dos algoritmos distintos:

- *Single-end*
- *Paired-end*

Las secuencias en estas bibliotecas fueron secuenciadas en modo *paired-end*, en el que hay dos secuencias semilla (fragmentos de alrededor de 24pb) en las orillas de los *reads* que tienen que empatar en un fragmento equidistante de ~70pb del genoma de referencia. La ventaja que proporciona el mapeo por *paired-end* es que aumenta la confianza de mapeo.

Al mapear *single-end* sólo una secuencia semilla tendría que mapear en el genoma de referencia para poder identificar la secuencia antigua. En principio, al mapear *single-end* aumentaría el número de secuencias identificadas al no tener la restricción de que sean los dos fragmentos semilla equidistante los que deban tener identidad con los fragmentos de la referencia.

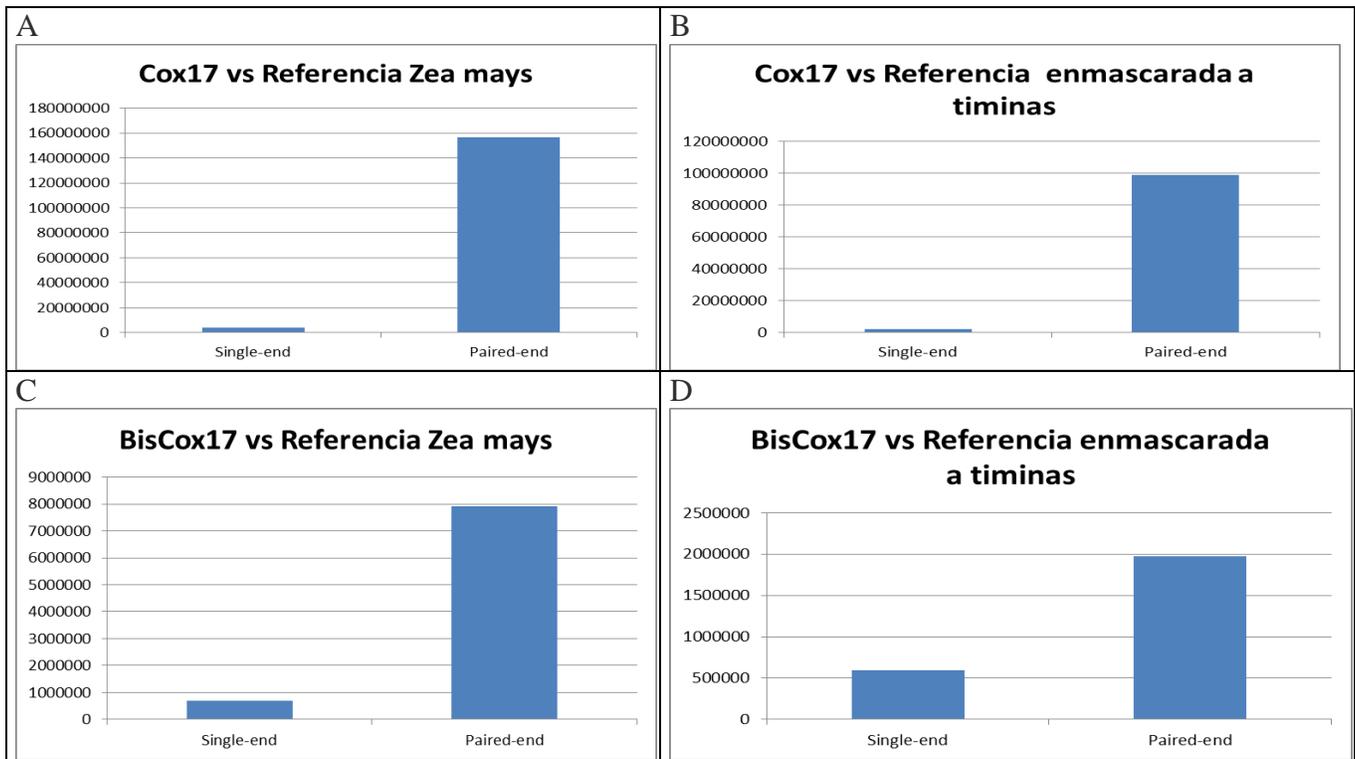


Figura 45. Gráficas con los valores de mapeo obtenidos con las dos modalidades de BWA mem

Figura 45 A- *Reads* mapeados de la biblioteca sin tratamiento con respecto al genoma de referencia, se obtuvieron 156 millones para *paired-end* y 3 millones para *single-end*. Figura 45 B- *Reads* mapeados de la biblioteca sin tratamiento con respecto al genoma enmascarado a timinas, se obtuvieron 98 millones para *paired-end* y alrededor de dos millones para *single-end*. Figura 45 C- *Reads* mapeados de la biblioteca con tratamiento con respecto al genoma de referencia, se obtuvieron alrededor de 8 millones para *paired-end* y alrededor de 700 mil para *single-end*. Figura 45 D- *Reads* mapeados de la biblioteca con tratamiento con respecto al genoma de referencia enmascarado a timinas, se obtuvieron alrededor de dos millones para *paired-end* y alrededor de 600 mil para *single-end*

A partir de la comparativa de mapeos *paired-end* vs *single-end* se eligió el algoritmo *paired-end* al obtener un mayor porcentaje de secuencias alineadas con el genoma de maíz.

En la biblioteca con tratamiento de bisulfito (BisCox17) surgía otra problemática: los polimorfismos causados por los cambios nucleotídicos de C > T con el tratamiento de bisulfito resultarían en *reads* erróneamente mapeados, por lo que determinar los parámetros para el mapeo en la biblioteca con tratamiento con respecto a la referencia fue de suma importancia para una correcta representación.



Figura 46. Esquema de la problemática del daño en muestras antiguas tratadas con bisulfito.

Dentro del recuadro rojo se muestran los polimorfismos nucleotídicos (marcados con X) encontrados en muestras antiguas. El recuadro naranja muestra que hay polimorfismos adicionales con el tratamiento de bisulfito.

Antes de obtener los resultados de la secuenciación, esta problemática se trató de resolver con una biblioteca ya secuenciada y mapeada de maíz antiguo: SM10. A esta biblioteca se le realizó tratamiento “in silico” de bisulfito. Conociendo el porcentaje aproximado de metilación que existe en las accesiones modernas de maíz se decidió que para todos los *reads* del cromosoma 1 de la biblioteca SM10 se seleccionaran las citosinas y de manera azarosa se modificara dicha base a timina o se mantuviera como citosina.

Se realizaron ocho eventos independientes en la biblioteca SM10. A priori se sabía que todos los *reads* correspondían a maíz pero después del tratamiento “in silico”, se perdía la eficiencia del mapeo con respecto a la secuencias de referencia. Por lo que una vez recibida las secuencias de Cox17 era necesario modificar parámetros de mapeo. El programa usado tiene el siguiente algoritmo para descartar o permitir que un *read* sea eliminado o se le penalice de acuerdo a su nivel de igualdad con la referencia:

$$\{.75 * \exp[-\log(4) * B/A]\} \text{ (Li. 2010)}$$

Donde A es el valor de igualdad (*match score*) de la secuencia mientras B es la penalización por *mismatch* (*mismatch penalty*). Cuando un *read* no cumple con el valor mínimo establecido, es descartado. Es por ello que se procedió a variar los parámetros al usar la biblioteca con tratamiento para determinar los valores óptimos para su secuenciación respecto a la penalización de mapeo.

Se tomaron como controles negativos una base de datos de hongos y bacterias para comparar cuántos *reads* mapeaban tanto a maíz como a bacterias y como control positivo otro cereal que se encuentra filogenéticamente cercano al maíz: el sorgo.

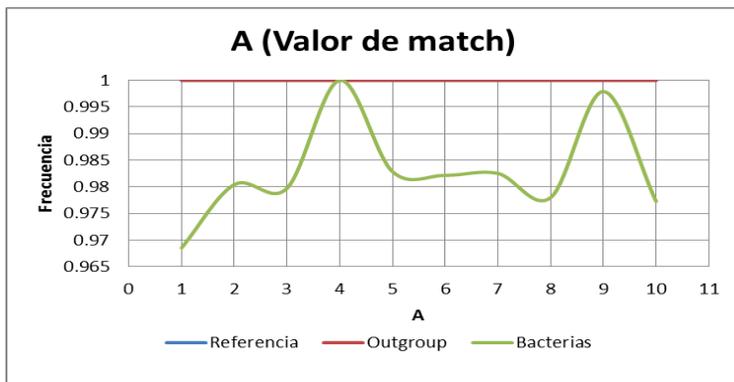
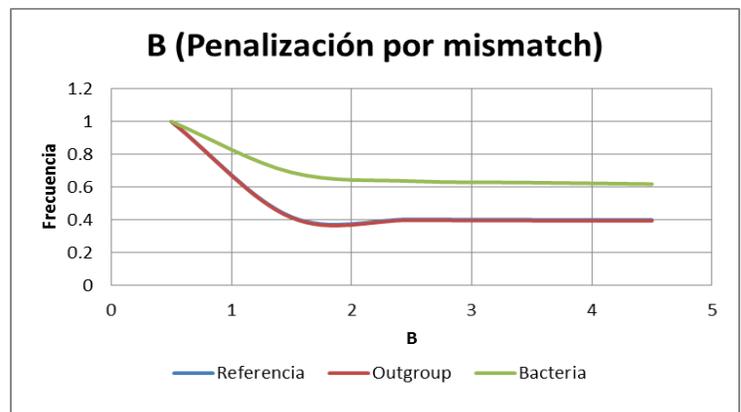


Figura 47. Variación del parámetro A que equivale a la igualdad o *match* de un *read* con respecto a la referencia.

Se varían los valores de igualdad al mapear los *reads* de BisCox17 a sorgo (*outgroup*), B73 (referencia) y bacterias con respecto a una frecuencia de mapeo. Los valores de referencia y el *outgroup* tienen la misma frecuencia.

