Herramienta web para la integración de sistemas de predicciones inmunoinformáticas (REVXINE)

Juan Luis Salazar Mendoza

Proyecto de residencias profesionales

ASESORES

Dr. Marco Antonio Castro Liera (ITLP)

Dr. Carlos Eliud Angulo Valadez (CIBNOR)

La Paz B.C.S, miércoles 4 de junio del 2014

Contenido

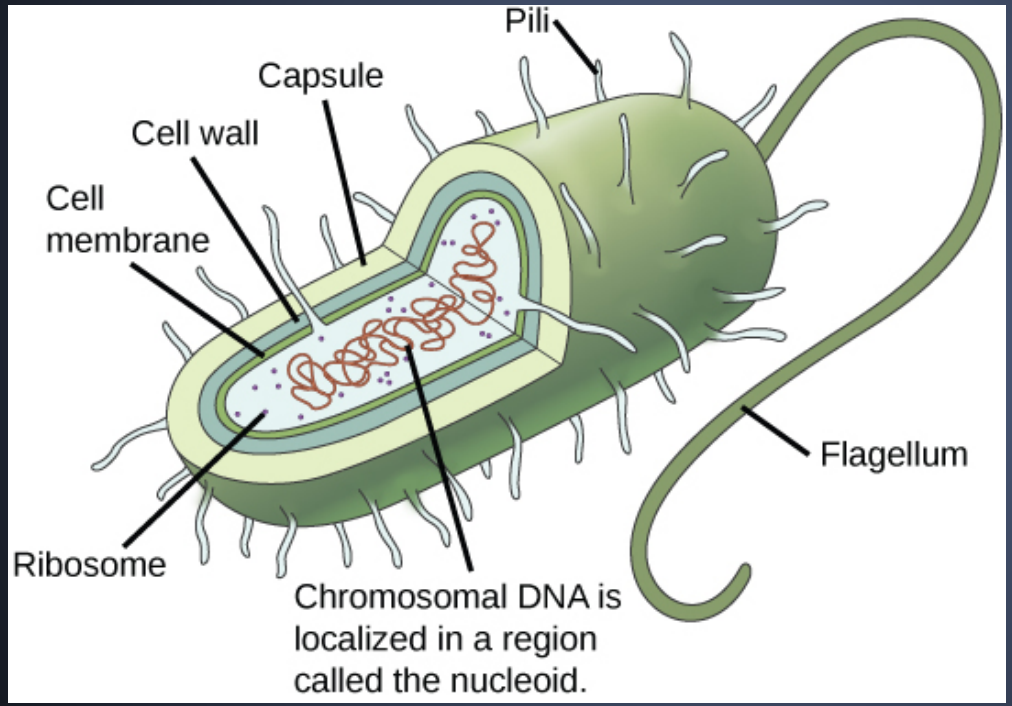
1. Introducción
2. Antecedentes
3. Problemas
4. Posibles soluciones
5. Justificación
6. Objetivo general y particulares
7. Materiales y métodos
8. Resultados y manejo de la herramienta
9. Conclusiones
10. Perspectivas
11. Agradecimientos

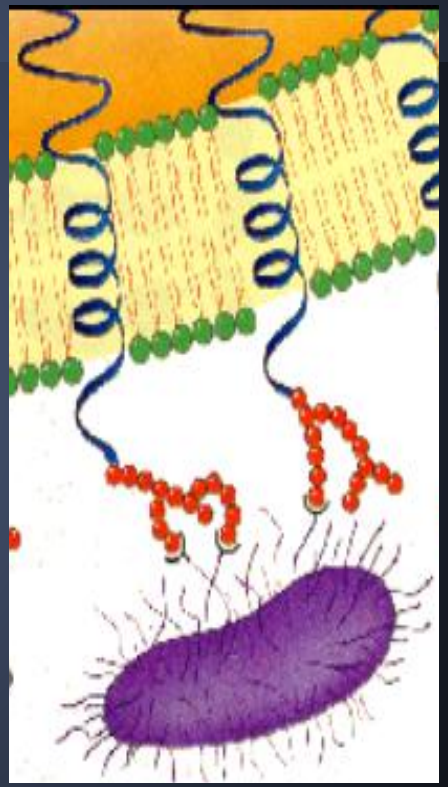
Introducción

- **Las vacunas tradicionales**
  - Empleando patógenos muertos

- **Finalidad de las vacunas:**
  - Respuesta para eliminar al patógeno.
  - Memoria inmune adaptativa.
  - Prevenir reinfecciones.

