CEDAR
Cancer Epitope Database and Analysis Resource

A meta-analysis of defined epitopes from prostate specific antigens

Opportunities and knowledge gaps

cedar.iedb.org

Presented by: Alessandro Sette, IEDB Principal Investigator

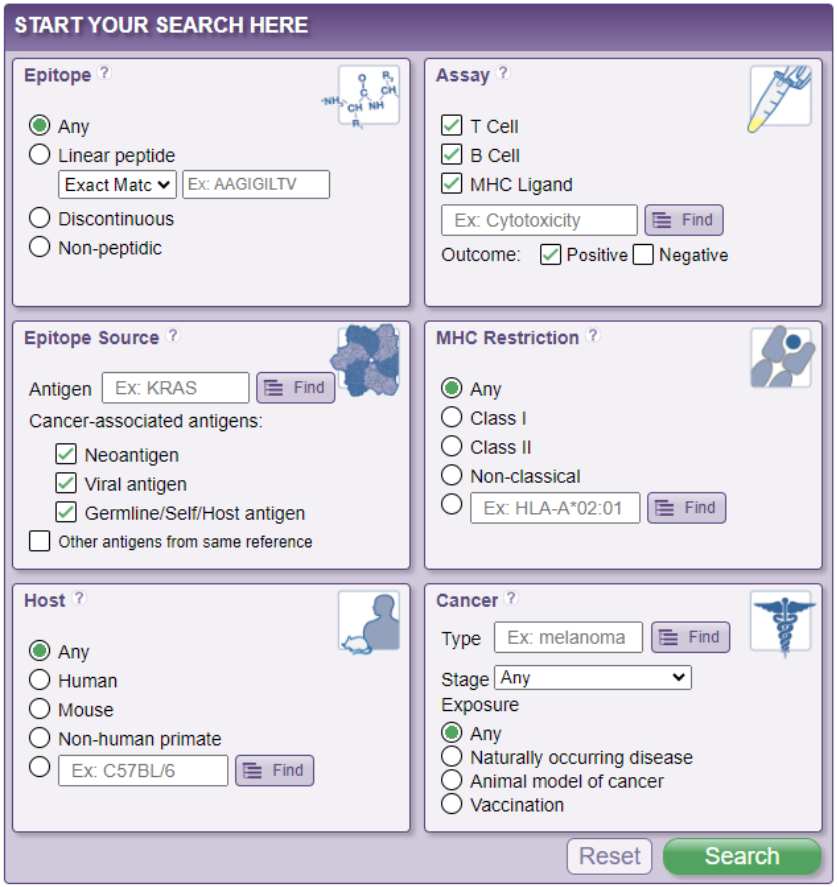
Introduction

- Cancer Epitope Database Analysis Resource (CEDAR):
  - Novel database initiative, to eventually host all cancer-related epitope data (Koşaloğlu-Yalçın et al., Front Immunol. 2021)
  - Initiated in 2021, leveraging experience of nearly 20 years by the sister database IEDB (Vita et al., Nucleic Acids Res. 2019)
- In parallel with database development, we decided to curate two prototype categories:
  - Prostate cancer antigens
  - Neoepitopes
- Orthogonally distinct data sources, one derived from well-defined and non-mutated antigens associated with "cold" tumors, and one associated with highly mutated tumors, targeted by neoepitope vaccines
- Here we present a meta-analysis of prostate associated antigens
- Goals:
  - Illustrate CEDAR capabilities at this early stage and to solicit feedback
  - Identify strengths and knowledge gaps related to the prostate antigens

