



Title: [EAGER: Computationally Predicting and Characterizing the Immune Response to Viral Infections](#)

[Marc Riedel Database Profile](#)

NSF Award #: [2036064](#)

[YouTube Recording with Slides](#)

[July 2021 CIC Webinar Information](#)

Transcript Editor: Shikhar Johri

Transcript

स्लाइड 1

तो, सभी को नमस्कार। मैं आज पेप्टाइड बाइंडिंग की समस्या पर कम्प्यूटेशनल कार्य के बारे में बात करूंगा और यह मेयो क्लिनिक से जॉर्ज वासमात्ज़िस के साथ एक सहयोग है।

स्लाइड 2

इसलिए, COVID-19 के साथ, हम सभी ने सुना है कि बीमारी है- यह किसी भी बीमारी की तरह व्यवहार करती है। लक्षणों की गंभीरता में एक सीमा है जो लोग अनुभव करते हैं, लेकिन यह शायद अन्य बीमारियों की तुलना में COVID-19 के साथ अधिक स्पष्ट है। वायरस को अनुबंधित करने वाले लोगों का एक अच्छा अंश कोई लक्षण नहीं दिखाता है। अधिकांश या तो कोई या हल्के लक्षण नहीं दिखाते हैं, लेकिन एक महत्वपूर्ण अंश गंभीर लक्षण दिखाता है। और निश्चित रूप से, एक छोटा सा अंश महत्वपूर्ण लक्षण दिखाता है या परिणामस्वरूप मृत्यु का अनुभव भी करता है। इसलिए, बीमारी की गंभीरता को प्रभावित करने वाले कई कारक हैं और विशेषज्ञ इस बारे में मुझसे बेहतर बात कर सकते हैं। आयु, लिंग और विशेष रूप से सहरुणता बहुत महत्वपूर्ण हैं और इसी तरह के वायरस के पिछले संपर्क में एक भूमिका निभाता है, लेकिन रोग की गंभीरता के संदर्भ में यहां समीकरण का एक हिस्सा हमारी प्रतिरक्षा प्रणाली में जन्मजात अंतर भी है। अनिवार्य रूप से, हमारे माता-पिता से विरासत में मिले जीन एक भूमिका निभाते हैं कि हमारे शरीर इस बीमारी या अधिकांश बीमारियों से कितनी अच्छी तरह लड़ते हैं। इसलिए, यहां हमारे अध्ययन का विषय कम्प्यूटेशनल रूप से किसी व्यक्ति की प्रतिरक्षा प्रणाली की तथाकथित सेलुलर प्रतिक्रिया में जन्मजात अंतर के एक पहलू की भविष्यवाणी करना है।

स्लाइड 3

तो, सेलुलर प्रतिक्रिया रक्षा की पहली पंक्ति है जो हमारे शरीर में किसी भी वायरल संक्रमण के जवाब में होती है। यह वह तंत्र है जिसके द्वारा कोशिकाओं में पेश किए गए विदेशी पेप्टाइड्स को पेप्टाइड्स नामक छोटे टुकड़ों में काट दिया जाता है। फिर इन्हें कोशिका की सतह पर ले जाया जा सकता है जहां वे एमएचसी [मेजर हिस्टोकम्पैटिबिलिटी कॉम्प्लेक्स] क्लास 1 अणुओं नामक सेल सतह रिसेप्टर्स से जुड़ सकते हैं। यदि इस तरह से बाध्य किया जाता है, तो संक्रमित कोशिकाएं हत्यारे टी कोशिकाओं के लिए लक्ष्य बन जाती हैं जो संक्रमित कोशिकाओं को बंद कर सकती हैं और मार सकती हैं- या तो वायरल संक्रमण से मारे गए या जैसा कि यह पता चला है, यह सबसे प्रभावी बचाव है जो हमारे पास कैंसर के खिलाफ है। कैंसर कोशिकाओं को भी इस तरह से मार दिया जाता है। यह रक्षा की पहली पंक्ति है क्योंकि अगर हम संक्रमित होने के बाद यह किक करता है, तो प्रतिरक्षा प्रणाली- हत्यारा टी कोशिकाएं सभी संक्रमित कोशिकाओं को मार सकती हैं, इससे पहले कि उनके पास जाने का मौका हो। लेकिन अगर सेलुलर प्रतिक्रिया विफल हो जाती है, तो संक्रमित कोशिकाएं कारखाने बन जाती हैं। वे वायरस की कई प्रतियों को मंथन करना शुरू कर देते हैं, और प्रतिरक्षा प्रणाली के अन्य पहलुओं को संभालना पड़ता है।