Figura 48. Variación en el parámetro B que equivale a la penalización por *mismatch* de un *read* con respecto a la referencia.

Al mapear BisCox17 a sorgo (*outgroup*), B73 (referencia) y bacterias con respecto a frecuencia de mapeo. La frecuencia del *outgroup* y la referencia son los mismos.



Dentro de los parámetros de mapeo determinado como óptimos en el experimento anterior se encuentran los valores por default del programa: 1 para A y 4 para B. Cuando A es 1, arroja la menor frecuencia de *reads* mapeados de bacteria, mientras que cuando B es 4, la frecuencia es menor en bacterias pero es consistente con sorgo. Se mapeó con parámetros por defecto para ambas bibliotecas.

4.6.4 Mapeo de *reads* contra genoma de referencia.

Las secuencias en las bibliotecas fueron secuenciadas en modo *paired-end*, utilizando los parámetros por default para caracterizar ambos clústeres como se muestra en la siguiente tabla:

Bibliotecas	Cox17		BisCox17	
Total de secuencias	235 034 771	100%	156 659 225	100%
Secuencias con calidad	230 639 203	98.12%	145 003 447	92.55%
Secuencias sin duplicados	157 299 472	66.92%	41 324 236	26.37%
Secuencias mapeadas a B73	156 458 666	66.89%	7 918 307	5.12%
Secuencias mapeadas a regiones repetitivas	145 260 719	61.80%	6 729 224	4.29%
Secuencias mapeadas a regiones únicas	11 197 947	4.76%	1 189 083	0.83%
Calidad promedio de mapeo	19.81		20.91	
Porcentaje GC	38.29%		58.76%	
Longitud promedio de las secuencias (pb)	75		70	
Cobertura total (Mb)	9.7		1.9	
Profundidad promedio	1.6x		0.09x	

Tabla 11. Caracterización de las bibliotecas genómicas.

4.6.5 Estimación del daño molecular.

Posteriormente se estimó el daño molecular para la biblioteca Cox17 para tener idea del daño molecular en ambas bibliotecas, ya que provienen del mismo extracto de ADN. Se usó el programa MapDamage 2.0 obteniendo los siguientes estimadores:

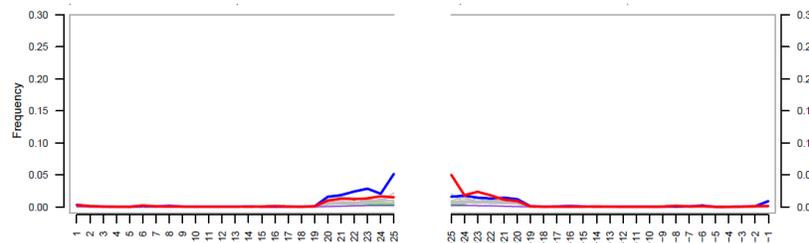


Figura 49. Gráfica con estimador de daño molecular.

Muestra los fragmentos en sus extremos 5' y 3' así como la frecuencia a la que ocurre el daño molecular siguiendo el modelado HYK para sustituciones nucleotídicas en ADN antiguo (Jónnson *et al.* 2014).

En la figura 49 se representa el nivel de daño molecular en los extremos de los fragmentos secuenciados, tal y como se espera en fragmentos de ADN antiguo, con la frecuencia de daño molecular en el eje Y, mientras que los valores positivos y negativos en el eje x representan los extremos 5' y 3' de los fragmentos antiguos. Se esperaría tener un aumento en la frecuencia en los extremos, aseverando que se trata de ADN antiguo; sin embargo, debido al algoritmo *paired-end* se obtiene la frecuencia del daño molecular en la parte central. Gracias a esta representación gráfica, se conoce el porcentaje global de daño en las secuencias (0.05%) en la biblioteca de Cox17, lo que es extrapolable a la biblioteca BisCox17 por provenir de la misma alícuota.

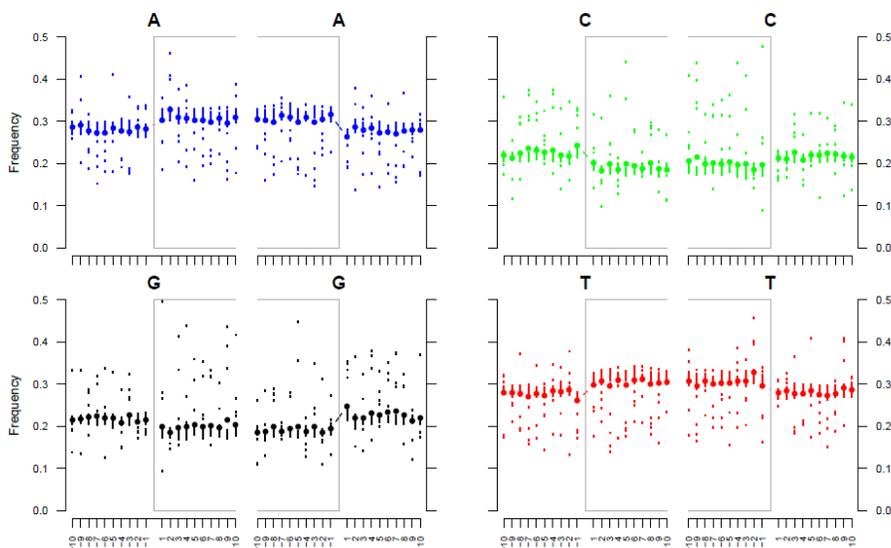


Figura 50. Gráficas con las frecuencias de bases nucleotídicas para regiones antiguas de Cox17 en yuxtaposición con regiones modernas. En el eje y se tiene la frecuencia de nucleótidos en las secuencias de Cox17. El eje de las x representa la posición de los fragmentos en 5' y 3'. Los puntos que se encuentran dentro del corchete corresponden a las secuencias antiguas (Jónnson *et al.* 2014).

Los valores iniciales de los porcentajes de GC para la biblioteca Cox17 son consistentes con lo que vemos en la gráfica de la figura 49 en donde la frecuencia se mantiene en valores muy similares a los porcentajes reportados en las secuencias aledañas, de esta manera aseveramos que se trata de ADN antiguo de maíz.

4.6.6 Búsqueda de sitios metilados.

Posteriormente, para la búsqueda de sitios metilados, a partir de los fragmentos filtrados con porcentajes menores a 45% de GC (figura 41) es en donde se obtiene una representación de los *reads* en los que fue efectivo el tratamiento de bisulfito. Se procedió a identificar aquellas bases que fueran diferenciales con respecto a las referencias enmascaradas (citosinas por timinas y guaninas por adeninas) para identificar sitios únicos en donde el tratamiento de bisulfito fuese exitoso.

Recordando que los porcentajes de GC son 44% para el set de *reads* no tratados y 35% para el set de *reads* que sí sufrió tratamiento (figura 41), se obtiene una caracterización de los sitios en los que fue efectivo el tratamiento de bisulfito tal y como se muestra en la siguiente tabla:

Biblioteca extraída de BisCox17	Datos de <i>reads</i> con porcentaje menor a 45% GC	
Reads extraídos con porcentaje menor a 45% GC	10,518,389	100%
Secuencias mapeadas a B73	1,758,469	16.72%
Secuencias mapeadas a regiones repetitivas	997,774	9.48%
Secuencias mapeadas a regiones únicas	760 695	7.23%
Calidad promedio de mapeo	19.81	
Porcentaje GC	14.09%	
Longitud promedio de las secuencias (pb)	70	
Cobertura total (Mb)	0.42	
Profundidad promedio	0.02x	

Tabla 12. Caracterización de los fragmentos tratados con bisulfito.

Como se puede apreciar en la tabla 12, los *reads* que fueron tratados con el bisulfito fueron alrededor de 760 mil para todo el clúster de *reads*. Se procedió a identificar en las regiones compartidas de las dos bibliotecas antiguas mapeadas tanto en el genoma enmascarado con timinas como con adeninas, para identificar los sitios que tuvieran cobertura y profundidad mínima de 3x. Se utilizó el visor IGV con los metilomas modernos (Tabla 7) como comparativa de metilación en los sitios encontrados para determinar su anotación en el genoma de B73 versión 3 (Figura 44).

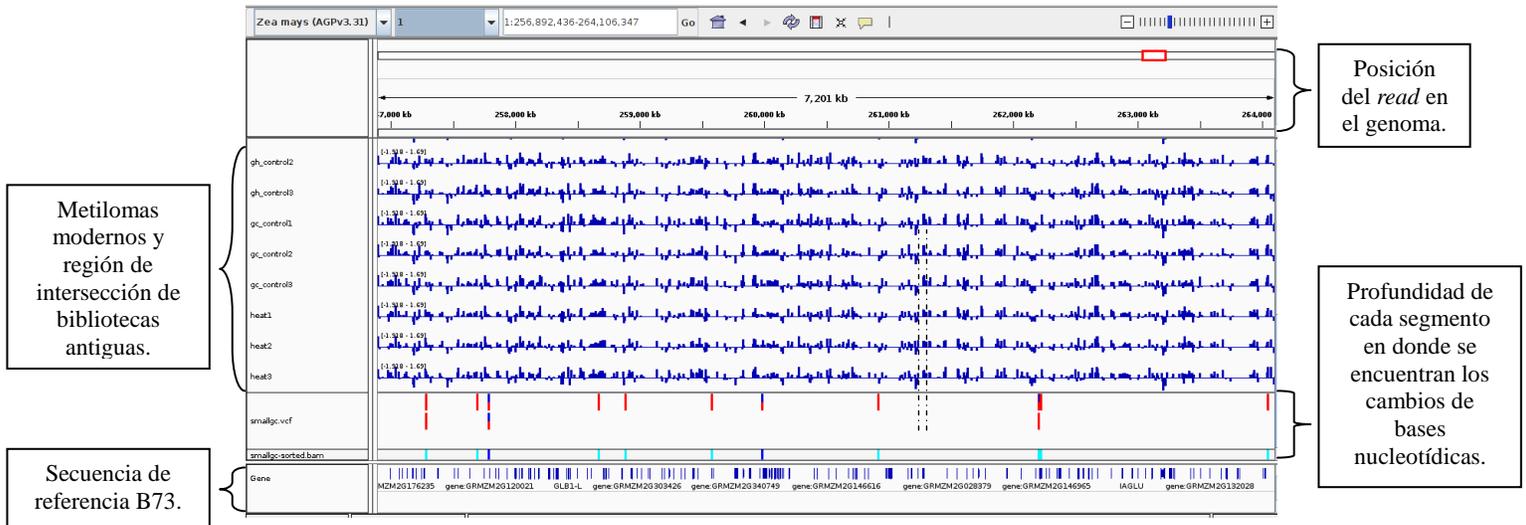


Figura 51. Esquema del visor IGV con un segmento del cromosoma 1 con cambios en bases nucleotídicas en *reads* con tratamiento de bisulfito. Regiones compartidas de las bibliotecas con y sin tratamiento con una profundidad que permite identificar cambios en bases nucleotídicas. La sección de metilomas muestra señal de hiper o hipometilación en los histogramas azules. En la sección de profundidad, para cada segmento se tiene una profundidad que va desde 18x hasta 1x; se toman en consideración aquellas que sean mayores a 3x. En la referencia (sección de secuencia de referencia), indica si hay genes anotados (barras azules) para esas posiciones.