Definition of reports associated with prostate-specific antigens

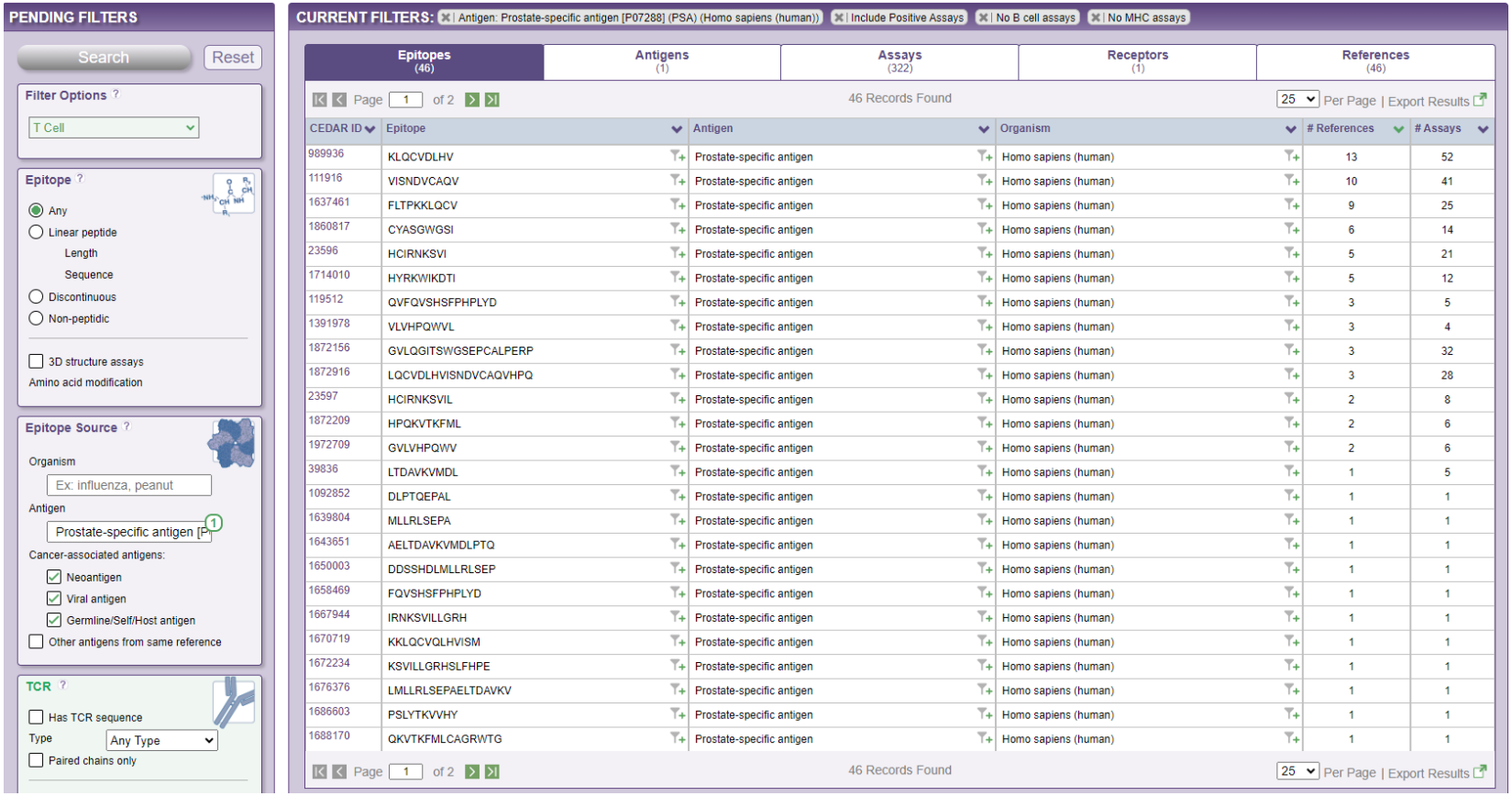
- Text classifier identifies relevant cancer-related papers (Wang, BMC Bioinformatics 2007)
- As of 9/15/22, of a total of 11,442 cancer-related papers, 395 were classified as related to prostate associated antigens
- Processed 96.5% of them; after abstract review 222 were potentially curatable and 165 were found to contain curatable information.
- A manual literature search for prostate-specific antigens identified eight major prostate-specific antigens
- All represented in the curated papers classified as prostate antigens
- Analysis focuses on these 8 prostate-specific antigens, and excludes antigens associated with, but not specific for, prostate cancer