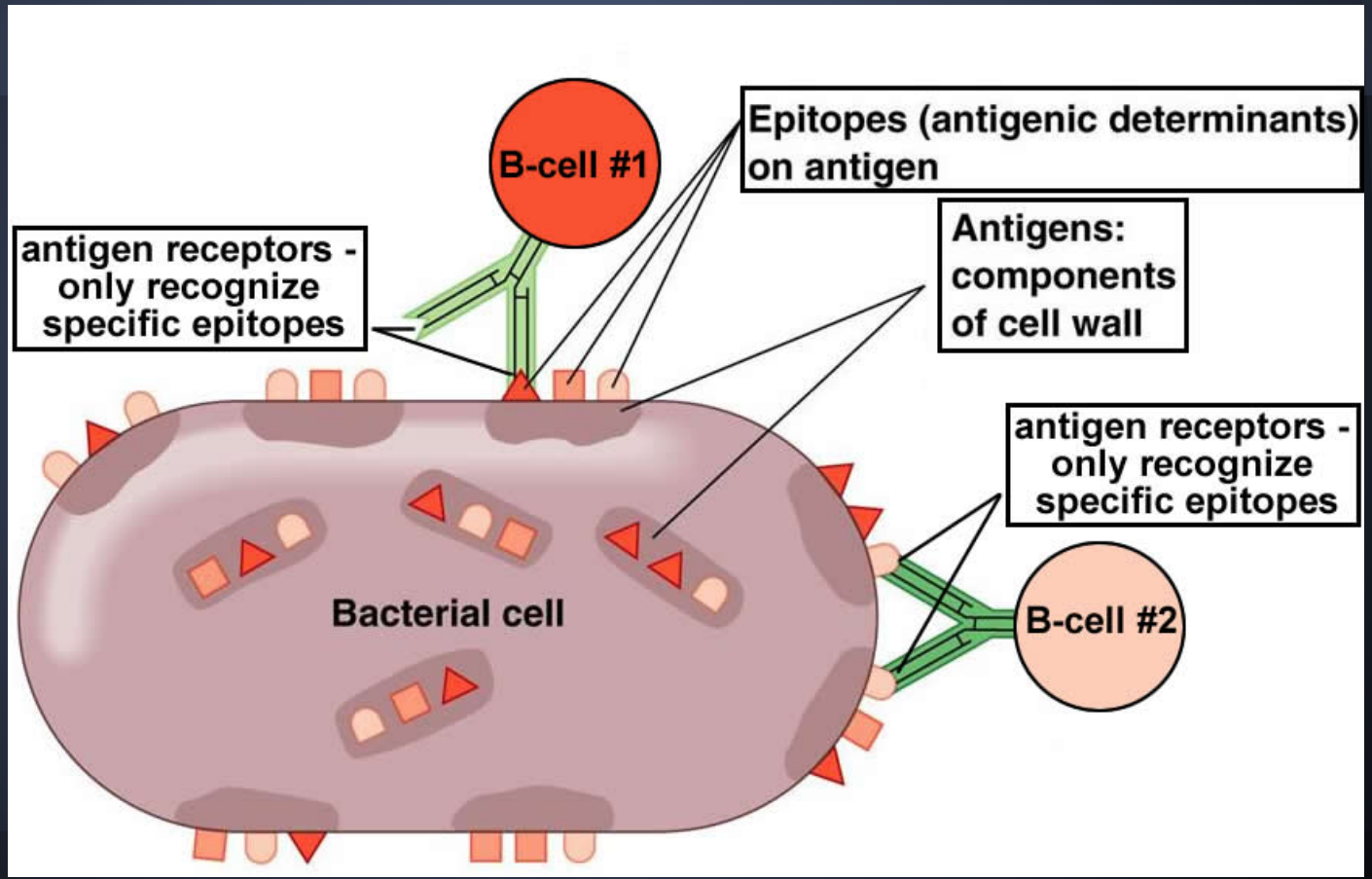
Patógenos: Patrones moleculares asociados a patógenos (PAMP´s).



