

*Clinical & Omics Data Archive*

# 임상유전체 생명정보시스템

국립보건연구원 생명정보연구과  
조성범



# 목차

## 1. 사업 소개

- 사업 개요 (~ 17년도)
- 사업 배경 및 목적

## 2. CODA 시스템

- 시스템 구성도
- 메뉴 구조도

## 3. 아카이브 현황

- 데이터 보유 현황
- 2018년도 운영 현황

## 4. CODA 기능 고도화 내용

- 대용량 파일 전송 체계 구축
- 응용프로그램 기능 개선
- 웹 디자인 개선
- 보안 취약점 개선 및 장애 대응

## 5. 요약



# I . 추진전략

## 1. 서비스 기반 마련

- Data Sharing을 위한 연구자 공감대 형성 및 관련제도 개선
- 운영규정 및 데이터 유형별 표준안 마련
- 시스템 구축 및 운영
- 관련위원회 구성 및 운영

## 2. 서비스 기반 강화

- Data Sharing 정책의 확대
- 시스템 고도화
- 대용량 연구데이터 활용을 위한 연구환경 마련 (컴퓨팅 자원, 네트워크 환경, 분석지원서비스 연계 등)

## 3. 서비스 안정화

- 고도화 시스템의 운영

구축기  
(‘14~’16)

강화기  
(‘17~’19)

안정화기  
(‘20~’21)



## Ⅱ. 추진경과

구 분	사업 내용	장비 구축
1년차('14년)	<ul style="list-style-type: none"> <li>- 임상연구 및 유전체 염기서열 데이터 관리시스템 개발</li> <li>- 공개된 연구데이터의 분양신청, 심의 및 분양 기능 개발</li> </ul>	방화벽, 침입방지시스템, 검증서버, 네트워크 스위치, 고속데이터 전송서버 및 솔루션, 스토리지(96TB)
2년차('15년)	<ul style="list-style-type: none"> <li>- 유전체 데이터 유형 확대               <ul style="list-style-type: none"> <li>· DNA microarray, epigenome-seq, targeted-seq</li> </ul> </li> <li>- 국가생명연구자원정보센터(KOBIC)와의 정보연계를 위한 유전체 메타데이터 체계 수정·보완 및 시스템 기능 개선</li> <li>- 등록 및 분양 절차 간소화               <ul style="list-style-type: none"> <li>· 운영기반 마련(데이터공유계획서 제출 의무화 및 운영가이드 마련)</li> </ul> </li> <li>- 보건의료종합정보시스템(HTDream)의 과제정보 연계</li> </ul>	<ul style="list-style-type: none"> <li>- 장비 이중화 : 방화벽, 침입방지시스템, 검증서버, 네트워크 스위치               <ul style="list-style-type: none"> <li>*장비 이중화 목적: 단일장비 구성으로 장애 발생시 복구 및 처리시간 과다 소요에 따른 업무 단절 최소화</li> </ul> </li> <li>- 백업관리서버, 백업소프트웨어, 백업스토리지(테이프방식,,225TB), 백업 라이선스 증설, SAN스위치 라이선스 증설, 스토리지(150TB)</li> </ul>
3년차('16년)	<ul style="list-style-type: none"> <li>- CODA 시스템 운영 개시 및 홍보</li> <li>- 포스트게놈 다부처 유전체 사업 등의 유전체정보 수집</li> <li>- 시스템 기능 고도화               <ul style="list-style-type: none"> <li>· 질량분석기술 기반 단백질체·대사체 데이터 관리기능 추가</li> <li>· 통계기능, 검증기능 고도화, 영문홈페이지 구축, Pubmed 문헌 정보 연계</li> <li>· 시료(specimen)단위 데이터 검색, 제공</li> </ul> </li> </ul>	- 스토리지 증설(200TB)
4년차('17년)	<ul style="list-style-type: none"> <li>- 임상오믹스 데이터의 체계적인 수집</li> <li>- 등록 분양 활성화</li> <li>- 시스템 개선·보완               <ul style="list-style-type: none"> <li>· 대용량 파일전송을 위한 고속 전용 네트워크 환경 구성</li> <li>· 응용프로그램 기능(사용자·관리자) 개선 및 고도화</li> <li>· 보안취약점 및 소스 품질 개선</li> </ul> </li> </ul>	- 스토리지 증설(150TB)



### Ⅲ. 2017년도 주요 추진내용

과제명	맞춤의료 통합정보 자원화
총 연구기간	2017.1.1.~2017.12.31. (1년)
총 연구비	280,000(천원)
예산과목	포스트게놈 다부처 유전체사업
연구목표	<ol style="list-style-type: none"> <li>1) 연구데이터의 체계적인 수집</li> <li>2) 연구 결과물 등록·분양 활성화</li> <li>3) CODA 시스템의 지속적인 개선·보완</li> </ol>



# <CODA 시스템 기능 고도화>

## 1) 메인페이지

### Before

**Before Main Page Features:**

- Search bar with a dropdown menu.
- Recent studies list with columns for Study and Sample.
- Navigation links for NGS, DNA microarray, and Clinical research.
- Footer with contact information and logos.

### After

**After Main Page Features:**

- Enhanced search bar with filters for NGS, DNA microarray, and Clinical study.
- Large data visualization (pie chart) showing the distribution of 1094 samples across different study types.
- Table with columns: Study, Participants, Samples, Files, and File size.
- Updated footer with contact information and logos.

<http://coda.nih.go.kr>



## 2) 영문페이지

### 소개페이지

[About CODA](#)   [Search](#)   [Usage guide](#)   [Related references](#)   [Statistics](#)

#### Introduction ➤ > About CODA > Introduction

The Clinical and Omics Data Archive (CODA), the national repository, was launched in 2016 to collect, deposit and provide clinical research and human omics data. The goal of operating CODA is to support healthcare studies to prepare the new era of personalised, predictive, preventive and participatory medicine, and encourage the data-sharing policy about the researches supported by the national R&D fund.

**Collection**: Raw and processed data of clinical and personal omics studies from public fund or donated data voluntarily from

**Preservation**: Central and public archive for clinical and human omics data

**Data sharing**: Provision of released data and establishment of research infrastructures required for data sharing

### 검색 및 상세화면

🔍 Search > All > All

All ▼

Search

☒ NGS
☒ DNA microarray
☐ Clinical study

NGS: 24 Read More ➤

More information	
R000209	<p>A clinical survey and cy condition</p> <p><b>Submission date</b> 2016-07-27</p> <p><b>Project</b> CODA accession number: R000209 Project number: HD1582756 Project title: A clinical survey and genomic analysis of alcohol-related death and injury in emergency condition Principal investigator: Jhn-Geun Choi Organization: Hallym University Industry Academy Cooperation Foundation Public release date: 2016-06-05</p>
R000186	Development of GW
DNA microarray: 6	
R000325	<p>Prospective obserati- stry for Organ Transp</p> <p><b>Experiment</b> Phenotype or Disease: Alcohol Data type: Whole exome-seq Sequencing platform: NGSex File format: Fastq Reference genome: Mapping algorithm: Calling algorithm:</p>
R000786	Prospective obserati- stry for Organ Transp
<p><b>Samples</b> Composition of samples: Patient : 11, All : 11 Age: -1 ~ 54</p>	

- CODA 인용논문, 해외저널(BMC bioinformatics) 게재(2017)

※ 논문제목: HLAscan: genotyping of the HLA region using next-generation sequencing data(2017), Sojeong Ka1†, Sunho Lee2†, Jonghee Hong1, Yangrae Cho2, Joohon Sung3, Han-Na Kim4, Hyung-Lae Kim4\*and Jongsun Jung2\*

#### Availability of data and materials

Sequencing data for families #5–#9 (37 individuals) used in this study are deposited in the Clinical Omics Data Archive (CODA, <http://codan.nih.go.kr>), but restrictions apply to the availability of these data, and they are not publicly available. However, all data obtained and/or analyzed during the current study are available from the authors upon reasonable request. HLAscan is available at [http://www.genomekorea.com/display/tools/HLA\\_SCAN](http://www.genomekorea.com/display/tools/HLA_SCAN).