स्लाइड 4

तो, सेलुलर प्रतिक्रिया है- जैसा कि इम्यूनोलॉजी कॉम्प्लेक्स के सभी पहलुओं के साथ है। इसे समझना बहुत जरूरी है। यह SARS-CoV-2 जैसे उपन्यास वायरस की गंभीरता को समझने और भविष्यवाणी करने के लिए महत्वपूर्ण है, वायरस को लक्षित करने वाले टीके के विकास के लिए, वायरल म्यूटेशन के प्रभावों को भी समझने के लिए। अलग-अलग वायरस अलग-अलग लोगों को उनकी जन्मजात सेलुलर प्रतिरक्षा प्रतिक्रिया के माध्यम से कैसे प्रभावित करेंगे, और जैसा कि मैं कैंसर इम्यूनोथेरेपी के विषय पर लौटता रहता हूं, जो मेरे सहयोगी की विशेषज्ञता का क्षेत्र है- कई लोगों की तरह हमने अपना ध्यान COVID-19 की ओर लगाया, और कौशल सेट को फिर से तैयार किया। तो, इस मामले में, कम्प्यूटेशनल भविष्यवाणियां मूल रूप से कैंसर इम्यूनोथेरेपी को लक्षित कर रही हैं। तो, और निश्चित रूप से यह सब सिद्धांत ऑटोइम्यून बीमारियों पर लागू होता है। तो, सेलुलर प्रतिरक्षा प्रतिक्रिया, इसके मूल में, एक कम्प्यूटेशनल समस्या है। यदि हमारे पास दाईं ओर की आकृति में नीला पेप्टाइड्स है, तो यह वायरस से जुड़ा प्रोटीन टुकड़ा है। सवाल यह है: क्या वह नीला पेप्टाइड एक नाली के अंदर बंधेगा या पीले सेल सतह अणु एमएचसी 1 अणु के अंदर एक फांक होगा? और प्रतिरक्षा प्रतिक्रिया तब इस बात पर निर्भर करेगी कि क्या यह प्रोटीन टुकड़ा बांधता है, और क्या यह हत्यारा टी कोशिकाओं को पहचानने के लिए पर्याप्त रूप से बांधता है। वायरस से आने वाले नीले टुकड़े- वे उपन्यास हैं। इसलिए, एक नया, वायरस दिया गया है हमारे पास पेप्टाइड्स का एक पूरी तरह से नया सेट होगा। पीले कोशिका की सतह के अणु- जो हमारे जीन द्वारा निर्धारित किए जाते हैं और प्रत्येक व्यक्ति में छह अलग-अलग एमएचसी 1 अणु होते हैं जो हमारी विरासत द्वारा निर्धारित होते हैं- प्रत्येक माता-पिता से तीन। और कोशिका की सतह के अणु, एमएचसी अणु, हमारे जीनोम में सबसे विविध हैं, मानव आबादी में लगभग 21,000 प्रकार हैं। यह कोई दुर्घटना नहीं है। विकास ने यह सुनिश्चित किया है कि हमारी प्रतिरक्षा प्रणाली का यह पहलू बहुत विविध है ताकि हम पिछले वायरल संक्रमणों से बचने में सक्षम हों।

