

**OsteoBLAST: Rotina computacional de análise molecular global
aliada à biologia sistêmica e aplicada à produção de biomateriais**

Marcel Rodrigues Ferreira

Dissertação apresentada ao Instituto de Biociências, *Campus* de Botucatu, UNESP, para obtenção do título de Mestre no Programa de Pós-Graduação em Biotecnologia.

Willian Fernando Zambuzzi

**BOTUCATU – SP
2017**

UNIVERSIDADE ESTADUAL PAULISTA

“Júlio de Mesquita Filho”

INSTITUTO DE BIOCÊNCIAS DE BOTUCATU

**OsteoBLAST: Rotina computacional de análise molecular global
aliada à biologia sistêmica e aplicada à produção de biomateriais**

MARCEL RODRIGUES FERREIRA

Dissertação apresentada ao Instituto de Biociências, *Campus* de Botucatu, UNESP, para obtenção do título de Mestre no Programa de Pós-Graduação em Biotecnologia.

Willian Fernando Zambuzzi

BOTUCATU – SP
2017

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSEMEIRE APARECIDA VICENTE-CRB 8/5651

Ferreira, Marcel Rodrigues.

OsteoBLAST : rotina computacional de análise molecular global aliada à biologia sistêmica e aplicada à produção de biomateriais / Marcel Rodrigues Ferreira. - Botucatu, 2017

Dissertação (mestrado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Willian Fernando Zambuzzi

Capes: 90400003

1. Biotecnologia. 2. Materiais biomédicos. 3. Análise de microarranjo. 4. Fosforilação. 5. Bioinformática.

Palavras-chave: Bioinformática; Biologia sistêmica; Biomateriais; Fosforilação; Osso.

Agradecimentos

Primeiramente aos meus pais por me guiarem e apoiarem em todos os momentos da minha vida , em especial durante o este período. Sem eles nada seria possível.

Aos meus amigos de Itapetininga por me permitirem dividir algumas das histórias mais fantásticas que vivi.

Aos meus eternos coleguinhas, sejam da VIII turma de Física Médica do Instituto de Biociências de Botucatu ou de outros cursos da nossa querida faculdade, por todos os momentos de aprendizados que vivemos juntos. Parece que os bixos de 2011 estão virando gente grande.

Aos meus colegas de equipe de tênis de mesa da atlética geral de Botucatu por todos os treinos e Inters juntos. Mesmo longe estamos juntos!

A todos os membros da república Põe na mão de Deus, meus veteranos e meus bixos, seja pelas discussões, seja pelas risadas. Foram sempre a família e a base que tive em Botucatu.

Aos meus colegas de cursinho do IB. Obrigado por compartilharem a paixão pela prática da docência, mesmo quando sem apoio algum.

Ao professor Dr. Maikel Peppenlenbosch, da Erasmus Medical Center, pelo auxílio nos experimentos com o PamChip®, além de auxiliar na análise dos dados.

Ao dr. Renato Milani por toda a orientação e paciência em todos os skypes e e-mails trocados. Sem isso não seria possível.

A todos meus colegas de pós graduação em Biotecnologia. Todos os perrengues que passamos foram fundamentais no processo de aprendizado durante este período.

Aos colegas de departamento, em especial a Gabi e a Elaine pelas risadas e por nos proporcionar nosso glorioso café de cada dia.

A professora Margarida Juri Saeki pelo auxílio nas análise de MEV e a todos os técnicos do CME-IBB pelo preparo das amostras.

Ao meu caro colega Rafael Innocenti por todas nossas discussões filosóficas, políticas e científicas. Me fez sair melhor a cada conversa que tivemos.

Aos membros do laboratório de bioensaios e dinâmica celular pela convivência durante este período. Vou levar o carinho de vocês por toda minha vida.

Ao meu orientador, professor Willian Zambuzzi, por ter me recebido em seu grupo, ter me guiado neste período, além de todo o apoio dado a todo momento.

A FAPESP pelo apoio financeiro dado a este projeto através dos processos 2015/03639-8 e 2014/22689-3, fundamental para a minha manutenção neste período.

Por último agradeço a todos que demonstraram carinho me incentivando e acreditando em mim durante todo este período. Foram responsáveis pela manutenção do meu sorriso. Não poderia citar todos os nomes aqui, mas espero que sintam-se todos abraçados.

“Os registros mostram, eu recebi as pancadas

E fiz tudo do meu jeito”

(Frank Sinatra)

Resumo

As tendências na terapia com implantes têm incluído a modificação de suas superfícies utilizando ferramentas de nanotecnologia e princípios de bioengenharia, aumentando seu desempenho quando implantado. Embora muito se tenha alcançado em ferramentas para desenvolvimento destes materiais, metodologias de avaliação biológicas não avançaram nesta velocidade. Amparados por ferramentas de bioinformática e utilizando conceitos de biologia sistêmica, o objetivo deste trabalho foi produzir uma metodologia computacional, alternativa ao uso de experimentação animal, capaz de detectar e analisar o quonoma da resposta da interação célula-biomaterial, obtida com ensaio de microarranjo de peptídeos. Estes dados servirão para a construção de um banco de dados para guiar a produção de biomateriais para a área médico-odontológica. Batizaremos este de OsteoBLAST. Para tanto, fizemos uso de superfícies de titânio com diferentes superfícies (maquinado e duplo ataque ácido), as quais foram desafiadas com o cultivo de células tronco indiferenciadas. As amostras biológicas foram utilizadas para avaliar quinases diferencialmente ativadas através de substratos sintéticos, cuja metodologia é conhecida como PamGene, as quais foram reunidas em um banco de dados chamado OsteoBLAST, o qual fora constituído através de um algoritmo em quatro etapas que selecionou os resultados confiados de PamGene, em seguida obtendo o quonoma diferencial e um nível de similaridade com superfícies amplamente usadas na rotina. Nossos resultados mostraram que o algoritmo adotado desempenhou eficientemente a comparação entre as superfícies avaliadas, revelando as proteínas EGRF, ENO2, EPHA4, FRK, KRT6B, NCF1, PDPK1, PDGFRB e KDR envolvidas neste cenário molecular de resposta. Em conjunto, nossos resultados mostram que o algoritmo OsteoBLAST pode ser usado como uma potente ferramenta *in silico* para investigar potenciais biomateriais destinados à aplicações biomédicas, servindo-se como ferramenta de análise, capaz de otimizar o setor produtivo e diminuir o uso de animais de experimentação.

Palavras-chave: osteoblastos; adesão celular; Transdução de sinais; Bioinformática; Novas tecnologias.

ABSTRACT

Current trends in implant therapy have included the modification of their surfaces using nanotechnology tools and principles of bioengineering, increasing their performance when deployed. Although much has been achieved in tools for developing these materials, biological analysis methodologies did not advance at this speed. Supported by bioinformatics tools and using concepts of systemic biology, the objective of this work was to produce a computational methodology, alternative to the use of animal experimentation, capable of detecting and analyzing the kinome of the response of the cell-biomaterial interaction, obtained with microarray peptides. These data will serve to build a database to guide the production of biomaterials for the medical-dental area. We will baptize this one from OsteoBLAST. To do so, we made use of titanium surfaces with different surfaces (machined and double acid etched), which were challenged with the cultivation of undifferentiated stem cells. Biological samples were used to evaluate differentially activated kinases through synthetic substrates, a methodology known as PamGene, as the banks were assembled in a database called OsteoBLAST, which consists of a four-step algorithm that selected the results confident of PamGene, then obtaining the differential kinome and a level of similarity with surfaces widely used in the routine. Our results showed that the EGRF, ENO2, EPHA4, FRK, KRT6B, NCF1, PDPK1, PDGFRB and KDR as proteins involved in this molecular response scenario were efficiently compared. Together, our results show that the OsteoBLAST algorithm can be used as a powerful in silico tool to investigate potential biomaterials for biomedical applications, serving as an analysis tool capable of optimizing the productive sector and reducing the use of experimental animals .

Keywords: osteoblasts; cell adhesion; transduction of signals; bioinformatics; new technologies.

SUMÁRIO

1. Introdução	19
1.1 Osso: aspectos gerais	20
1.2 Biomateriais	20
1.3 Mecanismos de transdução de sinais celulares: Balanço de quinases e fosfatases	22
1.4 Adesão de Células Sobre Materiais: Primeiro Passo Integração de Biomateriais e Regeneração Óssea	24
1.5 Quinoma	25
1.6 Animais em experimentação	26
1.7 Biologia sistêmica	29
1.8 Bioinformática	30
2. Objetivos	33
3. Materiais e métodos	34
3.1 Cultivo celular	34
3.2 Disco de titânio	34
3.3 Morfologia celular	34
3.4 PamChip	34
3.5 Bioinformática aplicada: OsteoBLAST	35
4. Resultados e Discussão	36
4.1 OsteoBLAST	36
4.2 PamChip® tirosina quinase	36
4.3 Algoritmo	37
4.3.1 Seleção	38
4.3.2 Normalização	41
4.3.3 Verificação	42
4.3.4 Comparação	44

4.4	Superfícies teste	45
4.5	Resultados para a superfície com duplo ataque ácido (DAA)	47
4.6	Resultados para a superfície com hidroxiapatita na escala manométrica (Ti).	49
4.7	Análise Biológica	51
4.7.1	Análise de redes proteicas para o grupo DAA-Maq	51
4.7.2	Análise de redes proteicas para os grupos Ti-Maq.	59
4.7.3	A via de sinalização ErbB	63
4.7.4	O ponto de adesão focal	65
4.7.5	Proteínas que não participam de redes	67
5.	Constatações	68
6.	Conclusão	69
7.	Perspectivas futuras	70
8.	Bibliografia	71

LISTA DE FIGURAS

- Figura 1 - Exemplo de scaffolds utilizados no tratamento de lesões ósseas. (https://horizon-magazine.eu/article/live-cells-could-be-used-reconstruct-faces_en.html, http://www.virtualscience.com/cad_rp_section/cad_rp_popups/bone_scaff_popup.htm, <http://www.medicaldevice-network.com/features/feature359/feature359-1.html>). Acessado em 15/01/2017..... 21
- Figura 2 – Mecanismos de transdução de sinais. Estímulos físico-químicos são captados pela célula através de receptores transmembranas. Estes darão início a uma cascata de eventos intracelulares controlados pelo balanço da fosforilação. Ao final desta cascata uma maquinaria gênica será ativada no núcleo da célula podendo resultar na modificação de eventos como proliferação, diferenciação, adesão, alteração do metabolismo celular, entre outras. 22
- Figura 3 – Balanço da fosforilação proteica. Quinases são enzimas que transferem um grupo fosfato do ATP para uma proteína X, assim fosforilando-a. Já fosfatases são enzimas que atuam no sentido contrário retirando o grupo fosfato da proteína X fosforilada, porém o fosfato não é devolvido ao ADP e sim ao citosol. 23
- Figura 4 - Mecanismos intracelulares decorrentes da adesão celular. Integrinas ligam o meio extracelular com o citoesqueleto da célula. Após captar sinais extracelulares as integrinas mobilizaram moléculas de sinalização para o ponto de adesão focal. O balanço da fosforilação de FAK, Src e paxilina e, ao final desta cascata, a cofilina levará um sinal intracelular que será responsável pelo rearranjo dos filamentos de actina. Estas alterações do citoesqueleto são responsáveis por adaptações celulares a superfícies diferentes. Fonte: ZAMBUZZI et al., 2011..... 25
- Figura 5 - O zoologista William M.S. Russell e o microbiologista Rex L. Burch publicaram em 1959, “The Principles of Humane Experimental Technique”. Este trabalho foi um marco para a experimentação em animais por inserir o conceito dos 3 Rs, Replacement, Reduction e Refinement. Embora diversos avanços na área tenham ocorrido, o cerne da ética e da busca a métodos alternativos se encontra nesta obra..... 27
- Figura 6 – Esquema das etapas necessárias para um medicamento, vacina ou, como em nosso caso, um biomaterial chegar ao ser humano. Primeiramente são realizados testes em culturas de células. A próxima etapa será experimentos em animais de pequeno porte, como ratos ou camundongos. O cãozinho Snoopy do desenho “Peanuts”,

representa o passo seguinte, experimento em animais de maior porte como cães. Só aprovado em todos estes passos será levado ao ser humano.	29
Figura 7 – Criado em 1986 o número de sequências depositadas cresce ano após ano desde então. Atualmente são 551.987 diferentes sequências depositadas. Fonte: http://www.uniprot.org/statistics/Swiss-Prot	31
Figura 8 - Gráfico do número de trabalhos adicionados no PubMed por ano de 1950 até 2012. Fonte: VARDAKAS et al.	32
Figura 9 - Esquema de PamChip(r) tirosina quinase. XXXXXXYXXXXXX respresenta uma sequência qualquer de resíduos de aminoácidos que contenha um resíduo tirosina (Y). Anticorpo PY20 conjugado com FITC serão transferidos ao sequência através da fosforilação do resíduo tirosina, catalisada pela atuação de uma quinase.	37
Figura 10 - Fluxograma do algoritmo de OsteoBLAST.....	38
Figura 11 - Exemplo de arquivo saída da etapa de seleção. A primeira coluna, protein, contém o nome das proteínas de acordo com o código do spot do PamChip®. Já nas colunas M_m, A_m e SD_m, contém respectivamente a mediana, média e desvio padrão de cada proteína. P1_m e P2_m contém os valores dos parâmetros P1 e P2, e cP1_m e cP2_m contém a classificação. cE_m contém os valores da classificação de estrelas baseada nos parâmetros P1 e P2.	41
Figura 12 – Exemplo da etapa de normalização. Na primeira coluna temos o código do spot, nas 2, 3 e 4 temos os valores de intensidades obtidos para as triplicatas para cada spot. Nas colunas 5, 6 e 7 temos os valores do quinto percentil para cada coluna de triplicatas, nas colunas 8, 9 e 10 temos os valores do nonagésimo quinto percentil de cada coluna de triplicatas e nas colunas 11, 12 e 13 temos os valores de intensidade já normalizados de acordo com a equação 1.	42
Figura 13 - Modelo saída de resultados da verificação de fosforilações significativamente diferentes entre dois grupos, neste caso a comparação entre os grupos Maq e DAA. A primeira coluna contém o código dos spots em PamChip®, a coluna var_m_daa contém os valores do teste F, caso eles sejam menores que 0,05 a coluna sigvar_m_daa recebe o valor 1, caso contrário recebe 0. A coluna pv_m_daa recebe o valor de P para o teste t, caso o valor da coluna sigvar_m_daa seja 1, ou do teste U, caso o valor da coluna sigvar_m_daa seja 0. Por ultimo, a coluna sigpv_m_daa apresenta valores de 1 ou 0 de acordo com a significancia dos valores da coluna pv_m_daa, 1 caso <0,05 e 0 caso >0,05.	43

Figura 14 - Modelo de saída do arquivo rptProtein. A primeira coluna contém o código dos spots de PamChip®, a segunda contém o nome da proteína segundo UniProt, a terceira contém o gene que codifica e a quarta o código de acesso da proteína em UniProt.	44
Figura 15 - Comportamento celular frente a diferentes tratamentos na superfície. Superfície maquinada em com magnificação de 500x (A) e de 5000x (B), e superfície com cobertura de hidroxiapatita em escala nanométrica (HANano). É possível perceber que células estão com um processo de “spreading” avançado na superfície com HANano em comparação com a superfície maquinada. Imagens obtidas no CME – IBB.....	46
Figura 16 - Localização de carbono e titânio nos discos. As figuras A e D contém imagens da microscopia eletrônica para as superfícies maquinada e com hidroxiapatita em escala nanométrica. As figuras B,C,E e F foram obtidas com EDX. B e E mostram a localização de carbono nas superfícies, e C e F mostram a localização de titânio na amostra. A maior dissipação de carbono no grupo com superfície modificada (E) pode indicar um processo de “spreading” mais acelerado que na superfície maquinada (B). 47	47
Figura 17 - Gráfico de pizza para as frequências de Ce do grupo DAA. 40 spots (27,78%) possuem Ce = 5, estes passarão pelas etapas de normalização e verificação. 48	48
Figura 18 - Gráfico de pizza para as frequências de Ce do grupo Ti. Apenas 17 spots (11,81%) possuem Ce = 5, estes passarão pelas etapas de normalização e verificação. 50	50
Figura 19 - Rede obtida através de STRING database para EGFR, ENO2, EPHA4 FRK, KRT6B, KDR, NCF1, e PDGFRB.	52
Figura 20 - Rede expandida buscando mais interações. A partir das seis primeiras proteínas a rede foi expandida, utilizando os seguintes parâmetros: Não mais do que 10 interações na primeira camada, não mais do que 200 interações na segunda camada e índice de confiança superior a 0.900, considerado altíssimo.	53
Figura 21 - Gráfico de Betweenness centrality para a rede DAA - Maq. EGFR(0.079), Src(0.071) e GRB2(0.065) foram nodos que apresentaram os maiores valores.....	54
Figura 22 - Gráfico de Closeness centrality para a rede de DAA - Maq. Src(0.682), GRB2(0.670) e EGFR(0.666) foram nodos que apresentaram maiores valores.	54
Figura 23 - Cluster 1 para a rede DAA-Maq, com 26 nodos e score de 17,92.....	56
Figura 24 - Cluster 2 para a rede DAA-Maq, com 65 nodos e score de 14,688.....	56
Figura 25 - Layout de DAVID para a rede inicial da rede DAA-Maq	57
Figura 26 - Layout de DAVID tools para funções presentes no cluster 1.....	58

