# 1. 유전체 정보를 활용한 개인 맞춤형 치료가 실현될 수 있는지 설명하고 예들 들라.

|  |
| --- |
| 여러 약물에서 전혀 효과를 보지 못하는 환자가 약 30~60% 정도 된다는고 하며 이렇듯 약물의 부작용과 같이 각 개인에 의해서 다르게 나타난다. 최근 연구에 따라 유전적 요인에 따라서 개인 간의 약물반응에 대한 차이를 나타낸다는 사실이 알려지게 되고 약물의 부작용과 같이 개인간의 차이에 의해서 발생하는 시행착오를 피함으로써 각 개인에 맞는 처방을 할 수 있음.  와파린은 심방세동과 심부정맥혈전증, 심장판막치환술 등에서 혈액 응고로 혈관이 막히는 혈전, 색전증 등에 사용하는 약물이며, 와파린을 적절하게 사용하지 못하여 실제 필요한 양보다 부족하게 사용할 경우에는 혈관이 막혀서 중풍을 초래하게 될 수 있는 반면에 과도한 양을 사용하게 되면 뇌출혈을 일으킬 수 있다. 개인의 유전변이에 따라 와파린에 대한 약물반응이 다르게 나타날 수 있다는 사실을 밣혀냄에 따라 각 개인에 대한 유전정보를 이용하여 와파린에 대한 적정 용량을 조절함으로써 개인에 대한 맟춤 처방이 가능하게 됨. |

# 2. 유전체변이의 종류와 그 예를 들어 설명하라.

단일염기 단위에서 발생하는 유전변이형으로는 단일염기에 대한 삽입, 결손, 돌연변이와 단일염기당형성( Single Nulceodite Polymerphism, SNP )이 있으며, 변이가 눈 색깔, 맛을 느끼는 감각, 알코올 반응에 따른 피부색 변화부터 암이나 당뇨와 같은 복합질환에 영향을 준다는 사실이 알려줌.

유전체 구조변이는 두개 이상의 염기서열의 변화를 말하며, 부수체, 소부수체, 역전, 전좌, 복제수변이, 염색체 단위의 구조변이 등이 나타나며, 많은 사람들에게 알려진 유전 질환인 다운증후군은 염색체 단위의 구조변이에 의해서 발생하게 되며, 녹말섭취에 따른 인종별 아밀라아제 효소양의 차이가 발생함.

|  |  |  |
| --- | --- | --- |
| SNP | 설명 | 단일염기다형성( Single Nulceodite Polymerphism, SNP )는 개인 특정 뉴클레오티드가 다른 뉴클레뉴클레오 치환되는 것. |
| 예 | 눈 색깔,  맛을 느끼는 감각,  알코올 반응에 따른 피부색 변화  암이나 당뇨와 같은 복합질환에 영향 |
| CNV | 설명 | 유전자 복제수 변이( Copy Number Variation, CNV)는 사람의 모든 유전자가 1쌍이라고 생각했던 기존의 유전학의 가정과는 달리 특정유전자의 수는 사람마다 다르다는 것 |
| 예 | (1) 녹말섭취에 따른 인종별 아밀라아제 효소양의 차이  (2) 다운증후군, 암세포  (3) CCL3L1 chemokine은 chemokine receptor 5(CCR5)의 리간드로 CCR5와 결합한다. 이 CCR5는 HIV-1이 T세포로 들어가는 co-receptor의 역할도 하게 된다. 따라서 HIV-1과 CCL3L1은 CCR5와 경쟁적으로 결합하기에 CCL3L1의 DNA copy number가 상대적으로 낮은 사람은 HIV-1이 결합할 기회가 늘어나므로 감염에 취약하며 반대로 CCL3L1의 copy number가 상대적으로 높은 사람은 HIV감염에 저항성이 높게 된다 |