※CODA 시스템이 해외저널에서 공공 DB로 인정받은 첫번째 논문



### 3) 샘플단위 검색

#### 샘플단위 기본검색화면

NGS

기본검색

상세검색

샘플 단위

연구 단위

전체

데이터ID, 데이터 타입, 나이, 성별, 샘플구분, 질환명, 파일포맷, 공개일에서 검색어가 포함되는 항목

검색

(전체: 1,094) 전체선택: ☐

관심 데이터

정렬: 최근승인일

↺

<input type="checkbox"/>	데이터ID	데이터유형	나이	성별	샘플구분	질환명	검체	파일포맷	공개일	상세
<input type="checkbox"/>	R000209-00000001	Whole exome-seq	나이 불명	남(M)	Patient	당뇨	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000005	Whole exome-seq	1	남(M)	Patient	당뇨	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000009	Whole exome-seq	31	여(F)	Patient	당뇨	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000011	Whole exome-seq	54	여(F)	Patient	당뇨	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000003	Whole exome-seq	34	여(F)	Patient	당뇨	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000008	Whole exome-seq	23	여(F)	Patient	당뇨	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000002	Whole exome-seq	46	여(F)	Patient	당뇨	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000004	Whole exome-seq	45	남(M)	Patient	당뇨	혈액	Fastq	2016-06-05	
<input checked="" type="checkbox"/>	R000209-00000010	Whole exome-seq	나이 불명	여(F)	Patient	당뇨	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000007	Whole exome-seq	54	남(M)	Patient	당뇨	혈액	Fastq	2016-06-05	

⏪

⏩

1

2

3

4

5

6

7

8

9

10

⏪

⏩

#### 샘플단위 상세검색화면

NGS

기본검색

상세검색

샘플 단위

연구 단위

검색조건

질환명

당뇨

+

OR

검체

혈액

+

-

AND

파일포맷

Fastq

+

-

검색

(전체: 311) 전체선택: ☐

관심 데이터

정렬: 최근승인일

↺

<input type="checkbox"/>	데이터ID	데이터유형	나이	성별	샘플구분	질환명(국문)	질환명(영문)	검체	파일포맷	공개일	상세
<input type="checkbox"/>	R000209-00000001	Whole exome-seq	나이 불명	남(M)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000005	Whole exome-seq	1	남(M)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000009	Whole exome-seq	31	여(F)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000011	Whole exome-seq	54	여(F)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000003	Whole exome-seq	34	여(F)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000008	Whole exome-seq	23	여(F)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	
<input checked="" type="checkbox"/>	R000209-00000002	Whole exome-seq	46	여(F)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000004	Whole exome-seq	45	남(M)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	
<input checked="" type="checkbox"/>	R000209-00000010	Whole exome-seq	나이 불명	여(F)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	
<input type="checkbox"/>	R000209-00000007	Whole exome-seq	54	남(M)	Patient	당뇨	Diabetes	혈액	Fastq	2016-06-05	

⏪

⏩

1

2

3

4

5

6

7

8

9

10

⏪

⏩



## 4) 데이터신청

### 데이터신청 목록

#### 오믹스

🏠 > 나의CODA > 데이터 신청 > 오믹스

전체

신청접수번호,과제명,과제번호 에서 검색어가 포함되는 항목을 조회합니다

검색

- 데이터 신청을 완료(취소)하려면, 신청 건 선택 후 신청완료(취소) 버튼을 누르세요.  
- 신청취소는 심의대기 상태 일까만 가능합니다

전체: 34 ( 데이터신청중 : 13 , 심의대기 : 9 , 심의중 : 4 , 처리완료 : 8)

신청완료

신청취소

10개씩 보기

<input type="checkbox"/>	S000327	국내 인플루엔자 예방 및 관리 기술 지원을 위한 마우스 병원성 인플루엔자 바이러스주 개발 ( Development of mouse adapted influenza virus for technical support in korea influenza control and prevention )	데이터신청대기
		과제번호 : HD1480027 데이터 이용 동의서 혹은 연구계획 IRB 심의 결과서가 누락 되었습니다.	X
<input type="checkbox"/>	S000326	장기이식프로토포트 구축 및 운영 ( Prospective observational study of patients with organ transplantation ( National Registry for Organ Transplantation ) )	데이터신청대기
		과제번호 : HD14A0006 데이터 이용 동의서 혹은 연구계획 IRB 심의 결과서가 누락 되었습니다.	X
<input type="checkbox"/>	S000324	대기오염에 대한 건강피해 연계성 연구 ( A study on the national and local health outcome from ambient air quality )	심의대기
		과제번호 : HD14A1132	

### 데이터 신청상세

#### 데이터 신청서-오믹스

##### 데이터 신청 접수정보

신청접수번호	S000326	신청자	규리전라지
신청일		신청상태	데이터 신청중
승인일	심의완료 후 내역이 출력 됩니다	데이터 다운로드 기간	심의완료 후 내역이 출력 됩니다
접수수	2		
파일사이즈	등록된 파일이 없습니다.		

##### 신청자 정보

\*필수입력

##### 과제(연구)정보

☐ 직접입력 (직접 입력한 정보는 내과에 갱신해 자동저장 됩니다.)

과제번호 HD14A0006

##### 과제상태

##### 과제/연구 제목(타

- 데이터 이용 동의서, 연구계획 10% 심의결과서는 필수 첨부 사항입니다.  
(관련서식)데이터 이용 동의서.docx

##### 과제/연구 제목(연구

##### 연구책임자(국문)

##### 연구책임자(영문)

##### 연구수행기관(국)

##### 연구수행기관(영)

##### 승인구간

##### 당해년도연구기간

##### 수행연도

##### 보안과제여부

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 연구요약

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 데이터 활용기간

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터

##### 신청 데이터



## 5)Metabolomics 등록

### Metabolomics 목록화면

Metabolomics ☞ 나의CODA > 오믹스 > Metabolomics

전체

CODA 등록번호,과제명,과제번호,연구기관 에서 검색어가 포함되는 항목

검색

등록을 시작하려면 아래의 새로운등록 버튼을 누르세요.

저출생등록 신청 후, 저출생등록을 누르면 자동적으로 검증과정을 거쳐 승인요청이 됩니다.

검출 또는 승인 중에는 등록내용을 수정할 수 있습니다.

검출승인나 승인반려됨에이하는 상세정보에서 수정 후 다시 제출할 수 있습니다.

검출승인나 내용은 "기" 버튼을 클릭하면 확인할 수 있습니다.

전체: 14

새로등록

제출

공개

10개씩 보기

R000607

장기이식코hort 구축 및 운영 ( Prospective observational study of patients with organ transplantation (National Registry for Organ Transplantation) )

승인일(최근 수정일) : 2016-06-27

오믹스데이터(Data Set,Etc.Identifications/characterizations/modifications/Search engine output files),Peak, Quantification(Quantification software output files),kik,mzData,mzML,mzXML을 등록해 주세요.