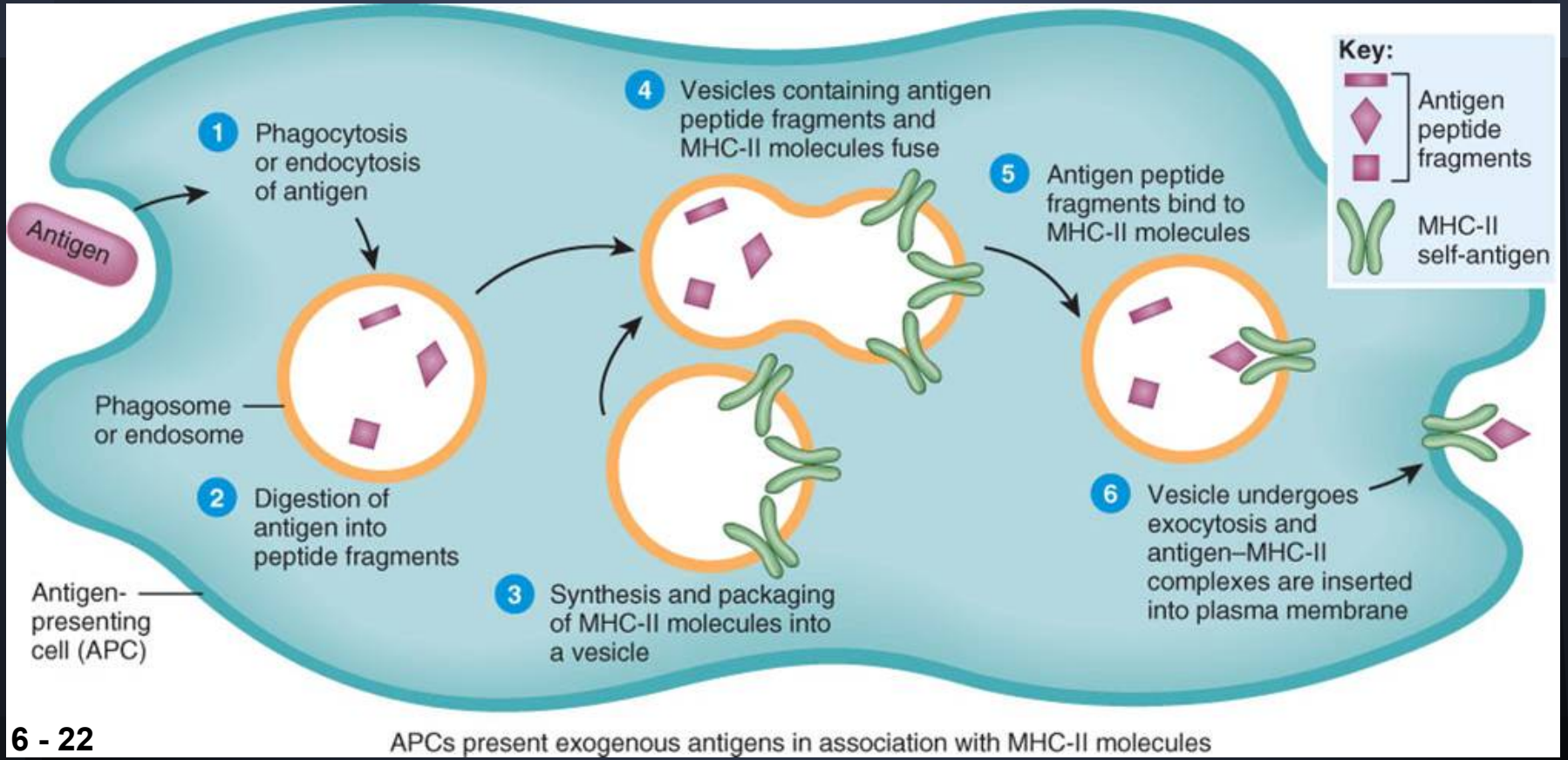


Proteínas en superficie de bacterias patógenas

Epitopes

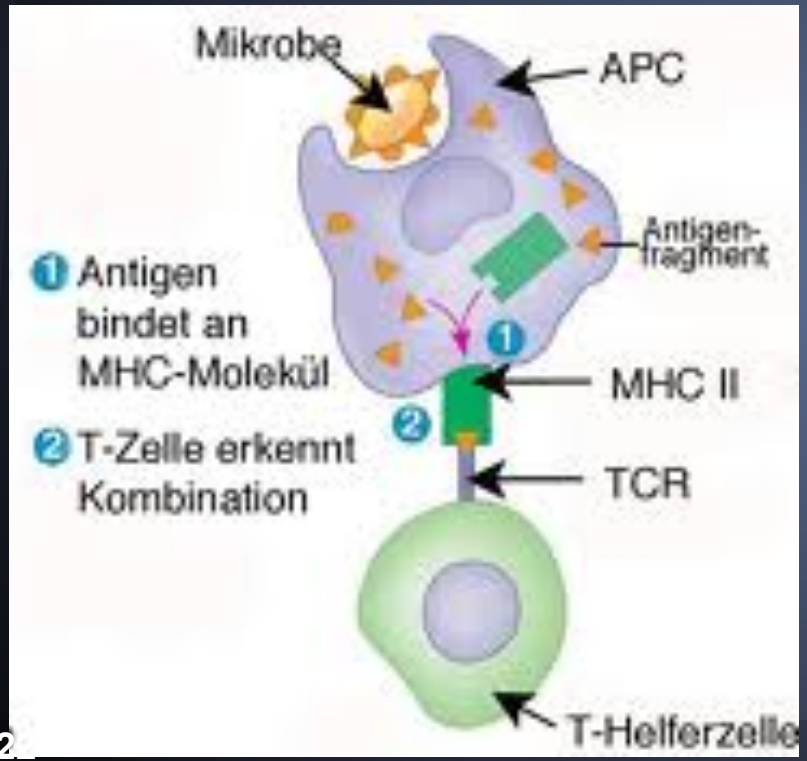

B-cell #1
Epitopes (antigenic determinants) on antigen
antigen receptors - only recognize specific epitopes
Antigens: components of cell wall
antigen receptors - only recognize specific epitopes
Bacterial cell
B-cell #2

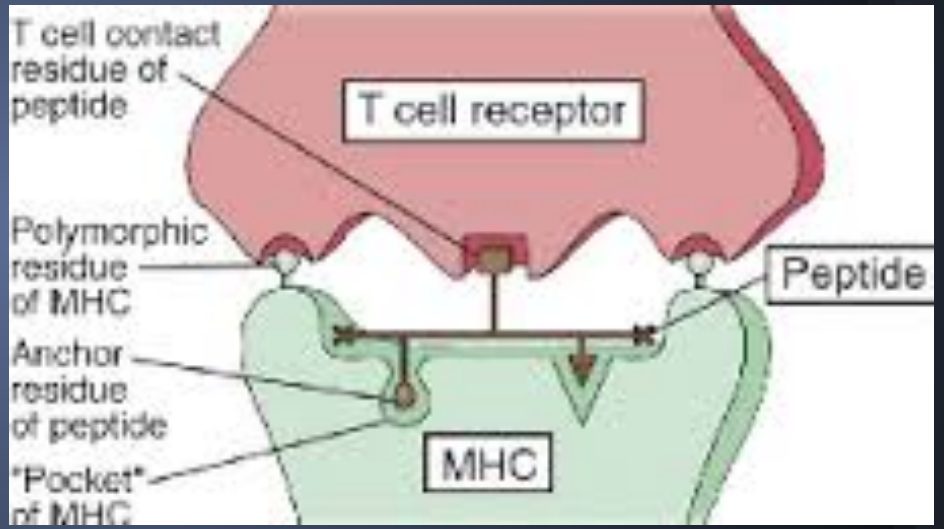
Procesamiento y presentación de epitopes