Figura 27 – Layout de DAVID tools para funções biológicas obtidas com <i>Funcional Annotation Chart</i> do DAVID tools para o cluster 2.....	59
Figura 28 - As proteínas que apresentaram fosforilação diferencial para o grupo Ti-Maq não apresentam um comportamento de rede.	59
Figura 29 - Rede expandida para o grupo Ti-Maq. A rede foi construída a partir das proteínas NCF1 e PDPK1 que não possuíam ligações entre si.	60
Figura 30 - Gráfico de Betweenness centrality para a rede de Ti-Maq. HSP90AA1(0.168), AKT1(0.164) e UBC(0.157) foram os nodos que apresentaram maiores valores para esta centralidade.	61
Figura 31 - Gráfico de Closeness centrality para a rede de Ti-Maq. UBC(0.685), AKT1(0.683), HSP90AA1(0.632) e Src(0.583) foram os nodos que apresentaram maiores valores para esta centralidade.	61
Figura 32 - Cluster para a rede de Ti-Maq com score de 12,333 e 31 nodos.....	62
Figura 33 - Layout de Functional Annotation Chart para o cluster de Ti-Maq.....	63
Figura 34 - Via de sinalização de ErbB. Fonte: KEGG pathway.....	64
Figura 35 - Mecanismo de dimerização de membros da família ErbB é fundamental para a ativação da via. O processo de dimerização da ErbB-2 não depende de estímulos externos e dependente do desligamento das chaperonas HSP90 e CDC37 de sua cauda citosólica. Ao ser desligada das chaperonas, ErbB-2 é capaz de se dimerizar com outros membros ErbB, e ao se dimerização fosforilar resíduos presentes na cauda citosólica que darão início ao mecanismo de sinalização.....	65
Figura 36 - Complexo de adesão focal. Fonte: KEGG pathway	67

LISTA DE TABELAS

Tabela 1- Tabela de valores de Ce (classificação por estrelas) para "spots" dependendo do nível dos parâmetros P1 e P2.....	39
Tabela 2 - Resumo das frequências de Ce para cada combinação em cada grupo.....	40
Tabela 3 – Frequência de spots por nível de confiança no processo de seleção de OsteoBLAST para a superfície com duplo ataque ácido (DAA).	48
Tabela 4 - Lista de proteínas com diferenças estatísticas detectadas na etapa de verificação. As informações sobre o gene que as codificam, o resíduo fosforilado e o código de acesso no Uniprot completam a tabela.	49
Tabela 5 - Frequência de spots por nível de confiança no processo de seleção de OsteoBLAST para a superfície com hidroxiapatita em escala nanométrica (Ti).	50
Tabela 6 - Lista de proteínas com diferenças estatísticas detectadas na etapa de verificação para o grupo Ti. As informações sobre o gene que as codificam, o resíduo fosforilado e o código de acesso no Uniprot completam a tabela.	51

LISTA DE ABREVIACOES

A	Mdia
ADP	Adenosina difosfato
AKT1	Protena serina/treonina quinase RAC-alfa
APK	Protenas quinases atpicas
ATP	Adenosina trifosfato
BLAST	Basic Local Alignment Search Tool
CAMPK	Protena quinase dependente de Ca ²⁺ /calmodulina
CDC37	Chaperona CDC37
Ce	Classificao de estrelas
CME	Centro de microscopia eletrnica
CONCEA	Conselho Nacional de Controle de Experimentao Animal
DAA	Superfcie com Duplo Ataque cido
DAVID	Database for Annotation, Visualization and Integrated Discovery
DNA	cido desoxirribonucleico
EBI	Instituto Europeu de Bioinformtica
EDX	Espectroscopia por disperso de raios x
EGF	Fator de crescimento endotelial
EGFR	Receptor do fator de crescimento endotelial
ENO2	Gama enolase
EPHA4	Receptor de efrina do tipo-A 4
EPK	Protenas quinases eucariotas
ErbB	Receptor protena tirosina quinase ErbB
EROS	Espcies reativas de oxignio

FAK	Quinase de adesão focal
FIOCRUZ	Fundação Oswaldo Cruz
FITC	Isotiocianato de fluoresceína
FRK	Proteína tirosina quinase FRK
GRB2	Proteína ligante do fator de crescimento 2
HSP90AA1	Chaperona HSP90
IBB	Instituto de Biociências de Botucatu
KDR	Receptor do fator de crescimento endotelial vascular 2
KEGG	Enciclopédia de genes e genomas de Kyoto
KRT6B	Queratina citoesquelética, do tipo II, 6B
LWM-PTP	Proteína tirosina fosfatase de baixo peso molecular
M	Superfície Maquinada
M	Mediana
MAPK	Proteína quinase ativada por mitógeno
MAPKp38	Proteína quinase ativada por mitógeno p38
NCF1	Fator citosólico de neutrófilos 1 (National Institutes of Health – National Center for Biotechnology Information)
NIH – NCBI	
NLM	Biblioteca Nacional EUA de Medicina
P1	Parâmetro 1
P2	Parâmetro 2
p5	Quinto percentil
p95	Nonagésimo percentil
Pax	Paxilina
PDB	Banco de dados de proteína

PDGFRB	Receptor do factor de crescimento derivado de plaquetas beta
PDPK1	Proteína quinase 3-fosfoinositida-dependente 1
PIR	Recurso de Informação de Proteína
PTP	Proteína tirosina fosfatases
PTP1B	Proteína tirosina fostase 1B
PY20	Fluoresceína de ligação
RNA	Ácido ribonucleico
S	Serina
Sd	Desvio padrão
SIB	Instituto Suíço de Bioinformática
SOS1	Son of Sevenless
Src	Proteína tirosina quinase Src (sarcoma)
T	Treonina
TGF- α	Fator de crescimento transformante alfa
Ti	Superfície com HAnano
UBC	Poli-ubiquitina C
Y	Tirosina

1. Introdução

Defeitos ósseos constituem um dos problemas mais prementes na área médico-odontológica. A reconstrução do tecido lesado é um dos maiores desafios dos cirurgiões, sendo muitas vezes necessária a aplicação de um biomaterial que permita uma boa regeneração tecidual. Nesta área, os materiais osteo-substitutos constituem uma das mais comuns e antigas formas de terapia regenerativa. Para se ter uma ideia, arqueólogos encontraram, no Peru, um crânio de um chefe tribal, datado de 2000 a.C., em que uma placa de 1 mm de espessura de ouro batido cobria um defeito ósseo (HENKEL et al., 2013). Até o presente momento, não foi desenvolvido nenhum material sintético que reúna as propriedades e capacidades ofertados pelo material autógeno, conhecidamente como material ideal. Os materiais atualmente disponíveis, carecem de estabilidade e retenção, o que implica em dificuldades na mastigação, podem causar dores por se moverem e se soltarem (SOHRABI et al., 2012).

A pesar de ser uma área com alto impacto médico, social e econômica, e de serem amplamente estudados, biomateriais tem recebido pouco investimento em novas ferramentas capazes de prever e ranquear novos produtos. Desta forma, as pesquisas realizadas por companhias e universidades acabam por utilizar-se de modelos experimentais tradicionais que façam uso de modelos com animais, o que compromete os avanços da área. Neste sentido conhecer os mecanismos intracelulares surge como uma alternativa. Uma vez detectados biomarcadores envolvidos na resposta da interação célula-biomaterial será possível monitorar estes em novos biomateriais podendo prever o sucesso de sua implantação comparando-os a biomateriais já conhecidos.

Ferramentas computacionais são fundamentais na organização de dados, como por exemplo bancos de dados biológicos para as mais diversas áreas, além de estarem em constante desenvolvimento de novos métodos de bioanálises. Posto isto, a obtenção e organização de biomarcadores da interação célula-biomaterial pode guiar o processo de produção destes. Neste sentido, temos centrado esforços na elaboração de um banco de dados que possa reunir dados de proteínas intracelulares, disparadas em resposta aos mais variados biomateriais.

1.1 Osso: aspectos gerais

O tecido ósseo é um tecido altamente dinâmico capaz de se remodelar ao longo de toda a vida. Além de fornecer suporte e proteção para tecidos moles, o tecido ósseo atua como reserva de minerais, como cálcio, magnésio e fosfato necessários para a manutenção de diversos processos fisiológicos (GEMINI-PIPERNI et al., 2014; ZAMBUZZI; MILANI; TETI, 2010). Devido a capacidade de armazenar e mobilizar estes íons minerais para o corpo se faz necessária uma alta vascularização do tecido ósseo.

O tecido ósseo é composto por três tipos celulares: osteócitos, osteoblastos e osteoclastos. Osteoblastos são responsáveis pela síntese dos componentes orgânicos da matriz óssea, como colágeno tipo I, proteoglicanos e glicoproteínas, além de promover a mineralização do tecido. Ao terminar de sintetizar a matriz o osteoblasto, aprisionado, é chamado de osteócito. Já os osteoclastos são células responsáveis pela reabsorção da matriz, passo fundamental no processo de remodelagem do tecido ósseo.

Uma vez lesado o tecido ósseo, a cicatrização pode ocorrer por dois processos: reparação e regeneração. A regeneração ocorre quando o tecido danificado é povoado com base na proliferação e diferenciação das células originais do tecido. Assim sendo, o tecido regenerado possui as mesmas propriedades funcionais e aspectos morfológicos do tecido primitivo. Já no processo de reparação o tecido danificado é substituído por um tecido conjuntivo fibroso, resultando em uma cicatriz de colágeno. Isto acarreta em uma perda das características e funções originais. Atualmente a regeneração do osso é vista como prioritária, em relação a reparação, em casos de traumas. Desta forma recuperando a forma e as funções originais do osso (ZAMBUZZI et al., 2011b).

1.2 Biomateriais

A engenharia de tecidos para o osso tem se apoiado principalmente em três aspectos básicos: 1) o recrutamento de células competentes, 2) moléculas de sinalização capazes de estimular as células no sentido de uma resposta eficaz e 3) matrizes de suporte, capazes de direcionar e sustentar a formação do osso durante sua regeneração, os biomateriais.

Recentemente, foi constatada a necessidade da inclusão de um novo pilar: a estabilidade mecânica (MANISCALCO; CAFORIO, 2016; WEE et al., 2016). Pesquisas mostraram que a estabilidade do microambiente no qual a lesão se encontra influencia no sucesso da resposta biológica dos outros três aspectos já citados.

Idealmente o biomaterial deve ser capaz de recrutar células osteoprogenitoras para a região da lesão, além de estimular sua proliferação e diferenciação. Caso o biomaterial não induza a uma cascata de eventos biomoleculares que acelerem a regeneração óssea, deve ao menos servir como “scaffold” para a proliferação e migração de células hospedeiras (ZAMBUZZI et al., 2011b). O biomaterial deve também garantir uma boa estabilidade a região da lesão, assim promovendo o bem estar do paciente (SOHRABI et al., 2012).

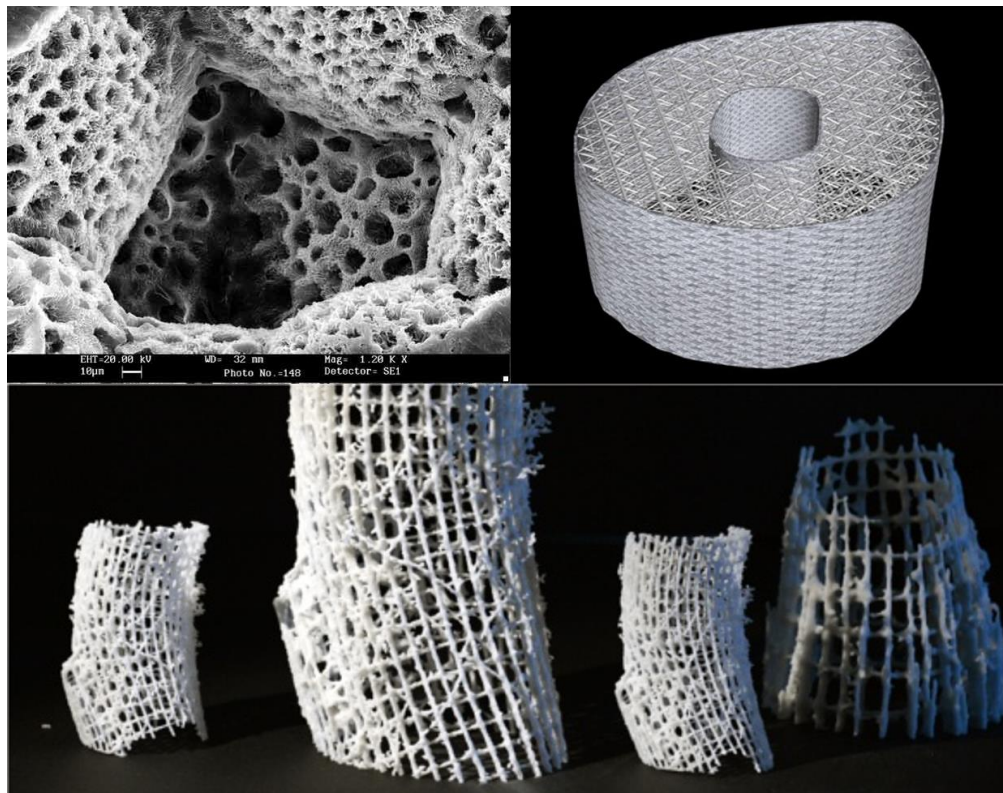


Figura 1 - Exemplo de scaffolds utilizados no tratamento de lesões ósseas. (https://horizon-magazine.eu/article/live-cells-could-be-used-reconstruct-faces_en.html, http://www.virtualscience.com/cad_rp_section/cad_rp_popups/bone_scaff_popup.htm, <http://www.medicaldevice-network.com/features/feature359/feature359-1.html>). Acessado em 15/01/2017.

Um fator preponderante para o sucesso da osteointegração do biomaterial é sua superfície (LONGO et al., 2016). A adição de uma camada de hidroxiapatita, com escala micrométrica ou nanométrica, na superfície do biomaterial reforça propriedades

biomecânicas do mesmo, além de tais modificações físico-químicas na superfície afetarem certos comportamentos celulares como formação de componentes de adesão focal e organização do citoesqueleto com efeitos na maturação dos osteoblastos e subsequente mineralização (ZAMBUZZI et al., 2014).

1.3 Mecanismos de transdução de sinais celulares: Balanço de quinases e fosfatases

O processo no qual uma célula reage a estímulos externos e modifica seu estado e atividade é chamado de transdução de sinais. Por meio de receptores transmembrana, a célula é capaz de captar estímulos físico, como por exemplo estresse mecânico ou a luz, ou químicos, como hormônios e fatores de crescimento. Estes estímulos alteram a conformação espacial destes receptores, e esta alteração é responsável por disparar uma cascata de eventos intracelulares. Estes eventos são regidos por duas classes de enzimas: quinases e fosfatases.

