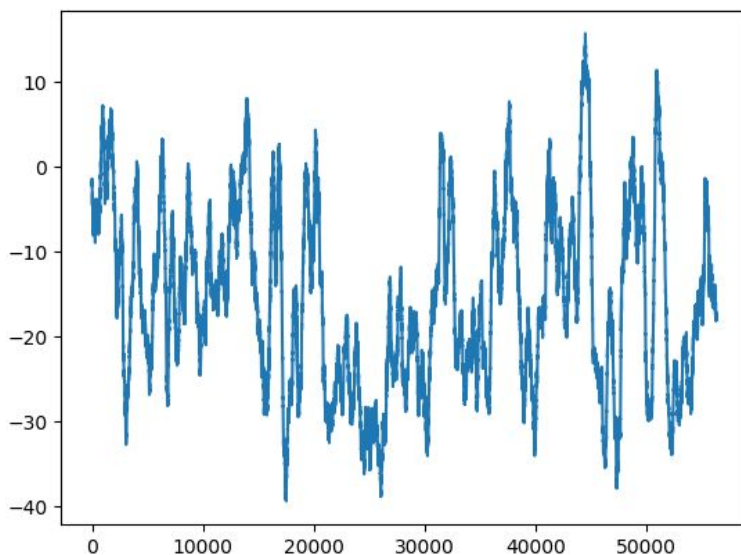


Ergebnisse der Studienleistung ALBI

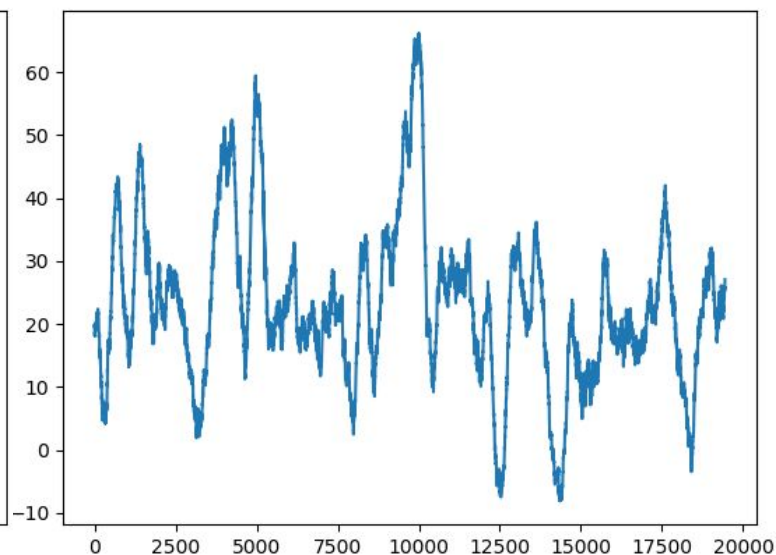
SoSe 2019, Constantin Zackl

Im Rahmen der Studienleistung habe Ich mit den von mir geschriebenen Programmen menschliche mitochondriale DNA Sequenzen (**CLUH**) untersucht und die folgenden Plots erstellt. Dabei sind für die Fenstergröße 401 jeweils die log likelihood ratios für die gegebenen background und CpG Sequenzdateien und für die mitochondriale Sequenz dargestellt.

Background



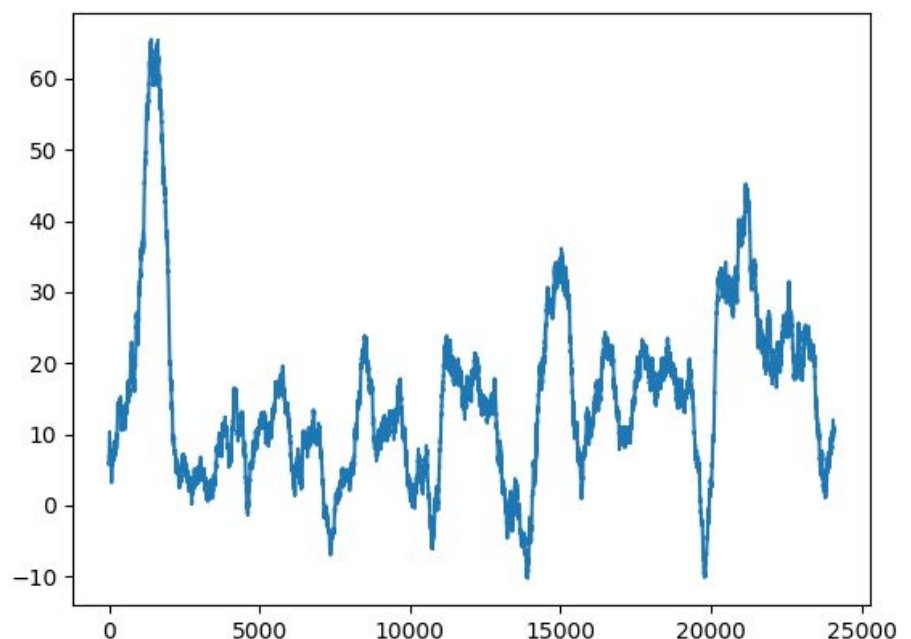
CpG Islands



Auffällig sind die wesentlich höheren Werte im CpG Islands Plot. Die Werte werden kaum negativ. Auch sind klare einzelne Spitzen nach oben zu erkennen während die Background Datei eher einem Hintergrundrauschen gleicht. Ich vermute, dass sich signifikant höhere Werte und klar abgegrenzte Spitzen in den Plots zur Charakterisierung von CpG Islands nutzen lassen.