APCs present exogenous antigens in association with MHC-II molecules

Respuesta inmunológica

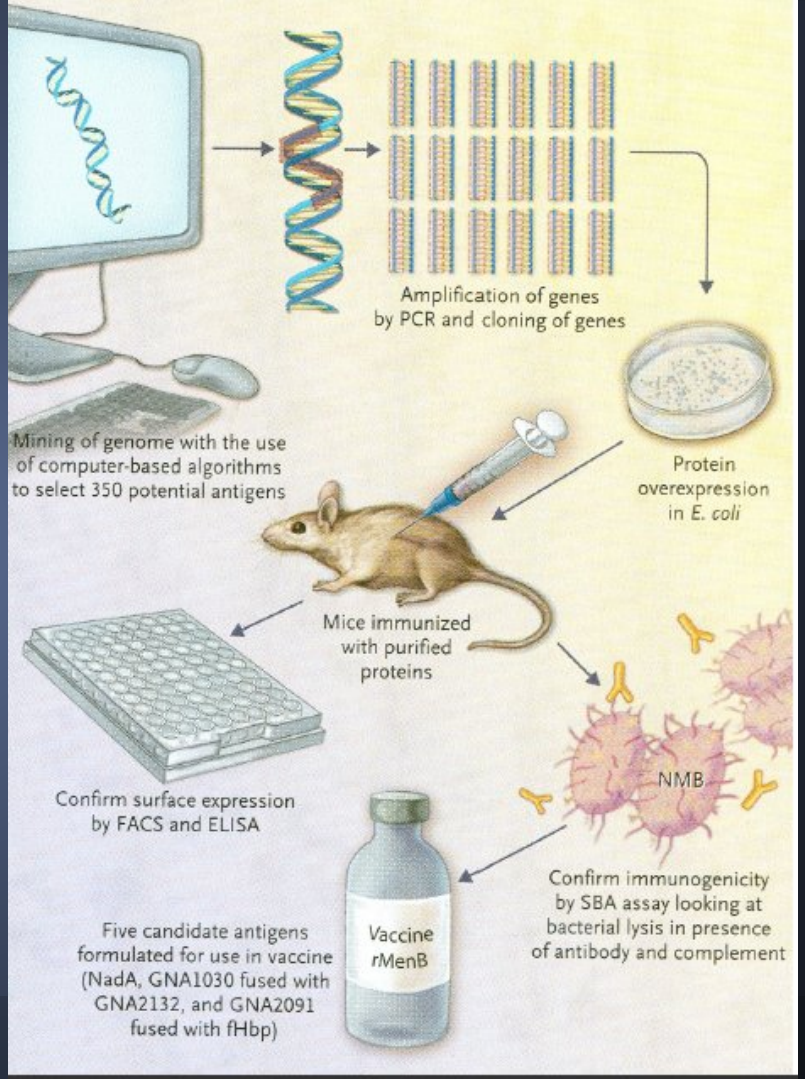




Antecedentes

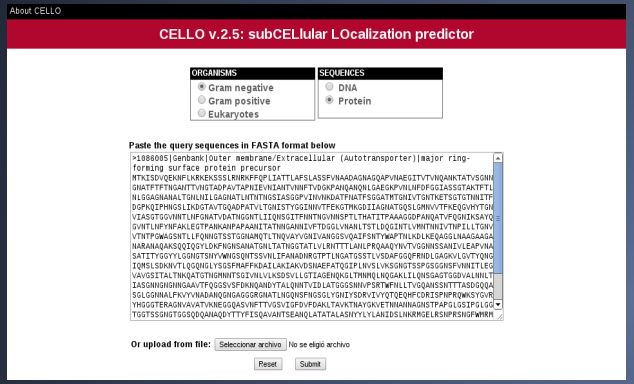
- Estrategias computacionales
  - Extensas bases de datos
  - Modelado matemático

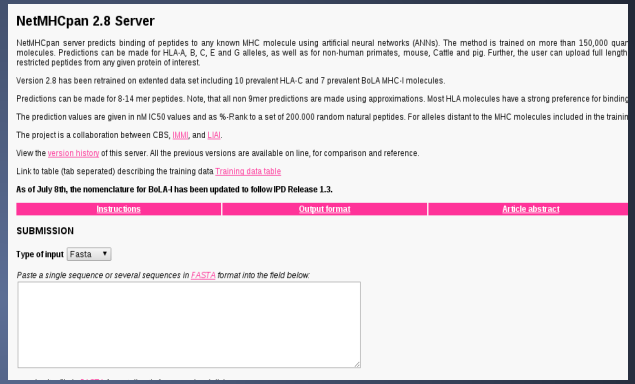
- Herramientas inmunoinformáticas
  - Localización subcelular
  - Unión de péptidos a moléculas MHCI
  - Unión de péptidos a moléculas MHCII
  - Localización de epitopos de Células B