CEDAR Home Page – *cedar.iedb.org*



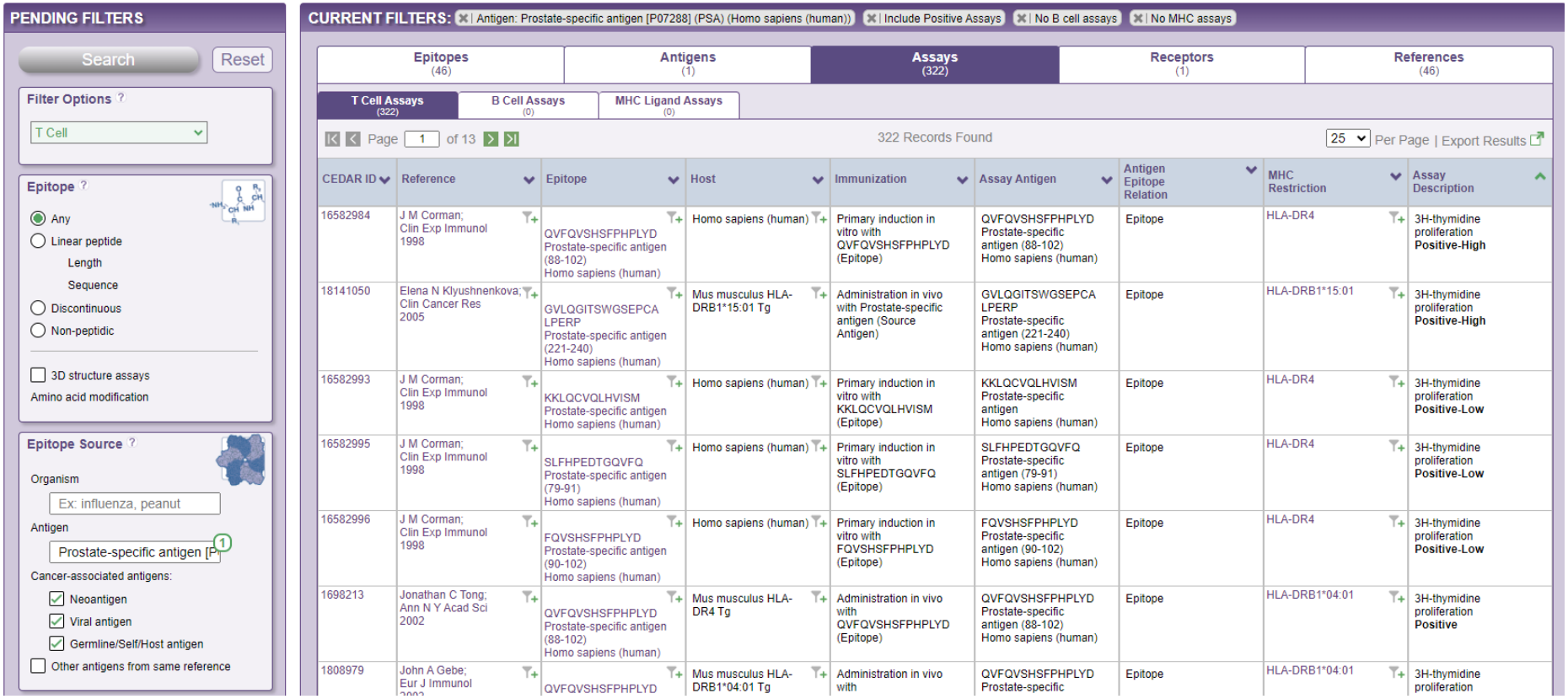
- Six main search panes allow for the selection of specific criteria
- Queries can be performed by epitope(s) of interest (e.g., linear B-cell epitopes from specific protein)
- Queries can be used to narrow down the data to certain subsets of the data (e.g., humans with pancreatic cancer)

CEDAR Result Page
Query: T cell epitopes of prostate antigens



- Five tabs show the number of records for Epitopes, Antigens, Assays, Receptors, and References.
- Result details can be accessed by clicking the relevant tabs.
- Query parameters can be adjusted by changing the search parameters on the left side of the page.

