



질병관리청

보도참고자료

다시 도약하는 대한민국
함께 잘사는 국민의 나라

보도 일시	2023.3.27.(월) 배포 즉시	배포 일시	2023.3.27.(월) 08:50		
담당 부서 <총괄>	미래의료연구부 헬스케어인공지능연구과	책임자	과 장	이주희	(043-719-8850)
		담당자	연구관	김상철	(043-719-8861)
담당 부서	미래의료연구부 바이오뱅크과	책임자	과 장	전재필	(043-719-6550)
		담당자	연구관	조상연	(043-719-6526)
담당 부서	미래의료연구부 바이오빅데이터과	책임자	과 장	채희열	(043-249-3003)
		담당자	연구관	김정은	(043-249-3072)

코로나19 연구 자료 및 인체 자원 추가 공개

- 코로나19 백신 및 치료제 개발 연구 활용 기대 -

주요 내용

- 코로나19 확진자(459명) 및 백신 접종자를 포함한 일반인(161명)의 다중오믹스* 연구 자료와 인체 자원을 추가 공개함으로써 질병 기전 이해, 백신 및 치료제 개발 등 연구 활성화 기대
- * 다중오믹스(Multi-Omics) : 유전적, 기능적, 환경적, 면역반응 등의 이해를 위해 최신의 기술을 활용하여 생산된 종합 데이터를 의미함(붙임 1 참고).

- 질병관리청(청장 지영미) 국립보건연구원(원장직대 장희창)은 코로나19 확진자 다중오믹스 자료, 임상 정보와 인체 자원을 연구목적으로 추가 공개·분양한다고 밝혔다.
- 국립보건연구원은 코로나19 확진자(300명) 및 일반인(120명) 다중오믹스 분석결과와 임상 정보, 인체 자원을 연구목적으로 2022년 1월 3일 공개·분양한 바 있으며,

- 생명·윤리적으로 적절한 동의 절차를 거쳐 코로나19 확진자 459명과 백신 접종자 57명을 포함한 일반인의 임상 정보와 인체 자원을 추가 확보하였다(붙임2 참고).
- 특히, 백신접종자는 백신 접종 전·후 등 5시점*의 연구자료가 포함되어 감염에 의한 변화와 비교 분석도 가능하다.
 - * 1차 접종 전, 1차 접종 후 1주 후, 2차 접종 전, 2차 접종 후 1주 후, 2차 접종 후 4주 후
- 코로나19 확진자 459명은 중증도에 따라서 경증(329, 72%)은 3시점에서, 중증(130명, 28%)은 최대 7시점에서 인체 자원을 추가 확보*하였다.
 - * (3시점) 입원 시, 입원 7일 후, 퇴원 시, (5시점) 입원 시, 입원 7일 후, 약물 투여 전후, 퇴원 시 등
- 수집된 인체 자원을 활용하여 질병관리청에서는 일반혈액검사, 191개의 면역 인자(사이토카인) 및 다중오믹스 자료를 추가 생산하였으며(붙임3 참고),
 - 추가 생산된 자료에는 전장유전체분석(WGS), 단일세포 전사체 분석(scRNA-seq), T세포/B세포 수용체 발현정보(bulk TCR/BCR RNAseq) 등 다중오믹스와 인간백혈구항원 유형(HLA typing)을 포함한다(붙임4 참고).
 - 이렇게 생산된 자료는 코로나19 감염 후 인간의 면역 반응에 심도 있는 분석에 이용되어 중증도 예측, 치료제 개발에 중요한 자료로 활용 가능하다.
- 국립보건연구원은 후속 사업을 통해서 코로나19 확진자의 단백질* 분석 자료를 추가 확보하여 신속하게 공개할 예정이다.
 - * 단백질 : 체내에서 기능을 하는 단백질의 총합을 의미함 (붙임 1 참고)
- 코로나19 확진자 및 일반인 다중오믹스 데이터와 인체 자원은 국립보건연구원 국립중앙인체자원은행을 통해 분양하며, '23년 3월 27일부터 국립중앙인체자원은행 누리집*에서 신청 접수가 가능하다.
 - * 누리집 바로가기 : <http://www.nih.go.kr/biobank>
- 신청한 과제는 국립중앙인체자원은행 분양위원회 심의*를 거쳐 연구자에게 제공된다.
 - * 인체자원이용계획서, IRB 심의용 연구계획서 등 분양신청서류 제출(붙임5 참고)

○ 코로나19 다중오믹스 자료에 대한 상세한 정보는 다중오믹스 정보 시스템*에서 검색 가능하며(붙임 6 참고),

* 시스템 바로가기 : <http://www.coda.nih.go.kr>

- 데이터 분석은 연구자가 네트워크를 통한 원격 분석(한국과학기술정보연구원 슈퍼컴퓨터) 또는 현장 방문 분석(충북 오송)으로 이루어질 예정이다.

□ 지영미 질병관리청장은 “본 사업의 코로나19 확진자 다중오믹스 연구 자료 및 인체 자원 개방이 향후 신종 감염병 발생 시 백신 및 치료제 개발 역량 강화의 기반이 될 것으로 기대한다.”라고 밝혔다.