स्लाइड 5

लेकिन यह भी एक बहुत ही महत्वपूर्ण कम्प्यूटेशनल चुनौती बन गया है जैसा कि मैं वर्णन करूंगा। इसलिए, हम जिस मुख्य समस्या से निपट रहे हैं, वह यह है कि कंप्यूटर विज्ञान को यह अनुमान लगाने के लिए कैसे लागू किया जाए कि क्या नीला पेप्टाइड पीले सेल की सतह के अणु के अंदर बंधेगा, और पेप्टाइड्स की एक बहुत विस्तृत श्रृंखला के लिए ऐसा करने के लिए, वायरस से जुड़े सभी, और सेल सतह रिसेप्टर्स एमएचसी 1 अणुओं की बहुत विस्तृत श्रृंखला के लिए। अब बहुत सारे पूर्व कार्य-कम्प्यूटेशनल पूर्व कार्य हो चुके हैं। प्रयोगात्मक कार्य किया गया है जो निश्चित रूप से इन अणुओं की संरचना पर आणविक जानकारी प्रदान करता है, और कम्प्यूटेशनल कार्य जो बहुत सफल रहा है, मशीन सीखने को लागू करने के लिए किया गया है- उन लोगों के लिए कोई आश्चर्य की बात नहीं है जिनके पास कंप्यूटर विज्ञान और तंत्रिका नेटवर्क की पृष्ठभूमि है। तो, नेटएमएचसी नामक एक पैकेज है जिसे नीले पीले अणुओं के ज्ञात जोड़े के लिए बाध्यकारी ताकत के साथ प्रयोगात्मक डेटा पर बड़े पैमाने पर प्रशिक्षित किया गया है- एमएचसी 1 पेप्टाइड जोड़े। और एक तंत्रिका नेटवर्क संरचना के आधार पर, यह कार्यक्रम भविष्यवाणी कर सकता है, एक नया पेप्टाइड दिया गया है, यह कितना मजबूत भविष्यवाणी करेगा। और यह अनुमान केवल पेप्टाइड अनुक्रम पर आधारित है, इसलिए अक्षरों का अनुक्रम, ये पेप्टाइड में अमीनो एसिड हैं, एक लेबल के साथ जोड़ा जाता है जो कोशिका सतह अणु से मेल खाता है, एमएचसी 1 अणु। एक ताकत है। यह सब प्रयोगात्मक डेटा है। और इसलिए, कई हजारों ऐसे जोड़े दिए गए, तंत्रिका नेटवर्क को प्रशिक्षित किया जा सकता है। और एक बार जब यह प्रशिक्षित हो जाता है, तो एक उपन्यास पेप्टाइड अनुक्रम, अमीनो एसिड का एक उपन्यास अनुक्रम दिया जाता है, यह भविष्यवाणी कर सकता है कि पेप्टाइड किसी दिए गए एमएचसी 1 अणु से कितनी अच्छी तरह बंधेगा। तो यहाँ, पेप्टाइड अनुक्रम ड्राइंग में नीला पेप्टाइड अणु होगा, और लेबल पीले सेल सतह रिसेप्टर, एमएचसी 1 अणु के अनुरूप होगा। यह महान है और तंत्रिका नेटवर्क इस अर्थ में शक्तिशाली है कि उन्हें आसानी से प्रशिक्षित किया जा सकता है और वे दिए गए डेटा के आधार पर बहुत प्रभावी ढंग से भविष्यवाणियां कर सकते हैं। लेकिन यहाँ सीमा बिल्कुल कोई आणविक डेटा जो भी है है, हम बस प्रशिक्षण लेबल और पत्र कर रहे हैं, और यह भी प्रशिक्षण डेटा प्रयोगात्मक डेटा का एक बहुत ही विविध सेट के लिए लिया जाता है। उदाहरण के लिए इसमें से बहुत कुछ एचआईवी, एक अलग वायरस और एचआईवी से जुड़े पेप्टाइड्स से आता है, और समस्या को SARS-CoV-2 जैसा एक नया वायरस दिया जाता है, अधिकांश पेप्टाइड्स पहले कभी नहीं देखे गए हैं, और एक तंत्रिका नेटवर्क भविष्यवाणियां करेगा जो नकली हैं क्योंकि यह पेप्टाइड अंतरिक्ष के एक बहुत अलग क्षेत्र से निष्कर्ष निकाल रहा है।

स्लाइड 6

और इसलिए, हमारा दृष्टिकोण आणविक-स्तरीय सिमुलेशन करना है, और निश्चित रूप से इस विषय पर बहुत सारे पूर्व कार्य हुए हैं। बहुत परिष्कृत आणविक सिमुलेशन तकनीकों को जाना जाता है और व्यापक रूप से उपयोग किया जाता है। एक को आणविक गतिकी कहा जाता है। मोंटे कार्लो आधारित सिमुलेशन भी है। वे सिमुलेटेड एनीलिंग नामक एक तकनीक का उपयोग करते हैं, और आणविक डॉकिंग एक और दृष्टिकोण है। ऐसे आणविक स्तर सिमुलेशन के लिए उपलब्ध सॉफ्टवेयर का व्यापक रूप से उपयोग किया जाता है, लेकिन वे विशेष हैं। वे सामान्य उद्देश्य हैं। वे अणुओं के बंधन के व्यापक वर्गों के लिए विकसित किए गए हैं और वे बेहद कम्प्यूटेशनल रूप से गहन हैं। एक पेप्टाइड, एक एमएचसी 1 जोड़ी लेने के लिए, और इसे अनुकरण करने के लिए मौजूदा सॉफ्टवेयर का उपयोग करने के लिए, एक बाध्यकारी घटना को अनुकरण करने में दिन, कभी-कभी सप्ताह लगते हैं। तो, एक ही भविष्यवाणी करने के लिए वास्तव में सुपरकंप्यूटिंग समय के सप्ताह। और हम जिस समस्या का सामना कर रहे हैं, उसका दायरा यह है कि हमारे पास पीले सेल की सतह के अणुओं, MHC21,000 अणुओं के 1 प्रकार हैं, और SARS-CoV-2 के लिए, अगर हम केवल स्पाइक प्रोटीन पर ध्यान केंद्रित करते हैं और हम इसे काटते हैं पेप्टाइड्स के लिए