기탁동일서 및 제3자 계공동의증빙서류가 누락 되었습니다.

승인중

R000605

항암발현 저조합 메르스 바이러스 전장플로 유효성 평가 ( Development of fluorescence expressing MER S-CoV for drug and vaccine evaluation )

승인일(최근 수정일) : 2016-06-05

오믹스데이터(Data Set,Etc.Quantification(Quantification software output files),mzData,mzML,ument output files)를 등록해 주세요.

기탁동일서 및 제3자 계공동의증빙서류가 누락 되었습니다.

작성중

R000604

경년기 여성에서 신체7장신 건강 상태 및 위험요인 규명을 위한 전향적 연구 ( Prospective study on the risk factors for physical and mental disorders in post-menopausal women )

승인일(최근 수정일) : 2016-06-23

오믹스데이터(Data Set,Etc.Quantification(Quantification software output files),es,mzData,mzML)을 등록해 주세요.

기탁동일서 및 제3자 계공동의증빙서류가 누락 되었습니다.

작성중

R000588

국내 국소저체중아 관리지표 생산을 위한 레지스트리 구축 및 운영 ( Development of Korean neonatal network to generate treatment indicators for the very low-birth-weight infants )

승인일(최근 수정일) : 2016-06-22

오믹스데이터( Instrument output files,Etc,HMD8,Identifications/characterizations/modifications/Search engine output files),Quantification(Quantification software output files),mzXML,other file formats)을 등록해 주세요.

기탁동일서 및 제3자 계공동의증빙서류가 누락 되었습니다.

승인완료 (비공개)

### Metabolomics 등록화면

Metabolomics 등록 + 필수입력

과제정보

실험정보

오믹스데이터

기본 정보(직접입력 또는 HTDream 검색 후 자동입력, 역)

과제번호

과제명

과제/연구 제목(국문)

과제/연구 제목(영문)

연구책임자(국문)

연구책임자(영문)

연구수행기관(국문)

연구수행기관(영문)

출연기관

당해년도 연구기간

수행연도

보전과제여부

키워드

연구요약 정보

공개

공개일

연락처

개인정보 수집 여부 동의(필수)

통제되지 서비스 및 행사 참여(선택)

실험정보

실험제목

실험설계

연구대상 질환

Sample preparation protocol

Mass spectrometry

데이터 파일 유형

Reference database

Analysis software & version

Data processing protocol

실험제목

실험설계

연구대상 질환

Sample preparation protocol

Mass spectrometry

데이터 파일 유형

Reference database

Analysis software & version

Data processing protocol

제출

연구



## 6) Proteomics 등록

### Proteomics 목록화면

#### Proteomics

🏠 > 나의CODA > 오믹스 > Proteomics

전체

CODA 등록번호, 과제명, 과제번호, 연구기관 에서 검색어가 포함되는 항목

검색

- 등록을 시작하려면 이전의 새로운 등록 버튼을 누르세요.
- 제출할 등록 신청 후, 제출 버튼을 누르면 자동으로 검토과정을 거쳐 승인요청이 합니다.
- 검토 또는 승인 중에는 등록내용을 수정할 수 없습니다.
- 검토되거나 승인반려된 데이터는 상세정보에서 수정 후 다시 제출할 수 있습니다.
- 검토되거나 내용은 '기' 버튼을 클릭하면 확인할 수 있습니다.

전체: 11

새로등록 제출 공개 10개씩 보기 ↻

<input type="checkbox"/>	R000606	장기이식코호트 구축 및 운영 ( Prospective observational study of patients with organ transplantation (National Registry for Organ Transplantation ) )	작성중
		승인일(최근 수정일) : 2016-06-23	
		실험정보를 등록해 주세요	
<input type="checkbox"/>	R000589	데이터 표준화에 기반한 omics 연구용 질량분석기 자원확보 ( Collecting Biospecimen Panels with Standardized Formats for Omics Researches )	승인완료 (비공개)
		승인일(최근 수정일) : 2016-06-22	
		오믹스데이터 (output filePeak list output fileIdentifications.ce database.mzXML)를 등록해 주세요.	
		기탁동의서 및 제3자 제공동의용발서류가 누락 되었습니다.	
<input type="checkbox"/>	R000580	경년기 여성에서 신체?정신 건강 상태 및 위험요인 규명을 위한 전향적 연구 ( Prospective study on the risk factors for physical and mental disorders in post-menopausal women )	승인중
		승인일(최근 수정일) : 2016-06-23	
		오믹스데이터(Data Set, Identifications, mzData, txt,txt1)를 등록해 주세요.	
		기탁동의서 및 제3자 제공동의용발서류가 누락 되었습니다.	

### Proteomics 등록화면

#### Proteomics 등록

• 필수입력

기본 정보(의입력 또는 HTDREAM 검색 후 자동입력, 직접 입력한 정보는 내과제에 자동저장 됩니다.)

직접입력

과제정보

실험정보

오믹스데이터

과제명

실험명

오믹스데이터

#### Proteomics 수정

• 필수입력

과제정보

실험정보

오믹스데이터

과제명

실험명

오믹스데이터

과제정보

실험정보

오믹스데이터

과제명

실험명

오믹스데이터

입력완료 닫기

저장 닫기



## 7) 등록성과물

### 논문 목록 및 등록화면

🔍 > 나의CODA > 등록활용성과 > 논문

전체

제목,저널명,저자,디지탈ID,PubMedID,관 호/페이지 에서 검색어가 포함!

검색

전체: 4

새로등록

삭제

다운로드

10개씩 보기

↺

■ The acute and sub-acute toxicity of C60/PVP complex in vivo

Baek

The acute and sub-acute toxicity of C60/PVP complex in vivo

PubMedID : DOI : NCMKID : CODA등록번호 : R000009

논문 상세조회

✕

• 등록신청

○ 논문

• 제목

The acute and sub-acute toxicity of C60/PVP complex in vivo

• 저널명

Adv Nano Res

• 저자

Marina A. Dumpis, Viktor V. Ilin, Elena V. Litavova, Dmitriy N. Nikolsaev, Valentina. Krylova, Irina V. Okunevich

• 관호/페이지

167-179

DOI

10.1802/gpe.22032

PubMedID

27990689

디지탈아이덴티티ID

• 초록

The detailed study of acute and sub-acute toxicity of the complex polyvinylpyrrolidone (PVP 20 kDa)-wrapped fullerene C60 after intraperitoneal (ip) administration was carried out on adult male Wistar rats. The LD50 value of C60/PVP complex was found to be 7, 8 g/kg. In sub-acute study which lasted for 30 days the rats were exposed to daily administration of the complex in the doses of 350 or 700 mg/kg. All animals survived during the study and had no significant changes in clinical signs, organ weight, hematological and biochemical parameters of blood. The electrophysiological properties of myocardium and the excretory function of kidney remained normal. Histological analysis of liver, kidney and spleen at the end of the study also did not demonstrate toxic alterations. It was thus established that intraperitoneal administration of complex C60/PVP has no toxic effect. These results suggest that C60/PVP has no acute and sub-acute toxicity and is a perspective substance for potential application in biology and medicine.