- <붙임>
1. 다중오믹스 등 용어 설명
 2. 코로나19 확진자 및 일반인 공개 인체 자원
 3. 코로나19 확진자 및 일반인 공개 정보
 4. 코로나19 확진자 및 일반인 다중오믹스 분석정보
 5. 코로나19 연구용 인체 자원 및 데이터 분양 절차 안내
 6. 다중오믹스 정보시스템 안내

붙임 1

다중오믹스 등 용어 설명

- 다중오믹스(Multi-Omics)란,
 - 오믹스(Omics)란, 생명과학 분야에서 대용량 분석기법이 발전하면서 분석결과로 나오는 많은 분자들이나 세포 등의 집합체 전부를 의미
 - 다중오믹스(Multi-Omics)란, 동일 인체 자원에서 생산되는 다양한 형태의 오믹스 분석결과를 동시에 분석하기 위한 오믹스의 집합을 의미
- 공개분양 대상 다중오믹스 종류별 정의
 - 유전체(Genome) : 약 24,000개 유전자로 구성된 유전자들의 집합으로, 부모로부터 물려받은 고유의 특성
 - 전사체(Transcriptome) : 유전자로부터 발현되는 전사물(transcript)의 집합으로, 인체내 기관별(뇌, 심장, 위 등)로 같은 유전체에서 다른 유전자 발현 양상을 나타내 기관별 구분되는 기능을 수행
 - 단일세포 전사체(single cell Transcriptome) : 단일세포 유전자로부터 발현되는 전사물(transcript)의 집합으로, 면역세포인 T세포 또는 B 세포에서 단일세포에서 유전자 발현 양상을 오믹스 형태로 분석하면 면역세포의 기능적 특성 연구에 활용 가능
 - 단백질체(Proteome) : 체내에서 발현하는 단백질의 총합. 특정 환경에서 체내 기능을 하는 단백질의 총합

붙임 2

코로나19 확진자 및 일반인 공개 인체 자원

○ (수집대상자) 코로나19 확진자와 일반인에서 인체 자원 수집

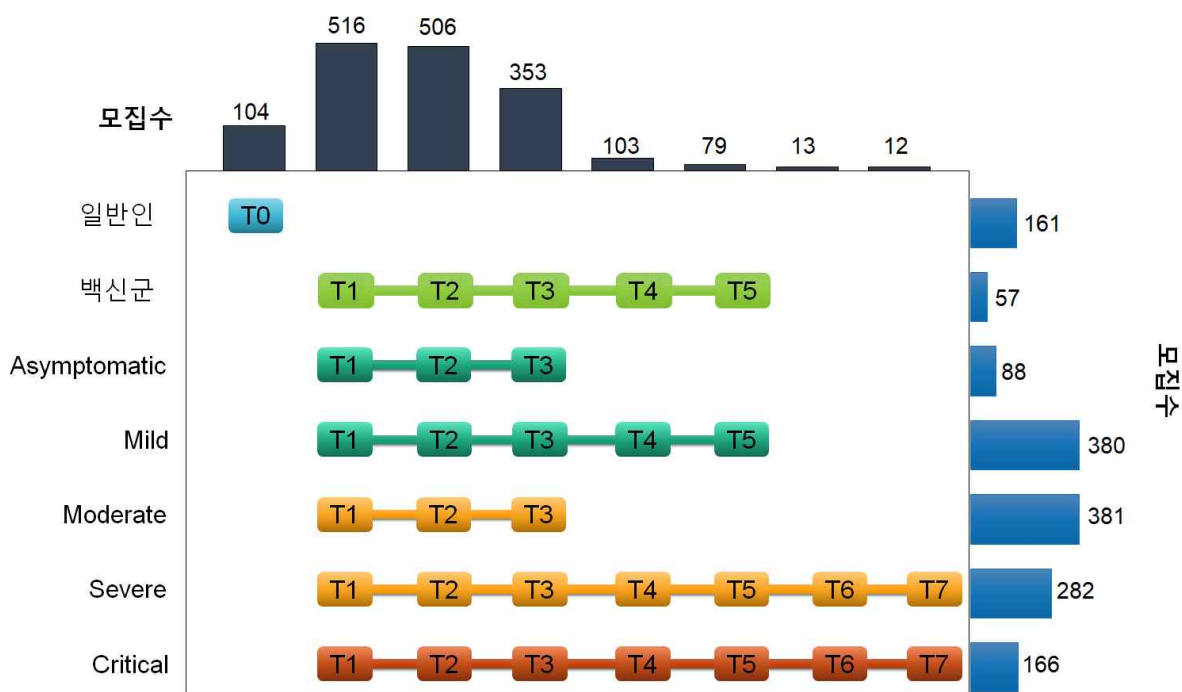
- (일반인) 기본 시점 인체 자원만 수집하여 대조군으로 활용

* 백신접종자의 경우 최대 5시점 수집(1차 접종 전, 1차 접종 후 1주 후, 2차 접종 전, 2차 접종 후 1주 후, 2차 접종 후 4주 후)

- (코로나19 확진자) 중증도*에 따라 구분하여 수집하였으며, 경증은 3시점(입원시, 입원 후 7일, 퇴원시), 중증은 최대 7시점(입원시, 입원 후 7일, 약물투여시, 퇴원시 등)에서 인체 자원 수집

* 경증(Asymptomatic, Mild(5시점 1명), Moderate), 중증(Severe, Critical)

< 수집 대상자별 인체 자원 수집시점 개요 >



* T는 시점을 의미함(T1은 1시점, T7은 7시점)

- 중증도는 폐렴과 함께 다음 증상 중 하나가 동반되는 경우로, 호흡수 30회 이상, 산소포화도 93% 이하, 산소화 지표*(PaO2/FiO2) 300mmHg 이하, 흉부 방사선 혹은 흉부 컴퓨터 단층촬영(CT)상 폐의 침윤 50% 이상 여부로 분류