# 3. Bioconductor를 설치하고 ALL자료를 이용하여 이상치를 찾고, t-test를 실시하여 유의한 유전자를 찾고 Heatmap과 군집분석을 실시하고 해석하라.

|  |
| --- |
| (1) ALL 자료에서 t-test를 실시하여 유의한 유전자 검색 |
| source("http://bioconductor.org/biocLite.R")  biocLite()  biocLite("ALL")  library(ALL)  data(ALL)  dat=exprs(ALL)  dim(dat)  cl <- c( rep(0, 95), rep(1,33))  biocLite("multtest")  library(multtest)  resT <- mt.maxT( dat, classlabel=cl, B=1000 )  head( resT, 10 )  tail( resT, 6 )  ================== 결과 ==================  index teststat rawp adjp  37344\_at 7414 -20.75682 0.01 0.01  38147\_at 8225 20.71285 0.01 0.01  33238\_at 3268 20.69660 0.01 0.01  41609\_at 11719 -20.23509 0.01 0.01  2059\_s\_at 1174 19.40021 0.01 0.01  36773\_f\_at 6838 -19.38354 0.01 0.01  35010\_at 5058 -1.552503e-03 1 1  507\_s\_at 12073 1.535708e-03 1 1  40322\_at 10420 -8.960036e-04 1 1  41712\_at 11823 -7.175577e-04 1 1  1462\_s\_at 499 -1.843578e-04 1 1  AFFX-BioB-5\_at 12561 -2.455781e-05 1 1  **Head()를 이용해서 출력된 유전자들은 보정 전 p-값(rawp)와 보정후 p-값(adjp)값이 유의수준 0.05 보다 작으므로 출력된 모두 유전자는 T-Cell과 B-Cell에서 다르게 발현이 되며,**  **Tail() 를 이용해서 출력된 유전자들은 p-값이 1이므로 T-cell과 B-cell에서 모두 발현되는것으로 판단됩니다.** |
| (2) ALL자료를 Heatmap과 군집분석 |
| x <- dat[1:50, 96:128]  standardize <- function(x) {  (x - mean(x, na.rm=T)) / sqrt(var(x, na.rm=T))  }  x <- t( apply(x, 1, standardize) )  heatmap(x, main="Heatmap")    **상단의 계층적 군집을 보면, 유전자들이 크게 5개로 군집을 나누어질 수 있으며,**  **왼쪽의 계층적 군집을 보면, 유전자칩들은 3개의 군집으로 나누어짐.** |
| (3) ALL자료를 이상치를 검색 |
| biocLite("OutlierD")  library(OutlierD)  x <- log2(dat)  fit1 <- OutlierD(x1=x[,1], x2=x[,2], method="constant")  fit2 <- OutlierD(x1=x[,1], x2=x[,2], method="linear")  fit3 <- OutlierD(x1=x[,1], x2=x[,2], method="nonlin")  fit4 <- OutlierD(x1=x[,1], x2=x[,2], method="nonpar")  Outlier\_t = fit4$x[ fit4$x$Outlier==TRUE,]  Outlier\_f = fit4$x[ fit4$x$Outlier==FALSE,]  head(Outlier\_t)  outlier\_plot <- function( fit, title\_name ) {  plot(fit$x$A, fit3$x$M, pch=".", xlab="A", ylab="M")  i <- sort.list(fit3$x$A)  lines(fit$x$A[i], fit$x$Q3[i], lty=2); lines(fit$x$A[i], fit$x$Q1[i], lty=2)  lines(fit$x$A[i], fit$x$LB[i]); lines(fit$x$A[i], fit$x$UB[i])  title( title\_name )  }  outlier\_plot(fit4, "nonpar") # 비모수  par(mfrow=c(2,2))  outlier\_plot(fit1, "constant")  outlier\_plot(fit2, "linear")  outlier\_plot(fit3, "nonlin")  outlier\_plot(fit4, "nonpar")  ##############################################################  ## 결과  Outlier X1 X2 A M Q1 Q3 LB  1 TRUE 2.925491 2.902931 2.914211 0.022559867 -0.005228838 -0.005228838 -0.005228838  2 TRUE 2.335196 2.302330 2.318763 0.032865879 -0.004949674 -0.004949674 -0.004949674  3 TRUE 1.963647 2.073188 2.018417 -0.109541170 -0.006557477 -0.006557477 -0.006557477  4 TRUE 2.561658 2.625042 2.593350 -0.063384637 -0.010111172 -0.010111172 -0.010111172  5 TRUE 2.566878 2.563837 2.565357 0.003041861 -0.030083619 -0.030083619 -0.030083619  6 TRUE 3.099462 3.382432 3.240947 -0.282970057 0.024992625 0.024992625 0.024992625      **비모수 분위회귀를 이용한 이상치 탐색에서 1, 2, 3, 4, 5, 6 유전자가 이상치이며, 이상치 분포는 위의 M-A ScatterPlot과 같이 나타남.** |