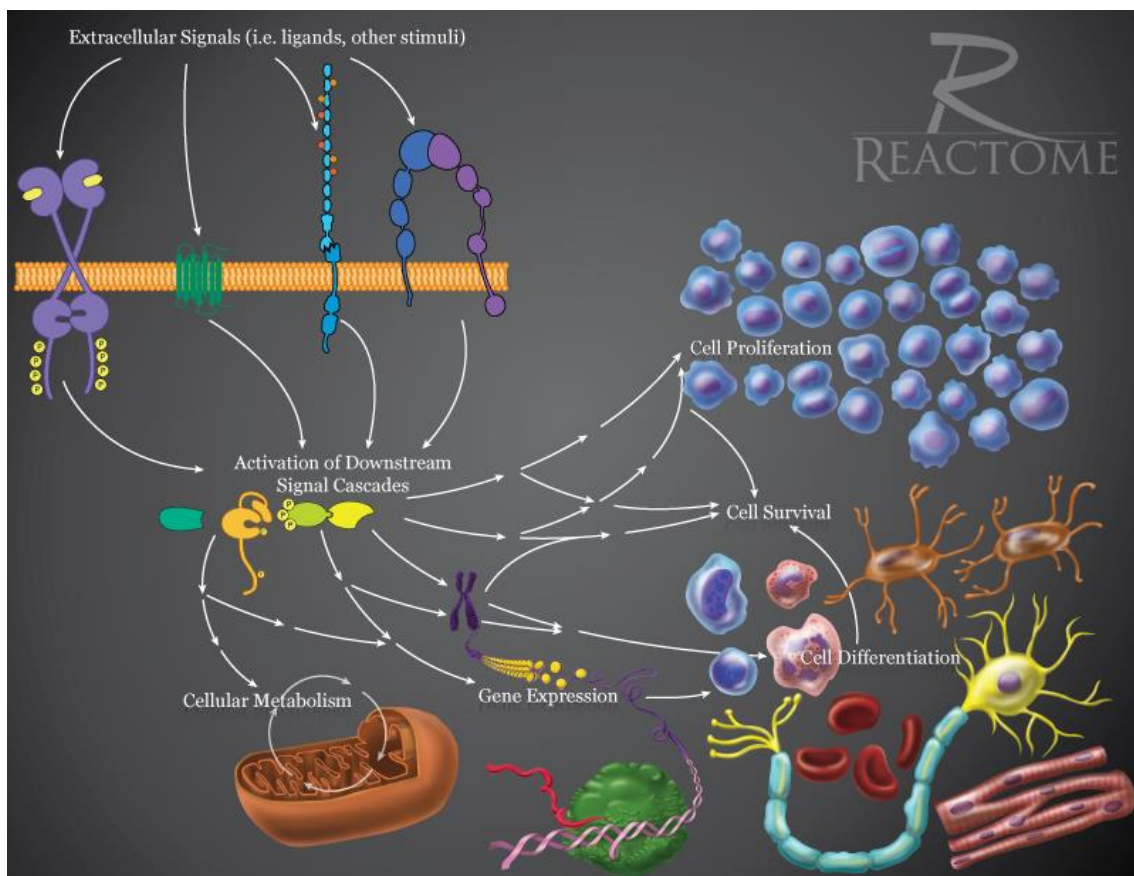


Figura 2 – Mecanismos de transdução de sinais. Estímulos físico-químicos são captados pela célula através de receptores transmembranas. Estes dão início a uma cascata de

eventos intracelulares controlados pelo balanço da fosforilação. Ao final desta cascata uma maquinaria gênica será ativada no núcleo da célula podendo resultar na modificação de eventos como proliferação, diferenciação, adesão, alteração do metabolismo celular, entre outras.

Quinases e fosfatases são classes de enzimas transferases, responsáveis por transferir o grupo fosfato, assim catalisando a fosforilação. Porém seus papéis são antagônicos. Enquanto proteínas quinases transferem um grupo fosfato do ATP, transformando-o em ADP, para uma determinada proteína, as fosfatases retiram o grupo fosfato da mesma. É importante ressaltar que ao ser retirado, pelas fosfatases, o grupo fosfato não retorna ao ADP, formando ATP.

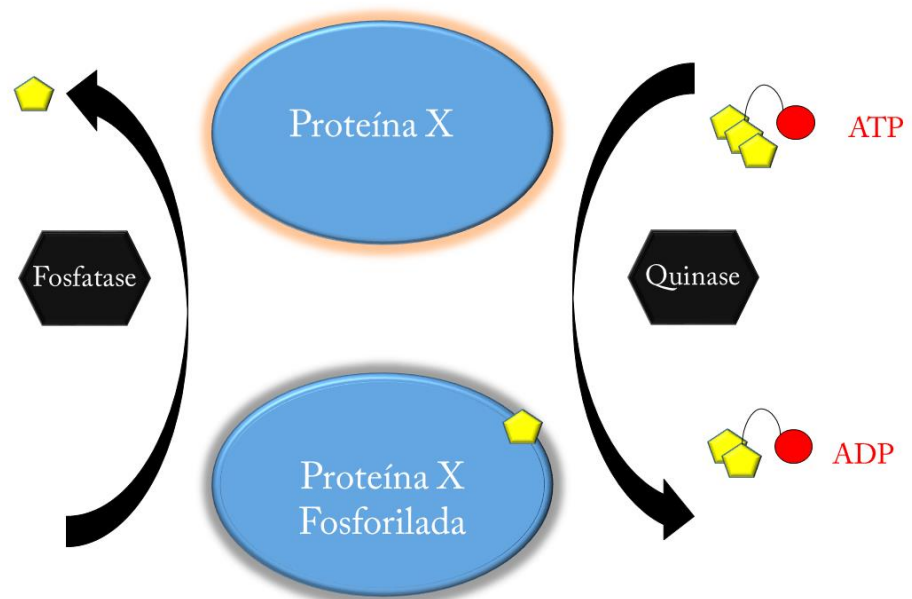


Figura 3 – Balanço da fosforilação proteica. Quinases são enzimas que transferem um grupo fosfato do ATP para uma proteína X, assim fosforilando-a. Já fosfatases são enzimas que atuam no sentido contrário retirando o grupo fosfato da proteína X fosforilada, porém o fosfato não é devolvido ao ADP e sim ao citosol.

O fosfato possui cargas negativa e ao ser transferido a uma proteína ele provocara uma modificação conformacional na estrutura da mesma. É sabido que estrutura e função de proteínas são intimamente ligadas, logo tal mudança conformacional pode ativar ou inibir o sítio ativo da proteína fosforilada.

Em eucariotos a fosforilação pode ocorrer apenas em três resíduos de aminoácidos: Serina (S), Treonina (T) e Tirosina (Y). O sítio de fosforilação reconhecido por uma quinase, sequência de cerca de treze resíduos de aminoácidos na proteína a ser fosforilada, é altamente específico. Isto faz com que quinases, e por

consequência fosfatases, sejam importantíssimas no controle de reações biológicas através da fosforilação. Eventos celulares como proliferação, diferenciação, adesão, controle do ciclo celular, entre outros, são exemplo de reações controladas através de mecanismo de transdução de sinais.

1.4 Adesão de Células Sobre Materiais: Primeiro Passo Integração de Biomateriais e Regeneração Óssea

Ao se utilizar um biomaterial para o tratamento de uma lesão óssea o desejável é que ocorra o evento da osteointegração. Como dito anteriormente, a superfície do biomaterial é um fator decisivo neste processo. Modificações físico-químicas na superfície do biomaterial disparam mecanismo de transdução de sinais que serão responsáveis pelo sucesso da osteointegração. É fundamental que estes induzam a adesão, proliferação e diferenciação celular sobre a superfície do biomaterial.

A adesão celular é um evento altamente estudado hoje em dia, seja nos processos biológicos seja em condições patológicas. Um dos exemplos da importância do estudo da adesão, temos o desenvolvimento e a progressão do câncer. Nestes processos diversas moléculas relacionadas a adesão atuam promovendo a invasibilidade e metástases do tumor. Isto devido ao fato de estarem envolvidas em interações célula-célula e célula-matriz (OKEGAWA et al., 2004).

Recentes estudos têm demonstrado que a adesão celular participa de maneira ativa na regulação da diferenciação, proliferação e morte celular. Logo é altamente desejável que um biomaterial induza uma boa adesão celular. Nos últimos anos nosso grupo vem investindo na busca por biomarcadores da interação célula material em biomateriais, e assim ser possível conhecer as propriedades adesivas de novos biomateriais. Em estudo recente, nosso grupo mostrou que nas primeiras 2 horas pós-plaqueamento o metabolismo dos osteoblastos culmina em diferentes vias de sinalização responsáveis pelo rearranjo do citoesqueleto, com sinais oriundos da ativação de proteínas como MAPKp38 e FAK (ZAMBUZZI et al., 2009) temos mostrado, ainda, que a presença de EROs modula a ação de proteínas tirosina fosfatases (PTPs) durante mecanismos celulares envolvidos com a adesão de osteoblastos (FERNANDES et al., 2014).

Embora algum progresso tenha sido alcançado neste sentido, muito ainda precisa ser descoberto sobre estes mecanismos em outras condições, para que haja uma transferência direcionada de conhecimento ao setor de desenvolvimento de materiais.

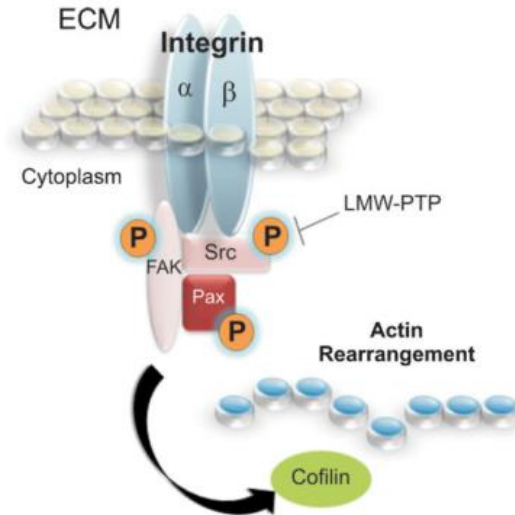


Figura 4 - Mecanismos intracelulares decorrentes da adesão celular. Integrinas ligam o meio extracelular com o citoesqueleto da célula. Após captar sinais extracelulares as integrinas mobilizaram moléculas de sinalização para o ponto de adesão focal. O balanço da fosforilação de FAK, Src e paxilina e, ao final desta cascata, a cofilina levará um sinal intracelular que será responsável pelo rearranjo dos filamentos de actina. Estas alterações do citoesqueleto são responsáveis por adaptações celulares a superfícies diferentes. **Fonte:** ZAMBUZZI et al., 2011.

1.5 Quinoma

Em 1958, Francis Crick postulou o dogma central da biologia molecular. Nele Crick postulou como ocorre o fluxo de informações do código genético. Segundo este dogma, o fluxo da informação genética segue o seguinte sentido: DNA → RNA → PROTEÍNAS. Com a descoberta de enzimas capazes de sintetizar DNA através de uma fita de RNA, este dogma sofreu atualizações, porém ele contém uma importante mensagem desde sua postulação: a receita para se fazer proteínas se encontra no genoma.

O conjunto de genes que codificam proteínas quinases é conhecido como quinoma. Em seu trabalho, MANNING et al. identificaram 518 genes que codificam proteínas quinase. Estes genes são classificados em duas famílias de acordo com a presença, ou não, do domínio catalítico quinásico. A primeira família é denominada de proteínas quinases eucariotas (EPK), e contém todos os genes que codifiquem proteínas

quinases que possuam o domínio catalítico quinásico, totalizando 478 genes. Já a segunda família contém 40 genes que não possuem o domínio catalítico quinásico, porém apresentam atividade quinásica por isto são denominadas proteínas quinases atípicas (APK).

Como dito anteriormente, quinases são enzimas que catalisam a fosforilação. Devido a cada um destes genes codificarem um conjunto distinto de substratos, o conhecimento do quinoma em diferentes situações, como por exemplo no processo de interação célula-biomaterial, é de grande valia na busca pelo entendimento destes processos.

Classicamente métodos de estudos de fosforilação são muito trabalhosos, como Espectroscopia de massa e Western Blotting. Neste sentido microarranjo de peptídeos vem se mostrando uma ferramenta poderosa no estudo da fosforilação. Por meio desta técnica é possível a obtenção de informações referentes a quais quinases estão ativas em uma determinada condição celular (Quinoma diferencial), assim podendo relacionarmos a processos celulares que estas participem. Vale a ressalva que, por mais que esta técnica seja similar à de microarranjo de DNA, porém biologicamente é mais interessante o estudo da atividade enzimática do que na expressão gênica (ROSENBERGER et al., 2015).

1.6 Animais em experimentação

Na madrugada do dia 18/10/2012 militantes dos direitos dos animais invadiram o Instituto Royal, em São Roque-SP, e retiraram cerca de 170 cães da raça beagle, além de camundongos e coelhos. Os ativistas alegavam que o laboratório praticava maus-tratos contra os animais com a finalidade de produzir experiências científicas. O caso ganhou repercussão na mídia e reacendeu o debate sobre a experimentação em animais (fonte: <http://g1.globo.com/sao-paulo/sorocaba-jundiai/noticia/2014/10/invasao-ao-predio-do-instituto-royal-em-sao-roque-completa-um-ano.html>. Acessado em 23/09/2016).

No Brasil a lei que regulamenta os testes em animais é a lei de nº 11794, de 8 de outubro de 2008, sancionada pelo então presidente da república Luís Inácio Lula da Silva. Conhecida como lei Arouca, ela regulamenta os procedimentos para uso da experimentação em animais em território nacional, além de criar o Conselho Nacional

de Controle de Experimentação Animal – CONCEA, com a competência de formular e zelar pelo cumprimento das normas na experimentação, credenciar instituições, monitorar e avaliar a introdução de novas técnicas que possam substituir a utilização de animais em ensino e pesquisa, dentre outras.

Mundialmente há uma tendência na diminuição da experimentação animal, bem como na criação, validação e implementação de métodos alternativos. Com o objetivo de diminuir o número de animais utilizados na pesquisa e minimizar a dor e o desconforto e buscar alternativas para a substituição dos testes *in vivo*, houve o surgimento do programa dos três R's. Batizado assim devido as iniciais de suas principais ações visadas, **R**eplacement, **R**eduction e **R**efinement, em português, substituição, redução e refinamento, o programa foi criado em 1959 em um estudo publicado por William Russell e Rex Burch. (CAZARIN; CORRÊA; ZAMBRONE, 2004)



Figura 5 - O zoologista William M.S. Russell e o microbiologista Rex L. Burch publicaram em 1959, “The Principles of Humane Experimental Technique”. Este trabalho foi um marco para a experimentação em animais por inserir o conceito dos 3 Rs, Repleacement, Reduction e Refinement. Embora diversos avanços na área tenham ocorrido, o cerne da ética e da busca a métodos alternativos se encontra nesta obra.

O primeiro dos três princípios, substituição refere-se a métodos que permitem evitar ou substituir o uso de animais. As substituições podem ocorrer de duas formas: substituições relativas ou absolutas. Substituições absolutas é quando substituímos animais por sistemas inanimados, tais como programas de computador. Já as substituições relativas animais mais sensíveis, como os vertebrados, com animais que a

evidência científica atual indica um potencial significativamente mais baixo para a percepção da dor, tais como alguns invertebrados.

Já a alternativa de redução, refere-se a qualquer estratégia que irá resultar em menor número de animais a ser utilizado para obter dados suficientes para responder à questão de pesquisa, ou para maximizar a informação obtida por animal e, portanto, potencialmente limitando ou evitando o uso subsequente de animais adicionais, sem comprometer o bem-estar animal.

O último dos Rs, o refinamento, refere-se à modificação da criação ou procedimentos experimentais para minimizar a dor e angústia, e para melhorar o bem-estar de um animal usado na ciência a partir do momento que nasce até sua morte. (RUSSELL; BURCH, 1959).