* 혈액 내 산소 수치가 낮거나 이산화탄소 수치를 측정할 때 사용하는 지표

- (수집자원 종류) 코로나19 확진자 및 일반인으로부터 gDNA, PBMC (세포), 혈청, 혈장, 뇨, 상하기도검체(객담) 인체 자원 수집

< 코로나19 확진자 및 일반인 인체 자원 공개대상 >

	명수	최대 추적 회수	인체 자원 종류 (단위: 바이알)						
			gDNA	혈청 (0.3mL vial)	혈장(EDTA) (0.3mL vial)	혈장(CPT) (0.3mL vial)	PBMC	Urine (1mL vial)	객담
일반인	104	1회	243	632	804	1,215	441	1,144	0
백신군	57	5회	652	1,379	2,178	3,300	1,468	892	0
확진자	459	7회	3,825	11,601	12,668	19,401	6,052	18,380	154
계	620		4,720	13,612	15,650	23,916	7,961	20,416	154

붙임 3

코로나19 확진자 및 일반인 공개 정보

○ 일반 임상 정보 및 진단검사결과(변수 총 210개 항목)

연번	구분(변수 수)	변수 설명
1	등록정보(4)	성별, 만 나이, 내원/입원일, 임상정보수집일
2	인구학적 정보(10)	대상자 그룹, 거주지, 임신 혹은 출산 여부, COVID19 검사결과, 환자상태 병명, 내원경로 등
3	내원 시 소견(16)	발열 유무, 기침 유무, 객담 유무, 호흡곤란 유무, 인후염 유무, 콧물 유무, 근육통 유무, 피로/권태 유무, 두통 유무, 의식장애 유무, 구토/오심 유무, 설사 유무, 기타 증상 등
4	내원 시 기저질환(18)	고혈압 유무, 당뇨 유무, 관상동맥질환 유무, 뇌졸중 유무, 심부전 유무, 만성폐쇄성폐질환 유무, 결핵 유무, 천식 유무, 만성신장병 유무, 악성종양 유무, 만성간염/간경화 유무, HIV 감염 및 치료 유무, 흡연 유무 등
5	내원 당시 약물력(17)	항바이러스 투여 여부 및 종류, 항생제 투여 여부 및 종류, 스테로이드 투여 여부 및 종류
6	전원 시 약물력(10)	혈장 치료제 처방 여부, 항체치료제 처방 여부, 레보비르 처방 여부, 승압제 처방 여부, 기타 약물 등
7	내원 시 접종력(1)	인플루엔자 백신 접종 여부
8	신체계측정보(3)	키, 몸무게 BMI
9	내원 시 생체 징후(9)	체온, 심박수, 호흡수, 수축기 및 이완기 혈압, 산소포화도, 의식수준 등
10	내원 시 영상의학적 검사(2)	Chest X-ray/CT 시행 여부, Chest X-ray/CT 폐렴 소견 여부
11	입원 중 생체 징후 (가장 나쁜 수치)(9)	체온, 심박수, 호흡수, 수축기 및 이완기 혈압, 산소포화도 의식수준 등
12	비약물적 치료 (평가일 중 가장 나쁜 상태)(18)	고유량 산소치료 종류 적용여부 및 최고 산소유량, 인공호흡기 종류 적용 여부 및 최고 흡입산소 농도, ECMO 적용 여부, CRRT 적용 여부 등
13	약물적 치료(27)	항바이러스 투여 여부 및 종류, 항생제 투여 여부 및 종류, 스테로이드 투여 여부 및 종류, 혈장 치료제, 항체치료제 등
14	합병증(11)	Shock, Bacteremia, Acute respiratory distress syndrom, Cardiac arrhythmia, Cardiomyopathy, Cardiac arrest, Acute renal injury, Liver dysfunction emd
15	입원 중 혈액검사(47)	혈액검사 시행여부, Hemoglobin, WBC, Neutrophil, Lymphocyte, Platelet, Sodium, Potassium, Urea, Creatinine, Albumin, AST, ALT, Total bilirubin, LDH, D-dimer, Lactate, Troponin I /T, Ferritin 등
16	퇴원 시 정보(4)	모니터링 종료시점의 최종상태, 모니터링 종료일, 모니터링 종료시점의 격리해제 여부, 격리해제일
17	일반진단검사(37)	Albumin, ALT, AST, r-GTP, Calcium, BUN, Creatinine, Uric acid, Homocysteine, Iron, UIBC, hs-CRP, Cholesterol, TG, HDL, LDL, Ferritin, Vit B12, Folate, Cystatin C, Apo A I, Apo A II, Apo B, Lipoprotein(a) 등
18	기타(2)	특이사항1, 2

○ 면역인자(사이토카인, Cytokine) 분석 결과(변수 총 191개)

- 면역인자 분석 결과는 경증, 중증 일부 환자에 대해서만 생산

· 일반인의 경우, 1시점에서 면역인자 분석결과 생산

* 백신접종자의 경우 최대 5시점 생산(1차 접종 전, 1차 접종 후 1주 후, 2차 접종 전, 2차 접종 후 1주 후, 2차 접종 후 4주 후)