Los *reads* únicos que fueron tratados con bisulfito tienen un porcentaje de 14.09% (tabla 10), al contabilizar el número de bases nucleicas obtenidas, únicamente se cuenta con un 0.5% de citosinas metiladas y en contraste, el 55.35% es contenido de timinas. Por lo que se confirma que esta selección de *reads* sí sufrió del tratamiento.

Finalmente, se presentan ejemplos de sitios que contaba con la profundidad mínima de 3x en donde se encontraron perfiles puntuales de metilación además de calidad de mapeo lo suficientemente robusta como para aseverar que las citosinas que se presentan a continuación son citosinas metiladas provenientes del paleoepigenoma de Cox17.

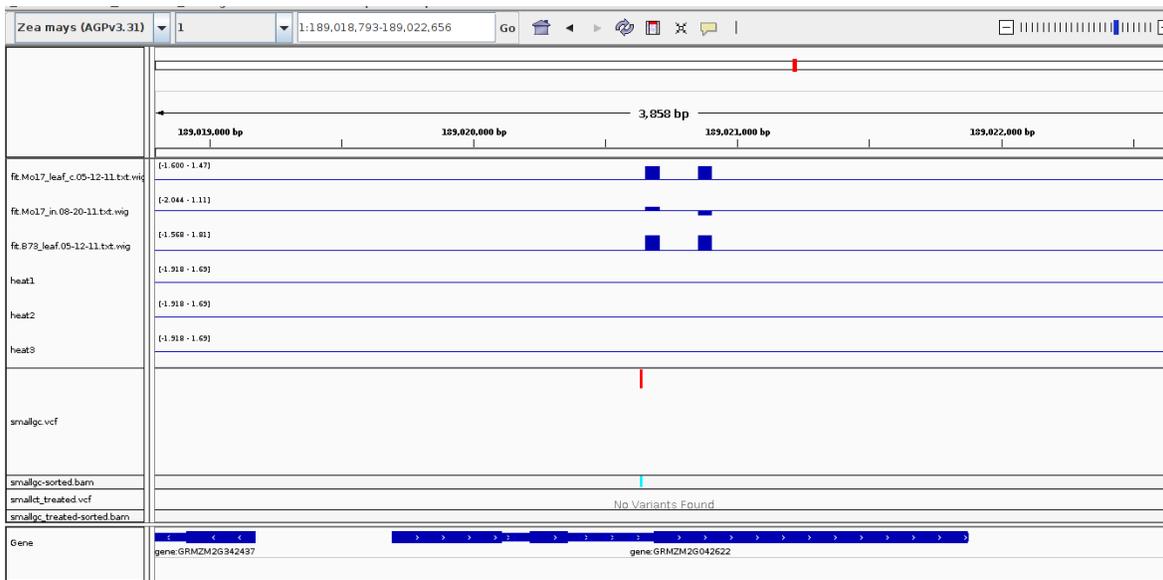


Figura 52. Visor IGV con un sitio del cromosoma 1 con tratamiento de bisulfito.

La barra roja muestra cambios nucleotídicos de interés: citosina con una profundidad de 6x cuando se tiene timina en la referencia. La barra azul clara representa los datos condensados de esos sitios en el mapeo, calidad y probabilidad de encontrar uno o varios polimorfismos. Las barras horizontales azules de los metilomas bajo condiciones de calor (heat 1-3) no muestran metilación en ese sitio. La citosina metilada se encuentra en el gen GRMZM2G042622.

Sitios con citosinas metiladas que se encontraran en contexto génico y que al compararlo con metilomas actuales mostraran metilación en la misma posición no permiten una adecuada interpretación biológica. En la Figura 52 se muestra un ejemplo de manera descriptiva de un sitio metilado con una profundidad de 6x y calidad de mapeo de 20 en escala phred score. Se encuentra en un intrón del gen GRMZM2G042622 el cual no se encuentra anotado. En los metilomas de las plantas de maíz que fueron sometidas a estrés, no se muestra metilación en ese sitio, solamente en tejido de hoja en donde hay hipermetilación con diversos gradientes tanto en la accesión B73 como Mo17.

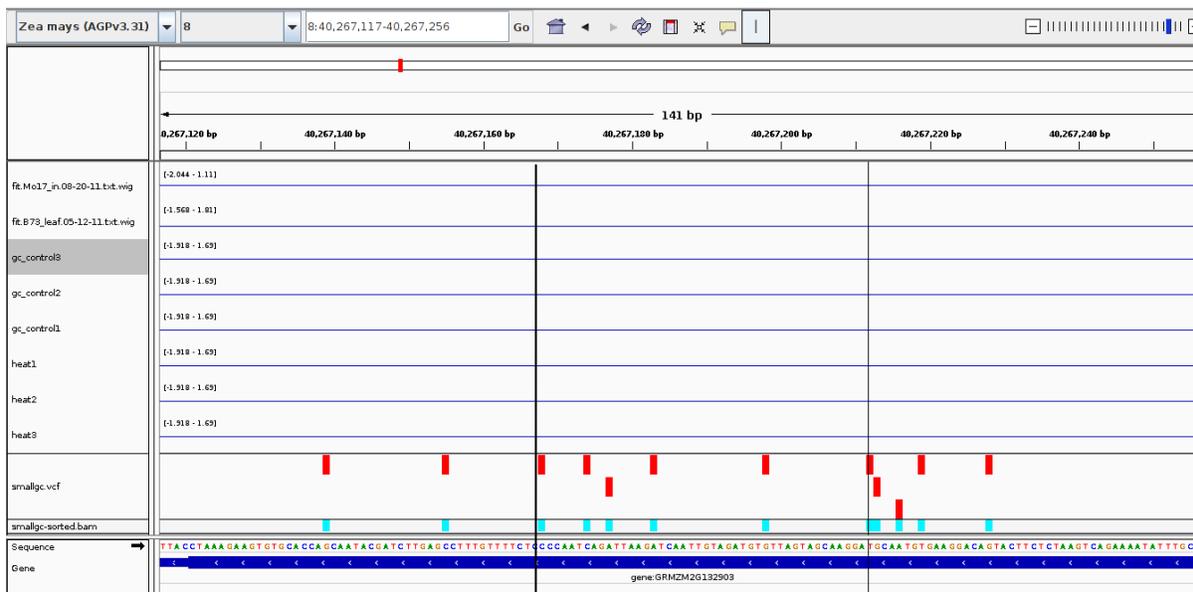


Figura 53. Visor IGV con un sitio del cromosoma 8 con tratamiento de bisulfito.

Las barras rojas muestran cambios polimórficos. Los sitios con línea negra a la derecha muestra un cambio nucleotídico de interés: citosina en el genoma antiguo y timina en la referencia con una profundidad de 8x. La línea negra a la izquierda muestra un cambio nucleotídico en donde la referencia es una citosina y el genoma antiguo tiene una timina con profundidad 5x. Las barras horizontales azules representan los datos condensados de esos sitios en el mapeo, calidad y probabilidad de encontrar uno o varios polimorfismos. Las barras horizontales azules de los metilomas bajo condiciones de calor (heat 1-3) y los controles no muestran metilación en algunas réplicas para el exón 4 del gen GRMZM2G132903.

En la Figura 54 se muestra un ejemplo de manera descriptiva de un sitio no metilado (línea negra izquierda) y un sitio metilado (línea negra derecha) con una profundidad de 5x y 8x respectivamente. Las calidades de mapeo son de 55 y 60 respectivamente en escala phred score. Se encuentra en un exón del gen GRMZM2G132903 el cual no se encuentra anotado. En los metilomas actuales no se muestra metilación en ese sitio pero este ejemplo es representativo debido a su alta calidad y profundidad con lo que hay una mayor certeza probabilística que se trata de un sitio metilado en el genoma antiguo de Cox17. Los demás polimorfismos indicados en rojo no son informativos con respecto a perfiles de metilación en este sitio.



Figura 54. Visor IGV con un segmento del cromosoma 8 con tratamiento de bisulfito.

Las barras rojas/azules muestran cambios polimórficos. Los sitios con línea negra a la derecha muestra un cambio nucleotídico de interés: citosina en el genoma antiguo y timina en la referencia con una profundidad de 65x mientras que la línea a la izquierda señala un cambio de citosina en genoma antiguo y timina en la referencia con una profundidad de 77x. Ambos sitios tienen calidad de mapeo de 29 y 32 en escala phred. Las barras azules representan los datos condensados de esos sitios en el mapeo, calidad y probabilidad de encontrar uno o varios polimorfismos. Las barras horizontales azules de los metilomas bajo condiciones de calor (heat 1-3) y los controles no muestran metilación en para el sitio de interés, sin embargo, hay sitios metilados en todas las muestras modernas a una distancia de 20pb. En las cercanías (100Kb) se encuentra el gen PARP2.

