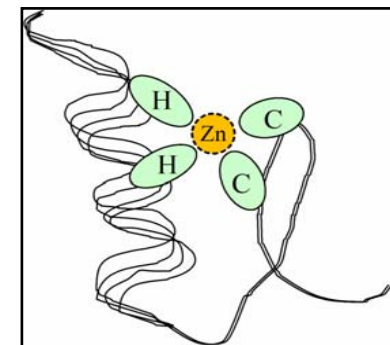


頻出パターンの抽出

- テキストデータ(日本語や英語), 購買履歴データ, 時空間データ, 分子配列データなどから頻繁に出現する規則性を**正規パターン**として抽出

< C - x(2,4) - C - x(3) - [LIVMFYWC] - x(8) - H - x(3,5) - H >

細胞組織名	配列データ
EVI1_HUMAN/75-97	...QE CK ...E CD ...QV F PDL...QSLEK H M L S... H .
TSH_DROME/533-557	...LT CK ...V CD ...K F NSL...GDLSN H MAKN. NH .
ZG5A_XENLA/90-112	...F S CT...V CG ...EM F TYR...AQF S K H M L K... H .
TSH_DROME/466-490	...L K CM...R CG ...ES F RS L ...GEMTK H M Q ET. QH .
YL57_CAEEL/344-367	...D H CB...R CV ...IK F PRA...RDY F A H MIK.. HH .
TF3A_BUFAM/134-158	...F K CS H .E G CD...K C YAS P ...SRL K R H EKT... H .
SDC3_CAEEL/2117-2141	...D D C Q ...D C Y...ET L T S S...FEVIV H RIN H .. HH .
YL57_CAEEL/26-49	...Y L CB...Y CG ...K T LS D R...LE Y Q Q H M L K .. VH .
ZFA_MOUSE/542-564	...F K CD...I CL ...L T FSD T ...KEV Q Q H AL V ... H .
P43_XENBO/163-187	...Y R CS Y .E D C Q ...T V S P T W ...TAL Q T H L K K... H .
SNAI_XENLA/232-255	...Y Q CB...S CS ...R T F S RM...S L L H K H EET.. GC .
SDC1_CAEEL/117-139	...L T CA H ...C D ...W S F D N V ...M K L V R H R G V... H .



zinc fingerモチーフ