· 코로나19 확진자 경증인 경우, 3시점(입원시, 입원 후 7일, 퇴원시)에서 면역인자 분석결과 생산

· 코로나19 확진자 중증인 경우, 최대 7시점(입원시, 입원 후 7일, 약물투여 전·후, 퇴원시 등)에서 면역인자 분석결과 생산

사이토카인 분석항목			
alpha 2-Macroglobulin	SCF/c-kit Ligand	beta 2-Microglobulin	Thrombopoietin/Tpo
Fetuin A/AHSG	u-Plasminogen Activator (uPA)	MBL	VCAM-1/CD106
Protein S/PROS1	CCL22/MDC	Serpin A4/Kallistatin	CCL23/MPIF-1
CD14	CCL26/Eotaxin-3	Component C2	CD117/c-kit
CXCL4/PF4	CD163	LBP	CEACAM-1/CD66a
LRG1	Coagulation Factor III/Tissue Factor	Properdin	Follistatin-related Gene Protein/FLRG
Serpin C1	Furin	Serpin A10/ZPI	IFN-beta
Aldehyd Dehydrogenase 1-A1	IGFBP-1	BCMA/TNFRSF17	IL-11
CCL14	IL-28B/IFN-lamgda 3	CCL18/PARC	IL-4R alpha
Coagulation Factor XIV/Protein C	IL-6R alpha	Collagen I alpha 1	MCAM/CD146
Complement Factor D/Adipsin	Oncostatin M/OSM	DPPIV/CD26	Osteoactivin/GPNMB
Fibroblast Activation Protein alpha/FAP	Reg3A	Galectin-3	Resistin
Galectin-3BP/MAC-2BP	S100A9	IGFBP-2	SCGF/CLEC11a
Lumican	Serpin B3/SCCA1	Myeloperoxidase/MPO	TACI/TNFRSF13B
VAP-1/AOC3	Thrombomodulin/BDCA-3	CA15-3/MUC-1 (Unit)	Thrombospondin-2
CCL5/RANTES	TRAIL R2/TNFRSF10B	Complement component C9	TRNACE/TNFSF11/RANKL
EN-RAGE/S100A12	ADAMTS13	M-CSF R/CD115	APRIL/TNFSF13
MMP-2	CA125/MUC16 (Units)	MMP-9	Chitinase 3-like 1
Myoglobin	CXCL1/GRO alpha	Lipocalin-2/NGAL	CXCL10/IP-10
TIMP-1	Dkk-1	CCL19/MIP-3 beta	EGF
CCL2/MCP-1	ENPP-2/Autotaxin	CCL3/MIP-1 alpha	GDF-15
CCL4/MIP-1 beta	Growth Hormone	CD23/Fc epsilon RII	HGF
CD31/PECAM-1	IFN-alpha	CX3CL1/Fractalkine	IL-13

사이토카인 분석항목			
CXCL9/MIG	IL-18/IL-1F4	ErbB3/Her3	IL-23
Flt-3 Ligand/FLT3L	IL-33	Granzyme B	M-CSF
IL-4	G-CSF	IFN-gamma	IL-31
IL-5	IL-2	IL-1 beta	IL-1 alpha/IL-1F1
IL-10	IL-17/IL-17A	IL-12 p40	IL-12 p70
IL-6	uPAR	IL-8	IL-1ra/IL-1F3
Kallikrein 6/Neurosin	Angiopoietin-1	LIGHT/TNFSF14	Angiopoietin-like Protein 3/ANGPTL3
MMP-12	MMP-13	MMP-3	MMP-7
MMP-8	CCL24/Eotaxin-2/MPIF-2	CD40 Ligand/TNFSF5	CCL11/Eotaxin
Complement Component C5a	CXCL2/Gro Beta/MIP-2	CXCL6/GCP-2	FABP4/A-FABP
Fas Ligand/TNFSF6	Granzyme A	TGF-alpha	TNF-alpha
TREM-1	TSLP	VEGF	Alpha 1-Microglobulin
Angiopoietin-2	BAFF	BMP-10	Contactin-1
CXCL11/I-TAC	CXCL13/BCA-1	CXCL5/ENA-78	Endocan/ESAM-1
ErbB2/Her2	FGF basic/bFGF	Galectin-9	L-Selectin/CD62L
MIA	Nectin-4	Nephrin	IL-15
IL-17E/IL-25	IL-3	IL-36 beta/IL-1F8	IL-7
LIF	Osteopontin/OPN	PDGF-AA	PD-L1/B7-H1
ST2/IL-33R	GM-CSF	ICAM-1/CD54	Neuregulin-1 beta/NRG1 beta 1
Prolactin	SP-D	vWF-A2	Pentraxin 3/TSG-14
IGFBP-4	MIF	MICA	HE4/WFDC2
NT-4	Leptin/OB	Proprotein convertase 9/PCSK9	MAdCAM-1
TRAIL/TNFSF10	Syndecan-1/CD138	P-Selectin/CD62P	TFF3
BDNF	VEGFR1/Flt-1	TGF-beta 1	TGF-beta 2
PDFG-DD	Complement C1q	SDF-1	

○ 코로나19 바이러스 유전체 정보

Pango 계통*	국제보건기구 (WHO) 분류	주요 발생국가	수집건수
B.1.497		대한민국	160
B.1.1.7	alpha	영국	20
B.1		미국	14
B.1.620		대한민국	19
B.1.619		독일	60
B.1.427/429	epsilon	미국	14
B.1.470		인도네시아	6
B.1.351	beta	남아프리카	1
B.1.351.3	beta	방글라데시	1

B.1.1.25		방글라데시	1
P.1	gamma	브라질	1
A.23.1		캐나다	1
B		미국	1
AY.69	Delta	대한민국	57
AY.122.5	Delta	대한민국	2
B.1.617.2	Delta	미국	1
B.1.619.1		대한민국	34
BA.2.68	Omicron (BA.2-like)	대한민국	1
BA.2	Omicron (BA.2-like)	영국	1
BA.2.3	Omicron (BA.2-like)	미국	2
BA.2.3.14	Omicron (BA.2-like)	대한민국	1
BA.5.2	Omicron (BA.5-like)	미국	8
BE.1.1	Omicron (BA.5-like)	독일	1
BF.1	Omicron (BA.5-like)	영국	1
BA.5.2.1	Omicron (BA.5-like)	미국	1
BA.5.1	Omicron (BA.5-like)	미국	1
BA.5.5	Omicron (Unassigned)	미국	3

* Pango 계통 : 2020년 4월 영국과 호주에서 시작된 코로나19 바이러스의 계통 (Lineages)을 명명하기 위한 문자와 숫자로 된 시스템으로 새로운 바이러스를 추적하는데 사용됨(<https://cov-lineages.org>).