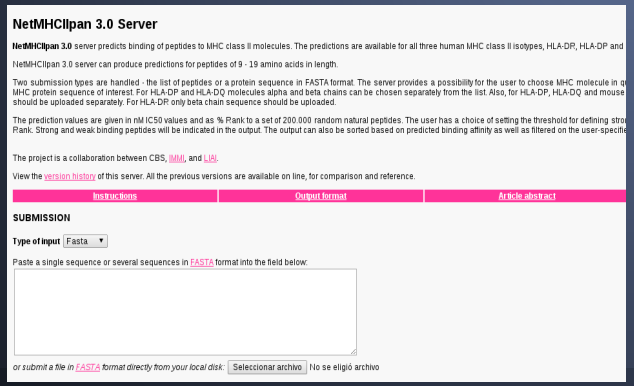


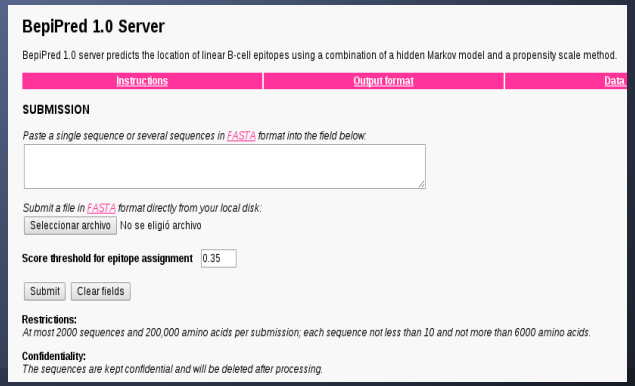
Problemas

Análisis independientes









Problemas

Análisis de resultados laborioso

NetMHCI

```
# NetMHCpan version 2.8

# Input is in FSA format

# Peptide length 9,10

HLA-A01:01 : Estimated prediction accuracy  0.853 (using nearest neighbor HLA-A01:01)

# Affinity Threshold for Strong binding peptides  50.000
# Affinity Threshold for Weak binding peptides 500.000
# Rank Threshold for Strong binding peptides   0.500
# Rank Threshold for Weak binding peptides   2.000
---------------------------------------------------------------------------------
  pos          HLA       peptide        Identity 1-log50k(aff) Affinity(nM)    %Rank  BindLevel
---------------------------------------------------------------------------------
  114   HLA-A*01:01    QSSFYADWY        MAP1609C          0.539       146.17     0.17 <= SB
   93   HLA-A*01:01    NTPAFEWYY        MAP1609C          0.504       215.03     0.25 <= SB
  274   HLA-A*01:01   SSNLKFQDAY        MAP1609C          0.502       217.75     0.25 <= SB
   40   HLA-A*01:01   FSRPGLPVEY        MAP1609C          0.351      1120.65     0.80 <= WB
  139   HLA-A*01:01    FLTSELPSY        MAP1609C          0.325      1491.37     0.80 <= WB
    0   HLA-A*01:01   MTDLSEKVRA        MAP1609C          0.323      1524.49     1.00 <= WB
  296   HLA-A*01:01   NANGTHSWEY        MAP1609C          0.318      1594.56     1.00 <= WB
  175   HLA-A*01:01   AVNHPDQFIY        MAP1609C          0.301      1920.82     1.00 <= WB
  140   HLA-A*01:01   LTSELPSYLA        MAP1609C          0.257      3106.04     1.50 <= WB
  140   HLA-A*01:01    LTSELPSYL        MAP1609C          0.246      3504.76     2.00 <= WB
  205   HLA-A*01:01    LAMGDAGGY        MAP1609C          0.233      4008.24     2.00 <= WB
  141   HLA-A*01:01    TSELPSYLA        MAP1609C          0.230      4157.81     2.00 <= WB
  204   HLA-A*01:01   GLAMGDAGGY        MAP1609C          0.216      4841.76     3.00
  110   HLA-A*01:01    PVGGQSSFY        MAP1609C          0.205      5450.75     3.00
   91   HLA-A*01:01   DINTPAFEWY        MAP1609C          0.193      6168.38     3.00
  118   HLA-A*01:01    YADWYQPAC        MAP1609C          0.190      6394.88     3.00
  113   HLA-A*01:01   GQSSFYADWY        MAP1609C          0.180      7146.84     3.00
   92   HLA-A*01:01   INTPAFEWYY        MAP1609C          0.179      7241.56     3.00
  297   HLA-A*01:01    ANGTHSWEY        MAP1609C          0.178      7288.73     4.00
  166   HLA-A*01:01   MSGSSAMILA        MAP1609C          0.178      7304.02     4.00
  241   HLA-A*01:01    GHNTRLWVY        MAP1609C          0.176      7437.22     4.00
  118   HLA-A*01:01   YADWYQPACG        MAP1609C          0.176      7437.83     4.00
    0   HLA-A*01:01    MTDLSEKVR        MAP1609C          0.171      7824.19     4.00
  164   HLA-A*01:01    ISMSGSSAM        MAP1609C          0.170      7978.85     4.00
  275   HLA-A*01:01    SNLKFQDAY        MAP1609C          0.169      8027.68     4.00
  176   HLA-A*01:01    VNHPDQFIY        MAP1609C          0.167      8234.95     4.00
  168   HLA-A*01:01     GSSAMILAV       MAP1609C          0.166      8260.57     4.00
```