Query: T cell epitopes of prostate antigens



- Nine columns show the most relevant information about the individual assays
- Clicking the IEDB ID shows the detail of entire experimental process
- T-cell assay results can be downloaded in Excel by clicking the icon “export results” in the top right

Inventory of Epitopes from Prostate Antigens

Protein	Abbreviation	Species	Proteome Accession	References[1]	Protein Lengths	Epitopes[2]							Epitopes		
						Total[4]	Host Human	Host Mouse	B Cell	T Cell	T Cell Class I	T Cell Class II	MHC Binding	MHC Elution	Grand Total
Prostatic acid phosphatase	PAP; PAcP	human, mouse, rat	P15309, Q8CE08, A0A0G2K4B4	**30**	**386 381 381**	**67**	**35**	**34**	**11**	**64**	**31**	**33**	**20**	**2**	**73**
Prostate-specific antigen	PSA, KLK3	human	P07288	**44**	**261**	**78**	**39**	**47**	**52**	**51**	**33**	**18**	**22**	**1**	**82**
Prostate-specific membrane antigen	PSMA	human, mouse	Q04609 O35409	**32**	**750 752**	**51**	**39**	**17**	**34**	**34**	**27**	**6**	**23**	**1**	**60**
Prostein (Solute carrier family 45 member 3)	PROS[3]	human, mouse, rat	Q96JT2 Q8K0H7 D3ZPP5	**5**	**553 553 564**	**42**	**42**	**0**	**38**	**4**	**4**	**0**	**6**	**2**	**48**
Prostate stem cell antigen	PSCA	human	O43653	**6**	**114**	**18**	**18**	**1**	**7**	**17**	**17**	**0**	**12**	**2**	**20**
Metalloreductase STEAP1	STEAP	human, mouse	Q9UHE8 Q9CWR7	**8**	**339 339**	**15**	**7**	**10**	**0**	**15**	**13**	**2**	**2**	**5**	**19**
Transient receptor potential cation channel subfamily M member 8	Trp-p8	human, mouse	Q7Z2W7 Q8R4D5	**2**	**1104 1104**	**2**	**1**	**1**	**0**	**2**	**2**	**0**	**5**	**9**	**14**
TCR gamma alternate reading frame protein	TARP	human	A2JGV3.1	**5**	**58**	**5**	**6**	**0**	**0**	**5**	**3**	**2**	**3**	**1**	**6**
Total				**132**		**278**	**187**	**110**	**142**	**192**	**130**	**61**	**93**	**21**	**322**

[1] Does not include papers with MHC assays only

[2] Total includes B cell and T cell epitopes only

[3] PROS is used an abbreviation for Prostein (Solute carrier family 45 member 3)

[4] Papers might describe more than one protein

Data refer to epitope structures associated with positive data

Conclusions I: Inventory of epitopes derived from eight main prostate antigens

- CEDAR contains a total of 322 positive epitopes studied in various assays (each epitope can be studied in more than one context):
  - 192 T cell ssays
  - 142 B cell assays
  - 93 MHC binding assays
  - 21 MHC elution assays
- Number of epitopes vary by protein (highest with PAP, PSA and PSMA):
  - Number of epitopes correlates with number of publications (r_s = 0.85032, p = 0.007)
  - No correlation between protein length and epitopes (r_s = -0.02381, p = 0.955)