La figura 52 muestra un ejemplo de un par de sitios con mucha profundidad y calidad en ambos pero además demuestra que para esa región se permite discriminar sitios de heterocigosidad. En el caso de los sitios marcados con la línea negra, ambos conservaron la metilación en el genoma antiguo. Los metilomas hipermetilados modernos reportados se encuentran a una distancia de 20pb. En cuanto al contexto genómico, se encuentra el gen PARP2 y su transcrito a aproximadamente 100Kb río arriba.

La polimerasa 2 poli (ADP-ribosa) es una proteína post-transcripcional involucrada en la arquitectura de la cromatina. Responde a estrés abiótico permitiendo defensa o muerte celular en plantas. Se han realizado diversos estudios que asocian la transcripción y metilación de este gen con el estrés osmótico y de salinidad (Ming-pu, 2008; Vanderauwera *et al.* 2007; Doucet-Chabeaud *et al.* 2001). Si bien la cercanía al gen de dos sitios metilados antiguos plasma una posible perspectiva de estudio para correlacionar paleoepigenética con adaptación ambiental, en este proyecto no se cuenta con soporte estadístico suficiente para correlacionarlos.

CAPÍTULO V. DISCUSIÓN

5.1 Paleoclima en el Valle de Tehuacán

Se ha demostrado que la energía extraterrestre recibida por la tierra ejerce un fuerte control en el clima a escalas de milenios o submilenios (Lozano-García *et al.* 2007). En escalas menores, los patrones regionales de variabilidad climática que afectan la evolución de los ecosistemas son mayoritariamente modulados por las dinámicas internas del planeta. Un ejemplo de esto durante el Holoceno se conoce como el fenómeno de “El Niño” (Conroy *et al.* 2008; Moy *et al.* 2002; Toth *et al.* 2015).

Durante el evento de El Niño-Oscilación Sur (ENOS) las temperaturas superficiales del mar en las costas del Ecuador y al norte de Perú causan cambios convectivos en las regiones costeras del noroeste Sudamericano. Gracias a las evidencias geológicas se ha estimado la frecuencia a la que suceden los eventos ENOS, siendo éstos muy débiles o inexistentes hasta los 7000 años antes del presente, posteriormente llegarían a su máximo entre 2000-1000 años antes del presente. En fechas modernas la frecuencia del evento ha decaído (Moy *et al.* 2002).

El calentamiento climático global durante el periodo medieval cálido pudo ser causado por condiciones climáticas globales similares a las de El Niño, aunadas a la radiación solar que propiciaron el aumento de temperatura superficial en los océanos cercanos al Ecuador, obteniendo como resultado un sistema dinámico que mantuvo bajo condiciones cálidas a la región mesoamericana (Cooke, 2005; Piperno, 2006).

A pesar de ser un comportamiento climático global, la evidencia se vuelve particular para cada zona, por ejemplo, en la franja occidental de México, se tiene evidencia de clima más húmedo debido a la entrada de la celda alta del pacífico este durante la temporalidad de interés (Moy *et al.* 2002; Steinhilber *et al.* 2009).

La evidencia de los espeleotemas sugiere la prevalencia de condiciones secas y cálidas entre ~950 a 1400 CE (Lachniet *et al.* 2004). Como soporte de esta inferencia, se encontraron durante la revisión bibliográfica otros sitios con características paleoclimáticas similares en esa misma temporalidad: en la península de Yucatán de ~950 a 1250 CE (Curtis & Hodell, 1996), en Nuevo León (Roy *et al.* 2014), en Panamá (Correa-Metrio *et al.* 2011) y en el caribe (Peterson *et al.* 1991), muestran que ésta época representa el más grande periodo de sequía de los últimos dos milenios (Hodell *et al.* 2001). Por lo que representa una ventaja la cercanía del lago de Aljojuca a la zona de estudio al ser un importante indicador paleoclimático para el Valle de Tehuacán (Bhattacharya *et al.* 2015).

A pesar de ello, hay ciertas características propias del Valle que hay que tomar en cuenta para una correcta interpretación paleoclimática: Al norte se ha registrado una menor precipitación que al sur del Valle en los últimos 100 años (McNeish, 1964), lo cual es consistente con la evidencia arqueológica encontrada en cuestión de asentamientos humanos durante las diferentes etapas de desarrollo poblacional. En la zona sur del Valle no sólo se tiene evidencia de un mayor temporal de lluvias, sino también de sistemas de riego. Específicamente a 3 Km de la cueva de San Marcos y en zonas cercanas a la cueva de Coxcatlán, se encuentra un sistema de recolección de agua para el posible riego de hortalizas (Figura 8).

Se sabe que durante la temporalidad de interés (1500 años antes del presente), ya se practicaba el cultivo de plantas domesticadas. Las primeras evidencias de la existencia de una sociedad agrícola comienzan a partir de 7000 años antes del presente, con varios tipos de hortalizas y a partir de la época de Abejas con el cultivo de Maíz en la región (McNeish, 1964).

Aunando la evidencia arqueológica junto con la evidencia paleoclimática, se infiere que hubo condiciones de sequía pronunciadas hace 1450 años antes del presente en la región de interés.



Figura 55. Mapa de las zonas lacustres aledañas al Valle de Tehuacán.

Mapa del noreste de Tehuacán en donde las estrellas indican la ubicación de las lagunas de San Bernardino: Laguna chica $18^{\circ}35'52.4''N$ $97^{\circ}16'22.0''W$ y Laguna grande $18^{\circ}36'03.7''N$ $97^{\circ}15'48.0''W$.

Específicamente en el Valle de Tehuacán, sería interesante realizar estudios paleoclimáticos en zonas lacustres aledañas que permitan discernir fluctuaciones puntuales para esa misma temporalidad. En la figura 45 se proponen dos posibles sitios para un futuro estudio paleoclimático del sitio de interés, las Lagunas chica y grande de San Bernardino. Sería necesario hacer una salida de campo con investigadores que puedan aseverar que ese sitio podría arrojar evidencia paleoclimática, pero se parte de la idea de que los núcleos sedimentarios encontrados en los lagos proveen información de cambios climáticos (Cohen, 2003; Smol, 2008).

5.2 ADN antiguo en muestras paleobotánicas.

El daño que sufre el ADN una vez que los procesos diagenéticos comienzan a ocurrir han sido ampliamente estudiados, aunque se desconocen las cinéticas de las reacciones: La longitud de los fragmentos se ve comprometida al degradarse las hebras de ADN hasta un tamaño de 40-500 pares de bases (pb) (Pääbo, 1989). Esta fragmentación no decae consistentemente con el paso del tiempo (Lindahl & Anderson, 1972). Se sugieren procesos autolíticos que fragmentan el ADN inmediatamente después de la muerte del organismo (Lindahl & Nyberg, 1972). Además se presenta desnaturalización de las hebras de ADN así como deaminación en los extremos 5-fosfato y 3-aldehído como producto de la depurinización hidrolítica (Pääbo & Wilson, 1991; Lindahl, 1993; Jones *et al.* 1968).

Gracias a éstas marcas de daño podemos diferenciar ADN endógeno de una muestra antigua de la contaminación, ya sea de hongos, bacterias o como consecuencia del manejo de la muestra. A pesar de que se tienen protocolos establecidos para la excavación y procesamiento de las muestras arqueológicas, la contaminación con ADN proveniente de microorganismos es inevitable, pues puede

ocurrir desde la muerte del organismo y durante todo el tiempo que los restos permanecen enterrados. En contraste con las muestras previamente estudiadas (como SM10), en donde la cantidad de ADN endógeno fue informativa, otras muestras no han resultado tan fructíferas. Se ha demostrado que la preservación del ADN en una muestra ocurre como evento independiente incluso entre muestras del mismo sitio, pues las variaciones de temperatura, salinidad y pH del suelo pueden ser ligeramente distintas, además de los factores bióticos que puedan intervenir (Dabney *et al.* 2013).

En el caso de las muestras del Valle de Tehuacán, durante la extracción se siguieron protocolos estrictos para su protección, con lo que la contaminación moderna resultó ser nula o imperceptible en los análisis posteriores. En este proyecto, la calidad de ADN extraído de SM1 era mejor que la de Cox17. No obstante, no fue posible obtener una biblioteca de ADN tratado con bisulfito del extracto de SM1, por lo que la muestra tuvo que ser descartada.

La calidad y cantidad de ADN extraído de Cox17 permitió tener una primera aproximación a las problemáticas del análisis de metilación en genomas antiguos, lo que constituye la principal aportación de este trabajo, además del meta-análisis paleoclimático.

5.3 Bibliotecas genómicas

La elaboración de bibliotecas presentó un par de complicaciones técnicas. La primera fue durante el tratamiento de bisulfito, pues el protocolo estipula el uso de una columna con filtro en donde se realizan las reacciones y la elución final del producto con tratamiento (Anexo 3). Al finalizar el tratamiento, se espera una pérdida teórica del 30% de la concentración original, sin embargo, para la alícuota de Cox17, se realizó el tratamiento de bisulfito en varias ocasiones debido a que la pérdida era de hasta un 60%.

El segundo problema se reflejó en la disminución de la concentración de ADN en los pasos con purificación de la elaboración de bibliotecas. En el caso de SM1 y BisCox17 no hubo problema, pero para las bibliotecas de BisSM1 y Cox17 fue necesario repetir el proceso hasta obtener una biblioteca cualitativamente aceptable. En el caso de Cox17 fue posible (Figura 32), pero para BisSM1, la concentración impidió su uso para una posterior secuenciación (Anexo 5).

La propuesta para tratamientos de bisulfito en muestra antiguas sucesivos es utilizar perlas Ampure en todos los pasos del tratamiento de bisulfito, en vez de usar la columna de filtrado. Es importante llevar a cabo los pasos extra especificados en el Anexo 3 para obtener una mayor elución y continuar con el volumen deseado de ADN para los pasos sucesivos al tratamiento.