○ 등록받은 정보

필요내역선택

CODA등록번호	수정연도	과제번호	과제명	실험명	데이터(연구) 유형	상태
R000000	2016	T001088	신변중 병용제 및 인체유동...		Clinical study	✗
R000246	2016	T001374	노인전식 환자의 인체유동...	32세 10월13일 ~ 10월14일	NGS	✗
R000269	2016	T001390	노인전식 환자의 인체유동...	ddddd	NGS	✗
R000325	2016	T001471	장기미시프로호트 구획 및 분...	321	DNA-microarray	✗
R000785	2016	T001086	장기미시프로호트 구획 및 분...		Clinical study	✗
R000786	2016	T001087	장기미시프로호트 구획 및 분...	321	DNA-microarray	✗

합계 : 개

작성

삭제

### 학회/심포지엄 목록 및 등록화면

🔍 > 나의CODA > 등록활용성과 > 학회/심포지엄

전체

제목,발표자,페이지 에서 검색어가 포함되는 항목을 조회합니다.

검색

전체: 2

새로등록

삭제

다운로드

10개씩 보기

↺

■ The effect of glucagon-like peptide-1 and glucagon-like peptide-2 on microcirculation: a systematic review.

W. Kim Dong-hoon

2017-03-07

CODA등록번호 : R000246, R000206

학회/심포지엄 상세

✕

• 등록신청

○ 학회/심포지엄

• 제목

The effect of glucagon-like peptide-1 and glucagon-like peptide-2 on microcirculation: a systematic review.

• 게재일자

2017-03-07

• 발표자

Kim Dong-hoon

발표형태

Poster \* Spoken

페이지

17

• 초록

Glucagon-like peptide-1 (GLP-1) and -2 (GLP-2) are gut-derived hormones used in the treatment of diabetes type-2 and short bowel syndrome, respectively. GLP-1 attenuates insulin resistance and GLP-2 reduces enterocyte apoptosis and enhances crypt cell proliferation in the small intestine. In addition, both hormones have vasodilatory effects and may be useful in situations with impaired microcirculation. The aim of this systematic review was to provide an overview of the potential effects of GLP-1 and GLP-2 on microcirculation.

○ 등록받은 정보

필요내역선택

CODA등록번호	수정연도	과제번호	과제명	실험명	데이터(연구) 유형	상태
R000246	2016	T001374	노인전식 환자의 인체유동...	32세 10월13일 ~ 10월14일	NGS	✗
R000269	2016	T001390	노인전식 환자의 인체유동...	ddddd	NGS	✗
R000325	2016	T001471	장기미시프로호트 구획 및 분...	321	DNA-microarray	✗
R000785	2016	T001086	장기미시프로호트 구획 및 분...		Clinical study	✗
R000786	2016	T001087	장기미시프로호트 구획 및 분...	321	DNA-microarray	✗

합계 : 개

작성

삭제



## 8) 통계조회-연도별

### Before

연도별

홈 > 통계 > 등록 > 연도별

성분종류 NGS 수합연도 2015 ~ 2016 검색

다운로드

데이터 유형	등록 건	과제 수	연구대상자 수	샘플 수	데이터 크기
ChIPSeq	0	0	0	0	0
MedipSeq	0	0	0	0	0
RNASeq	1	1	192	384	0
TargetedSeq	1	1	107	107	0
Whole exome	0	0	0	0	0
Whole genome	1	1	75	450	0
합계	3	3	374	941	0

### After

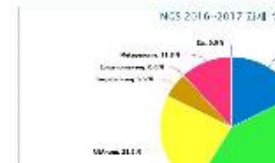
연도별

연도별 > 통계 > 등록 > 연도별

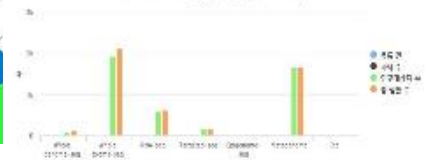
성분종류 NGS 수합연도 2015 ~ 2016 검색



다운로드



NGS 2015-2016 등록 분포 (8,800 Sample)



다운로드

데이터 유형	등록 건	과제 수	연구대상자 수	샘플 수	데이터 크기
ChIPSeq	0	0	0	0	0
MedipSeq	0	0	0	0	0
RNASeq	1	1	192	384	0
TargetedSeq	1	1	107	107	0
Whole exome	0	0	0	0	0
Whole genome	1	1	75	450	0
합계	3	3	374	941	0