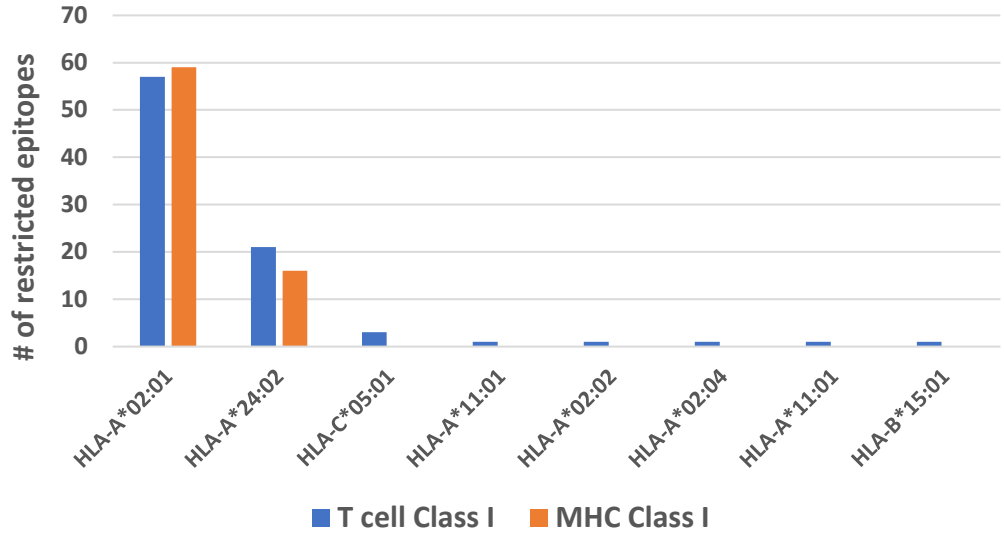
The value for correlation was calculated using https://www.socscistatistics.com/tests/spearman/default2.aspx
Note: the numbers will not add up to 322, some epitopes are measured in T cell and MHC assays.

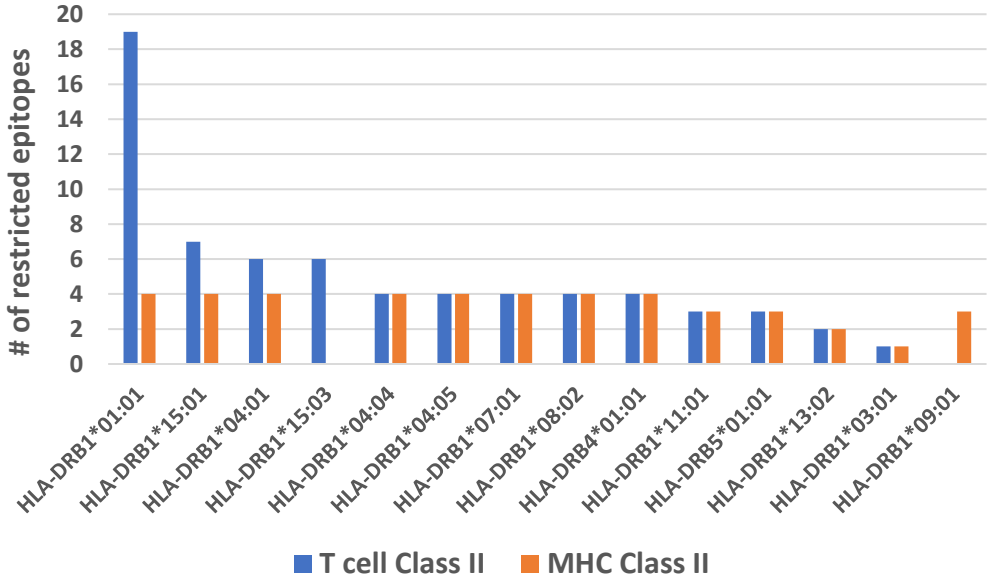
T cell assay types for epitopes from top antigens