छोटे टुकड़ों में, हमारे पास उनमें से लगभग 38,000 हैं। इसलिए, हम 1 बिलियन संयोजनों के बारे में बात कर रहे हैं जिन्हें हम तनाव के संदर्भ में अनुकरण करना चाहते हैं, और यदि इसमें सुपरकंप्यूटिंग समय का एक सप्ताह लगता है तो हमारे पास स्पष्ट रूप से इसका अध्ययन करने के लिए एक अरब सप्ताह नहीं हैं।

स्लाइड 7

अतः हमारा दृष्टिकोण दुगुना है। एक ओर, हम आणविक सिमुलेशन के लिए अत्यधिक अनुकूलित सॉफ्टवेयर बना रहे हैं और हम उस डोमेन के विवरण का उपयोग कर रहे हैं जिसमें हम काम कर रहे हैं। तो, हम पेप्टाइड के साथ शुरू कर रहे हैं। पेप्टाइड्स उनकी लंबाई या उनके आकार के मामले में इतना भिन्न नहीं होते हैं। हम पेप्टाइड्स के साथ शुरू कर रहे हैं जो कोशिका की सतह के अणु, एमएचसी 1 अणु के फांका के अंदर सही ढंग से संरेखित हैं, इसलिए हम अंतरिक्ष में पूरे पेप्टाइड को घुमाने में बहुत समय नहीं बिताते हैं। हम इसे ठीक उसी स्थान पर रखते हैं जहां यह होना चाहिए, और हम मरोड़ वाले स्थान में पूरी खोज करते हैं। इसलिए अणु को त्रि-आयामी स्थान के चारों ओर ले जाने के बजाय हम इष्टतम कॉन्फिगरेशन खोजने की कोशिश करने के लिए इसके बंधनों को मोड़ते और मोड़ते हैं। दूसरा योगदान यह है कि हम इसे बड़े पैमाने पर तैनात कर रहे हैं। इसलिए, हम जीपीयू का उपयोग कर रहे हैं, और फिर अंततः क्लाउड कंप्यूटिंग इन्फ्रास्ट्रक्चर, वास्तव में समस्या पर कंप्यूटिंग शक्ति फेंकने के लिए। और हमारा लक्ष्य, जैसा कि मैंने शीर्षक में कहा है, एक अरब दिनों को एक मिलियन मिनट या शायद एक महीने के क्लाउड कंप्यूटिंग समय में बदलना है।

स्लाइड 8

तो, चुनौती यहाँ है और मैं इस पर चमक डालूंगा- मैं लगभग समय से बाहर हूँ। एक बात: प्रयोगात्मक डेटा पूरा नहीं है। हमारे पास पूर्ण आणविक मॉडल नहीं हैं, न केवल सभी 21,000 वेरिएंट के, बल्कि हमारे पास एक अच्छा भौगोलिक प्रसार भी नहीं है। यह पिछली कुछ वार्ताओं से संबंधित है। अधिकांश प्रयोगात्मक डेटा पश्चिमी कोकेशियान जनसांख्यिकी से एमएचसी 1 अणुओं के विवरण के लिए है। तो, दूसरा दृष्टिकोण यह है कि हमें अणुओं की संरचना का अनुमान लगाना चाहिए और फिर उनका अनुकरण करना चाहिए।

स्लाइड 9

और इसलिए, अंतिम स्लाइड: इस काम का प्रभाव। वैसे यह एक व्यक्ति के लिए निर्धारित करने के लिए अपेक्षाकृत सरल है कि उनके पास कौन से जीन हैं- प्रासंगिक अणुओं के लिए वह कोड, एमएचसी 1 अणु। यह एचएलए टाइपिंग के माध्यम से किया जा सकता है, जो पितृत्व परीक्षण के लिए किया जाता है। उस जानकारी को देखते हुए और हमारे कम्प्यूटेशनल बुनियादी ढांचे को देखते हुए, हम एक व्यक्ति के लिए भविष्यवाणी करने में सक्षम होंगे, वह व्यक्ति एक नए रोगजनक का जवाब कैसे देगा। एक वायरस के लिए, अलग-अलग व्यक्तियों के लिए वायरस के वेरिएंट के लिए, और अलग-अलग व्यक्तियों के लिए अलग-अलग वायरस पर अलग-अलग टीकों के प्रभाव के लिए। और निश्चित रूप से, जैसा कि मैंने शुरुआत में कहा था, यह काम न केवल वायरस पर लागू होगा, बल्कि संभावित रूप से कैंसर इम्यूनोथेरेपी और ऑटोइम्यून बीमारियों पर भी लागू होगा। तो, मैं वहाँ बंद कर दूँगा। बहुत-बहुत धन्यवाद।