Otra mejora que se podría hacer para disminuir los problemas de mapeo es la preparación de las bibliotecas con la eliminación previa de uracilos, siempre y cuando la calidad inicial de la biblioteca lo permita. Es decir, debe ser eficiente el proceso de filtrado para poder implementar este paso. De esta manera las secuencias pueden ser diferenciadas más cómodamente pues los sitios en donde el genoma antiguo tiene un timina, pero no será mapeado si en el moderno se encuentra una citosina (Prüfer *et al.* 2014).

Los resultados obtenidos en la tabla 10 son consistentes con la metodología aplicada para cada biblioteca, en el caso de Cox17, se realizaron diez ciclos de amplificación mientras que para la biblioteca con tratamiento de bisulfito, aumentaron los ciclos de amplificación dada su baja concentración en las bibliotecas que se enviaron a secuenciar, que fue de un total de 0.86 ng/μl. Por lo

que las secuencias sin duplicados disminuyeron en un 70% para la biblioteca con tratamiento.

Finalmente, otro método propuesto para encontrar citosinas metiladas sin utilizar alícuotas adicionales de la extracción de ADN sin tratamiento de “wet lab”, se basa en un análisis “in silico” utilizando el software epiPALEOMIX (Hanghoj *et al.* 2016). Consiste en la comparativa entre metilomas modernos de los tejidos de interés con respecto a las regiones cubiertas del genoma antiguo con profundidad mayor a 2x para estimar sitios con posible metilación en los genomas antiguos. Sin embargo, en el caso de maíz que sufrió una transformación rápida en una escala temporal de milenios, no existen garantías de que los niveles de metilación actuales correspondan con los metilomas antiguos, por lo que esta aproximación debería aplicarse con las debidas reservas.

5.4 Secuenciación masiva.

5.4.1 Problemática durante el análisis bioinformático.

Diversas problemáticas se presentaron en ambas bibliotecas durante el análisis bioinformático: después de la secuenciación, los archivos catalogados como Cox17 y BisCox17 que se recibieron directamente del servicio de secuenciación fueron analizados siguiendo la metodología establecida. Debido a las múltiples amplificaciones por PCR las cuales fueron producto de una baja concentración inicial del extracto, el porcentaje disminuyó considerablemente en ambas bibliotecas, obteniendo el 66% para Cox17 y 29% para BisCox17 (tabla 9) para proceder con el pipeline informático. El daño molecular calculado por el algoritmo de MapDamage 2.0 fue menor a 0.05% para la biblioteca sin tratamiento. Tomando en cuenta que ambas bibliotecas provienen de la misma alícuota, el daño debería de ser el mismo para ambas bibliotecas (figura 47).

La principal problemática gira en torno a la interpretación de los porcentajes de GC tan lejanos de los valores teóricos esperados, pues en otros ensayos de bisulfito para maíz, el genoma tratado debería de tener en promedio una disminución del 20% con ligeras variaciones dependiendo de la accesión. Este es el primer indicio de que el tratamiento no tuvo el éxito esperado.

Se continuó trabajando a pesar de esta problemática debido a que era posible que los valores desproporcionados en la biblioteca con tratamiento se debiesen a contaminación de bacterias y hongos o al sesgo debido a las diversas amplificaciones que tuvieron lugar durante la elaboración de la biblioteca.

Durante el mapeo, el algoritmo BWA mem fue más eficiente para *reads* pareados que los no pareados o *single-end*, se realizaron diversos mapeos tanto con la referencia sin enmascarar como con las referencias enmascaradas de citosina a timinas y en todos los casos, el mapeo *paired-end* fue el más efectivo. Los resultados respecto a los parámetros usados para ambas bibliotecas cuando se tenían los *reads* mapeados resultaron ser robustos por lo que se implementaron de la misma manera durante el mapeo de los *reads* tanto a la referencia como a las referencias enmascaradas para identificar secuencias únicas mapeadas en maíz.

Siguiendo el consejo de la Dra. Gretrud Lund de revisar secuencias descartadas por el mapeo, se realizó una inspección más detallada del porcentaje de GC en la biblioteca con tratamiento. Se observó una curva de 64% muy diferenciada pero, a su vez hay una suave curva que tiene un porcentaje menor (figura 39). Se procedió a filtrar los *reads* con el menor porcentaje y fue ahí en donde se encontraron la fracción de aproximadamente 10 millones de *reads* en los cuales pudo ser efectivo el tratamiento de

bisulfito (tabla 10). Se procedió a encontrar *reads* compartidos en las bibliotecas con y sin tratamiento que cumplieran con una profundidad mínima de 3x y contexto biológico (CG, CGH, CHH).

Se utilizó un visor con los metilomas modernos como comparativa de metilación en regiones interesantes de la referencia (B73 versión 3). En este mismo visor se podía determinar si la región encontrada estaba anotada en el genoma o si tenía relevancia biológica. Encontrando un total de alrededor 260 mil citosinas metiladas en el genoma antiguo de maíz. Esto representa el 0.5% de las citosinas presentes en la biblioteca. Si bien el número no es significativo, es un primer indicio de que el tratamiento de bisulfito es efectivo en muestras paleobotánicas de maíz, especialmente si sufrieron procesos diagenéticos propios de climas tan áridos como es el caso de Tehuacán, Puebla.

En resumen, el porcentaje de GC que se observa en la figura 37 con tratamiento, se trata de un universo de *reads* compuesto por aquellos en los que hubo un tratamiento real de bisulfito, indicado por un porcentaje bajo de GC mientras aquellos que tienen un porcentaje mayor de CG no fueron tratados por bisulfito. Esto podría explicarse de dos maneras: la presencia de inhibidores en la muestra antigua que no permitieron completar la reacción de bisulfito de manera satisfactoria. Como segunda explicación, las múltiples amplificaciones de la librería para lograr la concentración mínima requerida para la secuenciación permitieron que se sesgara la amplificación a aquellos *reads* con alto contenido de GC, generando dímeros y por ende truncando un tratamiento efectivo. Esto debido a la propia naturaleza de la técnica en una muestra con baja concentración inicial.

La representación de la biblioteca fue muy pobre dada su concentración inicial y posterior disminución durante los diversos pasos de purificación. Es recomendable utilizar extractos con una mayor concentración inicial para los estudios de metilación antigua. Queda como perspectiva analizar de manera probabilística cada uno de los sitios donde se encuentran las citosinas metiladas y sus posibles genes anotados en la referencia moderna para tener una descripción más detallada de las citosinas metiladas encontradas y no sólo tener un porcentaje global de metilación.

Por lo que sería necesario investigar si una alícuota de aADN es capaz de inhibir la acción del bisulfito cuando se agrega a una muestra de ADN moderno, para confirmar si existe un factor inhibitorio más allá del propio sesgo que pueda tener el aADN.

CAPÍTULO VI. CONCLUSIONES

- La revisión de bibliografía paleoclimática indicó temporalidades de sequía a lo largo de Mesoamérica alrededor de ~800 a 1450 años antes del presente, incluyendo la región del Valle de Tehuacán.
- Se logró obtener aADN que permitió generar librerías genómicas. En el caso de Cox17, también se pudo generar la librería genómica tratada con bisulfito.
- A partir del aADN de maíz se obtuvo una representación del metiloma que permitió identificar sitios metilados en un contexto genómico.
- Encontramos 7 sitios de metilación reportados con metilación actual.
- Es posible realizar estudios de metilación en muestras de maíz antiguo provenientes de Tehuacán, Puebla a nivel genómico.

CAPÍTULO VII. BIBLIOGRAFÍA

- Ahmad, P., & Prasad, M. N. V. (2012). *Environmental adaptations and stress tolerance of plants in the era of climate change. Environmental Adaptations and Stress Tolerance of Plants in the Era of Climate Change*. <https://doi.org/10.1007/978-1-4614-0815-4>
- Allaby, R. G., Gutaker, R., Clarke, A. C., Pearson, N., Ware, R., Palmer, S. A., ... Smith, O. (n.d.). Using archaeogenomic and computational approaches to unravel the history of local adaptation in crops. <https://doi.org/10.1098/rstb.2013.0377>
- Alonso, C., Pérez, R., Bazaga, P., & Herrera, C. M. (2015). Global DNA cytosine methylation as an evolving trait: Phylogenetic signal and correlated evolution with genome size in angiosperms. *Frontiers in Genetics*. <https://doi.org/10.3389/fgene.2015.00004>
- Bhattacharya, T., Byrne, R., Böhnelt, H., Wogau, K., Kienel, U., Ingram, B. L., ... Butzer, K. W. (n.d.). Cultural implications of late Holocene climate change in the Cuenca Oriental, Mexico. <https://doi.org/10.1073/pnas.1405653112>
- Briggs, A. W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., & Pääbo, S. (2010). Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gkp1163>
- Chandler, V. L., Eggleston, W. B., & Dorweiler, J. E. (2000). Paramutation in maize. *Plant Molecular Biology Plant Gene Silencing*, 43, 121–145.
- Dabney, J., Meyer, M., & Pääbo, S. (n.d.). Ancient DNA Damage. <https://doi.org/10.1101/cshperspect.a012567>
- Dubin, M. J., Zhang, P., Meng, D., Remigereau, M. S., Osborne, E. J., Casale, F. P., ... Nordborg, M. (2015). DNA methylation in Arabidopsis has a genetic basis and shows evidence of local adaptation. *eLife*. <https://doi.org/10.7554/eLife.05255>
- Eichten, S. R., Swanson-Wagner, R. A., Schnable, J. C., Waters, A. J., Hermanson, P. J., Liu, S., ... Springer, N. M. (2011). Heritable epigenetic variation among maize inbreds. *PLoS Genetics*. <https://doi.org/10.1371/journal.pgen.1002372>
- Eichten, S. R., Briskine, R., Song, J., Li, Q., Swanson-Wagner, R., Hermanson, P. J., ... Springer, N. M. (n.d.). Epigenetic and Genetic Influences on DNA Methylation Variation in Maize Populations. <https://doi.org/10.1105/tpc.113.114793>
- Eichten, S. R., Briskine, R., Song, J., Li, Q., Swanson-Wagner, R., Hermanson, P. J., ... Springer, N. M. (2013). Epigenetic and genetic influences on DNA methylation variation in maize populations. *The Plant Cell*. <https://doi.org/10.1105/tpc.113.114793>
- Feng, S., & Jacobsen, S. E. (2011). Epigenetic modifications in plants: An evolutionary perspective. *Current Opinion in Plant Biology*. <https://doi.org/10.1016/j.pbi.2010.12.002>