# 2018 사업 개요 및 목적

이용자의 만족도 향상 및 효율적 운영



## 임상유전체 생명정보시스템(CODA)



네트워크 중심의 데이터  
수집·공유를 위한  
전산체계 구축

데이터 연계 및  
업무 자동화를 통한  
손쉬운 업무 처리 체계 구축

사용자 편의성 및  
관리 업무 효율성 향상을 통한  
연구지원 활성화



보안 취약점 개선 및  
장애 대응을 통한  
안정적인 시스템 운영

### "임상유전체생명정보시스템 (CODA) 개선 및 기능고도화"

#### 본 사업의 추진 배경

기존 질병관리본부  
인터넷망을 통한 대용량  
파일 전송의 어려움

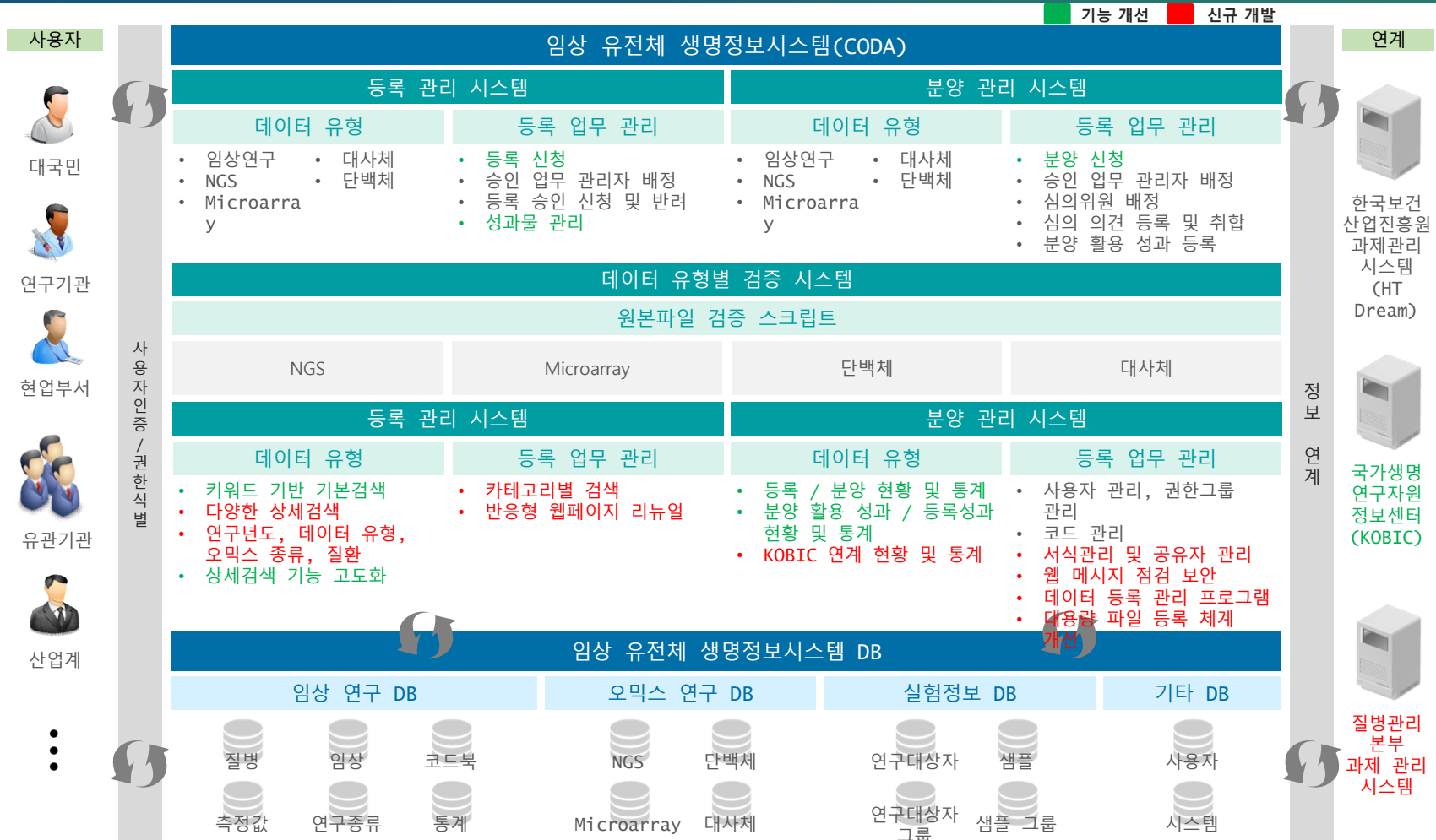
검색, 등록, 신청 등의  
사용자 편의성 개선 필요

자동화 프로세스 보완을  
통한 외부 데이터 수집 및  
연계 효율성 증대 필요

안정적인 시스템 운영을 위한  
장애처리 및 정보시스템  
취약점, 품질 개선 필요

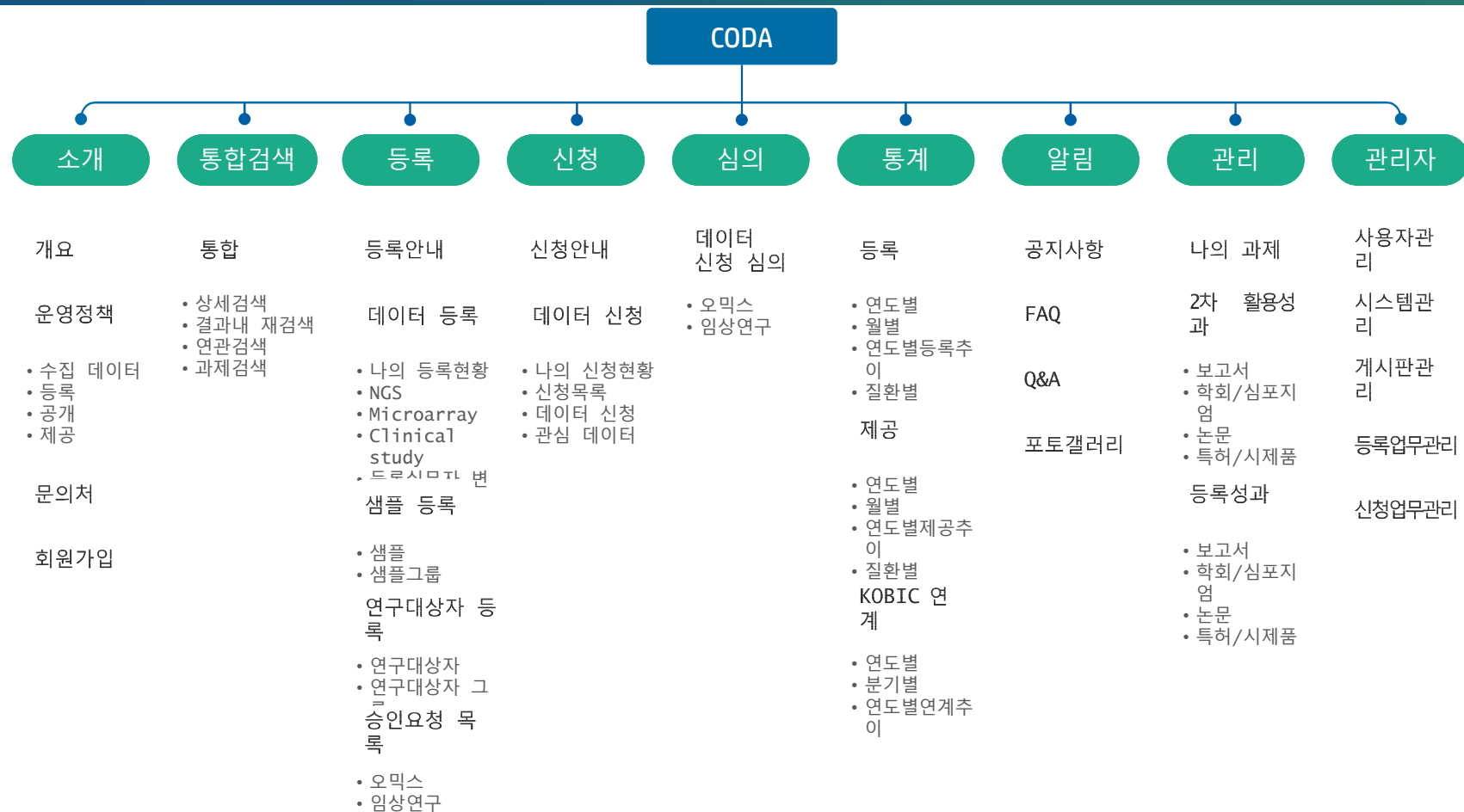


# 시스템 구성도





# 메뉴구조도





# 보유 데이터 현황

- 웹기반 유전체데이터 등록신청 시스템, CODA(Clinical & Omics Data Archive) 운영
  - 포스트게놈 다부처 유전체 사업 등에서 유전체데이터 수집
  - 공개 데이터는 신청을 받아 연구목적에 한하여 제공
  - 누적데이터 기간: '17년 1월 ~ 현재까지
  - CODA신청번호가 다르면 동일 데이터라도 별개의 제공 건으로 처리

	Data type	Project	Datasets	Participants	Samples	Data size
Sequence	WGS	5	7	388	467	139.1TB
	WES	12	20	2,104	2,703	35.8TB
	RNA-seq	5	8	596	639	2.1TB
	Targeted-seq	5	5	543	543	728.6GB
	Epigenome-seq	1	0	0	12	38.7GB
	Metagenome	2	5	1,668	1,668	393.9GB
	Subtotal	30	45	5,299	6,032	178.2TB
Microarray	Gene expression	2	4	292	292	7.3GB
	SNP detection	3	2	37,012	37,012	916.2GB
	Gene CNV	1	2	94	94	811.6MB
	DNA methylation	2	3	123	131	1.1GB
	Subtotal	8	11	37,521	37,529	925.4GB
	Total	38	56	42,820	43,561	179.1TB



# 2018년도 운영 현황

## 데이터 등록

등록기관명	데이터유형	샘플 수
이화여대, 강북삼성병원	WGS	55
	Metagenome	1290
서울대병원	Targeted-seq	157
	WES	765
한국식품연구원	Epigenome-seq	20
	WES	14
국립보건연구원	RNA-seq	19
	Epigenome-seq	15
전남대학교병원	WES	82
서울대학교	RNA-seq	22
	WES	22
연세대학교	Targeted-seq	209
카톨릭병원	WES	17
합계		2,687