붙임 4

코로나19 확진자 및 일반인 다중오믹스 분석정보

- 다중오믹스 분석정보 생산 현황
 - 전장유전체분석(WGS), 인간백혈구항원 유형분석(HLA typing)은 코로나19 확진자 및 일반인 1시점에서만 생산
 - 단일세포 전사체 분석(scRNA; scTCR/TCR-seq, Bulk TCR/BCR RNAseq)은 경증, 중증 일부 환자에 대해서만 생산
 - 일반인의 경우, 1시점에서 scRNA-seq 분석결과 생산
 - * 백신접종자의 경우 최대 5시점 생산(1차 접종 전, 1차 접종 후 1주 후, 2차 접종 전, 2차 접종 후 1주 후, 2차 접종 후 4주 후)
 - 코로나19 확진자 경증인 경우, 3시점(입원시, 입원 후 7일, 퇴원시)에서 scRNA-seq 분석결과 생산
 - 코로나19 확진자 중증인 경우, 최대 7시점(입원시, 입원 후 7일, 약물 투여 전후, 퇴원시 등)에서 scRNA-seq 분석결과 생산

< 코로나19 확진자 및 일반인 다중오믹스 분석정보 공개대상 >

오믹스 분석종류	확진자(명)	백신접종군(명)	일반인(명)	총계
전장유전체 분석(WGS)	459	57	104	620
HLA typing	459	57	104	620
scRNAseq(+scTCR/BCRseq)	199	54	64	317
Bulk TCR seq	184	37	49	270
Bulk BCR seq	184	37	49	270

○ 인간백혈구항원 유형분석(HLA typing)(17개 항목)

연번	변수명	상세설명
1	IMGT/A	Classical MHC Class I , HLA-A 시퀀싱
2	IMGT/B	Classical MHC Class I , HLA-B 시퀀싱
3	IMGT/C	Classical MHC Class I , HLA-C 시퀀싱
4	IMGT/DPA1	MHC Class II, HLA-DPA1 시퀀싱
5	IMGT/DPB1	MHC Class II, HLA-DPB1 시퀀싱
6	IMGT/DQA1	MHC Class II, HLA-DQA1 시퀀싱
7	IMGT/DQB1	MHC Class II, HLA-DQB1 시퀀싱
8	IMGT/DRB1	MHC Class II, HLA-DRB1 시퀀싱
9	IMGT/DRB3	MHC Class II, HLA-DRB3 시퀀싱
10	IMGT/DRB4	MHC Class II, HLA-DRB4 시퀀싱
11	IMGT/DRB5	MHC Class II, HLA-DRB5 시퀀싱
12	IMGT/E	non-Classical MHC Class I , HLA-E 시퀀싱
13	IMGT/F	non-Classical MHC Class I , HLA-F 시퀀싱
14	IMGT/G	non-Classical MHC Class I , HLA-G 시퀀싱
15	IMGT/H	non-Classical MHC Class I , HLA-H 시퀀싱
16	IMGT/MICA	MHC Class I chain-related gene A 시퀀싱
17	IMGT/MICB	MHC Class I chain-related gene B 시퀀싱

붙임 5

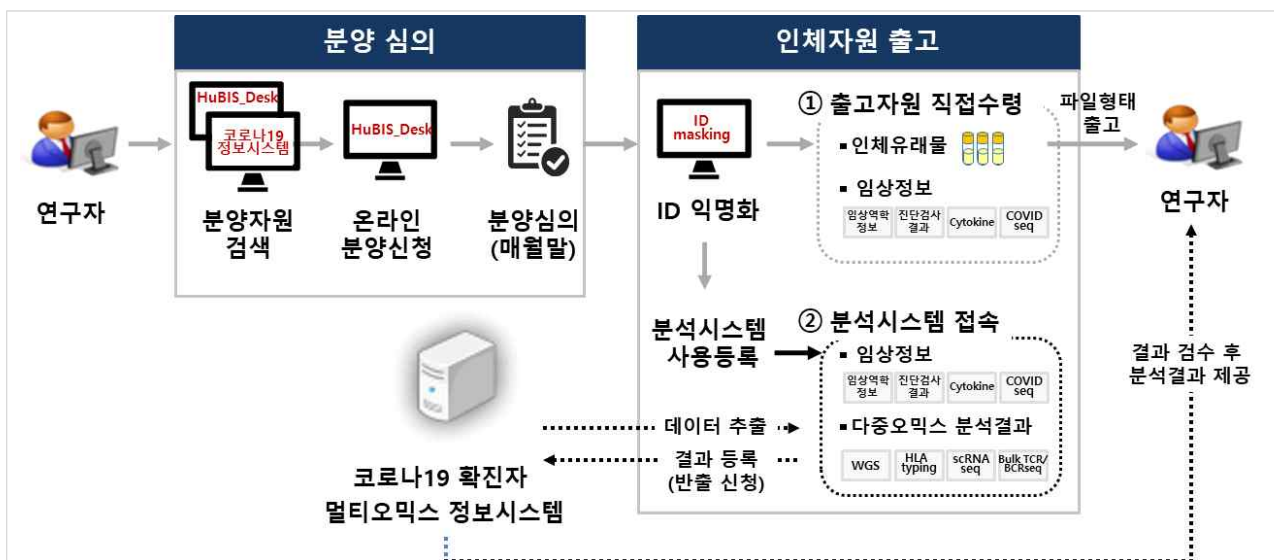
코로나19 연구용 인체 자원 및 데이터 분양 절차 안내

질병관리청 국립보건연구원 국립중앙인체자원은행에서는 코로나19 관련 연구를 지원하고자 아래와 같이 분양절차를 안내드립니다. 관심있는 기관의 많은 신청 바랍니다.