NetMHCII

```
# NetMHCIIpan version 3.0

# Input is in FASTA format

# Peptide length 15

# Threshold for Strong binding peptides (IC50)  50.000 nM
# Threshold for Weak binding peptides (IC50)    500.000 nM

# Threshold for Strong binding peptides (%Rank) 0.5%
# Threshold for Weak binding peptides (%Rank)   2%



# Allele: HLA-DPA1*01:03-DPB1*01:01
--------------------------------------------------------------------------------------------------------------------
  pos              Allele           peptide    Identity  Pos        Core 1-log50k(aff)  Affinity(nM)  %Rank   BindingLevel
--------------------------------------------------------------------------------------------------------------------
  131 HLA-DPA1*01:03-DPB1*01:01   CSTYKWETFLTSELP   MAP1609C_    3    YKWETFLTS         0.783         10.45   0.50  <=SB
  130 HLA-DPA1*01:03-DPB1*01:01   GCSTYKWETFLTSEL   MAP1609C_    4    YKWETFLTS         0.777         11.14   0.80  <=SB
  132 HLA-DPA1*01:03-DPB1*01:01   STYKWETFLTSELPS   MAP1609C_    2    YKWETFLTS         0.760         13.41   0.80  <=SB
  129 HLA-DPA1*01:03-DPB1*01:01   AGCSTYKWETFLTSE   MAP1609C_    5    YKWETFLTS         0.747         15.48   1.50  <=SB
  128 HLA-DPA1*01:03-DPB1*01:01   KAGCSTYKWETFLTS   MAP1609C_    6    YKWETFLTS         0.740         16.63   1.50  <=SB
  133 HLA-DPA1*01:03-DPB1*01:01   TYKWETFLTSELPSY   MAP1609C_    1    YKWETFLTS         0.730         18.47   1.50  <=SB
  134 HLA-DPA1*01:03-DPB1*01:01   YKWETFLTSELPSYL   MAP1609C_    0    YKWETFLTS         0.711         22.69   3.00  <=SB
  135 HLA-DPA1*01:03-DPB1*01:01   KWETFLTSELPSYLA   MAP1609C_    4    FLTSELPSY         0.669         35.90   5.00  <=SB
  136 HLA-DPA1*01:03-DPB1*01:01   WETFLTSELPSYLAS   MAP1609C_    3    FLTSELPSY         0.664         37.72   5.00  <=SB
   95 HLA-DPA1*01:03-DPB1*01:01   PAFEWYYQSGLSVIM   MAP1609C_    2    FEWYYQSGL         0.656         41.52   5.00  <=SB
   96 HLA-DPA1*01:03-DPB1*01:01   AFEWYYQSGLSVIMP   MAP1609C_    1    FEWYYQSGL         0.652         43.15   6.00  <=SB
   94 HLA-DPA1*01:03-DPB1*01:01   TPAFEWYYQSGLSVI   MAP1609C_    3    FEWYYQSGL         0.649         44.46   6.00  <=SB
   97 HLA-DPA1*01:03-DPB1*01:01   FEWYYQSGLSVIMPV   MAP1609C_    0    FEWYYQSGL         0.636         51.52   7.00  <=WB
  137 HLA-DPA1*01:03-DPB1*01:01   ETFLTSELPSYLASN   MAP1609C_    2    FLTSELPSY         0.626         57.08   8.00  <=WB
  179 HLA-DPA1*01:03-DPB1*01:01   PDQFIYAGSLSALLD   MAP1609C_    5    YAGSLSALL         0.608         69.60   9.00  <=WB
  127 HLA-DPA1*01:03-DPB1*01:01   GKAGCSTYKWETFLT   MAP1609C_    4    CSTYKWETF         0.608         69.63   9.00  <=WB
  266 HLA-DPA1*01:03-DPB1*01:01   EFLENFVRSSNLKFQ   MAP1609C_    5    FVRSSNLKF         0.607         70.02   9.00  <=WB
  265 HLA-DPA1*01:03-DPB1*01:01   AEFLENFVRSSNLKF   MAP1609C_    6    FVRSSNLKF         0.607         69.92   9.00  <=WB
  178 HLA-DPA1*01:03-DPB1*01:01   HPDQFIYAGSLSALL   MAP1609C_    3    QFIYAGSLS         0.606         70.76   9.00  <=WB
  180 HLA-DPA1*01:03-DPB1*01:01   DQFIYAGSLSALLDP   MAP1609C_    4    YAGSLSALL         0.604         72.43   9.00  <=WB
   93 HLA-DPA1*01:03-DPB1*01:01   NTPAFEWYYQSGLSV   MAP1609C_    4    FEWYYQSGL         0.599         76.34  10.00  <=WB
  181 HLA-DPA1*01:03-DPB1*01:01   QFIYAGSLSALLDPS   MAP1609C_    3    YAGSLSALL         0.593         81.73  15.00  <=WB
  267 HLA-DPA1*01:03-DPB1*01:01   FLENFVRSSNLKFQD   MAP1609C_    4    FVRSSNLKF         0.591         83.55  15.00  <=WB
  268 HLA-DPA1*01:03-DPB1*01:01   LENFVRSSNLKFQDA   MAP1609C_    3    FVRSSNLKF         0.587         87.40  15.00  <=WB
    4 HLA-DPA1*01:03-DPB1*01:01   SEKVRAWGRRLLVGA   MAP1609C_    3    VRAWGRRLL         0.583         90.90  15.00  <=WB
    3 HLA-DPA1*01:03-DPB1*01:01   LSEKVRAWGRRLLVG   MAP1609C_    4    VRAWGRRLL         0.576         98.19  15.00  <=WB
  138 HLA-DPA1*01:03-DPB1*01:01   TFLTSELPSYLASNK   MAP1609C_    1    FLTSELPSY         0.574         99.97  15.00  <=WB
    2 HLA-DPA1*01:03-DPB1*01:01   DLSEKVRAWGRRLLV   MAP1609C_    5    VRAWGRRLL         0.570        104.36  15.00  <=WB
    5 HLA-DPA1*01:03-DPB1*01:01   EKVRAWGRRLLVGAA   MAP1609C_    2    VRAWGRRLL         0.564        111.68  15.00  <=WB
  114 HLA-DPA1*01:03-DPB1*01:01   QSSFYADWYQPACGK   MAP1609C_    3    FYADWYQPA         0.561        115.01  15.00  <=WB
  113 HLA-DPA1*01:03-DPB1*01:01   GQSSFYADWYQPACG   MAP1609C_    4    FYADWYQPA         0.557        120.98  15.00  <=WB