	Antigen Name	Parent Protein	Total Tcell Assays	Assays with class I restriction	Assays with class II restriction	Class I restricted epitopes	Class II restricted epitopes
T cell assays[1]	PSA	P07288	314	222	92	30	19
	PSMA	Q04609	115	89	20	25	6
	PAP	P15309	265	146	119	28	33
	PSCA	O43653	49	49	0	17	0
	Other[3]		121	75	45	16	4
Tetramer	PSA	P07288	28	26	2	7	1
	PSMA	Q04609	2	2	0	1	0
	PAP	P15309	5	5	0	3	0
	PSCA	O43653	0	0	0	0	0
	Other		4	4	0	3	0
MHC ligand[2]	PSA	P07288	93	65	28	25	8
	PSMA	Q04609	102	86	16	42	5
	PAP	P15309	134	82	61	41	28
	PSCA	O43653	23	22	1	12	1
	Other		143	80	63	50	59
Total			**1398**	**953**	**447**	**300**	**164**

[1] Includes tetramer assays
[2] Includes MHC binding assays and MHC ligand elution (MHCLE) assays
[3] Includes numbers for PSCA, STEAP, Trp-p8, and TARP

Class I and class II restrictions of epitopes from prostate antigens





Cytokine assays by antigen

Protein	IFNg	TNFa	IL-10	IL-12	IL-13	IL-2	IL-4	IL-5	GrB
PSA	208	9	6	8	2	5	7	6	2
PSMA	101					2			
PAP	274	1							4
PSCA	43								

Note: The total number of assays can be greater than the number of epitopes as there are multiple assays for each cytokine (e.g. ELISA, ELISPOT, ICS, etc.).

Assay types of B cell epitopes

Protein	Monoclonal Response	Polyclonal Response	ELISA/ Western	Antigen Inhibition	SPR	Immuno Staining	IHC	Micro Array	In Vivo Assay	X-ray Crystallo-graphy	Biological Activity
PSA	64	26	40	11	6					5	4
PSMA	27	53	56	3		1	2	2	3		
PAP	2	37	24	3		1		2			
PSCA	0	24	23			1					

Note: The total number of assays can be greater than the number of epitopes

Assays involving recognition/killing of Prostate Tumor Cells

- Number of cytotoxicity assays vary by protein (5 to 24 assays)
- It is possible to search for tumor cell as antigen by exporting a search for cytotoxicity assays

Protein	Number of Cytotoxicity Assays
PSA	24
PSMA	5
PAP	18
PSCA	10

In vivo assays of epitopes from prostate antigens

- Very limited number of in vivo assays, measured by decreased disease, tumor burden, survival
- Most of the assays in xenograft models
- TRAMP is the only genetically engineered mouse model captured in CEDAR
- Possibly reflective of different anatomy of mouse and human prostate

Protein	Decreased Disease	Tumor Burden	Survival
PSA	3		
PSMA	6	2	
PAP	1	3	1
PSCA		1	

Receptor data captured for prostate antigens

TCR sequences

Epitope	Antigen	Antigen	Species	Response Type	MHC Allele	Type	Chain 1 CDR3	Chain 2 CDR3
KLQCVDLHV	PSA	P07288	Homo sapiens	T cell	HLA-A*02:01	αβ	CAVREEDYKLSF	ASSFRGPNLYTEAF

BCR sequences

Epitope	Antigen	Antigen	Species	Response Type	Type	Chain 1 CDR3	Chain 2 CDR3
H98, P99, L100, Y101, D116, H119, T143, K191, K194, F195, H248, Y249, R250, K251, W252, K254	PSA	P07288	Mus musculus	B cell	HL	ARADYGFNSGEAMDY	QQSNEDPYT
I25, W29, E30, C31, E32, K33, S35, Q36, W38, H82, D134, K137, Y153, K169, K170, C173, D175, G204, G205, N220, G221, V222	PSA	P07288	Mus musculus	B cell	HL	ARDGYRYYFDY	MQHLEYPVT
R45, G46, A48, N69, K70, L74, R77, L80, P83, G87, Q88, V89, Q91, V92, S93, H94, R125, S127, E128, F165	PSA	P07288	Mus musculus	B cell	HL	ARSGRLYFDV	QQTHEDPYT