Considerados estes pontos, ainda há o fato que não é possível relacionar dados obtidos com experimentação animal com seres humanos. Em um texto se posicionando sobre experimentação animal, a Fundação Oswaldo Cruz – FIOCRUZ, afirma:

“...A experimentação funciona como uma escada. É preciso passar por vários degraus até se chegar ao ser humano. Quando se começa a desenvolver um novo medicamento contra o câncer, por exemplo, é necessário primeiramente utilizar a droga em culturas de células para verificar se ela é eficiente na eliminação de células cancerosas. O próximo passo é testar a droga em animais de laboratório de pequeno porte, que estejam com câncer, para se avaliar como o este organismos reagem. Posteriormente, a mesma droga será testada em animais de maior porte, que sejam mais próximos evolutivamente do homem, a fim de verificar se os efeitos da droga se mantêm quando lidamos com organismos mais parecidos com o ser humano.” (<http://agencia.fiocruz.br/perguntas-e-respostas>. Acessado em 17/09/2016)

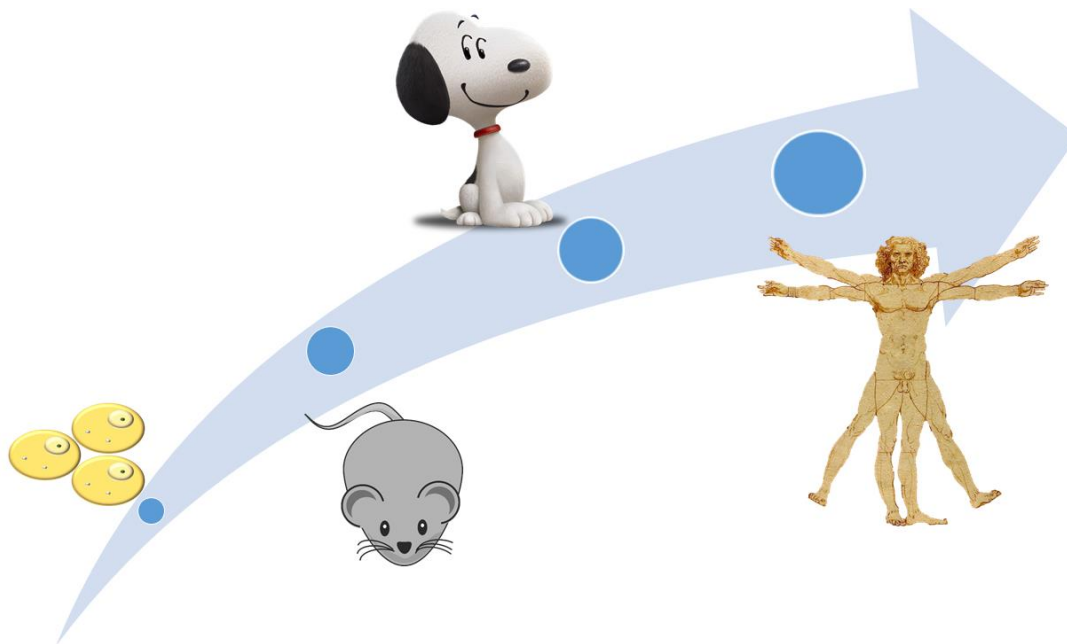


Figura 6 – Esquema das etapas necessárias para um medicamento, vacina ou, como em nosso caso, um biomaterial chegar ao ser humano. Primeiramente são realizados testes em culturas de células. A próxima etapa será experimentos em animais de pequeno porte, como ratos ou camundongos. O cãozinho Snoopy do desenho “Peanuts”, representa o passo seguinte, experimento em animais de maior porte como cães. Só aprovado em todos estes passos será levado ao ser humano.

No caso da pesquisa com biomateriais medico-odontológicos a despeito da importância ainda faltam alternativas à pesquisa com animais. Poucas são as metodologias de análise capazes de analisar e ranquear novos materiais a partir de ensaios *in vitro* ou *in silico*. Consequentemente, universidades e companhias recorrem, quase sempre, a testes convencionais fazendo uso de modelos animais, comprometendo avanços necessários nesta área.

1.7 Biologia sistêmica

Historicamente a ciência vem sendo construída utilizando-se do reducionismo. Partimos de fenômeno completo e buscamos fragmentá-lo para assim tentarmos entender suas partes isoladas, para então, somente após isso, reunir o conhecimento das partes e tentar compreender o todo do fenômeno.

A biologia sistêmica veio na contramão deste paradigma. Baseando-se no entendimento de que o todo é maior que a soma de suas partes, é o ramo que busca a compreensão das interações em redes gênicas, proteicas, metabólicas e assim por diante.

Redes de interação são construídas partindo de informações como interações físicas ou funções correlacionadas (NANDAGOPAL; ELOWITZ, 2011), e estudo destas redes de interações se faz útil pois possibilita ver de forma simples e explicativa fenômenos biológicos e patológicos, possibilitando assim aplicações biotecnológicas.

No caso da interação célula-biomaterial, uma vez conhecidas as quinases que estejam atuando neste processo a biologia sistêmica surge como uma opção na busca por elucidar vias e processos que atuem neste sentido. A partir de um pequeno número de proteínas, que aparentemente não tenham interações, é possível construir redes de interações complexas que revelem novas funções biológicas.

KARR et al. afirma que um dos problemas enfrentados na construção de redes, é a quantidade insuficiente de informações sobre estas interações. Embora nos últimos anos um grande número de dados tenha sido produzido, ainda carece de métodos matemáticos e computacionais para análise.

1.8 Bioinformática

Baseando-se em elementos da ciência da computação, matemática e estatística, a bioinformática é a área do conhecimento interdisciplinar que busca analisar e interpretar dados biológicos. Seu surgimento se deu com a necessidade de armazenar e manejar dados biológicos em rápida expansão, como por exemplo dados de biologia molecular.

A bioinformática foi uma ferramenta poderosa no desenvolvimento de diversas áreas da biologia. Na genética por exemplo, permitiu sequenciar e anotar, e com isto possibilitando a detecção de mutações nas mesmas. Na biologia estrutural, auxiliou a elucidar estruturas de moléculas como DNA, RNA e principalmente proteínas. Já na biologia sistêmica possibilitou o entendimento de redes biológicas. Isto além de outras áreas, como a evolução.

Um dos grandes trunfos da bioinformática na resolução de problemas biológicos está nos bancos de dados públicos online. O NIH – NCBI (National Institutes of Health – National Center for Biotechnology Information), por exemplo, abriga em seu site banco de dados de genes, genomas, RNAs, proteínas, artigos, livros, entre outros. Além disso, possui em seu site ferramentas como o BLAST (sigla em

inglês que significa: **B**asic **L**ocal **A**lignment **S**earch **T**ool) que possibilita ao usuário comparar sequências de resíduos de aminoácidos de diferentes proteínas ou de nucleotídeos de diferentes DNA.

Os bancos de dados biológicos crescem a cada ano. O UniProt (www.uniprot.org), por exemplo, um banco de dados gratuito sobre forma e funções de proteínas, cresce ano a ano em número de proteínas. UniProt é fruto de um consórcio gerado pelos Instituto Europeu de Bioinformática (EBI), Instituto Suíço de Bioinformática (SIB) e pelo Recurso de Informação de Proteína (PIR) (BATEMAN et al., 2015). Na figura 7 temos o gráfico do número de proteínas ano a ano no UniProt/Swiss-Prot. Outro banco de dados em constante expansão é o PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>). PubMed compreende mais de 26 milhões de citações de literatura biomédica, periódicos de ciências da vida, e os livros on-line. PubMed também fornece acesso à web sites adicionais relevantes e links para outros recursos de biologia molecular NCBI. É desenvolvido e mantido pelo Centro Nacional de Informações sobre Biotecnologia (NCBI), na Biblioteca Nacional EUA de Medicina (NLM), localizado no National Institutes of Health (NIH).

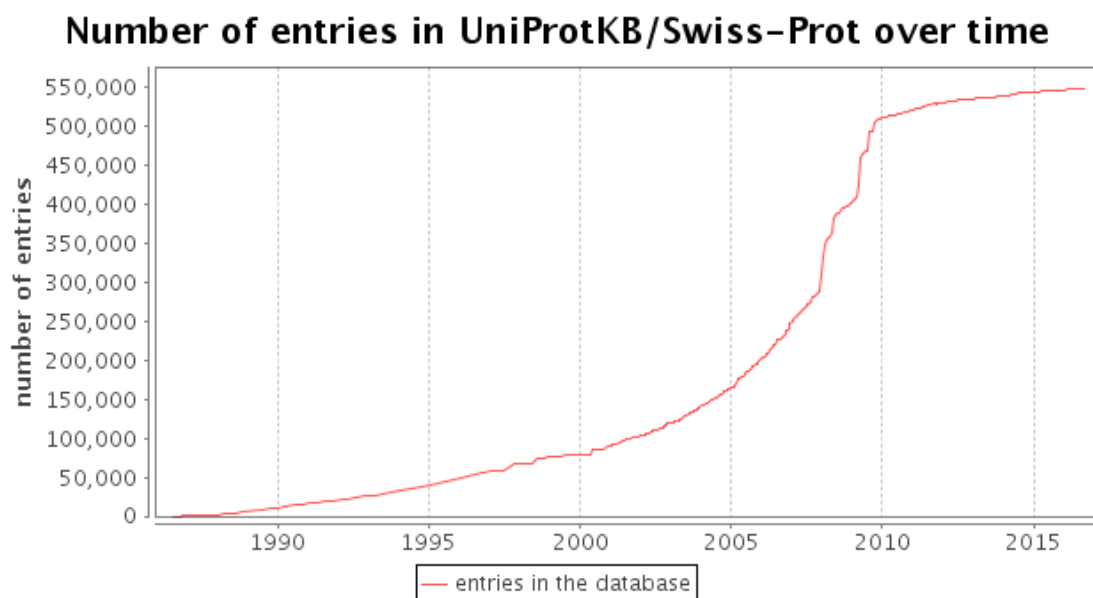


Figura 7 – Criado em 1986 o número de sequências depositadas cresce ano após ano desde então. Atualmente são 551.987 diferentes sequências depositadas. Fonte: <http://www.uniprot.org/statistics/Swiss-Prot>

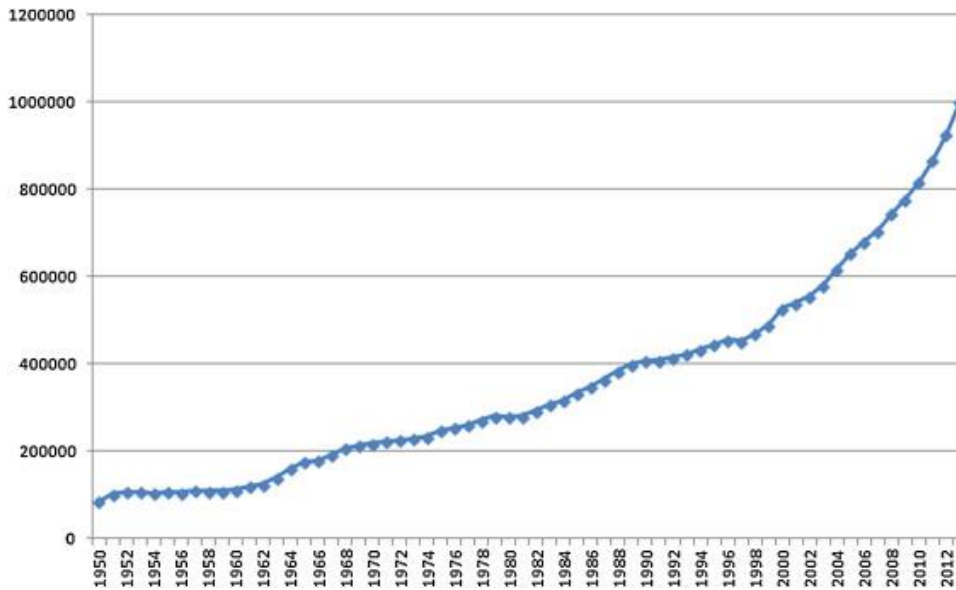


Figura 8 - Gráfico do número de trabalhos adicionados no PubMed por ano de 1950 até 2012. Fonte: VARDAKAS et al.

Com a consolidação dos bancos de dados de bioinformática novas ferramentas de análise foram criadas. STRING (SZKLARCZYK et al., 2015) é um banco de dados de interações proteicas. Alimentados por diversos outros bancos de dados como o próprio UniProt e o Pubmed, STRING é capaz de mostrar as interações entre proteínas isto permite a um pesquisador desenhar um mapa de interações que descreva um processo biológico observado a partir de algumas proteínas identificadas como relevantes em seus próprios experimentos.

Existem uma série de outros bancos de dados e ferramentas. Caso o pesquisador deseje trabalhar com drogas pode se utilizar do ChEMBL (<https://www.ebi.ac.uk/chembl>) ou STITCH (KUHN et al., 2014). Caso deseje elucidar estruturas de proteínas o Protein Data Bank (PDB) (<http://www.rcsb.org/pdb/home/home.do>) é uma arma poderosa. Gene Ontology (ASHBURNER et al., 2000; BLAKE et al., 2015) é uma ferramenta poderosa para estudo de funções moleculares, processos biológicos e componentes relacionados a um conjunto de genes. Enfim, bancos de dados estão constantemente sendo atualizados, expandidos e criados, o que torna cada vez mais útil e prático para o pesquisador o uso de técnicas de bioinformática em seus projetos.

Utilizando estas ferramentas e bancos de dados foi possível a construção de uma ferramenta de análise de biomateriais visando a redução do uso de experimentação animal na área médico-odontológica, contribuindo ao setor produtivo.

2. Objetivos

Amparados por ferramentas de bioinformática e utilizando conceitos de biologia sistêmica, o objetivo deste trabalho foi produzir uma metodologia computacional, alternativa ao uso de experimentação animal, capaz de detectar e analisar o quonoma da resposta da interação célula-biomaterial, obtida com ensaio de microarranjo de peptídeos. Estes dados servirão para a construção de um banco de dados para guiar a produção de biomateriais para a área médico-odontológica. Batizaremos este de OsteoBLAST.

3. Materiais e métodos

3.1 Cultivo celular

Foram utilizadas células tronco mesenquimais de medula óssea humana, obtidas da LONZA, as quais foram cultivadas em meio apropriado e recomendado pelo banco de células. Na rotina, o meio de cultivo era suplementado com antibióticos (100U/ml penicilina, 100mg/ml estreptomicina), e 10% Soro Fetal Bovino (SFB, Laborclin, Pinhais, PR, Brazil). Foram mantidos a 37°C e à atmosfera de 5% de CO₂.

3.2 Disco de titânio

Discos de três diferentes superfícies foram utilizados no presente trabalho: grupo Maquinado (Maq), grupo duplo ataque ácido (DAA) e tratado com nano-HA (Ti). Poliestireno foi utilizado como controle. Os discos de titânio foram doados pelo Diretor científico da S.I.N. – Sistema Nacional de Implantes, Sr. Fábio Bezerra.

3.3 Morfologia celular

As células foram semeadas nas diferentes superfícies de titânio a uma densidade de 50.000 células / disco. Após 4 h, as células foram fixadas com 2,5% de glutaraldeído em 0,1 M de tampão de fosfato de pH 7,3 durante 24 h. Amostras foram imersas em tetróxido de ósmio a 0,5% durante 40 minutos, e foram desidratadas através de uma série de álcoois, que secas em ponto crítico, foram metalizadas. O experimento foi realizado utilizando o microscópio eletrônico de varredura Quanta 200 – FEI Company com uma voltagem de 12,5 kV. As análises de EDX foram realizadas com uma tensão de 20 kV para avaliar os espectros de raios-X característicos de C e Ti nas amostras.

3.4 PamChip

Perfis de atividade de quinase foram determinados utilizando o sistema de microarray de tirosina-quinase PamChip (PamGene). Este microarray é constituído por 144 sequências únicas de peptídeos colocados em uma membrana porosa que permita o constante fluxo da mistura da reação. Cada peptídeo representa uma sequência de 15 aminoácidos correspondente a um local de fosforilação endógeno, que funciona como um substrato de tirosina quinase. A fosforilação é visualizada através da medição do sinal de ponto final fluorescente emitida como uma consequência de anticorpo anti-fosfotirosina marcado com fluoresceína de ligação (PY20).

A concentração de proteínas foi calculada pelo método de Bradford com entre 0,5 e 1,5 $\mu\text{g} / \text{mL}$, a variação na quantidade total de proteína aplicada ao array não influenciou a intensidade do sinal (dados não mostrados).

As amostras foram lisadas em reagente de M-PER contendo inibidores de fosfatase e de protease (Pierce). A mistura de reação consistiu em tampão de ABL 1 \times (Westburg), 100 $\mu\text{mol} / \text{L}$ de ATP (Sigma-Aldrich), 20 $\mu\text{g} / \text{mL}$ PY20 (Exalpha), e 5 μl de lisado. Antes de preencher os arrays, a mistura da reação foi bloqueada com 0,2% albumina sérica bovina (Calbiochem). O experimento foi repetido em triplicatas para cada amostra de cada grupo para maior reprodutibilidade técnica. Depois de carregar os lisados de células para as arrays de PamChip, a incubação foi iniciada durante 60 ciclos utilizando o PamStation96 seguido de lavagem e das medições de fluorescência FITC de todos os “spots” de peptídeos.