  263 HLA-DPA1*01:03-DPB1*01:01   MPAEFLENFVRSSNL   MAP1609C_    3    EFLENFVRS         0.555        123.26  15.00  <=WB
  112 HLA-DPA1*01:03-DPB1*01:01   GGQSSFYADWYQPAC   MAP1609C_    5    FYADWYQPA         0.552        126.85  15.00  <=WB
   98 HLA-DPA1*01:03-DPB1*01:01   EWYYQSGLSVIMPVG   MAP1609C_    1    WYYQSGLSV         0.551        129.11  15.00  <=WB
  264 HLA-DPA1*01:03-DPB1*01:01   PAEFLENFVRSSNLK   MAP1609C_    2    EFLENFVRS         0.540        144.80  32.00  <=WB
    6 HLA-DPA1*01:03-DPB1*01:01   KVRAWGRRLLVGAAA   MAP1609C_    1    VRAWGRRLL         0.537        149.30  32.00  <=WB
  111 HLA-DPA1*01:03-DPB1*01:01   VGGQSSFYADWYQPA   MAP1609C_    6    FYADWYQPA         0.536        151.14  32.00  <=WB
  261 HLA-DPA1*01:03-DPB1*01:01   ANMPAEFLENFVRSS   MAP1609C_    4    AEFLENFVR         0.534        154.82  32.00  <=WB
  269 HLA-DPA1*01:03-DPB1*01:01   ENFVRSSNLKFQDAY   MAP1609C_    2    FVRSSNLKF         0.533        155.81  32.00  <=WB
   92 HLA-DPA1*01:03-DPB1*01:01   INTPAFEWYYQSGLS   MAP1609C_    5    FEWYYQSGL         0.530        161.27  32.00  <=WB
  262 HLA-DPA1*01:03-DPB1*01:01   NMPAEFLENFVRSSN   MAP1609C_    3    AEFLENFVR         0.529        163.60  32.00  <=WB
  115 HLA-DPA1*01:03-DPB1*01:01   SSFYADWYQPACGKA   MAP1609C_    2    FYADWYQPA         0.521        177.50  32.00  <=WB
   91 HLA-DPA1*01:03-DPB1*01:01   DINTPAFEWYYQSGL   MAP1609C_    6    FEWYYQSGL         0.511        198.48  32.00  <=WB
  260 HLA-DPA1*01:03-DPB1*01:01   GANMPAEFLENFVRS   MAP1609C_    5    AEFLENFVR         0.510        199.76  32.00  <=WB
  182 HLA-DPA1*01:03-DPB1*01:01   FIYAGSLSALLDPSQ   MAP1609C_    2    YAGSLSALL         0.506        209.48  32.00  <=WB
    1 HLA-DPA1*01:03-DPB1*01:01   TDLSEKVRAWGRRLL   MAP1609C_    6    VRAWGRRLL         0.506        208.72  32.00  <=WB
  177 HLA-DPA1*01:03-DPB1*01:01   NHPDQFIYAGSLSAL   MAP1609C_    4    QFIYAGSLS         0.502        217.65  32.00  <=WB
   99 HLA-DPA1*01:03-DPB1*01:01   WYYQSGLSVIMPVGG   MAP1609C_    0    WYYQSGLSV         0.502        219.05  32.00  <=WB
   73 HLA-DPA1*01:03-DPB1*01:01   PAVYLLDGLRAQDDY   MAP1609C_    3    YLLDGLRAQ         0.497        232.06  32.00  <=WB
   72 HLA-DPA1*01:03-DPB1*01:01   SPAVYLLDGLRAQDD   MAP1609C_    3    VYLLDGLRA         0.490        249.62  32.00  <=WB
  270 HLA-DPA1*01:03-DPB1*01:01   NFVRSSNLKFQDAYN   MAP1609C_    1    FVRSSNLKF         0.489        251.06  32.00  <=WB
   71 HLA-DPA1*01:03-DPB1*01:01   GSPAVYLLDGLRAQD   MAP1609C_    4    VYLLDGLRA         0.481        273.53  32.00  <=WB
   74 HLA-DPA1*01:03-DPB1*01:01   AVYLLDGLRAQDDYN   MAP1609C_    2    YLLDGLRAQ         0.477        285.97  32.00  <=WB
  300 HLA-DPA1*01:03-DPB1*01:01   THSWEYWGAQLNAMK   MAP1609C_    3    WEYWGAQLN         0.475        292.22  32.00  <=WB
  126 HLA-DPA1*01:03-DPB1*01:01   CGKAGCSTYKWETFL   MAP1609C_    5    CSTYKWETF         0.475        293.67  32.00  <=WB
   70 HLA-DPA1*01:03-DPB1*01:01   NGSPAVYLLDGLRAQ   MAP1609C_    5    VYLLDGLRA         0.472        302.23  32.00  <=WB
  301 HLA-DPA1*01:03-DPB1*01:01   HSWEYWGAQLNAMKP   MAP1609C_    2    WEYWGAQLN         0.464        330.30  32.00  <=WB
  139 HLA-DPA1*01:03-DPB1*01:01   FLTSELPSYLASNKG   MAP1609C_    0    FLTSELPSY         0.463        335.32  32.00  <=WB
    7 HLA-DPA1*01:03-DPB1*01:01   VRAWGRRLLVGAAAA   MAP1609C_    0    VRAWGRRLL         0.459        350.02  32.00  <=WB
  259 HLA-DPA1*01:03-DPB1*01:01   GGANMPAEFLENFVR   MAP1609C_    6    AEFLENFVR         0.455        365.82  32.00  <=WB
  299 HLA-DPA1*01:03-DPB1*01:01   GTHSWEYWGAQLNAM   MAP1609C_    4    WEYWGAQLN         0.452        377.19  32.00  <=WB
  302 HLA-DPA1*01:03-DPB1*01:01   SWEYWGAQLNAMKPD   MAP1609C_    3    YWGAQLNAM         0.441        421.23  32.00  <=WB
   75 HLA-DPA1*01:03-DPB1*01:01   VYLLDGLRAQDDYNG   MAP1609C_    1    YLLDGLRAQ         0.436        447.88  50.00  <=WB
  271 HLA-DPA1*01:03-DPB1*01:01   FVRSSNLKFQDAYNA   MAP1609C_    0    FVRSSNLKF         0.435        454.01  50.00  <=WB
  116 HLA-DPA1*01:03-DPB1*01:01   SFYADWYQPACGKAG   MAP1609C_    1    FYADWYQPA         0.431        470.10  50.00  <=WB
  183 HLA-DPA1*01:03-DPB1*01:01   IYAGSLSALLDPSQG   MAP1609C_    1    YAGSLSALL         0.424        507.96  50.00
   41 HLA-DPA1*01:03-DPB1*01:01   SRPGLPVEYLQVPSA   MAP1609C_    3    GLPVEYLQV         0.419        538.96  50.00
  258 HLA-DPA1*01:03-DPB1*01:01   LGGANMPAEFLENFV   MAP1609C_    3    ANMPAEFLE         0.411        587.88  50.00
```