Conclusions II: *in vitro* and *in vivo* assays

- T cell assays:
  - Mostly IFNg assays
  - PSA is only protein with several other cytokine assays
  - Limited number of tetramer assays
- A total of 21 epitopes identified in MHC ligand elution (MHCLE) assays
- More diverse B cell assays:
  - Ratio of monoclonal/polyclonal varies by protein (PSA with most monoclonal)
  - Most assays are ELISA direct binding assays
  - Limited amount of in vivo assays (only PSMA has in vivo B cell assays)
- Limited number of in vivo tumor killing and in vivo assays
  - Lack of mouse models for prostate cancer
  - Most assays in xenograft models
- PSA only prostate-specific antigen with receptors (1 TCR and 3 BCR)
  - *Covid-2 has over 87000 TCR and 97 BCR captured*

Epitopes and Assays by Prostate Disease Type

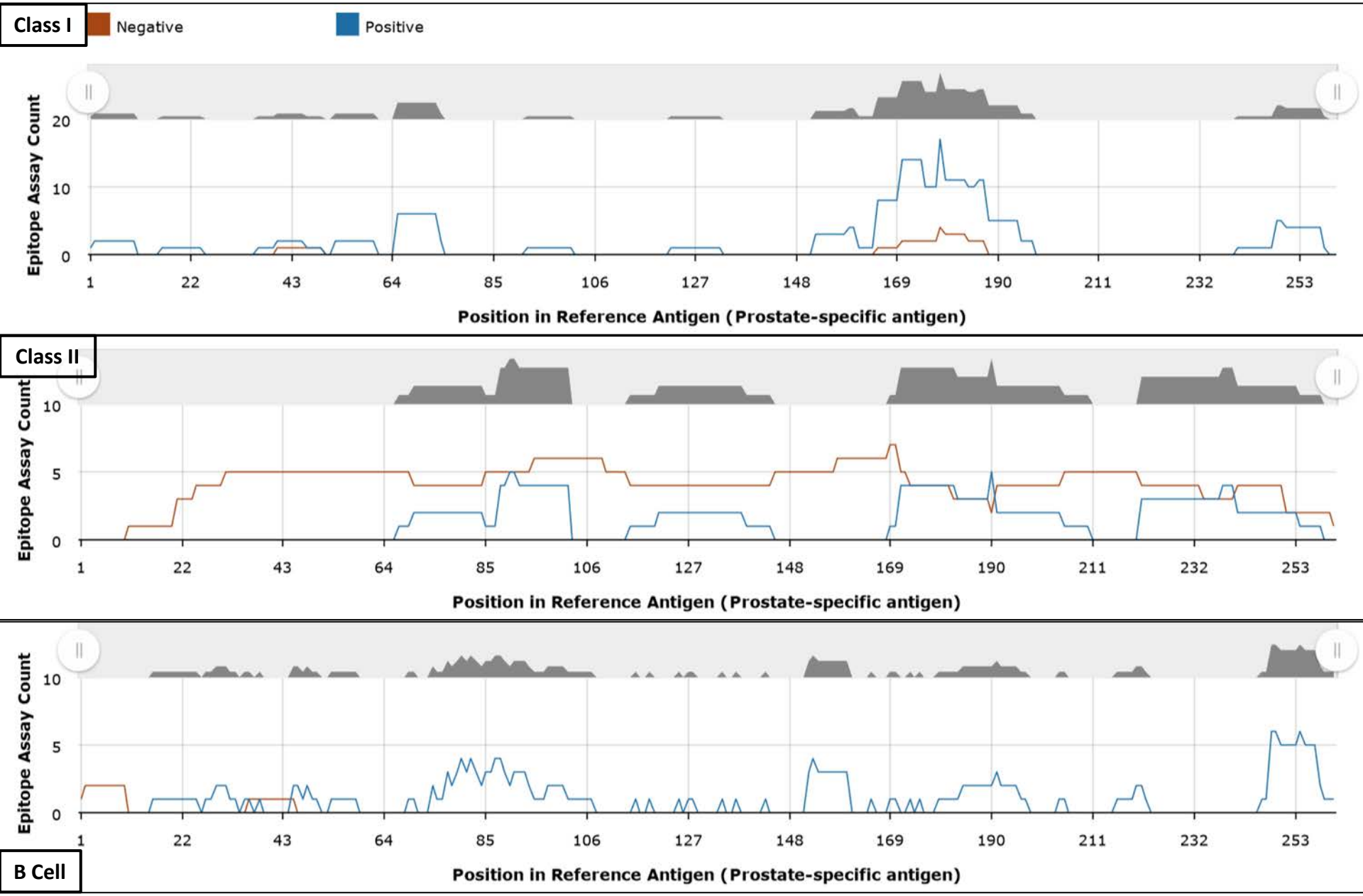
Disease State	Assay Type				Number of Epitopes			
	T Cell	B Cell	MHC	Total	T Cell	B Cell	MHC	Total
Prostate adenocarcinoma	78	41	3631	3747	31	40	3492	3565
Prostate cancer	415	264	0	679	167	193	0	360
Castration-resistant prostate carcinoma	72	90	0	162	35	29	0	64
Benign prostate phyllodes tumor	0	6	0	6	0	6	0	6
Lymphoepithelioma-like acinar prostate adenocarcinoma	0	6	0	6	0	6	0	6
Prostatitis	7	0	0	7	2	0	0	2
Prostate carcinoma	4	0	0	4	1	0	0	1