- Feng, S., & Jacobsen, S. E. (2011). Epigenetic modifications in plants: An evolutionary perspective. *Current Opinion in Plant Biology*. <https://doi.org/10.1016/j.pbi.2010.12.002>
- Gansauge, M.-T., & Meyer, M. (2014). genome sequencing Selective enrichment of damaged DNA molecules for ancient. <https://doi.org/10.1101/gr.174201.114>
- Gonzalzo, M. L., & Jones, P. A. (1997). Mutagenic and epigenetic effects of DNA methylation. *Mutation Research - Reviews in Mutation Research*. [https://doi.org/10.1016/S1383-5742\(96\)00047-6](https://doi.org/10.1016/S1383-5742(96)00047-6)
- Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F., & Orlando, L. (2013). MapDamage2.0: Fast approximate Bayesian estimates of ancient DNA damage parameters. In *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btt193>
- Kalisz, S., & Purugganan, M. D. (2004). Epialleles via DNA methylation: Consequences for plant evolution. *Trends in Ecology and Evolution*. <https://doi.org/10.1016/j.tree.2004.03.034>
- Kumar, S., Kumari, R., Sharma, V., & Sharma, V. (2013). Roles, and establishment, maintenance and erasing of the epigenetic cytosine methylation marks in plants. *Journal of Genetics*. <https://doi.org/10.1007/s12041-013-0273-8>
- Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. <https://doi.org/10.1038/nrg2719>
- Li, Q., Eichten, S. R., Hermanson, P. J., & Springer, N. M. (2014). Inheritance patterns and stability of DNA methylation variation in maize near-isogenic lines. *Genetics*. <https://doi.org/10.1534/genetics.113.158980>
- Li, Q., Song, J., West, P. T., Zynda, G., Eichten, S. R., Vaughn, M. W., & Springer, N. M. (2015). Examining the Causes and Consequences of Context-Specific Differential DNA Methylation in Maize. *Plant Physiology*. <https://doi.org/10.1104/pp.15.00052>
- Llamas, B., Holland, M. L., Chen, K., Cropley, J. E., Cooper, A., & Suter, C. M. (2012). High-resolution analysis of cytosine methylation in ancient DNA. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0030226>
- Llamas, B., Holland, M. L., Chen, K., Cropley, J. E., Cooper, A., & Suter, C. M. (2012). High-resolution analysis of cytosine methylation in ancient DNA. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0030226>
- Lu, Y., Rong, T., & Cao, M. (2008). Analysis of DNA methylation in different maize tissues. *Journal of Genetics and Genomics*. [https://doi.org/10.1016/S1673-8527\(08\)60006-5](https://doi.org/10.1016/S1673-8527(08)60006-5)
- Mascher, M., Schuenemann, V. J., Davidovich, U., Marom, N., Himmelbach, A., Korol, A., ... Stein, N. (2016). Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley. <https://doi.org/10.1038/ng.3611>

- Niederhuth, C. E., & Schmitz, R. J. (2014). Covering your bases: Inheritance of DNA methylation in plant genomes. *Molecular Plant*. <https://doi.org/10.1093/mp/sst165>
- Palmer, S. A., Clapham, A. J., Rose, P., Freitas, F. O., Owen, B. D., Beresford-Jones, D., ... Allaby, R. G. (2012). Archaeogenomic evidence of punctuated genome evolution in gossypium. *Molecular Biology and Evolution*. <https://doi.org/10.1093/molbev/mss070>
- Pedersen, J. S., Valen, E., Velazquez, A. M. V., Parker, B. J., Rasmussen, M., Lindgreen, S., ... Orlando, L. (2014). Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome. *Genome Research*. <https://doi.org/10.1101/gr.163592.113>
- Qing Li, W., Eichten, S. R., Hermanson, P. J., Zaunbrecher, V. M., Song, J., Wendt, J., ... Springer, N. M. (n.d.). Genetic Perturbation of the Maize Methylome. <https://doi.org/10.1105/tpc.114.133140>
- Reinders, J., Wulff, B. B. H., Mirouze, M., Marí-Ordóñez, A., Dapp, M., Rozhon, W., ... Paszkowski, J. (2009). Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes and Development*. <https://doi.org/10.1101/gad.524609>
- Russell, J., Mascher, M., Dawson, I. K., Kyriakidis, S., Calixto, C., Freund, F., ... Waugh, R. (2016). Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. <https://doi.org/10.1038/ng.3612>
- Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., & Pääbo, S. (2012). Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0034131>
- Schlumbaum, A., Tensen, M., & Jaenicke-Després, V. (2008). Ancient plant DNA in archaeobotany. *Vegetation History and Archaeobotany*. <https://doi.org/10.1007/s00334-007-0125-7>
- Schmidt, M., Hense, S., Minoche, A. E., Dohm, J. C., Himmelbauer, H., Schmidt, T., & Zakrzewski, F. (2014). Cytosine methylation of an ancient satellite family in the wild beet *beta procumbens*. *Cytogenetic and Genome Research*. <https://doi.org/10.1159/000363485>
- Shan, X., Wang, X., Yang, G., Wu, Y., Su, S., Li, S., Yuan, Y. (2013). Analysis of the DNA methylation of maize (*Zea mays* L.) in response to cold stress based on methylation-sensitive amplified polymorphisms. *Journal of Plant Biology*. <https://doi.org/10.1007/s12374-012-0251-3>
- Smith, O., Clapham, A. J., Rose, P., Liu, Y., Wang, J., & Allaby, R. G. (n.d.). Genomic methylation patterns in archaeological barley show de-methylation as a time-dependent diagenetic process. <https://doi.org/10.1038/srep05559>
- Balvanera, Patricia, Ernesto Arias-González, Ricardo Rodríguez-Estrella, Lucía AlmeidaLeñero, Juan J. Schmitter-Soto. 2016. Ecosistemas de México: una mirada a su conocimiento. Ciudad de México, Universidad Nacional Autónoma de México.

P. Ahmad and M.N.V. Prasad (2012), Environmental Adaptations and Stress Tolerance of Plants in the Era of Climate Change. Springer DOI 10.1007/978-1-4614-0815-4_1

Peterson, L. C., J. T. Overpeck, N. G. Kipp, and J. Imbrie (1991), A high-resolution Late Quaternary upwelling record from the anoxic Cariaco Basin, Venezuela, *Paleoceanography*, 6(1), 99–119, doi:10.1029/90PA02497.

M. Meyer, M. Kircher, M. T. Gansauge, H. Li, F. Racimo, S., ... S. Pääbo, A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338, 222–226 (2012). Medline doi:10.1126/science.1224344

K. Prüfer, F. Racimo, N. Patterson, F. Jay, S. Sankararaman, S. Sawyer, A.,... S. Pääbo, The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505, 43–49 (2014). Medline doi:10.1038/nature12886

J. S. Pedersen, E. Valen, A. M. Velazquez, B. J. Parker, M. Rasmussen, ... E. Willerslev, L. Orlando, Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome.

7.1 Bibliografía para revisión paleoclimática

Lozano-García, S., Torres-Rodríguez, E., Ortega, B., Vázquez, G., & Caballero, M. (2013). Ecosystem responses to climate and disturbances in western central Mexico during the late Pleistocene and Holocene. *Palaeogeography, Palaeoclimatology, Palaeoecology*.
<https://doi.org/10.1016/j.palaeo.2012.12.006>

Alcocer, J., & Bernal-Brooks, F. W. (2010). Limnology in Mexico. *Hydrobiologia*.
<https://doi.org/10.1007/s10750-010-0211-1>

Alley, R. B., & Ágústsson, A. M. (2005). The 8k event: Cause and consequences of a major Holocene abrupt climate change. *Quaternary Science Reviews*.
<https://doi.org/10.1016/j.quascirev.2004.12.004>

Almeida-Lenero, L., Hooghiemstra, H., Cleef, A. M., & Van Geel, B. (2005). Holocene climatic and environmental change from pollen records of lakes Zempoala and Quila, central Mexican highlands. *Review of Palaeobotany and Palynology*.
<https://doi.org/10.1016/j.revpalbo.2005.05.001>

Almeida-Lenero, L., Hooghiemstra, H., Cleef, A. M., & Van Geel, B. (2005). Holocene climatic and environmental change from pollen records of lakes Zempoala and Quila, central Mexican highlands. *Review of Palaeobotany and Palynology*.
<https://doi.org/10.1016/j.revpalbo.2005.05.001>

Arnaud, F., Révillon, S., Debret, M., Revel, M., Chapron, E., Jacob, J., Magny, M. (2012). Lake Bourget regional erosion patterns reconstruction reveals Holocene NW European Alps soil evolution and paleohydrology. *Quaternary Science Reviews*.
<https://doi.org/10.1016/j.quascirev.2012.07.025>