3.5 Bioinformática aplicada: OsteoBLAST.

Para a construção de uma rotina computacional de análise utilizou-se a linguagem R. A rotina utiliza o arquivo “_postwash.xlsx” saída da análise de PamChip. Os passos da rotina incluem a verificação da reprodutibilidade dos “spots”, normalização dos valores de intensidade para a retirada de valores discrepantes e por último comparação com os grupos controles. Posteriormente bancos de dados de bioinformática online, como UniProt, KEGG e Reactome foram consultados para inferências biológicas. Redes de interações proteicas foram construídas e analisadas com STRING e Cytoscape, respectivamente.

4. Resultados e Discussão

4.1 OsteoBLAST

O principal resultado deste trabalho é o algoritmo para a comparação de intensidades de fosforilação obtidas através do PamChip® tirosina quinase. Batizado de OsteoBLAST este algoritmo permite duas análises do material: 1) Permite ao usuário, conhecer proteínas que estejam diferencialmente fosforiladas em seu biomaterial. 2) Permite ao usuário comparar seu biomaterial com um banco de materiais depositados, a fim de dar ao usuário uma ideia do grau de semelhança com estes, levando a um melhor refinamento da análise.

O algoritmo foi escrito em R devido aos inúmeros pacotes de bioinformática que estão sendo feitos e atualizados nesta linguagem diariamente. A linguagem R foi concebida originalmente como um pacote estatístico, então diversas análises já vem programadas em comandos simples o que permite ao código ser mais compacto, rápido e elegante.

Para o melhor entendimento do OsteoBLAST é necessário um conhecimento sobre o chip PamChip® tirosina quinase.

4.2 PamChip® tirosina quinase

PamChip® tirosina quinase Array para PamStation®12 consiste de 144 peptídeos com sítios de fosforilação conhecidos, representando 100 proteínas diferentes. As sequências de peptídeos são derivados a partir da literatura ou previsões computacionais e são correlacionados com uma ou várias quinases.

As quinases presentes na amostra (de origem recombinante, celular ou de lisado de tecido) irão fosforilar ativamente o substratos dos spots do PamChip®, na presença de ATP. Anticorpo PY20, anti-fosfotirosina, conjugado com FITC, um composto fluorescente, presente na reação é utilizado para detectar a fosforilação de resíduos tirosina, assim sendo possível quantificar a fosforilação. As imagens são tomadas a cada 5 minutos, para gerar dados de cinética em tempo real (Figura 9).

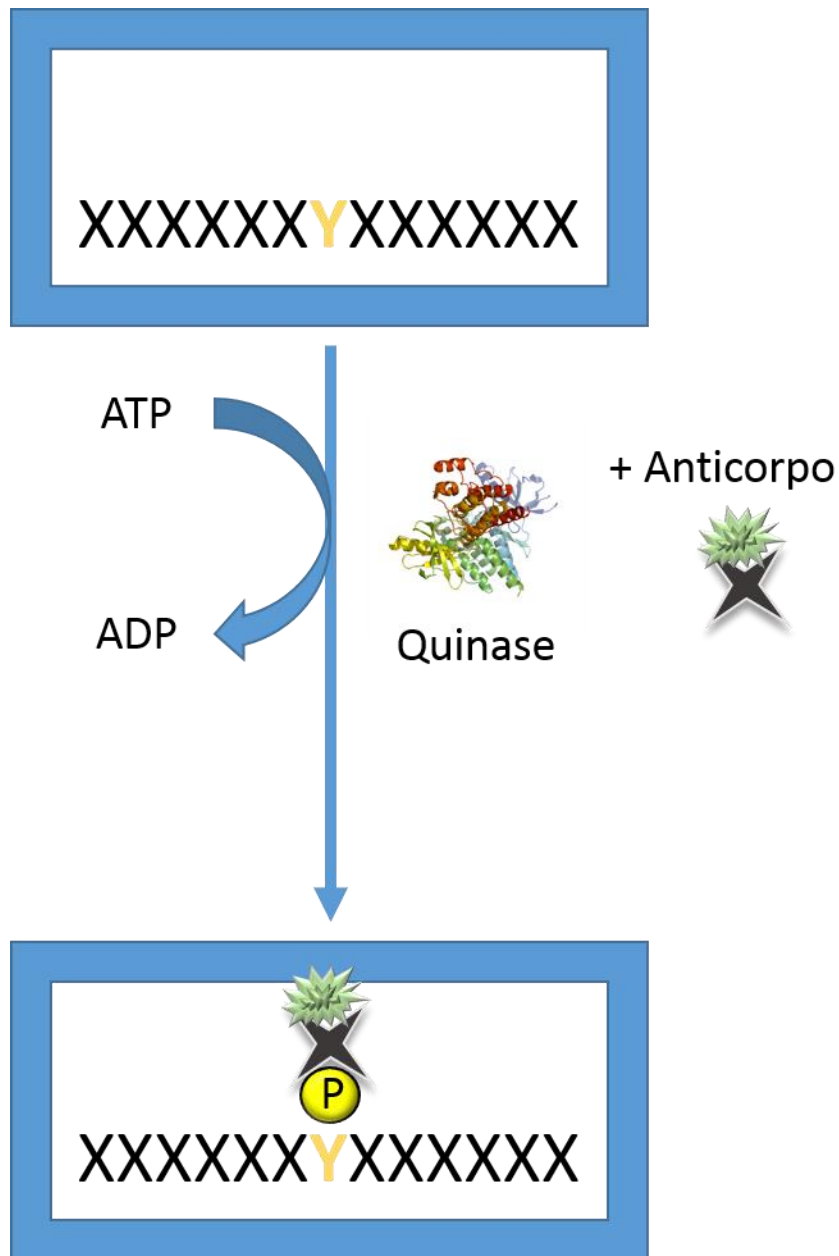


Figura 9 - Esquema de PamChip(r) tirosina quinase. XXXXXXXYXXXXXXXX representa uma sequência qualquer de resíduos de aminoácidos que contenha um resíduo tirosina (Y). Anticorpo PY20 conjugado com FITC serão transferidos ao sequência através da fosforilação do resíduo tirosina, catalisada pela atuação de uma quinase.

4.3 Algoritmo

OsteoBLAST parte dos dados de intensidade obtidos com PamChip® que estão no arquivo “postwash”, salvo após o término do experimento. Nele as intensidades do sinal já contém o background subtraído. É fundamental que o arquivo de entrada esteja no formato “csv” para que o algoritmo seja executado corretamente.

O algoritmo do OsteoBLAST considera como grupo controle, em sua análise, células tratadas em um disco de titânio sem tratamentos em sua superfície, chamamos de grupo Maquinado (Maq). Devido ao fato de não ter sua superfície modificada, este grupo controle é importante para entendermos os efeitos das alterações da topografia dos discos no quinoma das células.

Para análise do quinoma diferencial na amostra, bem como a comparação do biomaterial testado com outros de conhecida qualidade, o algoritmo OsteoBLAST utiliza-se de quatro etapas: Seleção, Normalização, Verificação e Comparação (Figura 10).

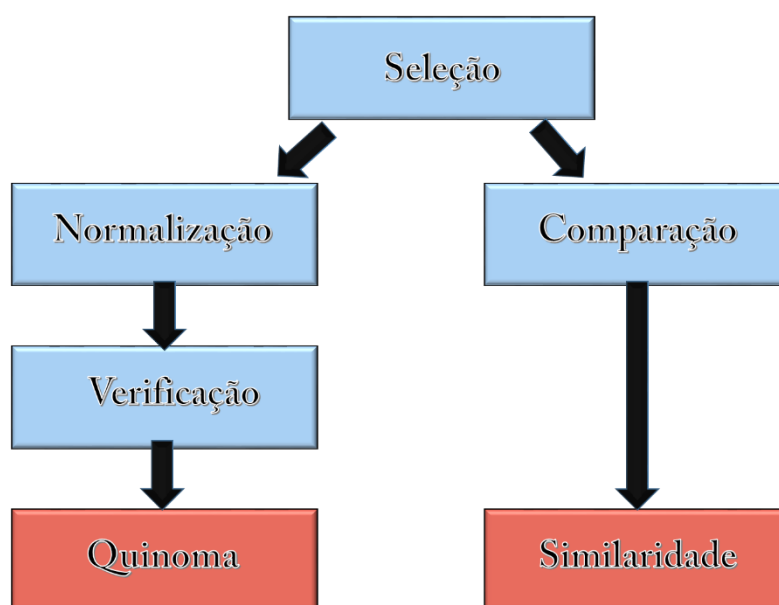


Figura 10 - Fluxograma do algoritmo de OsteoBLAST.

4.3.1 Seleção

Como dito anteriormente cada “spot” do chip contém a sequência de 13 resíduos de aminoácidos da região consenso da proteína que pode ser fosforilada. É sabido que a função de uma proteína está intimamente relacionada a sua estrutura, logo fosforilações inespecíficas são comuns em microarranjos de peptídeos. O primeiro passo para se trabalhar com os dados obtidos é selecionar os dados que possuam uma maior confiabilidade. Para isto usamos dois parâmetros: $P1 = Sd/A$, onde Sd é o desvio padrão e A é a média das triplicatas de sinais obtidos. E $P2 = A/M$, onde A é a média e M a mediana das triplicatas de sinais obtidos. Alguns métodos de análises para microarranjos de peptídeos, como o PepMatrix desenvolvido pelo Dr. Renato Milani (MILANI, 2010) sugerem que o nível de corte para “spots” confiáveis são $P1 < 20\%$ e

80% < P2 < 120%. Porém devido a algumas diferenças do chip utilizado no PepMatrix, que contém 1024 “spots”, com o utilizado pelo OsteoBLAST, que possui apenas 144 “spots”, algumas considerações foram feitas antes da escolha dos valores dos parâmetros.

ESCOLHA DOS PARAMETROS: A fim de se obter o maior número possível de resultados com uma alta confiabilidade as seguintes considerações foram feitas: 1) para cada parâmetro foram atribuídos três níveis de confiabilidade (“ALTO”, “MÉDIO” e “BAIXO”). 2) Uma classificação de 0 a 5 estrelas, chamada de Ce, conforme a combinação dos dois parâmetros (vide tabela a baixo). Por exemplo, caso um “spot” possua um valor P1 “ALTO” e P2 “ALTO” será atribuído um valor de 5 Ce. Por outro lado, caso um “spot” possua valores “BAIXO” para ambos os parâmetros, este receberá um valor de 0 Ce. A justificativa da escolha em criar esta classificação Ce é dar ao usuário do OsteoBLAST a possibilidade de escolher o nível de confiabilidade desejada para sua análise. O desejável é sempre escolher “spots” com Ce = 5, porém como a técnica de microarranjo de peptídeos possui um valor elevado é interessante dar ao usuário do OsteoBLAST a possibilidade de utilizar em sua análise dados de uma menor confiabilidade, porém sem ter a necessidade de repetir o experimento. A tabela 1 contém os valores de Ce para cada combinação de níveis dos parâmetros P1 e P2.

Tabela 1- Tabela de valores de Ce (classificação por estrelas) para "spots" dependendo do nível dos parâmetros P1 e P2.

P1	P2	Ce
ALTO	ALTO	5
ALTO	MÉDIO	4
MÉDIO	ALTO	4
ALTO	BAIXO	3
BAIXO	ALTO	3
MÉDIO	MÉDIO	2
BAIXO	MÉDIO	1
MÉDIO	BAIXO	1
BAIXO	BAIXO	0

A partir destas considerações, resta escolher os valores dos níveis de confiabilidade para ambos os parâmetros. Foram testados três combinação de valores para os dois parâmetros:

- 1) $P1 < 20\%$ e $80\% < P2 < 120\%$ - ALTOS/ $P1 < 50\%$ e $70\% < P2 < 140\%$ - MEDIOS. Valores fora destes limites foram considerados BAIXOS.
- 2) $P1 < 50\%$ e $80\% < P2 < 120\%$ - ALTOS/ $P1 < 75\%$ e $70\% < P2 < 200\%$ - MEDIOS. Valores fora destes limites foram considerados BAIXOS.
- 3) $P1 < 20\%$ e $80\% < P2 < 120\%$ - ALTOS/ $P1 < 75\%$ e $70\% < P2 < 140\%$ - MEDIOS. Valores fora destes limites foram considerados BAIXOS.

A combinação 1, que possui os valores mais rigorosos para a análise, retornou 57 “spots” $Ce = 5$ para o grupo Maq. Já a combinação 2, que possui valores menos rigorosos, retornou 91 “spots” com valor de $Ce = 5$ para o grupo Maq retornou 74 “spots” $Ce = 5$. A combinação 3 apresenta um rigor médio nos parâmetros e retornou 57 para o grupo Maq. As combinações 1 e 3 apresentaram a mesma quantidade de “spots” $Ce = 5$ para o grupo controle Maq. A diferença entre as duas combinações foi nos grupos de menor confiabilidade, na combinação 3 houve um aumento de 17 para 25 para o grupo Maq, “spots” de $Ce = 0$ comparados a combinação 1. Devido a seu maior rigor na análise, e ao fato de apresentar uma boa quantidade “spots” a combinação 1 foi a escolhida para o OsteoBLAST. A tabela 2 traz mais informações sobre estas combinações.

Tabela 2 - Resumo das frequências de Ce para cada combinação em cada grupo.

	Combinação 1			Combinação 2			Combinação 3		
	Ce	Frequências	%	Ce	Frequências	%	Ce	Frequências	%
Maquinado	0	17	11,80556	0	13	9,027778	0	25	17,36111
	1	23	15,97222	1	19	13,19444	1	18	12,5
	2	9	6,25	2	4	2,777778	2	6	4,166667
	3	12	8,333333	3	10	6,944444	3	13	9,027778
	4	26	18,05556	4	24	16,66667	4	25	17,36111
	5	57	39,58333	5	74	51,38889	5	57	39,58333

A etapa de seleção retorna um arquivo contendo as informações estatísticas do grupo analisado. Estes arquivos são nomeados na seguinte forma “stat_[nome do grupo].dat”, e contém as médias, medianas e desvio-padrão das triplicatas, também contém o cálculo dos parâmetros P1 e P2 e sua classificação como ALTO, MÉDIO ou BAIXO, e por último seu Ce. A figura 11 mostra um arquivo saída da etapa de seleção.

	A	B	C	D	E	F	G	H	I
1	protein	M_m	A_m	SD_m	P1_m	cP1_m	P2_m	cP2_m	cE_m
2	41_654_666	89	83.33333	10.69268	0.029728	Alto	0.93633	Alto	5
3	ACHD_383_395	1	3.666667	7.371115	0.69282	Baixo	3.666667	Baixo	0
4	AMPE_5_17	10	7.333333	4.618802	0.436839	Medio	0.733333	Medio	2
5	ANXA1_14_26	30	36.66667	14.22439	0.11024	Alto	1.222222	Medio	4
6	ANXA2_17_29	5	4.666667	6.506407	0.60924	Baixo	0.933333	Alto	3
7	ART_004_EAIYAAPFAKKKXC	447	473.3333	123.6217	0.111064	Alto	1.058911	Alto	5
8	B3AT_39_51	-4	-4	8	15.09967	Baixo	1	Alto	3
9	CTNB1_79_91	18	16.33333	9.609024	0.183303	Alto	0.907407	Alto	5
10	C1R_199_211	2	4.333333	4.041452	0.708872	Baixo	2.166667	Baixo	0
11	CALM_93_105	-1	-4	8.888194	1.5	Baixo	4	Baixo	0
12	PGFRB_1014_1028	40	42.33333	23.58672	0.260134	Medio	1.058333	Alto	4
13	ART_003_EAI(pY)AAPFAKKKXC	1305	1356	558.2499	0.244596	Medio	1.03908	Alto	4
14	CALM_95_107	4	4.333333	2.516611	0.391649	Medio	1.083333	Alto	4
15	CBL_693_705	21	23.66667	5.507571	0.072928	Alto	1.126984	Alto	5
16	CD3Z_116_128	0	3.666667	12.89703	0.314918	Medio	3.666667	Baixo	1
17	CD3Z_146_158	-9	-6.33333	7.371115	-1.3384	Baixo	0.703704	Medio	1
18	CD79A_181_193	1876	1907.667	296.7698	0.080553	Alto	1.01688	Alto	5
19	CDK2_8_20	183	184	20.51828	0.095996	Alto	1.005464	Alto	5
20	CDK7_157_169	27	28.66667	10.59874	0.204652	Medio	1.061728	Alto	4
21	CRK_214_226	17	15.66667	8.082904	0.264044	Medio	0.921569	Alto	4

Figura 11 - Exemplo de arquivo saída da etapa de seleção. A primeira coluna, protein, contém o nome das proteínas de acordo com o código do spot do PamChip®. Já nas colunas M_m, A_m e SD_m, contém respectivamente a mediana, média e desvio padrão de cada proteína. P1_m e P2_m contém os valores dos parâmetros P1 e P2, e cP1_m e cP2_m contém a classificação. cE_m contém os valores da classificação de estrelas baseada nos parâmetros P1 e P2.