Mitochondrium:

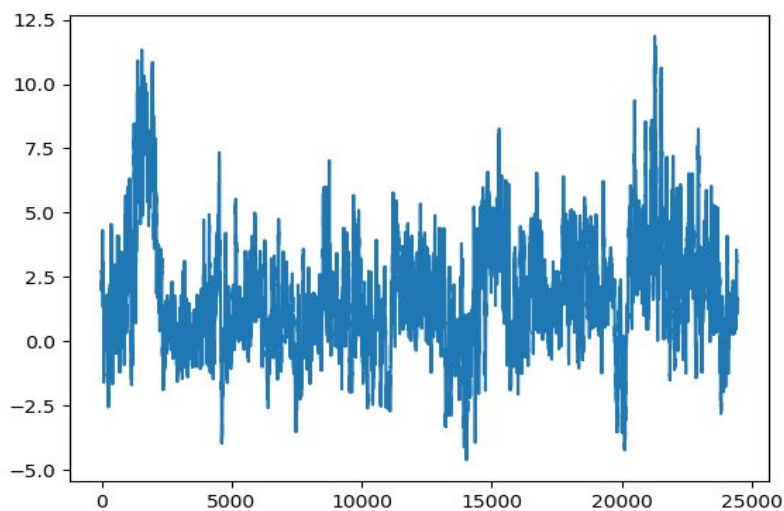
Betrachtet man den entsprechenden Plot für das menschliche Mitochondrium erkennt man ähnlich hohe Werte wie beim CpG Plot. Besonders auffällig ist die sehr hohe Spitze im Bereich der ersten 2500 Basenpaare. Auch in der zweiten Hälfte der Sequenz sind kleinere, aber klar abgegrenzte Spitzen in den Werten zu sehen.



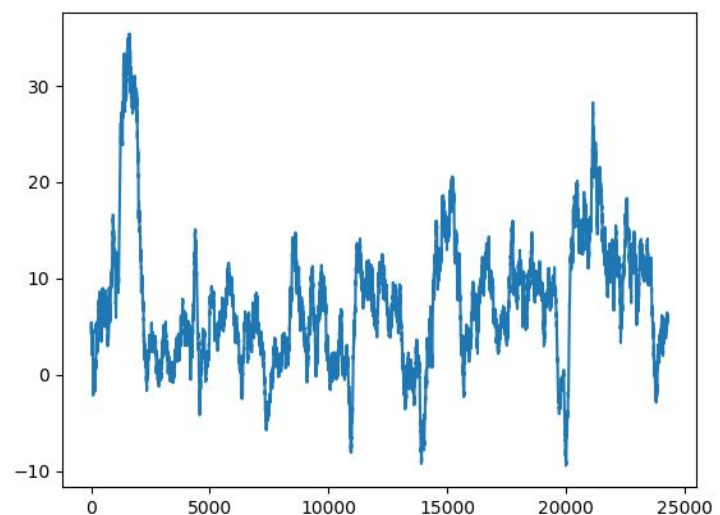
Meiner Meinung nach enthalten menschliche Mitochondrien CpG Inseln. Um sie klarer abzugrenzen wäre die Definition eines Schwellwertes sinnvoll, da die Background Datei zeigt, dass Werte über 0 nicht ausreichen um auf CpG Inseln hinzuweisen. Dieser Schwellwert ließe sich mit Hilfe des CpG Plots erstellen, wenn die genauen Inselgrenzen bekannt sind.

Im folgenden noch zwei Diagramme um kleinere Fenstergrößen beispielhaft am Mitochondrium zu betrachten:

Fenster Größe 51:



Fenster Größe 201:



Betrachtet man alle Plots des Mitochondriums fällt auf, dass sich effektiv nur für die Plots der Fenstergrößen 201 und 401 Aussagen zu CpG Islands treffen lassen, die kleinere Fenstergröße 51 gleicht eher einem Hintergrundrauschen. Am Effektivsten wäre ein Scan mit einer Fenstergröße die genau der gesuchten Island Größe entspricht. In einem Plot würde sich dies vermutlich als schmale und hohe Spitze zeigen, eine zu kleine Fenstergröße sorgt vermutlich eher für einen Plateaubereich, während ein zu großes Fenster den Ausschlag ebenfalls dämpft. So zeigt sich der Mitochondrien Plot im Bereich bei ca. 1000 Basenpaaren mit einer fast perfekten spitze (Fenster 401) während der Plot mit der halben Fenstergröße (201) eine abgeflachte und rundere Spitze an der gleichen Stelle besitzt. Mit vielen Scans unterschiedlicher Fenstergrößen könnte man die wahrscheinlichste Islandgröße also theoretisch annähern. Auch der höchste Ausschlag der Spitzen mit verschiedenen Fenstergrößen könnte bei der Bewertung der tatsächlichen CpG Island Größe hilfreich sein. Der meiner Meinung nach eindeutigste Plot mit der Fenstergröße 401 zeigt beim Mitochondrium auch den größten Ausschlag mit einem log likelihood ratio von fast 70.