Tabla de traslape

Tabla. Transposición epitopos de células T presentados por MHC I y MHC II.

Proteína	MHC I Epítopo Cel. T				MHC I Epítopo Cel. T			
	Secuencia	Posición	Allelo	Unión	Secuencia	Posición	Allelo	Unión
MAP1609C								
	DQFIYAGSL	180-188	BoLA-T2c	WB	HPDQFIYAGSLSALL	178-192	H-2-IAb	SB
	LLVGAAAAV	14-22	BoLA-T2c	WB	GRRLLVGAAAAVTLP	11-25	H-2-IAb	SB
	YQSGLSVIM	101-109	BoLA-T2c	WB	FEWYYQSGLSVIMPV	97-111	H-2-IAb	WB
	NAAGGHNAV	284-292	BoLA-T2c	WB	FQDAYNAAGGHNAVF	279-293	H-2-IAb	SB
	SMSGSSAMI	165-173	BoLA-T2c	WB	AVGISMSGSSAMILA	161-175	H-2-IAb	WB

Posibles soluciones

- Integración en una sola plataforma
- Resultados en formatos homogéneos y científicos

Justificación

- El presente trabajo aportaría una solución informática al área de vacunología.
- Aumentaría la velocidad y facilidad de selección de epitopes candidatos a vacunas.
- Evitaría los procesos de edición para formatos de presentación de resultados en tablas.

Objetivo general

Desarrollar una aplicación web integradora que permita predecir la ubicación subcelular de proteínas de patógenos y los péptidos de unión a células T y B, cuyos resultados se despliegen en formatos de artículos científicos.

Objetivo particulares

1. Analizar los mecanismos computacionales de las herramientas inmunoformáticas seleccionadas.
2. Construir una sola interfaz web integradora de las herramientas inmunoinformáticas.
3. Programar el análisis e interpretación de los datos.
4. Diseñar y desplegar los resultados en formatos de artículos científicos.

Materiales y métodos

- Materiales
  - Linux, IDE Eclipse, lenguajes; PHP, HTML y JavaScript.
  - Libreria cURL.

- Métodos
  - Se programó la interfaz de usuario y se aseguró la integración de cuatro herramientas inmunoinformáticas disponibles en la web:
    - CELLO (**cello**.life.nctu.edu.tw)
    - NetMHCI (www.cbs.dtu.dk/services/**NetMHCII**/)
    - NetMHCII (www.cbs.dtu.dk/services/**NetMHCII**/)
    - BepiPred (http://www.cbs.dtu.dk/services/BepiPred/)

Materiales y métodos

Métodos

- Se programó un intérprete homogéneo en PHP.

  - Expresiones regulares

- Se diseñaron las tablas de resultado en HTML.

- Se programó la creación automática de las tablas utilizando PHP.

Resultados

REVXINE es una herramienta informática para dar soporte a la vacunología

http://posgrado.itlp.edu.mx/revxine/index.php

Conclusiones

1. Se logró la integración de herramientas inmunoinformáticas.
2. Sus resultados pueden ser interpretados fácilmente con programación.
3. Es posible utilizar las respuestas individuales para armar tablas en formato científico.

Trabajo Futuro

1. Agregar otras herramientas inmunoinformáticas y generar nuevas tablas de reporte.
2. Control de usuarios y registro de experimentos.
3. Descargar los servidores y ejecutarlos de forma local.
4. Utilizar cómputo paralelo para agilizar los resultados.
5. Agregar la espera de resultados de forma interactiva.
6. Agregar una versión de descarga para Windows.

Agradecimientos

- Instituto Tecnológico de La Paz (ITLP)
  - Dr. Marco Antonio Castro Liera
- Centro de Investigaciones Biológicas del Noroeste (CIBNOR)
  - Dr. Carlos Eliud Angulo Valadez
  - Ibq. Perla Carlos Garcia
- Proyecto SEP-CONACYT (151818).
- A la Academia de Ciencias Agrícolas.

REVXINE

https://github.com/jlsalazar/revxine.git