울산과학기술원  
ULSAN NATIONAL INSTITUTE OF  
SCIENCE AND TECHNOLOGY



SNUH 서울대학교병원



가톨릭 병원



KFRRI 한국식품연구원



한림대학교



이화여자대학교  
EWHWA WOMANS UNIVERSITY



가천대 길병원  
Gachon University Gil Medical Center



연세대학교



전남대학교병원



국립보건연구원



강북삼성병원



Humanizing Genomics  
macrogen



서울대학교

## 데이터 제공

등록기관명	데이터유형	샘플 건수	데이터 크기
한림대학교	WES, WGS	809	32.1 TB
마크로젠	WES, WGS, RNA-seq	2,228	68 TB
연세대학교	WES, WGS	247	4 TB
가천대 길병원	Metagenome	1,463	94.9 GB
울산과학기술원	WGS, Gene CNV	539	56 TB
합계		5,286	249.2 TB



# 응용프로그램 기능 개선: 오믹스·임상연구 등록·신청메뉴 고도화

- NGS, 임상연구, 마이크로어레이:
  - 기본정보, 과제정보, 실험정보 등 등록 개선
  - 샘플 단위 데이터 신청 개선
- CDC과제관리시스템 연계
- 실험정보 등록 시 참조데이터 조회
- 샘플정보
  - 개별/일괄등록, 샘플추가 개선
- 데이터 파일 등록 개선
- 연관성과 등록·조회 개선
  - 보고서, 학회/심포지엄, 논문, 특허/시제품
- 첨부파일 등록 개선
  - 기탁등록동의서 및 전자문서
- 등록실무자 변경·조회 및 공유자 설정
- 임상연구: 연구대상자-개별/일괄등록



# 데이터 등록 기능

기본정보	과제정보	실험정보	샘플정보	데이터 파일	연관성과	첨부파일	공유자
* 필수입력							
* 실험제목 (Experiment title)							
* 실험설계 (Experiment design)		0 Byte # 4,000 Byte (한글 3byte, 영문 1byte)까지 가능합니다.					
		<div>ICD-10</div> <div>직접입력</div>					
* 연구대상 질환 (Targeted phenotype/disease)		<div>국문</div> <div>영문</div>					
* 데이터 유형 (Data type)		선택					
* 시퀀싱 플랫폼 (Sequencing platform)		선택					
* 시퀀싱 기관 (Sequencing center)		영문					
* 반복실험 유형 (Replicates type)		선택					
라이브러리 정보 (Library information)		* 라이브러리제작방법 (library construction description) 0 Byte # 4,000 Byte (한글 3byte, 영문 1byte)까지 가능합니다					
		* 어댑터 시퀀스 (scanning protocol) 0 Byte # 4,000 Byte (한글 3byte, 영문 1byte)까지 가능합니다					
		* Read type 선택					
		* Read length(bp)					
* Raw data		<div>Fastq</div> <div>BAM</div>					
Analysis data		<div>VCF</div> <div>BED</div>					
* 연관데이터 등록번호		Fastq 찾아 등록하기					

<실험정보 등록 화면>

CODA등록번호	수행연도 검색				
과제명					
데이터 유형	Whole genome-seq	Whole exome-seq	RNA-seq	Targeted-seq	
	Epigenome-seq	Metagenome	Etc		
파일 타입	Fastq	BAM	Etc	VCF	
	BED	ETC			
검색					

선택	CODA...	수행연도	과제명	데이터 종류	파일유형
<input type="radio"/>	R001585	2014	차세대 맞춤의료 유전체 사업단 태...	Whole exome-seq	Fastq,BAM
<input type="radio"/>	R001584	2017	유전체 빅데이터 기반 폐암 위험예측...	Targeted-seq	Fastq
<input type="radio"/>	R001562	2017	유전체 빅데이터 기반 폐암 위험예측...	Whole genome-seq	Fastq
<input type="radio"/>	R001542	2017	유전체 빅데이터 기반 폐암 위험예측...	Whole genome-seq	Fastq
<input type="radio"/>	R001541	2017	유전체 빅데이터 기반 폐암 위험예측...	Whole genome-seq	Fastq,BAM
<input type="radio"/>	R001540	2017	유전체 빅데이터 기반 폐암 위험예측...	RNA-seq	Fastq,BAM
<input type="radio"/>	R001517	2017	차세대 염기서열분석 기반의 다유전...	Whole genome-seq	VCF,Fastq
<input type="radio"/>	R001512	2017	유방암 예후예측 도구를 위한 포르말...	Whole exome-seq	Fastq

전체 : 110 개

선택 닫기

<Fastq 검색 및 선택 화면>



# 데이터 신청 기능

## 과제정보시스템 연계 구축

\* 필수입력    □ 직접입력

과제(연구)정보 (입력된 정보는 내과제에 자동저장 됩니다.)

과제번호			질병관리본부 HTDream 내과제
과제형태	선택		
* 과제/연구 제목	국문		
	영문		
* 연구책임자	국문		
	영문		
* 연구기관	국문		
	영문		
총 연구기간	YYYY-MM-DD	~	YYYY-MM-DD
당해 연구기간	YYYY-MM-DD	~	YYYY-MM-DD
수행연도			
보안과제여부	일반과제		

데이터 이용 동의서 IRB 심의결과서

업로드 데이터 이용 동의서 파일을 올려주세요. (PDF 파일)

연구계획 IRB 심의결과서

업로드 연구계획 IRB 심의결과서 파일을 올려주세요. (PDF 파일)

작성

데이터 전달방법 FTP/외장하드 선택

\* 데이터 전달방법 FTP HDD/NAS

\* FTP/PORT CODA가 접속해야 할 등록자의 FTP/PORT 정보입력

신청 파일포맷

\* 신청 파일포맷 BAM CEL Fastq IDAT TXT VCF

- 샘플을 선택하시면 파일 포맷 버튼이 생성됩니다.

신청데이터 샘플단위 검색

추가

제외

<input type="checkbox"/>	데이터ID	실험 유형	데이터 타입	나이	성별	샘플구분	파일포맷	공개일
<input type="checkbox"/>	R001512-000000...	NGS	Whole exome-seq	33	여(F)	Control	Fastq	2018-07-19
<input type="checkbox"/>	R000694-000000...	DNA microar...	Gene expression	12	남(M)	Control	CEL	2018-04-01
<input type="checkbox"/>	R000494-000000...	DNA microar...	Gene expression	24	남(M)	Patient	CEL,IDAT,...	2017-04-07