- Beer, J., & Wanner, H. (2012). Corrigendum to “Mid- to late Holocene climate change - an overview” [Quaternary Sci. Rev. 27 (2008) 1791-1828]. *Quaternary Science Reviews*.
<https://doi.org/10.1016/j.quascirev.2012.04.015>
- Bernal, J. P., Lachniet, M., McCulloch, M., Mortimer, G., Morales, P., & Cienfuegos, E. (2011). A speleothem record of Holocene climate variability from southwestern Mexico. *Quaternary Research*. <https://doi.org/10.1016/j.yqres.2010.09.002>
- Berrio, J. C., Hooghiemstra, H., Van Geel, B., & Ludlow-Wiechers, B. (2006). Environmental history of the dry forest biome of Guerrero, Mexico, and human impact during the last c. 2700 years. *The Holocene Edward Arnold (Publishers) Ltd*, 16(16), 63–80.
- Bhattacharya, T., Byrne, R., Böhnell, H., Wogau, K., Kienel, U., Ingram, B. L., ... Butzer, K. W. (n.d.). Cultural implications of late Holocene climate change in the Cuenca Oriental, Mexico. <https://doi.org/10.1073/pnas.1405653112>
- Bhattacharya, T., Byrne, R., Böhnell, H., Wogau, K., Kienel, U., Ingram, B. L., ... Butzer, K. W. (n.d.). Cultural implications of late Holocene climate change in the Cuenca Oriental, Mexico. <https://doi.org/10.1073/pnas.1405653112>
- Birks, H. J. B. (n.d.). CLIMATE CHANGE IN THE HOLOCENE -AN OVERVIEW.
- Blamey, N. J. F., Brand, U., Parnell, J., Spear, N., Lückey, C., Benison, K., ... Ni, P. (2016). Paradigm shift in determining Neoproterozoic atmospheric oxygen. *Geology*.
<https://doi.org/10.1130/G37937.1>
- Bradbury, J. P. (1989). Late Quaternary lacustrine paleoenvironments in the Cuenca de México. *Quaternary Science Reviews*. [https://doi.org/10.1016/0277-3791\(89\)90022-X](https://doi.org/10.1016/0277-3791(89)90022-X)
- Caballero, M., Lozano-García, S., Vázquez-Selem, L., & Ortega, B. (2010). Evidencias de cambio climático y ambiental en registros glaciales y en cuencas lacustres del centro de México durante el último máximo glacial. *Boletín de La Sociedad Geológica Mexicana*.
- Caballero, M., & Ortega, B. (n.d.). Evidencias de cambio climático y ambiental en registros glaciales y en cuencas lacustres del centro de México durante el último máximo glacial, 1–46.
- Carolina Ruiz-Fernández, A., Hillaire-Marcel, C., Páez-Osuna, F., Ghaleb, B., & Caballero, M. (2007). 210Pb chronology and trace metal geochemistry at Los Tuxtlas, Mexico, as evidenced by a sedimentary record from the Lago Verde crater lake. *Quaternary Research*.
<https://doi.org/10.1016/j.yqres.2006.11.003>
- Chen, C. W., Wei, K. Y., Mii, H. S., & Yang, T. N. (2008). A late Quaternary planktonic foraminiferal oxygen isotope record of the Banda Sea: Chronostratigraphy, orbital forcing, and paleoceanographic implications. *Terrestrial, Atmospheric and Oceanic Sciences*.
[https://doi.org/10.3319/TAO.2008.19.4.331\(IMGES\)](https://doi.org/10.3319/TAO.2008.19.4.331(IMGES))

- Correa-Metrio, A., Bush, M., Lozano-García, S., & Sosa-Nájera, S. (2013). Millennial-scale temperature change velocity in the continental northern neotropics. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0081958>
- Correa-Metrio, A., Elez, M. I. V, Escobar, J., St-Jacques, J.-M., Opez-P Erez, M. L., Curtis, J., & Cosford, J. (n.d.). Mid-elevation ecosystems of Panama: future uncertainties in light of past global climatic variability; Mid-elevation ecosystems of Panama: future uncertainties in light of past global climatic variability. <https://doi.org/10.1002/jqs.2899>
- Correa-Metrio, A., Vélez, M. I., Escobar, J., St-Jacques, J.-M., López-Pérez, M., Curtis, J., & Cosford, J. (n.d.). Mid-elevation ecosystems of Panama: future uncertainties in light of past global climatic variability. Retrieved from <http://mc.manuscriptcentral.com/jqs>
- Curtis, J. H., Hodell, D. A., & Brenner, M. (1996). Climate Variability on the Yucatan Peninsula (Mexico) during the Past 3500 Years, and Implications for Maya Cultural Evolution. *QUATERNARY RESEARCH*, 46, 37–47.
- Davies, S. J., Metcalfe, S. E., Mackenzie, A. B., Newton, A. J., Endfield, G. H., & Farmer, J. G. (n.d.). Environmental changes in the Zirahu e en Basin, Michoacán, Mexico, during the last 1000 years.
- Davies, S. J., Metcalfe, S. E., Mackenzie, A. B., Newton, A. J., Endfield, G. H., & Farmer, J. G. (n.d.). Environmental changes in the Zirahu e en Basin, Michoacán, Mexico, during the last 1000 years.
- Davies, S. J., Metcalfe, S. E., Mackenzie, A. B., Newton, A. J., Endfield, G. H., & Farmer, J. G. (n.d.). Environmental changes in the Zirahu e en Basin, Michoacán, Mexico, during the last 1000 years.
- Davies, S. J., Metcalfe, S. E., Bernal-Brooks, F., Chacó N-Torres, A., Farmer, J. G., Mackenzie, A. B., & Newton, A. J. (n.d.). Lake Sediments Record Sensitivity of Two Hydrologically Closed Upland Lakes in Mexico to Human Impact. *A Journal of the Human Environment*, 34(6), 470–475. <https://doi.org/10.1579/0044-7447-34.6.470>
- De Bruyn, M., Hall, B. L., Chauke, L. F., Baroni, C., Koch, P. L., & Hoelzel, A. R. (2009). Rapid response of a marine mammal species to holocene climate and habitat change. *PLoS Genetics*. <https://doi.org/10.1371/journal.pgen.1000554>
- Del, M., Lozano-García, S., Caballero, M., Ortega, B., Rodríguez, A., & Sosa, S. (n.d.). Tracing the effects of the Little Ice Age in the tropical lowlands of eastern Mesoamerica.
- De Bruyn, M., Hoelzel, A. R., Carvalho, G. R., & Hofreiter, M. (2011). Faunal histories from Holocene ancient DNA. *Trends in Ecology and Evolution*. <https://doi.org/10.1016/j.tree.2011.03.021>
- Lachniet, M. S., Bernal, J. P., Asmerom, Y., Polyak, V., & Piperno, D. (2012). A 2400 yr Mesoamerican rainfall reconstruction links climate and cultural change. *Geology*. <https://doi.org/10.1130/G32471.1>
- Del, M., Lozano-García, S., Caballero, M., Ortega, B., Rodríguez, A., & Sosa, S. (n.d.). Tracing the effects of the Little Ice Age in the tropical lowlands of eastern Mesoamerica.

- Fisher, C. T., Pollard, H. P., Israde-Alcá, I., Garduñ O-Monroy, V. H., Banerjee, S. K., & Marcus, J. (n.d.). A reexamination of human-induced environmental change within the Lake Pá tzcuaro Basin, Michoacán, Mexico.
- Foote, A. D., Kaschner, K., Schultze, S. E., Garilao, C., Ho, S. Y. W., Post, K., ... Gilbert, P. (2013). ARTICLE Ancient DNA reveals that bowhead whale lineages survived Late Pleistocene climate change and habitat shifts. *Nature Communications*, 4. <https://doi.org/10.1038/ncomms2714>
- García-Aguirre, M. C., Álvarez, R., Dirzo, R., Ortiz, M. A., & Eng, M. M. (2010). Delineation of biogeomorphic land units across a tropical natural and humanized terrain in Los Tuxtlas, Veracruz, Mexico. *Geomorphology*. <https://doi.org/10.1016/j.geomorph.2010.04.019>
- Grimm, E. C., Lozano-Garcia, S., Behling, H., & Markgraf, V. (n.d.). Holocene Vegetation and Climate Variability in the Americas.
- Hadly, E. A., Ramakrishnan, U., Chan, Y. L., Van Tuinen, M., O'Keefe, K., Spaeth, P. A., & Conroy, C. J. (2004). Genetic response to climatic change: Insights from ancient DNA and phylogenetics. *PLoS Biology*. <https://doi.org/10.1371/journal.pbio.0020290>
- Heine, K. (n.d.). Anthropogenic Sedimentological Changes during the Holocene in Mexico and Central America.
- Hoelzel, A. R. (n.d.). Looking backwards to look forwards: conservation genetics in a changing world. <https://doi.org/10.1007/s10592-010-0045-4>
- Hofreiter, M., & Stewart, J. (2009). Ecological Change, Range Fluctuations and Population Dynamics during the Pleistocene. *Current Biology*. <https://doi.org/10.1016/j.cub.2009.06.030>
- Hsing, P.-Y., Bo Fu, B., Elizabeth Larcom, B. A., Samantha Berlet, B. P., Timothy Shank, B. M., Annette Govindarajan, B. F., ... Charles Fisher, B. R. (n.d.). Evidence of lasting impact of the Deepwater Horizon oil spill on a deep Gulf of Mexico coral community. <https://doi.org/10.12952/journal.elementa.000012>
- Istrade-Alcántara, I., Garduño-Monroy, V. H., Fisher, C. T., Pollard, H. P., & Rodríguez-Pascua, M. A. (2005). Lake level change, climate, and the impact of natural events: The role of seismic and volcanic events in the formation of the Lake Patzcuaro Basin, Michoacan, Mexico. *Quaternary International*. <https://doi.org/10.1016/j.quaint.2004.10.022>
- Kienel, U., Bowen, S. W., Byrne, R., Park, J., Böhnelt, H., Dulski, P., ... Negendank, J. F. W. (2009). First lacustrine varve chronologies from Mexico: Impact of droughts, ENSO and human activity since AD 1840 as recorded in maar sediments from Valle de Santiago. *Journal of Paleolimnology*. <https://doi.org/10.1007/s10933-009-9307-x>
- Leng, M. J., Metcalfe, S. E., & Davies, S. J. (2005). Investigating Late Holocene climate variability in central Mexico using carbon isotope ratios in organic materials and oxygen isotope ratios from diatom silica within lacustrine sediments. *Journal of Paleolimnology*. <https://doi.org/10.1007/s10933-005-6748-8>

