

NCBI 계정 만들기과 API 키 발급

유전자·게놈 데이터베이스 이용 안내 · 최종 확인 2026-08-22

API 키가 왜 필요한가

NCBI는 누구나 로그인 없이 쓸 수 있습니다. 다만 **프로그램으로 자동 수집할 때** 속도 제한이 걸립니다.

	키 없음	키 있음
요청 속도	초당 3회	초당 10회
비용	무료	무료
발급 시간	—	2분

게놈 1,000건을 훑는다면 **10분 걸릴 일이 3분으로 줄어듭니다**. 발급이 무료이고 금방이니 받아두는 편이 낫습니다.

키가 필요 없는 경우

아래는 **로그인도 키도 필요 없습니다**. 그냥 쓰시면 됩니다.

- 브라우저로 NCBI 검색, BLAST 웹 화면
- EFI-EST / EFI-GNT (유사도 네트워크·유전자 이웃 분석)
- Foldseek, AlphaFold DB, MGnify

키가 필요한 건 **E-utilities**로 자동 수집하는 코드를 돌릴 때입니다.

1단계 · 계정 만들기

 NCBI 자체 아이디는 2022년에 없어졌습니다. 반드시 다른 계정으로 연결해 로그인합니다. 새로 아이디·비밀번호를 만드는 방식이 아닙니다.

<https://account.ncbi.nlm.nih.gov/> 로 접속합니다.

로그인 방법을 고르는 화면이 나옵니다. 편한 것을 고르시면 됩니다.

방법	추천 상황
Google ★	가장 간단. 구글 계정이 있으면 클릭 두 번
ORCID	연구자 식별번호가 있고 논문 실적과 연결하고 싶을 때
소속기관 로그인	대학 계정이 NCBI 연합인증에 등록된 경우
Microsoft	회사·학교 MS 계정을 쓰는 경우

처음 로그인하면 이름과 이메일 확인 화면이 한 번 나옵니다. 확인만 누르면 계정이 생성됩니다.

2단계 · API 키 발급

로그인한 상태에서 진행합니다.

1. 오른쪽 위 자기 이름(또는 이메일) 클릭
2. Account settings 선택
(바로 가기: <https://account.ncbi.nlm.nih.gov/settings/>)
3. 아래로 내려 API Key Management 항목을 찾습니다
4. Create an API Key 버튼 클릭

36자리 문자열이 화면에 나타납니다. 이게 API 키입니다.

예시 형식: a1b2c3d4e5f6a7b8c9d0e1f2a3b4c5d6e7f8

알아두실 점

- 계정당 키는 하나입니다. NCBI 사이트에 여러 서비스가 나열되어 있어도 키를 따로따로 받는 게 아닙니다. 이 키 하나가 E-utilities·Datasets API 등에 공통으로 쓰입니다.
- 키가 노출되었다면 같은 화면에서 재발급하면 됩니다. 이전 키는 즉시 무효가 됩니다.
- 언제든지 다시 확인할 수 있으니 못 외워도 괜찮습니다.

3단계 · 키 보관

키를 코드 안에 직접 써 넣지 마세요. 실수로 공유하거나 GitHub에 올리면 그대로 노출됩니다.

이 연구실은 `C:\Users\sdkpa\.secrets\ENVIRONMENT_VARIABLES.json` 에 모아 보관합니다.

```
{
  "data_api_keys": {
    "NCBI_API_KEY": "여기에 발급받은 키"
  }
}
```

지켜야 할 것은 세 가지입니다.

- `.secrets` 폴더는 **연구 폴더 바깥**에 둡니다
- GitHub에 올리는 폴더에 **절대 넣지 않습니다**
- 화면 공유·발표 자료에 **띄우지 않습니다**

4단계 · 쓰는 법

모든 NCBI 요청 주소 끝에 `&api_key=키` 를 붙이면 됩니다.

```
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=protein&term=hyaluronan+synthase&api_key=
```

파이썬 예시

```
import json, os, urllib.parse, urllib.request

KEY = json.load(open(os.path.expanduser('~/.secrets/ENVIRONMENT_VARIABLES.json'),
                    encoding='utf-8'))['data_api_keys']['NCBI_API_KEY']

params = {
    'db': 'protein',
    'term': 'hyaluronan synthase',
    'retmode': 'json',
    'api_key': KEY,
    'tool': 'my-research',      # 권장: 프로그램 이름
    'email': 'you@example.com', # 권장: 문제 시 NCBI가 연락할 주소
}

url = 'https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?' + urllib.parse.urlencode(params)
print(json.load(urllib.request.urlopen(url))['esearchresult']['count'])
```

`tool` 과 `email` 은 필수는 아니지만 NCBI가 권장합니다. 요청에 문제가 있을 때 차단 대신 연락을 받게 됩니다.

이 프로젝트에는 이미 `scripts/ncbi.py` 로 만들어져 있습니다.

`from ncbi import count, search, summary, fetch` 로 바로 쓰시면 됩니다.

5단계 · 확인

키가 제대로 동작하는지 확인합니다.

```
import sys; sys.path.insert(0, 'scripts')
from ncbi import count
print(count('protein', 'hyaluronan synthase[Protein Name]'))
```

숫자가 나오면 정상입니다.

지켜야 할 이용 규칙

NCBI는 공공 무료 서비스입니다. 아래를 어기면 IP가 차단될 수 있습니다.

상황	지켜야 할 것
일반 요청	초당 10회를 넘기지 않기 (코드에 간격을 두세요)
100건 이상 대량 작업	주말 또는 평일 밤(미국 동부 기준 21시~5시)에 실행
대량 BLAST	결과를 조회할 때 간격을 충분히 두기. 1초마다 확인하면 차단 대상
전체 데이터가 필요할 때	E-utilities 대신 FTP 일괄 다운로드 사용

문제가 생기면

증상	원인	조치
<code>API rate limit exceeded</code>	요청이 너무 빠름	요청 간 간격을 0.11초 이상으로
<code>Invalid api key</code>	키 오타·앞뒤 공백	키 앞뒤 공백 제거, 재확인
갑자기 안 됨	키 재발급으로 이전 키 무효	새 키로 교체
응답이 계속 비어 있음	검색어 문법 문제	브라우저에서 같은 검색어로 먼저 확인
접속 자체가 막힘	대량 요청으로 IP 차단	몇 시간 뒤 재시도, <code>email</code> 파라미터 추가

참고 주소

용도	주소
계정 로그인	https://account.ncbi.nlm.nih.gov/
계정 설정 (키 발급)	https://account.ncbi.nlm.nih.gov/settings/
E-utilities 안내	https://www.ncbi.nlm.nih.gov/books/NBK25501/
API 전체 목록	https://www.ncbi.nlm.nih.gov/home/develop/api/

마지막 주소에 여러 서비스가 나열되어 있는데, **각각 따로 키를 받는 게 아닙니다.**
2단계에서 받은 키 하나로 전부 쓰실 수 있습니다.