4.3.2 Normalização

Munido dos dados dos spots selecionados no passo um, o algoritmo do OsteoBLAST realiza então seu segundo passo, a normalização dos spots. A normalização é um passo fundamental para o sucesso da análise uma vez que ela retira tanto ausências completas de fosforilação como fosforilações inespecíficas, que não são o foco de análise do chip. No processo de normalização são definidos dois percentis como nível de corte. p5 , o quinto percentil, que será o nível de corte de dados com menores valores, e p95 , o nonagésimo quinto percentil, que será o nível de corte para dados com maiores valores. Estes valores de percentis são definidos para cada triplicata do experimento. Uma vez estes valores definidos a normalização procede de acordo com a equação

$$N_{ij} = \frac{n_{ij} - p5_j}{p95_j - p5_j} \times 100 \quad (1)$$

Onde n_{ij} é o valor original da spot, $p5_j$ é o valor do quinto percentil da triplicata, $p95_j$ é o nonagésimo quinto percentil da triplicata e N_{ij} será o valor normalizado (Figura 12).

Spot	ECM			p5			p95				ECM_norm		
2	90	86	91	26.5	28.9	27.1	982.5	785.4	977.9	6.642259414	7.547918	8.82414	
5	50	42	52							2.458158996	1.731659	4.7223	
16	2511	2139	2381							259.8849372	278.9293	249.6	
17	252	208	236							23.58786611	23.67482	24.074	
25	917	753	857							93.14853556	95.71712	89.387	
30	53	48	54							2.771966527	2.524785	4.9326	
34	978	759	1001							99.5292887	96.51024	104.5	
35	182	162	216							16.26569038	17.59418	21.970	
37	188	180	188							16.89330544	19.97356	19.026	
38	23	24	25							-0.36610879	-0.64772	1.8826	
51	28	31	25							0.156903766	0.277594	1.8826	
52	75	63	66							5.073221757	4.507601	6.1947	
53	83	82	89							5.910041841	7.019167	8.6137	
54	154	140	147							13.33682008	14.68605	14.713	
59	470	372	484							46.39121339	45.3536	50.157	
64	64	71	61							3.922594142	5.565102	5.668	
69	42	39	47							1.621338912	1.335096	4.1964	
70	41	34	40							1.516736402	0.674157	3.4602	
73	54	63	54							2.876569038	4.507601	4.9326	
74	21	20	28							-0.57531381	-1.17647	2.1981	
80	69	53	65							4.445606695	3.185724	6.0896	
84	75	59	64							5.073221757	3.97885	5.9844	

Figura 12 – Exemplo da etapa de normalização. Na primeira coluna temos o código do spot, nas 2, 3 e 4 temos os valores de intensidades obtidos para as triplicatas para cada spot. Nas colunas 5, 6 e 7 temos os valores do quinto percentil para cada coluna de triplicatas, nas colunas 8, 9 e 10 temos os valores do nonagésimo quinto percentil de cada coluna de triplicatas e nas colunas 11, 12 e 13 temos os valores de intensidade já normalizados de acordo com a equação 1.

4.3.3 Verificação

O terceiro passo do OsteoBLAST então parte dos valores obtidos com a normalização para verificar quais spots apresentam fosforilação significativamente diferentes. Para isso o biomaterial testado será comparado com o controle Maq, e um teste t bicaudal é realizado com valor de $p < 0,05$. Para a realização do teste t é importante verificar duas condições:

- 1) Se os dados são normais. No entanto, esta condição não interfere devido ao baixo número de replicações, o que impede determinações precisas de normalidade. Além do passo dois do algoritmo, a Normalização, já diminuir as variações das médias;
- 2) Os dados precisam ser homocedásticos, isso significa não podem apresentar grandes as diferenças nas variâncias das replicações. Para isso o

OsteoBLAST realiza um teste F. Se as variâncias não forem estatisticamente diferentes, o teste t é aplicado. Agora caso as variâncias sejam diferentes, é aplicado o teste U de Mann-Whitney, versão não paramétrica do teste t de Student.

Após o teste estatístico ser realizado, OsteoBLAST retorna ao usuário quatro arquivos no formato dat, dois para cada grupo controle. Os arquivos que contém as informações sobre a significância estatística do teste são batizados “ttest_[nome do grupo controle]_[nome do grupo teste].dat”. Neles existem cinco colunas, a primeira contém o número do spot no chip. A segunda coluna contém o valor do teste F, e na terceira contém se o valor da segunda coluna é estatisticamente significativa. Caso seja, a terceira coluna constará o valor 1, caso contrário constará o valor 0. A quarta coluna contém os valores de p obtidos com o teste t. E na quinta se o teste obteve significância estatística, utilizando a mesma lógica da terceira coluna (Figura 13).

A	B	C	D	E
	var_m_daa	sigvar_m_daa	pv_m_daa	sigpv_m_daa
2	0.661979882	1	0.089846441	0
16	0.952325163	1	0.226877747	0
17	0.32568383	1	0.635565282	0
25	0.816189788	1	0.979658597	0
30	0.072554385	1	0.031171569	1
34	0.67730115	1	0.045825162	1
35	0.490410757	1	0.265820128	0
37	0.586373959	1	0.527914785	0
38	0.252108596	1	0.033087334	1
51	0.775808271	1	0.792594769	0
52	0.48097071	1	0.502640082	0
53	0.713162759	1	0.709321675	0
54	0.987623346	1	0.228412271	0
59	0.938254084	1	0.015987487	1
64	0.8923699	1	0.026997127	1
69	0.191542863	1	0.556434509	0
70	0.060552513	1	0.139996305	0
73	0.826141815	1	0.408978538	0
80	0.11041283	1	0.017219254	1
84	0.211543284	1	0.111197807	0
87	0.514676079	1	0.427124058	0
88	0.259235385	1	0.55027736	0

Figura 13 - Modelo saída de resultados da verificação de fosforilações significantemente diferentes entre dois grupos, neste caso a comparação entre os grupos Maq e DAA. A

primeira coluna contém o código dos spots em PamChip®, a coluna var_m_daa contém os valores do teste F, caso eles sejam menores que 0,05 a coluna sigvar_m_daa recebe o valor 1, caso contrário recebe 0. A coluna pv_m_daa recebe o valor de P para o teste t, caso o valor da coluna sigvar_m_daa seja 1, ou do teste U, caso o valor da coluna sigvar_m_daa seja 0. Por último, a coluna sigpv_m_daa apresenta valores de 1 ou 0 de acordo com a significância dos valores da coluna pv_m_daa, 1 caso $<0,05$ e 0 caso $>0,05$.

O segundo arquivo gerado no passo da verificação fornece informações sobre as proteínas que apresentem fosforilação diferenciais, para cada grupo controle. Estes são batizados de “rptProtein_[nome do grupo controle]_[nome do grupo teste].dat”, e são compostos de 4 colunas. A primeira trás o nome do spot como consta no arquivo saída do PamChip. A segunda coluna trás o nome da proteína em questão conforme consta no Uniprot. A terceira e quarta colunas, trazem o nome do gene que codifica a proteína e o código de acesso no Uniprot, respectivamente (Figura 14). Através deste arquivo saída, o usuário pode partir para bancos de dados de bioinformática online para uma análise de redes proteicas.

B	C	D	E
EGFR_1165_1177	Epidermal growth factor receptor	EGFR	P00533
FRK_380_392	Tyrosine-protein kinase FRK	FRK	P42685
K2C6B_53_65	Keratin, type II cytoskeletal 6B	KRT6B	P04259
NCF1_313_325	Neutrophil cytosol factor 1	NCF1	P14598
PGFRB_572_584	Platelet-derived growth factor receptor beta	PDGFRB	P09619
PGFRB_709_721	Platelet-derived growth factor receptor beta	PDGFRB	P09619
VGFR2_944_956	Vascular endothelial growth factor receptor 2	KDR	P35968

Figura 14 - Modelo de saída do arquivo rptProtein. A primeira coluna contém o código dos spots de PamChip®, a segunda contém o nome da proteína segundo UniProt, a terceira contém o gene que codifica e a quarta o código de acesso da proteína em UniProt.

4.3.4 Comparação

O quarto e último passo do algoritmo OsteoBLAST tem como objetivo comparar um biomaterial em desenvolvimento com biomateriais com qualidade já reconhecida e que sejam de uso na rotina médico odontológica. Para isso, OsteoBLAST utiliza os resultados obtidos da etapa da seleção, em seguida realiza os cálculos utilizando a seguinte equação,

$$\chi^2 = \frac{1}{M} \sum_{i=1}^M \left(\frac{I_0[i] - I_t[i]}{1 - |\sigma_0[i] - \sigma_t[i]|} \right)^2 \quad (2)$$

Onde χ é o parâmetro de similaridade, M é o total de “spots” selecionados na etapa da normalização, I_0 e σ_0 são, respectivamente, a média e o desvio-padrão da intensidade de sinal do grupo controle, e I_t e σ_t são a média e o desvio-padrão da intensidade de sinal do grupo que está sendo testado.

As superfícies teste foram comparadas com o grupo Maq e os seguintes valores foram obtidos: 31,12 para o grupo DAA e 31,75 para o grupo Ti. O grupo Ti é um biomaterial em desenvolvimento, enquanto o grupo DAA é um biomaterial com uma superfície amplamente utilizado na rotina médico odontológica. A proximidade nos valores obtidos quando comparados ao grupo controle, demonstra que o tratamento feito na superfície do grupo Ti resultou em uma superfície comparável a uma de eficácia conhecida.

4.4 Superfícies teste

Para analisarmos a eficácia do OsteoBLAST foram testadas duas superfícies modificadas. A primeira é a superfície com duplo ataque ácido (DAA), uma superfície mundialmente utilizada com variações topográficas na escala micrométrica. A segunda superfície é um protótipo desenvolvido por parceiros. Ela contém uma cobertura de hidroxiapatita com variações na escala nanométrica em sua topografia (Ti). É sabido que modificações físico-químicas na superfície do biomaterial alteram o comportamento celular (ZAMBUZZI et al., 2014). As figuras 15 e 16 exemplificam diferenças no comportamento celular em superfícies modificadas.

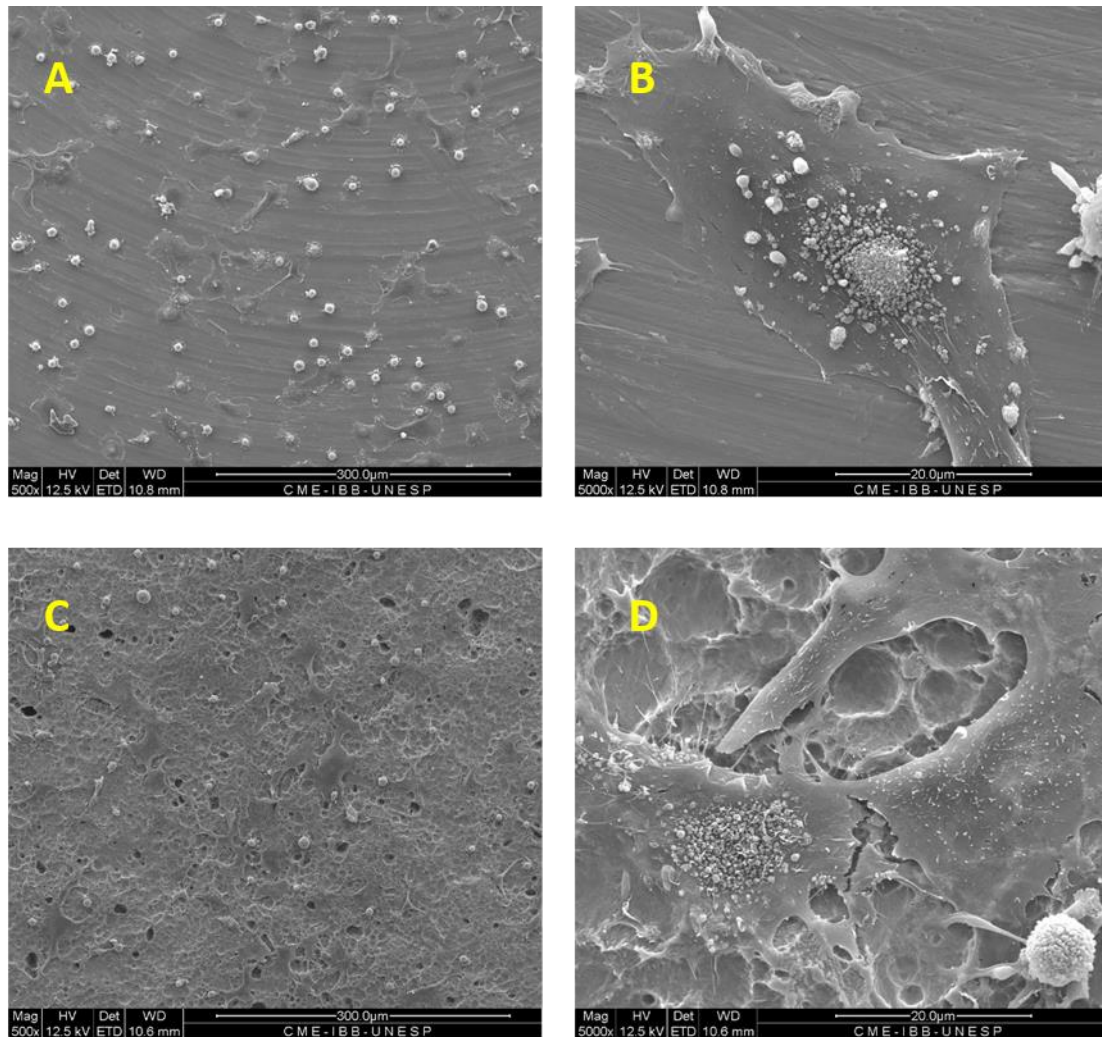


Figura 15 - Comportamento celular frente a diferentes tratamentos na superfície. Superfície maquinada em com magnificação de 500x (A) e de 5000x (B), e superfície com cobertura de hidroxiapatita em escala nanométrica (HANano). É possível perceber que células estão com um processo de “spreading” avançado na superfície com HANano em comparação com a superfície maquinada. Imagens obtidas no CME – IBB.

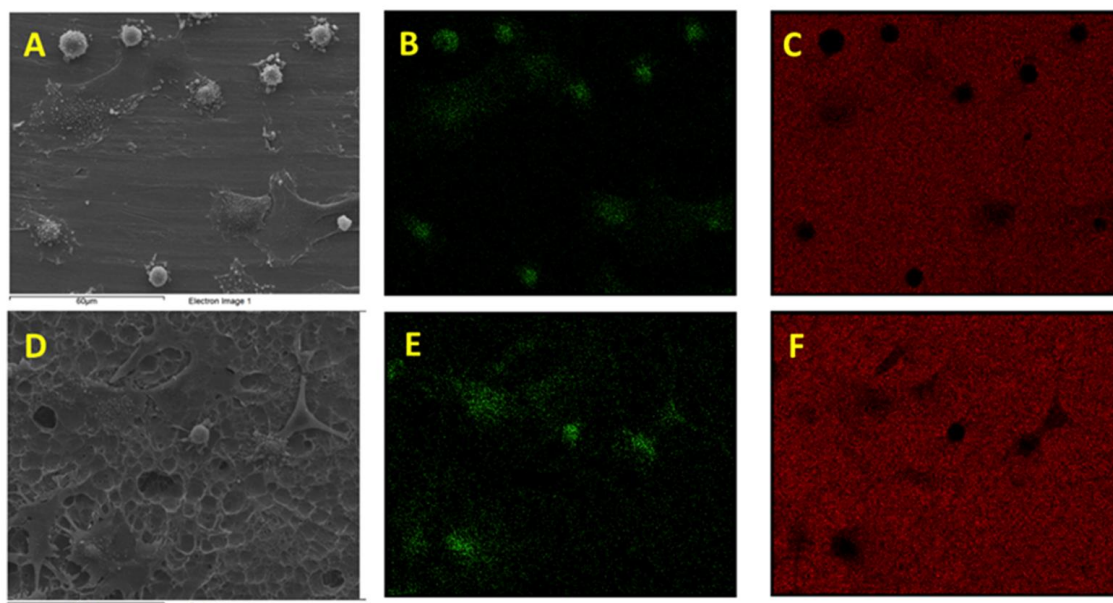


Figura 16 - Localização de carbono e titânio nos discos. As figuras A e D contêm imagens da microscopia eletrônica para as superfícies maquinada e com hidroxiapatita em escala nanométrica. As figuras B,C,E e F foram obtidas com EDX. B e E mostram a localização de carbono nas superfícies, e C e F mostram a localização de titânio na amostra. A maior dissipação de carbono no grupo com superfície modificada (E) pode indicar um processo de “spreading” mais acelerado que na superfície maquinada (B).

