

인체 내 SHBG와 내분비계 장애물질 간의 결합
분류를 위한 Machine learning-Deep learning 기반
계산독성학 QSAR 모델 개발

김동우, 유창규†

경희대학교

(ckyoo@khu.ac.kr†)

다양한 산업활동을 통해 발생하는 화학물질 중 인체 내에 유입되어 호르몬 작용에 영향을 미치는 내분비계 장애물질(EDCs: Endocrine disrupting chemicals)이 중요한 이슈로 떠오르고 있다. EDCs가 호르몬 적용에 영향을 미치기 위해서는 수송 단백질과 결합하여 세포 내로 유입되어야 한다. 본 연구에서는 EDCs와 체내 대표적인 수송 단백질인 성호르몬결합 글로불린(Sex hormone-binding globulin: SHBG)간의 결합 유무를 파악할 수 있는 QSAR (Quantitative structure-activity relationship) 모델을 개발하였다. QSAR 모델에 있어 독립변수로는 EDCs의 분자표현자가 사용되었으며 종속변수로는 SHBG와의 결합 유무에 대한 class(active/inactive)가 사용되었다. QSAR 모델에 사용될 주요 분자표현자를 선별하기 위해서 variable importance on projection (VIP) 기법과 K-평균 군집알고리즘을 적용하였으며, SHBG와 EDCs 간의 결합 유무를 파악하기 위해 기계 학습 방법인 로지스틱 회귀분석, 인공 신경망 및 서포트 벡터머신을 적용하고 예측 성능을 비교 분석하였다. 본 연구에서 제안된 기계학습 기반의 QSAR 모델은 EDCs의 인체 활성 유무에 대한 정보 제공 및 각종 위해성 평가에 기여할 것으로 판단된다. Acknowledgements: This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (No.2015R1A2A2A11001120)