

INFORME FINAL - GRUPO 8

Algoritmo CLONALG para Cobertura Óptima de Epítopos

Objetivo del Proyecto

Implementar el algoritmo de selección clonal (CLONALG) para resolver el problema de cobertura óptima de epítopos en un antígeno de 600 aminoácidos, simulando los mecanismos de selección del sistema inmune adaptativo.

2. INTRODUCCIÓN

2.1 Contexto Biológico

La identificación de epítopos es fundamental para el desarrollo de vacunas y diagnósticos. El desafío computacional radica en seleccionar un conjunto mínimo de péptidos que cubra exhaustivamente las regiones inmunogénicas de un antígeno, optimizando recursos experimentales. Para implementar CLONALG para cobertura óptima de epítopos, maximizando la cobertura única minimizando solapamientos., validando reproducibilidad con múltiples semillas. Y aso analizar distribución espacial de péptidos seleccionados

2.2 Analogía Biológica: Sistema Inmune Adaptativo

Selección clonal: Proliferación de linfocitos con alta afinidad. Hipermutación somática: Generación de diversidad y memoria inmunológica: Conservación de soluciones efectivas.

3. METODOLOGÍA

3.1 Datos y Configuración

Parámetro	Valor	Descripción
Antígeno	600 aa	Secuencia antigen_1
Péptidos (k=10)	591	Ventanas posibles
Población (N)	40	Individuos por generación
Iteraciones (T)	150	Ciclos de evolución
Semilla	6	Reproducibilidad

Parámetro	Valor	Descripción
Antígeno	600 aa	Secuencia antigen_1
Péptidos (k=10)	591	Ventanas posibles de longitud 10
Población (N)	40	Individuos por generación
Iteraciones (T)	150	Ciclos de evolución del algoritmo
Semilla	6	Control de reproducibilidad

3.2 Función de Fitness

Python def fitness(x, seq=antigen_seq, k=10, gamma=0.1): # Maximiza cobertura única # Penaliza solapamientos redundantes # Incluye regularización ($\gamma=0.1$)

3.3 Evaluación de Reproducibilidad: 3 réplicas independientes con semillas 6, 42, 123. Métricas: Media \pm DE del fitness final. Análisis: Consistencia en patrones de cobertura

3.4 Métricas de Evaluación: Cobertura única: Posiciones con exactamente un péptido. Eficiencia: Cobertura total vs. teórica máxima

Solapamiento: Número de péptidos por posición

Distribución espacial: Cobertura por segmentos

4. RESULTADOS Y ANÁLISIS

4.1 Desempeño del Algoritmo. Convergencia y Estabilidad. Fitness final: 23.40 ± 0.47 . Mejor fitness: 23.40 (semillas 42, 123). Peor fitness: 22.40 (semilla 6)

Estabilización: Iteración ≈ 100

Reproducibilidad

Baja variabilidad entre réplicas (DE = 0.47)

4.2 Análisis de la Mejor Solución

Eficiencia de Cobertura

Métrica Valor Interpretación. Péptidos seleccionados 4.96 0.84% del total

Cobertura única 24/600 (4.0%) Posiciones únicas. Cobertura total 30 posiciones incluye solapamientos. Eficiencia 80.0%. Cobertura por péptido, Solapamiento máximo 2 péptidos.

Distribución Espacial: Segmento 1 (0-120): 100% de cobertura (4.1% del segmento). Segmentos 2-5: 0% de cobertura. Brecha más larga: 576 posiciones. Porcentaje en brechas: 96.0% del antígeno

4.3 Comparación con Enfoques Alternativos

Análisis Comparativo

Métrica	CLONALG	Aleatorio	Máximo Teórico
Fitness	23.40	368.70	69.00
Cobertura	4.0%	99.0%	99.7%
Péptidos usados	5	288	591
Eficiencia relativa	33.9%	-93.7%	100%

Efectividad Operativa

Reducción de péptidos: 99.2% vs solución completa. Eficiencia vs máximo: 33.9% del potencial teórico. Optimización: -93.7% vs solución aleatoria

5. DISCUSIÓN Y PERSPECTIVAS

5.1 Interpretación Biológica

Efectividad en Selección

El algoritmo demostró capacidad para identificar subconjuntos mínimos de péptidos, reflejando la eficiencia del sistema inmune real en focalizar recursos en regiones relevantes mediante selección clonal.

La concentración en el segmento inicial sugiere: Posible sesgo algorítmico hacia regiones tempranas. Las características estructurales que favorecen estas regiones Y una necesidad de mecanismos de diversificación espacial.

5.2 Limitaciones Identificadas: Técnicas: Fitness simplificado: Solo cobertura geométrica. Longitud fija: $k=10$ no refleja diversidad biológica. Biológicas: No considera afinidad de unión a MHC, ignora propiedades fisicoquímicas

6. CONCLUSIONES

Se demostró capacidad para seleccionar conjuntos mínimos de péptidos con cobertura específica, optimizando el uso de recursos. Hay una alta reproducibilidad entre diferentes inicializaciones valida la confiabilidad del enfoque. El método establece bases sólidas para aplicaciones en diseño racional de herramientas inmunológicas.