- (대상검색) 국립중앙인체자원은행 온라인 분양시스템 또는 다중오믹스 시스템에서 분양대상자원 검색
 1. 국립중앙인체자원은행 온라인 분양시스템 이용방법
 - 질병관리청 질병보건통합관리시스템(is.kdca.go.kr) 회원가입 및 권한신청(문의 : 043-719-6534) → 온라인 분양데스크 접속 후 자원 검색
 2. 다중오믹스 정보 검색 및 분석 시스템 이용방법
 - 질병관리청 국립보건연구원 보건의료연구자원정보센터(CODA, coda.nih.go.kr) 회원가입(디지털 원패스 포함) 및 권한신청(문의: 043-249-3025) → 보건의료연구자원정보 센터 접속 후 분양신청 메뉴 선택 → 분양 대상 자원(COVID19) 검색
- (신청대상기관) 의료기관, 국공립 연구기관, 임상검사기관, 체외진단 의료기기 개발업체 등 기술개발(R&D) 목적의 인체 자원 이용에 적합한 연구시설(BL2 또는 BL3) 보유한 기관
- (신청방법) 온라인 분양신청
 - (절차) 질병관리청 질병보건통합관리시스템(is.kdca.go.kr) 회원가입 및 권한신청(문의: 043-719-6534) → 온라인 분양데스크 신청서 작성
 - (분양 신청서류)
 - 1) 인체 자원 이용계획서(온라인 작성)
 - 2) 서약서(온라인 작성)
 - 3) 개인정보수집이용동의서(온라인 작성)
 - 4) IRB 심의 결과서
 - 5) IRB 심의용 연구계획서
 - 6) 분양신청공문

- (분양심의) 국립중앙인체자원은행 분양위원회 심의를 거쳐 다음의 분양원칙에 따라 연구용 인체 자원 분양대상과제를 선정
 - (높은 활용성) 전체 참여자 대상 추가 데이터 생산하는 과제
 - (오믹스 데이터) 단백질, 대사체 등 오믹스 수준 데이터 생산하는 과제
 - (이용성과물 제출) 생산된 데이터 등 이용성과물 제출 의무 부여
- (오믹스데이터 분석·활용) 국립중앙인체자원은행 분양위원회 심의 완료 후 다중오믹스 정보시스템내에서 데이터 분석 진행
 - (연구자 등록) 질병관리청 다중오믹스 정보시스템 연구자 정보 등록 및 신청
 - (분석환경 구축) 분양심의 결과를 반영하여 코로나19 DB에서 데이터 추출하여 분석환경 구축 및 제공
 - 연구자 수요에 따른 분석환경 구축 제공
 - 한국과학기술정보연구원 슈퍼컴퓨터를 이용한 원격 접속기반의 데이터 분석진행
 - (분석결과 반출) 다중오믹스 정보시스템내에서 분석완료 후 외부로 반출되는 분석결과에 대한 검수 진행, 검수완료 후 연구자에게 분석결과 전송

< 인체 자원 및 데이터 분양·활용 절차 >



- 기타 문의사항은 아래 연락처로 문의
 - 담당 : 국립중앙인체자원은행 박슬기 (043-719-6522, biobank@korea.kr)

붙임 6

다중오믹스 정보시스템 안내

- (신청안내) 다중오믹스 데이터 활용신청 및 분석 인프라 신청은 보건의료 연구자원정보센터(CODA, <https://www.coda.nih.go.kr>)에서 가능하며,
 - 데이터 분석 인프라에서는 분석 파이프라인, Shell, R-Studio, Jupyter, 시각화를 이용하여 사용자의 원하는 분석을 수행할 수 있음

데이터 활용 신청

○ 다중오믹스 데이터 조회 및 분양 신청

인프라 활용 신청

○ 다중오믹스 데이터 분석 인프라 활용 신청

- (데이터분석 환경) 분석 파이프라인에서는 다중오믹스 데이터 (WGS/scRNA-seq /COVID-Seq/Bulk(TCR/BCR)RNAseq/HLA typing) 분석 파이프라인을 활용할 수 있으며, 분석 인프라 또한 활용하여 분석할 수 있도록 지원
- (분석 파이프라인) 웹을 통해서 사전정의된 파이프라인을 이용하여 기본분석 진행 가능(리눅스 명령어 및 간단한 프로그래밍 언어 능력 필요)

데이터 분석 파이프라인

□ 다중오믹스 데이터 분석

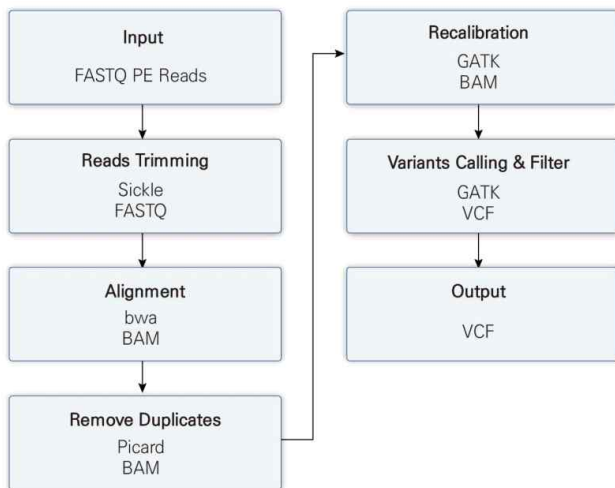
▶ 분석 정보

분석번호*	분석번호는 자동으로 생성됩니다.
분석명*	
참조서열*	선택하세요
분석구분*	선택하세요
분석 컨테이너*	분석 인프라 조회 버튼을 클릭하세요. 분석인프라조회