Estas superfícies foram então submetidas a análise com PamChip® tirosina quinase e seus resultados analisados com OsteoBLAST. Os resultados referentes ao grupo tratado com duplo ataque ácido recebeu o nome “DAA” e o grupo protótipo recebeu o nome “Ti”.

4.5 Resultados para a superfície com duplo ataque ácido (DAA)

A etapa de seleção retornou 40 spots com $C_e = 5$, ideais para análise por possuir ambos os parâmetros ALTOS, 42 spots com $C_e = 4$, spot que possui um parâmetro ALTO e outro MÉDIO. O usuário que desejasse poderia então selecionar a confiança de 4 ou mais e teria 82 spots para sua análise. Para este trabalho utilizou-se a confiança no processo de seleção de $C_e = 5$. A tabela 3 traz o número de spots por valor de C_e .

Tabela 3 – Frequência de spots por nível de confiança no processo de seleção de OsteoBLAST para a superfície com duplo ataque ácido (DAA).

Ce	Frequências de spots
0	34
1	9
2	4
3	15
4	42
5	40

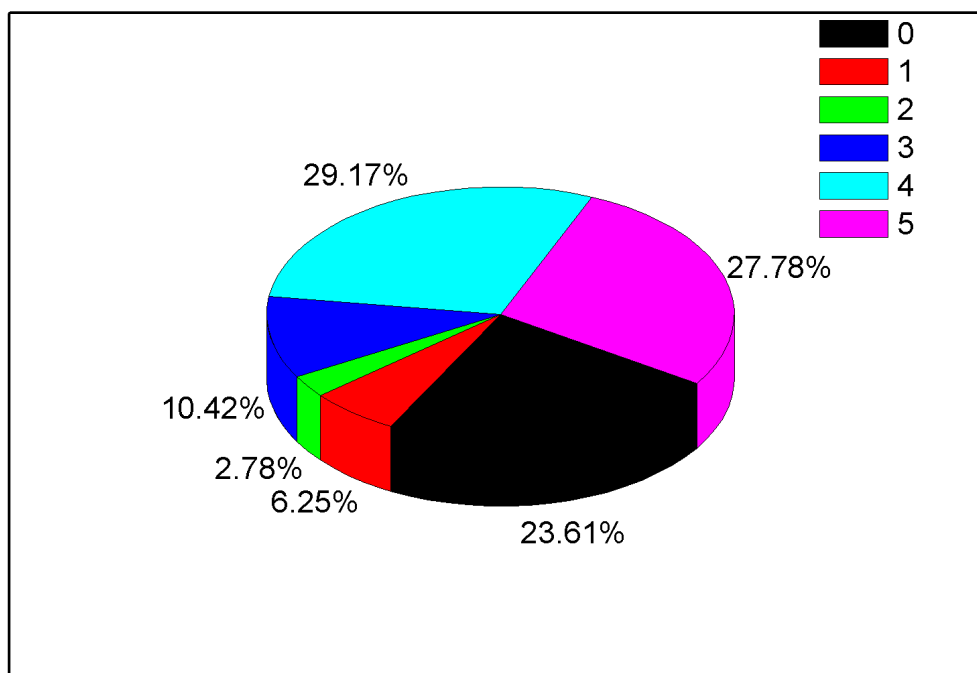


Figura 17 - Gráfico de pizza para as frequências de Ce do grupo DAA. 40 spots (27,78%) possuem Ce = 5, estes passarão pelas etapas de normalização e verificação.

Após os spots selecionados serem normalizados no passo dois do algoritmo OsteoBLAST, a verificação da fosforilação diferencial foi então realizada no passo três. O grupo DAA apresentou nove proteínas com fosforilação estatisticamente diferenciadas quando comparadas ao grupo controle Maq. As informações sobre as

proteínas detectadas com fosforilação diferencial em relação aos grupos controles estão na Tabela 4.

Tabela 4 - Lista de proteínas com diferenças estatísticas detectadas na etapa de verificação. As informações sobre o gene que as codificam, o resíduo fosforilado e o código de acesso no Uniprot completam a tabela.

Proteína	Gene	P	Uniprot
Maq Controle			
Epidermal growth factor receptor	EGRF	Y 1172	P00533
Gamma-enolase	ENO2	Y 44	P09104
Ephrin type-A receptor 4	EPHA4	Y 596	P54764
Tyrosine-protein kinase FRK	FRK	Y 387	P42685
Keratin, type II cytoskeletal 6B	KRT6B	S 60	P04259
Neutrophil cytosol factor 1	NCF1	S 320	P14598
Platelet-derived growth factor receptor beta	PDGFRB	Y 579/581	P09619
Platelet-derived growth factor receptor beta	PDGFRB	Y 716	P09619
Vascular endothelial growth factor receptor 2	KDR	Y 951	P35968

É importante ressaltar que embora a proteína *Platelet-derived growth factor receptor beta* (PDGFRB) apareça duas vezes na lista não se trata de um resultado duplicado pois os resíduos fosforilados são diferentes. As fosforilações em ambos os resíduos são devido a auto fosforilações, porém quando fosforilada no resíduo Y579/581 PDGFRB é ativada interagindo com proteínas da família das Src (BAXTER et al., 1998; PAUKKU et al., 2000; WARDEGA; HELDIN; LENNARTSSON, 2010). Já ao ser fosforilada no resíduo Y716, PDGFRB também se encontra ativada, assim participando da ativação da via Ras-Raf-Mek-Erk (ARVIDSSON et al., 1994).

4.6 Resultados para a superfície com hidroxiapatita na escala manométrica

(Ti).

A etapa de seleção retornou 17 spots com Ce = 5 para o grupo Ti, ideais para análise por possuir ambos os parâmetros ALTOS. Para este trabalho utilizou-se a confiança no processo de seleção de Ce = 5. A tabela 5 e a figura 18 trazem o número de spots por valor de Ce.

Tabela 5 - Frequência de spots por nível de confiança no processo de seleção de OsteoBLAST para a superfície com hidroxiapatita em escala nanométrica (Ti).

Ce	Frequências de spots
0	47
1	18
2	10
3	25
4	27
5	17

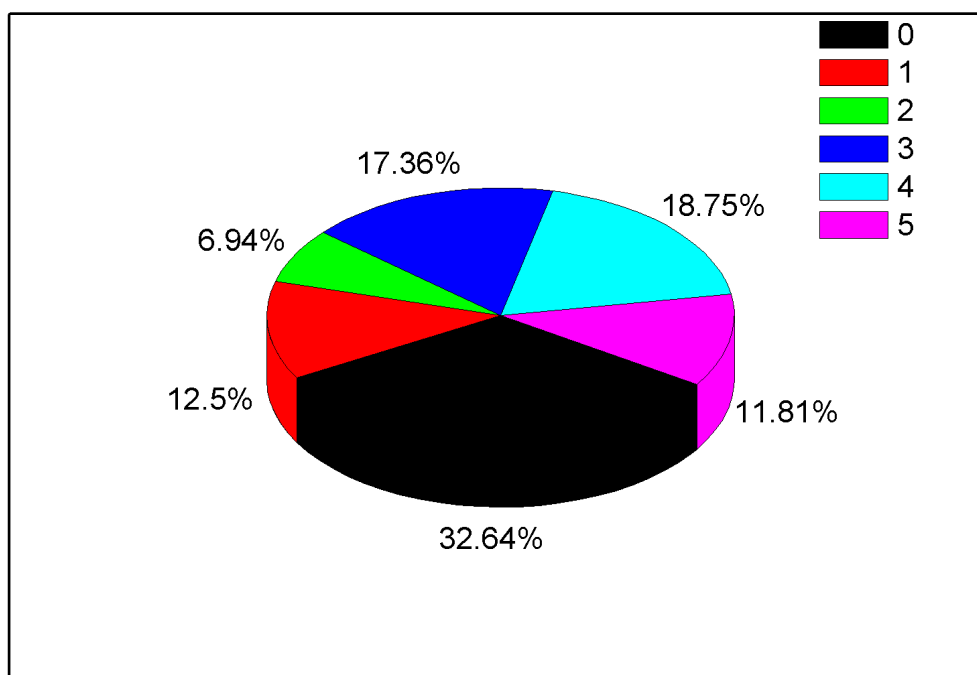


Figura 18 - Gráfico de pizza para as frequências de Ce do grupo Ti. Apenas 17 spots (11,81%) possuem Ce = 5, estes passarão pelas etapas de normalização e verificação.

Novamente, os spots selecionados foram normalizados na segunda etapa de OsteoBLAST, a etapa três verificou quais possuem fosforilação diferencial em relação aos grupos controles. O grupo Ti apresentou duas proteínas com fosforilação

estatisticamente diferenciadas em comparação aos dois grupos controle. As informações sobre as proteínas encontradas estão na tabela 6.

Tabela 6 - Lista de proteínas com diferenças estatísticas detectadas na etapa de verificação para o grupo Ti. As informações sobre o gene que as codificam, o resíduo fosforilado e o código de acesso no Uniprot completam a tabela.

Proteína	Gene	P	Uniprot
Maq Controle			
Neutrophil cytosol factor 1	NCF1	S 320	P14598
3-phosphoinositide-dependent protein kinase 1	PDPK1	Y 373/376	O15530

4.7 Análise Biológica

Para a análise biológica das proteínas diferencialmente fosforiladas nos grupos uma abordagem sistêmica foi adotada. Redes proteicas foram construídas a partir das proteínas diferencialmente utilizando STRING. Após isso Cytoscape (SHANNON et al., 2003) foi utilizado para buscar proteínas que apresentem centralidades importantes nas redes. As duas centralidades buscadas foram *Betweenness* e *Closeness*. O valor de *Betweenness* nos diz capacidade de interação e/ou sinalização com outras proteínas, processos biológicos, ou seja, quanto mais alto o valor de *Betweenness* maior será essa capacidade. Já os valores de *Closeness* nos indicam a possibilidade de um nodo fazer parte de uma regulação. Quanto menor o valor de *Closeness* maior a possibilidade de o nodo participar de uma regulação. Com o app MCODE clusters de proteínas foram obtidos, e após isso DAVID tools (HUANG; SHERMAN; LEMPICKI, 2009) foi utilizado para a obtenção de funções biológicas presentes nestes clusters.

4.7.1 Análise de redes proteicas para o grupo DAA-Maq

O grupo DAA apresentou sete proteínas com fosforilação diferenciada em relação ao controle Maq. São elas: EGFR (Y1172), ENO2 (Y44), EPHA4 (Y596), FRK (Y387), KRT6B (S60), KDR (Y951), NCF1 (S320), e PDGFRB (Y579/581 e Y716), esta última em dois sítios de fosforilação diferentes, este último para dois sítios de fosforilação diferentes.

Utilizando o banco de dados STRING uma rede de interação entre estas proteínas foi construída. Os parâmetros para a construção desta rede foram score ≥ 0.9 ,

para uma maior confiabilidade, como parâmetros para os Métodos de Predição Ativos utilizou-se as opções Experimentos, Bancos de Dados e Mineração de texto (busca na literatura). O score mostra a probabilidade aproximada que um ligação predita exista entre duas enzimas no mesmo mapa metabólico no banco de dados KEGG, ou seja, quanto maior o valor do score maior a confiabilidade no resultado. A escolha dos Métodos de Predição Ativos foi com o critério de serem parâmetros constantemente atualizados. Dessa forma tornando nosso resultado mais confiável.

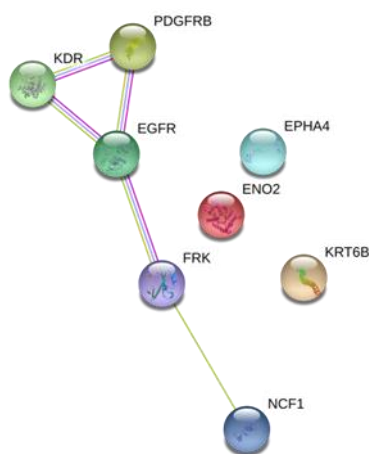


Figura 19 - Rede obtida através de STRING database para EGFR, ENO2, EPHA4 FRK, KRT6B, KDR, NCF1, e PDGFRB.

A rede foi expandida para 10 interações de primeira camada e 200 de segunda camada. Interações de primeira camada são interações diretas com as proteínas que estavam na rede inicial. Já interações de segunda camada são interações entre proteínas que não se encontravam na rede inicial. A figura 20 apresenta rede expandida. É importante notar que as proteínas ENO2 e KRT6B não apresentam comportamento de rede para em nenhum caso.

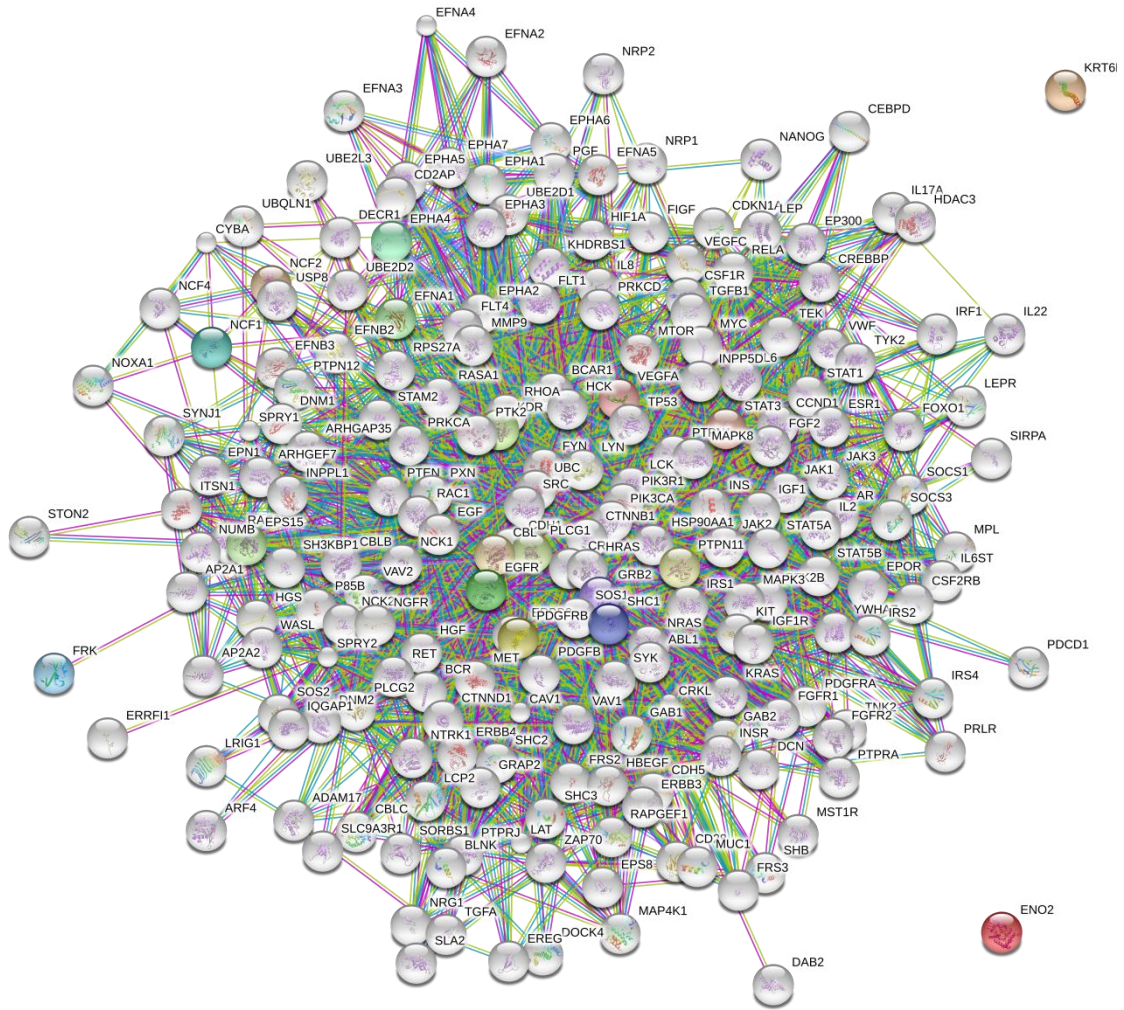


Figura 20 - Rede expandida buscando mais interações. A partir das seis primeiras proteínas a rede foi expandida, utilizando os seguintes parâmetros: Não mais do que 10 interações na primeira camada, não mais do que 200 interações na segunda camada e índice de confiança superior a 0.900, considerado altíssimo.