전체 : 7 개

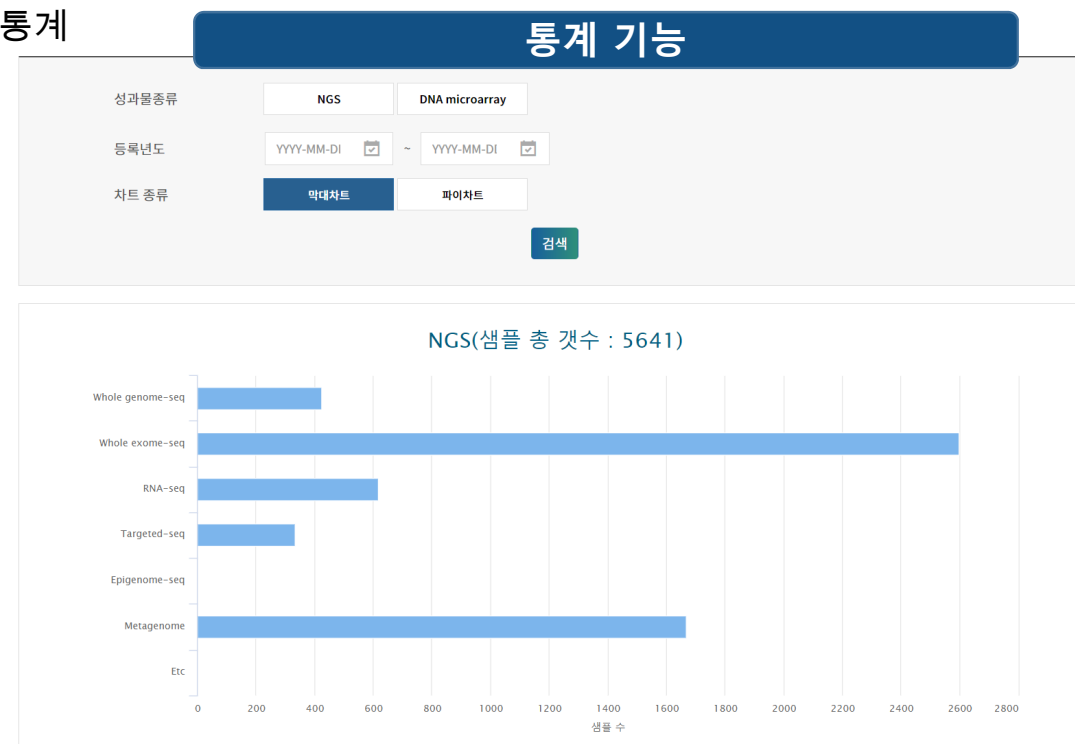
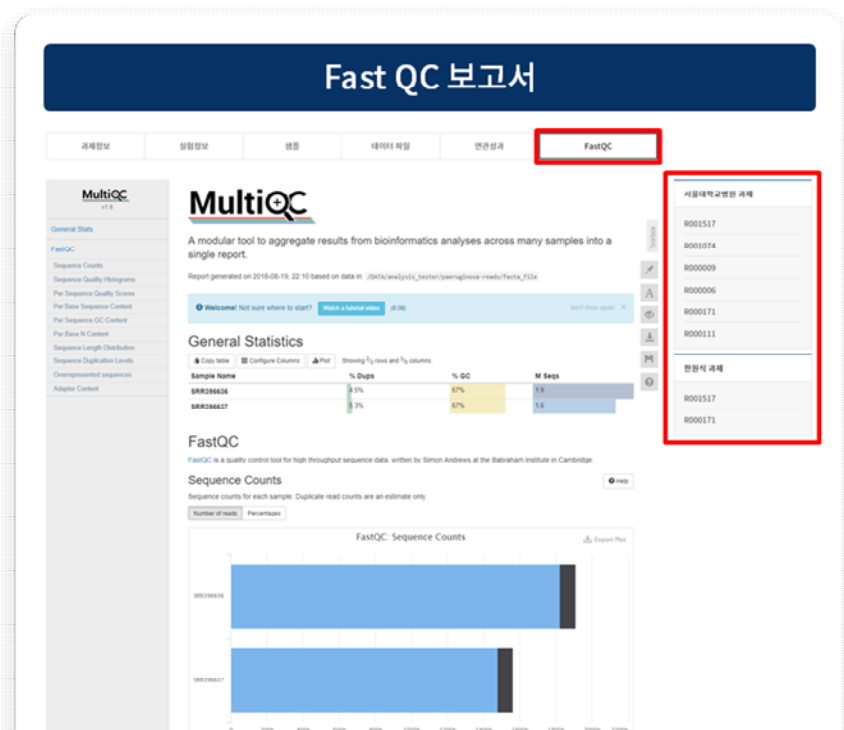
저장

목록



# 응용프로그램 기능 개선: 사용자 기능 고도화

- 검색 기능 개선
  - 통합검색, 결과내 재검색, 상세검색, 검색필터, 연관데이터
- QC 보고서 보기 기능 개선
- 통계 자료 개선
  - 연도별 통계, 월별 통계, 연도별 등록추이, 질환 별 통계





# 통합검색

## 상세검색

검색어를 입력해 주세요.

CODA등록번호

과정명

수행연도

YYYY - YYYY

과정번호

연구기관

연구책임자

연구대상질환

자동완성 기능 제공

공개여부

공개

비공개

유전체연구

Whole genome-seq

Whole exome-seq

RNA-seq

Targeted-seq

Epigenome-seq

Metagenome

Etc

Gene expression

SNP detection

Gene copy number variation

DNA methylation

Etc

임상연구

중재연구 (Interventional Study)

관찰연구 (Observational Study)

파일 유형 및 사이즈

Fastq

BAM

VCF

BED

1

CEL

IDAT

TXT

Flink binary

CSV/Excel

상세검색 열기

통합검색 56건 Clinical Study 0건 **NGS 48건** Microarray 8건

10개씩 보기

전체 56건, 현재페이지 1/6

R000694  
유전체자료 통합 분석 지원센터

R000614  
한국인 특이 유전변이 집을 이용한 한국인 연구집단 유전체정보 구축

## 검색필터

검색어를 입력해 주세요.

상세검색 열기

통합검색 56건 Clinical Study 0건 **NGS 48건** Microarray 8건

10개씩 보기

전체 48건, 현재페이지 1/5

유전체연구 MORE

Whole exome-seq  
Epigenome-seq  
RNA-seq  
Targeted-seq  
Metagenome

연구대상질환 MORE

만성폐쇄성폐질환  
폐선암  
뇌전증  
유방암  
계통 정보 없음

파일 유형 및 사이즈

Fastq  
BAM

연구기관 MORE

1. 서울대학교 의과대학 2. 서울대학교 자연과학대학 3. 삼성유전체연구소 4. 이화여자대학교 시스템생물학 연구소  
삼성서울병원  
한국식물연구회  
이화여자대학교(관자로 생산지 소속 기관:이화여자대학교 의과대학)

R001434  
유전체자료 통합 분석 지원센터

R001414  
K-루드의 글루코분해를 위한 견과기능 영양유전체 연구

R001354  
다만차상질환에 치료한 유전체데이터 분석기술개발을 통한 뇌전증 감수성 및 치료반응 예측인자 규명 및 검증

R001294  
유전체, 전사체 시퀀싱 기반 약물반응에 영향을 미치는 복제수변이, 발현이상 및 유전체원자 분석을 통한 개인 맞춤

R001274  
유전체, 전사체 시퀀싱 기반 약물반응에 영향을 미치는 복제수변이, 발현이상 및 유전체원자 분석을 통한 개인 맞춤

R001218  
약물유전체 다중시퀀싱을 이용한 항암 약물치료 기술개발

R001217

## 연관데이터

과제정보	담당자	생물	데이터 파일	연관문서
CODA 등록번호	R000694			
과제명	유전체자료 통합 분석 지원센터			
과제담당	홍준희			
과제연구 목적	(국립) 유전체정보 통합 분석 지원센터			
연구책임자	(국립) 유전체정보 통합 분석 지원센터			
연구기관	(국립) 유전체정보 통합 분석 지원센터			
총 연구기간	2014-04-01 - 2018-03-31			
남은 연구기간	2014-04-01 - 2018-03-31			
수행연도	2014			
연구목적	인간유전체			
연구자료	(국립) 유전체정보 통합 분석 지원센터			
연구내용	(국립) 유전체정보 통합 분석 지원센터			
자료유형 및 용량	(국립) 유전체정보 통합 분석 지원센터			
파일드	(국립) 유전체정보 통합 분석 지원센터			
연구비(단)	(국립) 유전체정보 통합 분석 지원센터			
공개일	2018-04-01			

과제정보 및 실험정보 PDF 변환



# 응용프로그램 기능 개선: 관리자 기능 고도화

데이터 신청 목록

검색

신청접수번호

과제번호

수행연도

과제명

데이터만료일

진행상태

신청대기

심의대기

심의중

처리완료

상세검색 닫기

안내

- 신청(취소)하려면 신청접수번호 선택 후 '신청(취소)' 버튼을 누르세요.
- 취소는 '심의대기' 상태 일때 만 가능합니다.