▶ 연구대상자 정보

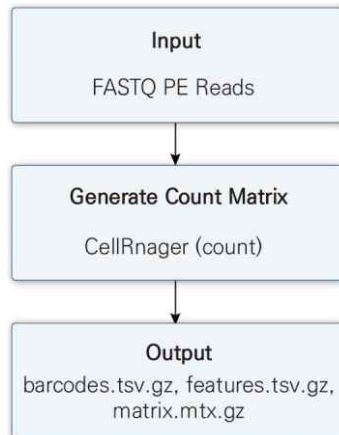
<input type="checkbox"/> 연구대상자 고유번호	연령	성별	건강상태	증상정도	기저질환	시점	WGS	scRNA-Seq	COVID-Seq	HLA typing	Cytokine Profiling	혈액검사	TCR,BCR	임상역학 정보
<input type="checkbox"/> COV-CCO-001	62	남성	COVID19	경증	무	1	0	0	0	0	0	0		
<input type="checkbox"/> COV-CCO-001	62	남성	COVID19	경증	무	2					0			
<input type="checkbox"/> COV-CCO-001	62	남성	COVID19	경증	무	3		0			0			
<input type="checkbox"/> COV-CCO-002	65	여성	COVID19	경증	무	1	0		0	0	0	0		
<input type="checkbox"/> COV-CCO-002	65	여성	COVID19	경증	무	2					0			
<input type="checkbox"/> COV-CCO-002	65	여성	COVID19	경증	무	3					0			

○ WGS 분석 프로세스



- WGS 분석 파이프라인은 Fastq 형식의 시퀀싱 데이터를 입력 파일로 받아 reads trimming 후 사용자가 선택한 참조서열에 alignment 수행 후 변이 정보를 탐색하여 VCF형식의 파일을 출력

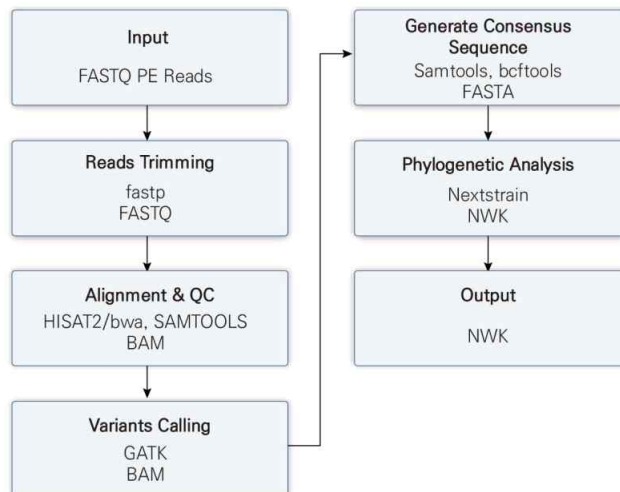
○ scRNA-seq 분석 프로세스



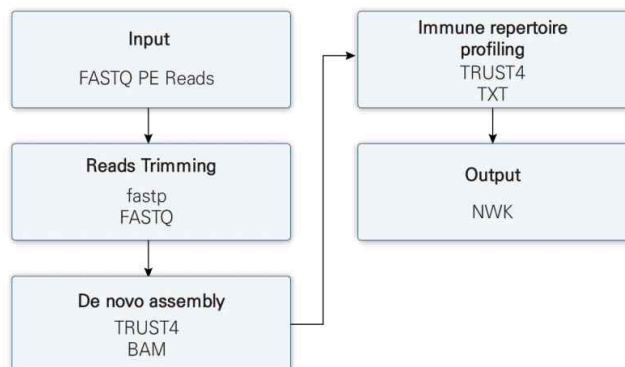
- Fastq 형식의 시퀀싱 데이터를 입력 파일로 받아 사용자가 선택한 참조서열에 alignment 수행후 각 바코드 서열에 대한 유전자 영역에 맵핑된 reads count 데이터를 텍스트 파일로 출력

○ COVID-Seq 분석 프로세스

- COVID 염기 서열 Fastq를 이용해 계통 트리 분석을 진행하여 NWK 형식의 계통분석 결과 파일(.vcf) 출력

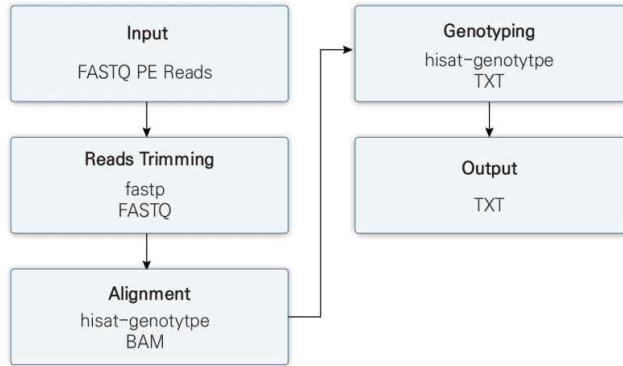


○ Bulk(TCR/BCR) RNAseq 분석 프로세스



- Bulk(TCR/BCR) RNAseq의 Fastq를 입력받아 사용자가 선택한 참조서열에 De novo assembly 수행 후 Immune repertoire profiling 수행 결과인 TSV 형식의 통계 파일 출력