- Epitopes associated with six different prostate diseases captured in CEDAR
- Prostate adenocarcinoma and castration-resistant carcinoma are most studied
- Higher level category “prostate cancer” has the most T-cell and B-cell assays
- Gaps:
  - Staging not searchable
  - Stages are bulked and not always captured
  - Prostate cancer staging is complex, might need new options

Epitope distribution as revealed by Immunome Browser analysis (Immunodominance Patterns)

- ImmunomeBrower tool (Dhanda et al., Bioinformatics. 2018)
  - Computes the number of assays and epitopes associated with each amino acid position along the antigen sequence
  - Visualizes all studied epitopes along the entire length of a prostate-specific antigen
  - Identifies immunodominant regions

PSA Immunodominance Pattern



Overall Conclusions

- Curation processed (96.5%) of prostate antigen papers and results from 165 papers are available in CEDAR
- The curated data provides the scientific community with a wealth of searchable information
- Knowledge gaps
  - Limited number of prostate epitopes with receptors available
  - Limited numbers of in vivo and tumor recognition assays captured
- Potential areas for CEDAR improvement:
  - Search function for cancer stages
  - Search function tumor cells as antigens
- Immunodominant regions identified for most prominent antigens with potential implication for vaccination strategies

Overall Conclusions

Review > Hum Immunol. 2023 Sep 5:S0198-8859(23)00313-0.
doi: 10.1016/j.humimm.2023.08.145. Online ahead of print.

A meta-analysis of epitopes in prostate-specific antigens identifies opportunities and knowledge gaps

Gabriele Foos [1], Nina Blazeska [1], Morten Nielsen [2], Hannah Carter [3], Zeynep Kosaloglu-Yalcin [1], Bjoern Peters [4], Alessandro Sette [5]

Affiliations + expand

PMID: 37679223 DOI: 10.1016/j.humimm.2023.08.145

Read the published article here - https://pubmed.ncbi.nlm.nih.gov/37679223/