신청상황 확인 가능

신청 신청취소 10개씩 보기

전체 32건, 현재페이지 1/4 ( 신청대기 : 15, 심의대기 : 1, 심의중 : 5, 처리완료 : 11)

S000327

✓ 심의대기

차세대 맞춤의료 유전체 사업단 데이터 통합지원  
(Data Center for Genomic Medicine)  
과제번호 : HD14A11321231231

S000326

✓ 신청대기

✕ 삭제

유전체 빅데이터 기반 폐암 위험예측 모델링 기법 개발  
(Development of prediction modeling methods for lung cancer using genomic big data)  
과제번호 : 2017R1D1A1B03029  
데이터 이용 동의서가 누락 되었습니다.

S000325

✓ 처리완료 (승인)

유전체 빅데이터 기반 폐암 위험예측 모델링 기법 개발  
(Development of prediction modeling methods for lung cancer using genomic big data)  
과제번호 : 2017R1D1A1B03029  
데이터 이용 동의서가 누락 되었습니다.

오믹스

등록업무관리

CODA 등록번호

등록자명

과제명

수행연도

과제번호

연구대상 질환

연구기관

연구책임자

데이터유형

진행상태

Whole genome-seq

Whole exome-seq

RNA-seq

Targeted-seq

Epigenome-seq

Metagenome

Etc

Gene expression

SNP detection

Gene copy number variation

DNA methylation

Etc

작성중

업무배정대기

업무배정완료

승인반려

승인완료

검색

다운로드

10개씩 보기

전체 153건, 현재페이지 1/16

R001599

✓ 작성중

✕ 삭제

승인일(최근 수정일) : 2018-09-11  
실험정보를 등록해 주세요.  
기타동의서 및 제3자 제공동의증빙서류가 누락 되었습니다.

R001593

표 임상·오믹스 등록 이력

✓ 승인완료

승인일(최근 수정일) : 2018-09-11

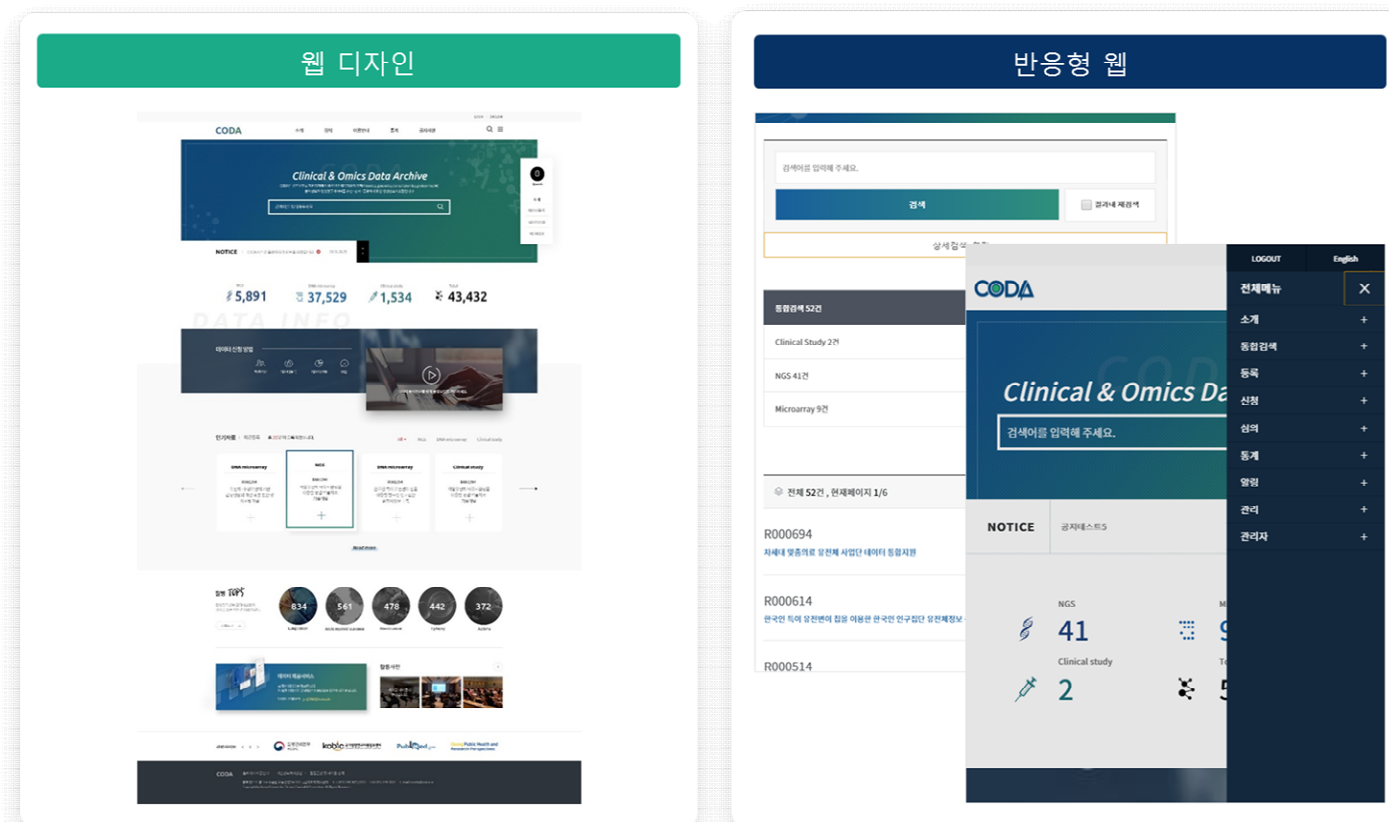
승인완료

- 승인일(최근 수정일): 2018-09-11
- 실험정보를 등록해 주세요.
- 기타동의서 및 제3자 제공동의증빙서류가 누락되었습니다.



# 웹 디자인 리뉴얼

- 홈페이지 개편 및 모바일웹 개발
  - 임상유전체 홈페이지 (<http://coda.nih.go.kr>) 콘텐츠 업데이트 및 웹접근성 인증 개선
  - 반응형 모바일 웹버전으로 개편하여 모바일 이용자에 대한 편의성 개선





# 요약

- 2018년도 CODA 기능 고도화

- 오믹스·임상연구 등록메뉴 고도화
- 사용자 기능 고도화
- 관리자 기능 고도화
- 웹 디자인 리뉴얼
- 보안취약점 개선 및 장애대응

- 기대효과

- 사용자 중심의 시스템 기능·환경 구축으로 사용자 만족도 향상
- 관리자 기능 고도화로 운영 효율성 제고
- 포스트게놈다부처 유전체사업 정보유통 활성화





# 질의응답



2018년 생명정보연구과 워크숍



2017년 이용자 교육(서울, 오송)



2018년 유전체학회 홍보부스