○ HLA typing 분석 프로세스

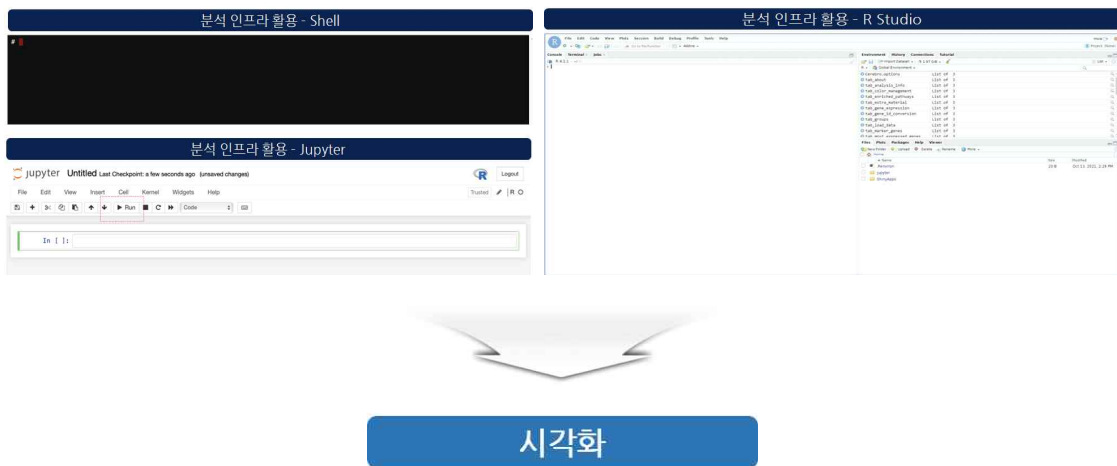


- HLA typing Fastq를 입력받아 사용자가 선택한 참조서열에 alignment 수행 후 유전자 영역에 맵핑된 csv 형식의 샘플별 HLA Allele 정보 파일 출력

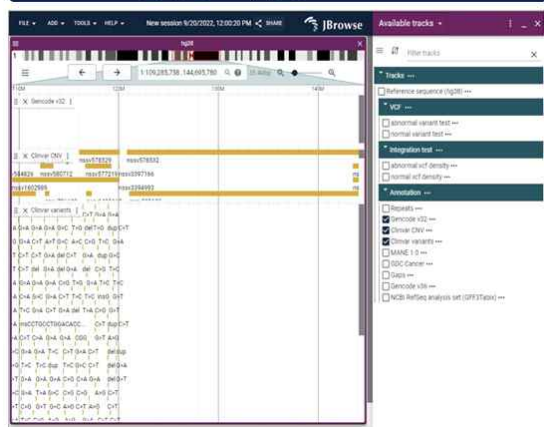
○ (분석 인프라) 분석 인프라(Linux shell, R studio, Jupyter, 시각화 등)를 활용하여 연구자가 주체적 분석 수행

데이터 분석 인프라

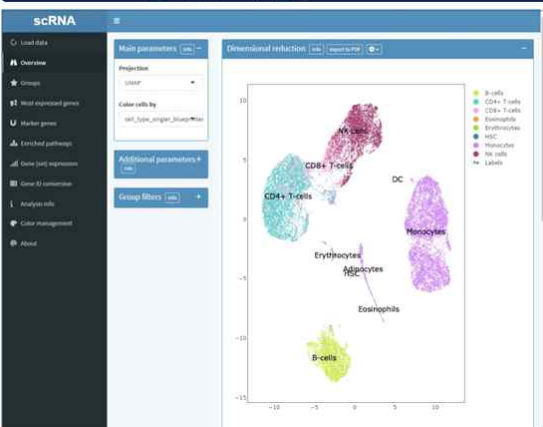
□ 다중오믹스 데이터 분석(WGS, scRNA, RNA, 통계분석, 시각화 등)



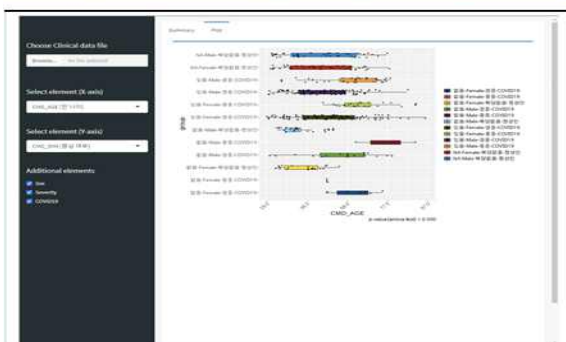
분석 인프라 활용 - JBrowse



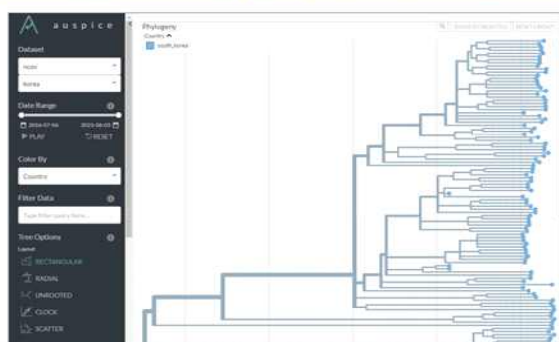
분석 인프라 활용 - scRNA



분석 인프라 활용 - 임상·역학정보



분석 인프라 활용 - COVID-Seq



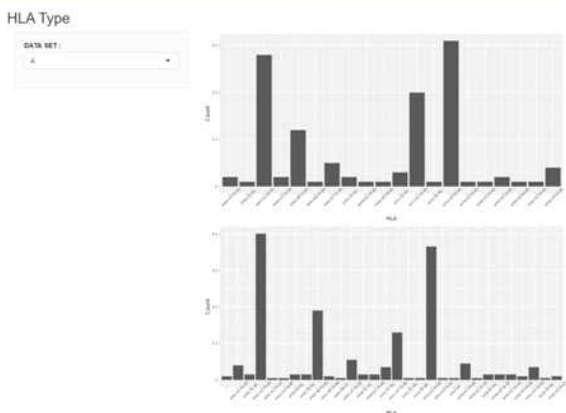
분석 인프라 활용 - 혈액 검사 정보



분석 인프라 활용 - Cytokine



분석 인프라 활용 - HLA



분석 인프라 활용 - Bulk BCR/TCR