- Leng, M. J., Metcalfe, S. E., & Davies, S. J. (2005). Investigating Late Holocene climate variability in central Mexico using carbon isotope ratios in organic materials and oxygen isotope ratios from diatom silica within lacustrine sediments. *Journal of Paleolimnology*. <https://doi.org/10.1007/s10933-005-6748-8>
- Ljung, K., Björck, S., Renssen, H., & Hammarlund, D. (2008). South Atlantic island record reveals a South Atlantic response to the 8.2 kyr event. *Clim. Past*, 4, 35–45. Retrieved from www.clim-past.net/4/35/2008/
- Lounejeva Baturina, E., Morales Puente, P., Cabadas Báez, H. V., Cienfuegos Alvarado, E., Sedov, S., Vallejo Gómez, E., & Solleiro Rebolledo, E. (2006). Late Pleistocene to Holocene environmental changes from $\delta \delta \delta \delta \delta$ C determinations in soils at Teotihuacan, Mexico. *Geofísica Internacional*, 45(2), 85–98.
- Lozano-García, S., Caballero, M., Ortega, B., Sosa, S., Rodríguez, A., & Schaaf, P. (2010). Late Holocene palaeoecology of Lago Verde: Evidence of human impact and climate change in the northern limit of the neotropics during the late formative and classic periods. *Vegetation History and Archaeobotany*. <https://doi.org/10.1007/s00334-010-0240-8>
- Lozano-García, S., Sosa-Nájera, S., Sugiura, Y., & Caballero, M. (2005). 23,000 yr of vegetation history of the Upper Lerma, a tropical high-altitude basin in Central Mexico. *Quaternary Research*. <https://doi.org/10.1016/j.yqres.2005.02.010>
- Lozano-García, S., Torres-Rodríguez, E., Ortega, B., Vázquez, G., & Caballero, M. (2013). Ecosystem responses to climate and disturbances in western central Mexico during the late Pleistocene and Holocene. *Palaeogeography, Palaeoclimatology, Palaeoecology*. <https://doi.org/10.1016/j.palaeo.2012.12.006>
- Lozano-García, S., Torres-Rodríguez, E., Ortega, B., Vázquez, G., & Caballero, M. (2013). Ecosystem responses to climate and disturbances in western central Mexico during the late Pleistocene and Holocene. *Palaeogeography, Palaeoclimatology, Palaeoecology*. <https://doi.org/10.1016/j.palaeo.2012.12.006>
- Lozano-García, S., & Vázquez-Selem, L. (n.d.). A high-elevation Holocene pollen record from Iztaccíhuatl volcano, central Mexico. <https://doi.org/10.1191/0959683605hl814rp>
- Ludlow-Wiechers, B., Almeida-Leñero, L., & Islebe, G. (2005). Paleoecological and climatic changes of the Upper Lerma Basin, Central Mexico during the Holocene. *Quaternary Research*. <https://doi.org/10.1016/j.yqres.2005.08.004>
- Mahowald, N. M., Yoshioka, M., Collins, W. D., Conley, A. J., Fillmore, D. W., & Coleman, D. B. (2006). Climate response and radiative forcing from mineral aerosols during the last glacial maximum, pre-industrial, current and doubled-carbon dioxide climates. *Geophysical Research Letters*. <https://doi.org/10.1029/2006GL026126>
- Mangerud, J., Andersen, S. T., Berglund, B. E., Donner Mangerud, J. J., Andersen, J., & Rner, D. (n.d.). Quaternary stratigraphy of Norden) a proposal for terminology and classification.

- Mayewski, P. A., Rohling, E. E., Stager, J. C., Karl, W., Maasch, K. A., Meeker, L. D., ... Steig, E. J. (2004). Holocene climate variability. *Quaternary Research*.
<https://doi.org/10.1016/j.yqres.2004.07.001>
- Metcalf, S. E. (1987). Review: Historical Data and Climatic Change in Mexico: A Review. *The Geographical Journal*. <https://doi.org/10.2307/634873>
- Metcalf, S. E. (1997). PALAEOCLIMATOLOGICAL RECORDS OF CLIMATE CHANGE IN MEXICO - FRUSTRATING PAST, PROMISING FUTURE? *Quaternary International*, 43144, 11–16.
- Metcalf, S. E., Davies, S. J., Braisby, J. D., Leng, M. J., Newton, A. J., Terrett, N. L., & O'Hara, S. L. (2007). Long and short-term change in the Tzucuaru Basin, central Mexico. *Palaeogeography, Palaeoclimatology, Palaeoecology*. <https://doi.org/10.1016/j.palaeo.2006.10.018>
- Metcalf, S. E., O'Hara, S. L., Caballero, M., & Davies, S. J. (2000). Records of Late Pleistocene-Holocene climatic change in Mexico - A review. *Quaternary Science Reviews*.
[https://doi.org/10.1016/S0277-3791\(99\)00022-0](https://doi.org/10.1016/S0277-3791(99)00022-0)
- Metcalf, S. E., Street-Perrott, F. A., Perrott, R. A., & Harkness, D. D. (1991). Palaeolimnology of the Upper Lerma Basin, Central Mexico: a record of climatic change and anthropogenic disturbance since 11 600 yr BP. *Journal of Paleolimnology*. <https://doi.org/10.1007/BF00200345>
- Metcalf, S., & Davies, S. (2007). Deciphering recent climate change in central Mexican lake records. *Climatic Change*. <https://doi.org/10.1007/s10584-006-9152-0>
- O'Hara, S. L., & Metcalf, S. E. (1997). THE CLIMATE OF MEXICO SINCE THE AZTEC PERIOD. *Quaternary International*, 43144, 25–3.
- O'Hara, S. L., Metcalf, S. E., & Street-Perrott, F. A. (1994). On the arid margin: The relationship between climate, humans and the environment. A review of evidence from the highlands of central Mexico. *Chemosphere*. [https://doi.org/10.1016/0045-6535\(94\)90163-5](https://doi.org/10.1016/0045-6535(94)90163-5)
- Ortega, B., Caballero, M., Lozano, S., Vilaclara, G., & Rodríguez, A. (2006). Rock magnetic and geochemical proxies for iron mineral diagenesis in a tropical lake: Lago Verde, Los Tuxtlas, East-Central Mexico. *Earth and Planetary Science Letters*. <https://doi.org/10.1016/j.epsl.2006.08.020>
- Ortega, B., Vázquez, G., Caballero, M., Israde, I., Lozano-García, S., Schaaf, P., & Torres, E. (2010). Late Pleistocene: Holocene record of environmental changes in Lake Zirahuén, Central Mexico. *Journal of Paleolimnology*. <https://doi.org/10.1007/s10933-010-9449-x>
- Park, J., Byrne, R., Böhnell, H., Garza, R. M., & Conserva, M. (2010). Holocene climate change and human impact, central Mexico: a record based on maar lake pollen and sediment chemistry. *Quaternary Science Reviews*. <https://doi.org/10.1016/j.quascirev.2009.10.017>
- Park, J., Byrne, R., Böhnell, H., Garza, R. M., & Conserva, M. (2010). Holocene climate change and human impact, central Mexico: a record based on maar lake pollen and sediment chemistry. *Quaternary Science Reviews*. <https://doi.org/10.1016/j.quascirev.2009.10.017>

- Piperno, D. R., Moreno, J. E., Iriarte, J., Holst, I., Lachniet, M., Jones, J. G., ... Castanzo, R. (n.d.). Late Pleistocene and Holocene environmental history of the Iguala Valley, Central Balsas Watershed of Mexico.
- Platt Bradbury, J. (2000). Limnologic history of Lago de Patzcuaro, Michoacan, Mexico for the past 48,000 years: Impacts of climate and man. *Palaeogeography, Palaeoclimatology, Palaeoecology*. [https://doi.org/10.1016/S0031-0182\(00\)00146-2](https://doi.org/10.1016/S0031-0182(00)00146-2)
- Ranere, A. J., Piperno, D. R., Holst, I., Dickau, R., & Iriarte, J. (n.d.). The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico.
- Roy, P. D., Rivero-Navarrete, A., Sánchez-Zavala, J. L., López-Balbiaux, N., Ic, M. E. X., So, A. A. C., ... G, L. Ó. (2014). Subsurface fire and subsidence at Valle del Potosí (Nuevo León, Mexico): Preliminary observations, *66*(3), 553–557.
- Stahle, D. W., Diaz, J. V., Burnette, D. J., Paredes, J. C., Heim, R. R., Fye, F. K., ... Stahle, D. K. (2011). Major Mesoamerican droughts of the past millennium. *Geophysical Research Letters*. <https://doi.org/10.1029/2010GL046472>
- Street-Perrott, F. A., Perrott, R. A., & Harkness, D. D. (1989). Society for American Archaeology Anthropogenic Soil Erosion around Lake Patzcuaro, Michoacan, Mexico, during the Preclassic and Late Postclassic-Hispanic Periods. *Source: American Antiquity*, *54*(4), 759–765. Retrieved from <http://www.jstor.org/stable/280680>
- Turner, C., & West, R. G. (1958). The subdivision and zonation of interglacial periods.
- Walker, M., Johnsen, S., Rasmussen, S. O., Popp, T., Steffensen, J. P., Gibbard, P., ... Schwander, J. (2009). Formal definition and dating of the GSSP (Global Stratotype Section and Point) for the base of the Holocene using the Greenland NGRIP ice core, and selected auxiliary records. *Journal of Quaternary Science*. <https://doi.org/10.1002/jqs.1227>
- Watts, W. A., & Bradbury, J. P. (1982). Paleoecological studies at Lake Patzcuaro on the west-central Mexican Plateau and at Chalco in the basin of Mexico. *Quaternary Research*. [https://doi.org/10.1016/0033-5894\(82\)90045-X](https://doi.org/10.1016/0033-5894(82)90045-X)
- Abrupt Climate Change Inevitable Surprises Library of Congress Cataloging-in-Publication Data. (2002). Retrieved from <http://www.nap.edu/catalog/10136.html>