A partir do resultado de nossa rede expandida utilizamos o programa *Cytoscape* para realizarmos análises baseadas em topografias da rede utilizando a ferramenta *NetworkAnalyzer*. As figuras 21 e 22 trazem os gráficos destas centralidades como função do número de vizinhos na rede de DAA-Maq.

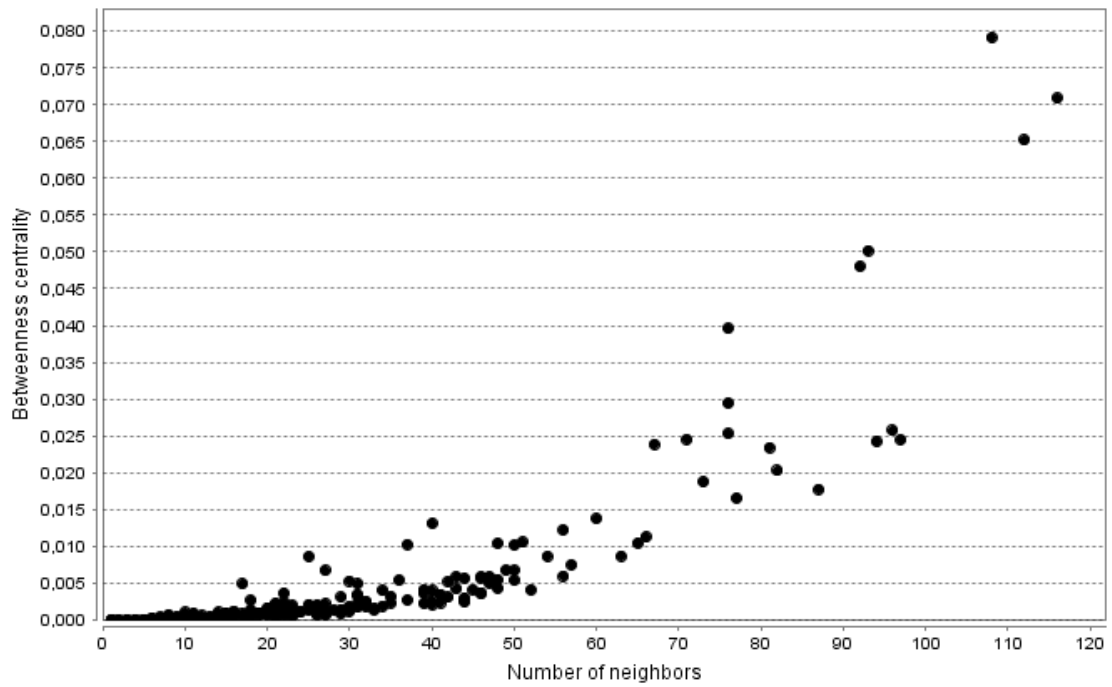


Figura 21 - Gráfico de Betweenness centrality para a rede DAA - Maq. EGFR(0.079), Src(0.071) e GRB2(0.065) foram nodos que apresentaram os maiores valores.

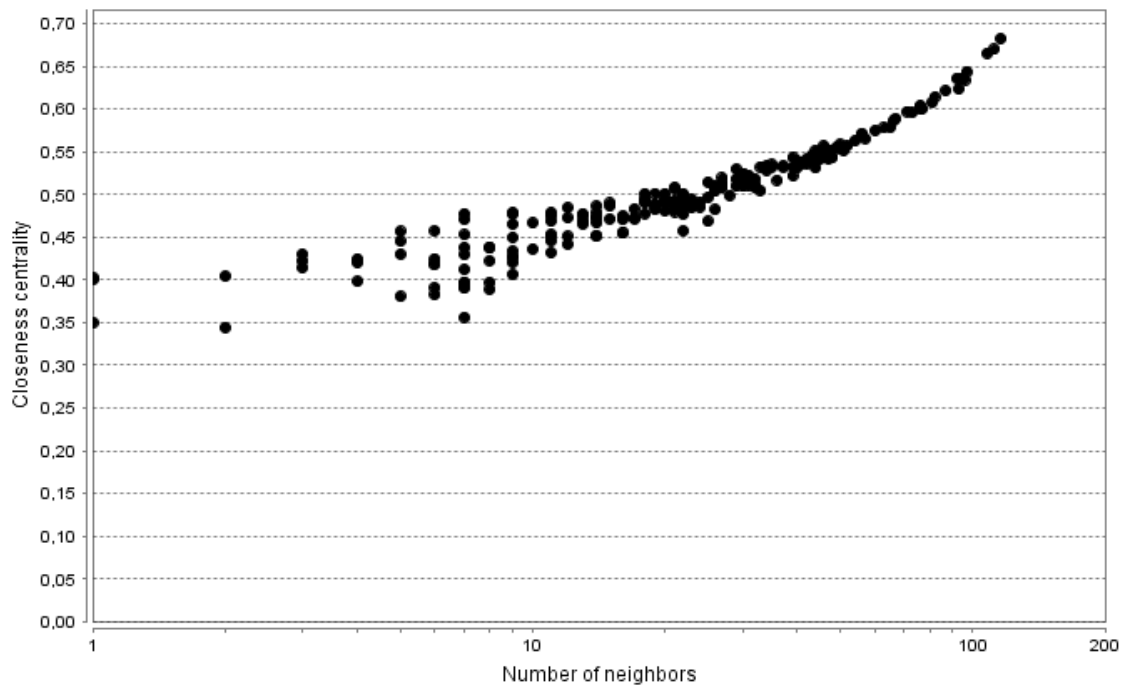


Figura 22 - Gráfico de Closeness centrality para a rede de DAA - Maq. Src(0.682), GRB2(0.670) e EGFR(0.666) foram nodos que apresentaram maiores valores.

EGFR, Src e GRB2 foram nodos que se destacam em ambas centralidade o que indica que são proteínas que sinalizam funções importantes dentro da rede obtida.

Em seguida foram obtidos clusters de proteínas, para nossa rede expandida. Para isso utilizou-se Cytoscape o app MCODE que fornece a opção de modificarmos alguns parâmetros para obtermos clusters com melhores scores, e dentro dessas opções estão:

- Degree Cutoff: controla o mínimo número de ligações necessários para um nó ter um valor de score. Por exemplo, nodos que compartilham apenas uma ligação com um outro nodo tem grau 1;
- Node Score Cutoff: esse é o parâmetro mais influente para o tamanho do cluster. Durante a expansão do cluster, os novos nodos são adicionados somente se a sua pontuação desvia da pontuação do nodo semente do cluster por menos que o ponto de corte definido. Esta é uma porcentagem, onde um valor de 0.2 para os novos nodos permite que não seja 20% menor do que o valor do nodo semente. Assim, valores menores criam clusters menores e vice versa;
- K-Core: esse parâmetro filtra clusters que não contêm um subconjunto maximamente intercalado de pelo menos k graus. O aumento desse parâmetro faz com que clusters menores sejam excluídos;
- Max. Depth from Seed: limita a distância a partir do nodo semente dentro do qual o MCODE pode procurar por membros do cluster.

Para a rede DAA-Maq utilizou-se Degree Cutoff: 20, Node Score Cutoff: 0.2, Haircut: true, Fluff: false, K-Core: 2, Max. Depth from Seed: 100.

Através destes parâmetros quatro clusters foram encontrados, com scores de 17,92, 14,688, 4,929 e 3,385. Seleccionamos os clusters com score acima de 10 para obtermos melhores resultados. As figuras 23 e 24 contém os clusters seleccionados para análise de função biológica.

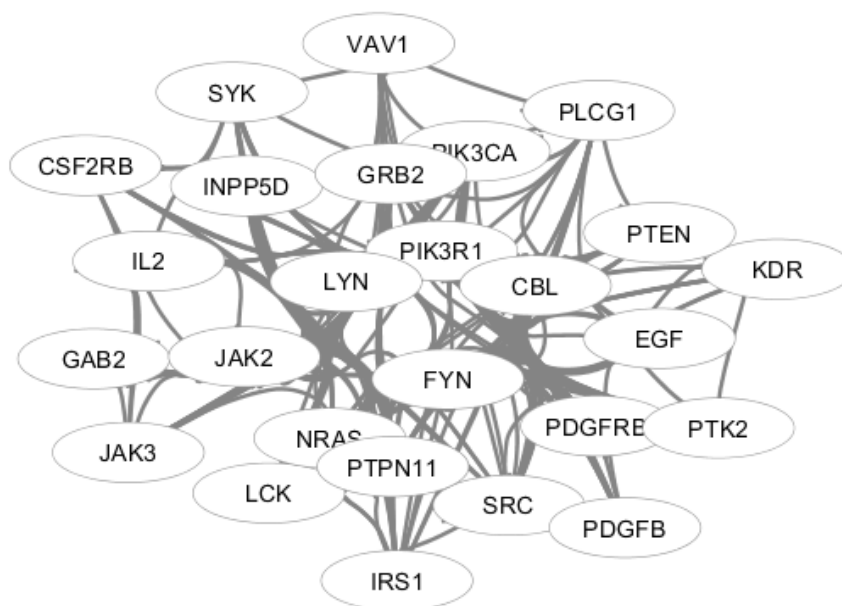


Figura 23 - Cluster 1 para a rede DAA-Maq, com 26 nodos e score de 17,92.

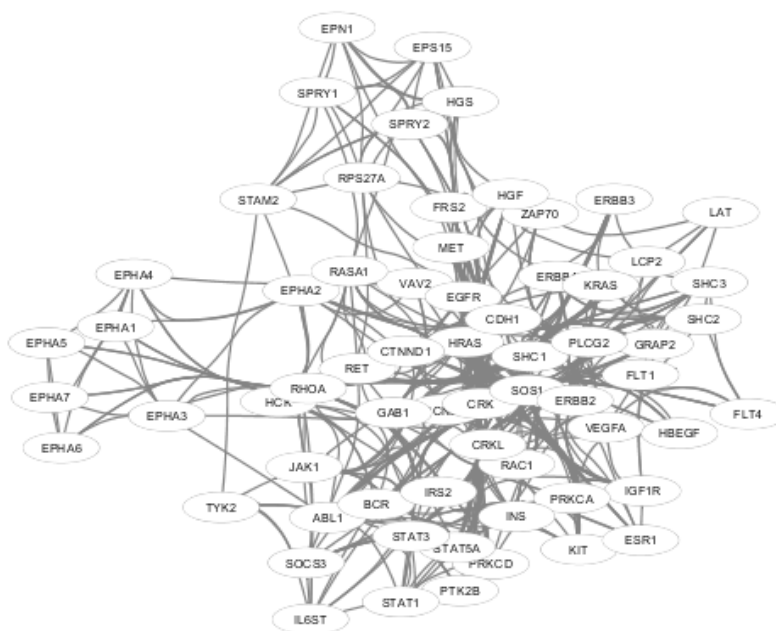


Figura 24 - Cluster 2 para a rede DAA-Maq, com 65 nodos e score de 14,688.

Em seguida a ferramenta DAVID tools foi utilizada. Para a busca em DAVID os seguintes bancos de dados foram selecionados: Gene Ontology (GOTERM_BP_FAT, GOTERM_CC_FAT e GOTERM_MF_FAT) e Pathway (BBID, BIOCARTA, KEGG e REACTOME).

Para princípio da análise das funções biológicas selecionamos a rede inicial, que foi obtida com as nove proteínas diferencialmente fosforiladas encontradas com OsteoBLAST. Foram encontrados 101 DAVID IDs para a rede inicial. A opção a *Funcional Annotation Chart* retornou funções biológicas relacionadas ao mecanismo de fosforilação, característico de proteínas tirosina quinases, além de funções na locomoção, migração e proliferação celular. DAVID tools fornece dois valores, p-value e Benjamini, estes são utilizados para que a resposta seja o mais confiável possível. Quanto menor estes parâmetros, mais confiável é a função biológica predita por DAVID. Neste trabalho utilizaremos valores de Benjamini menores quem 10^{-7} como valores confiáveis.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	GOTERM_MF_FAT	protein tyrosine kinase activity	RT		5	5,0	8,8E-7	4,8E-5
<input type="checkbox"/>	GOTERM_MF_FAT	transmembrane receptor protein tyrosine kinase activity	RT		4	4,0	4,5E-6	1,2E-4
<input type="checkbox"/>	GOTERM_MF_FAT	protein kinase activity	RT		5	5,0	1,5E-4	2,7E-3
<input type="checkbox"/>	GOTERM_BP_FAT	transmembrane receptor protein tyrosine kinase signaling pathway	RT		4	4,0	1,5E-4	4,0E-2
<input type="checkbox"/>	GOTERM_BP_FAT	protein amino acid phosphorylation	RT		5	5,0	1,8E-4	2,5E-2
<input type="checkbox"/>	GOTERM_BP_FAT	phosphorylation	RT		5	5,0	3,7E-4	3,3E-2
<input type="checkbox"/>	GOTERM_BP_FAT	enzyme linked receptor protein signaling pathway	RT		4	4,0	5,2E-4	3,5E-2
<input type="checkbox"/>	GOTERM_BP_FAT	phosphorus metabolic process	RT		5	5,0	7,8E-4	4,2E-2
<input type="checkbox"/>	GOTERM_BP_FAT	phosphate metabolic process	RT		5	5,0	7,8E-4	4,2E-2
<input type="checkbox"/>	GOTERM_BP_FAT	positive regulation of cell migration	RT		3	3,0	8,8E-4	4,0E-2
<input type="checkbox"/>	GOTERM_MF_FAT	ribonucleotide binding	RT		6	5,9	9,2E-4	1,3E-2
<input type="checkbox"/>	GOTERM_MF_FAT	purine ribonucleotide binding	RT		6	5,9	9,2E-4	1,3E-2
<input type="checkbox"/>	GOTERM_BP_FAT	positive regulation of cell motion	RT		3	3,0	1,1E-3	4,1E-2
<input type="checkbox"/>	GOTERM_BP_FAT	positive regulation of locomotion	RT		3	3,0	1,1E-3	4,1E-2
<input type="checkbox"/>	GOTERM_MF_FAT	purine nucleotide binding	RT		6	5,9	1,1E-3	1,2E-2
<input type="checkbox"/>	GOTERM_MF_FAT	nucleotide binding	RT		6	5,9	2,4E-3	2,2E-2
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of cell migration	RT		3	3,0	3,1E-3	1,0E-1
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of locomotion	RT		3	3,0	4,0E-3	1,2E-1
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of cell motion	RT		3	3,0	4,1E-3	1,1E-1
<input type="checkbox"/>	GOTERM_MF_FAT	vascular endothelial growth factor receptor activity	RT		2	2,0	4,3E-3	3,3E-2
<input type="checkbox"/>	GOTERM_MF_FAT	ATP binding	RT		5	5,0	4,4E-3	3,0E-2
<input type="checkbox"/>	GOTERM_MF_FAT	adenyl ribonucleotide binding	RT		5	5,0	4,6E-3	2,8E-2
<input type="checkbox"/>	GOTERM_MF_FAT	adenyl nucleotide binding	RT		5	5,0	5,6E-3	3,0E-2
<input type="checkbox"/>	GOTERM_BP_FAT	vascular endothelial growth factor receptor signaling pathway	RT		2	2,0	5,7E-3	1,3E-1

Figura 25 - Layout de DAVID para a rede inicial da rede DAA-Maq .

A rede inicial não apresentou muitas funções com estes parâmetros com valores baixos. Uma possível explicação para isso seja o baixo índice de clusterização da rede inicial. Ou seja, por mais que as proteínas obtidas com OsteoBLAST se organizem numa rede, esta não apresenta comportamento de grupo.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	enzyme linked receptor protein signaling pathway	RT		15	4,8	1,1E-16	1,0E-13
<input type="checkbox"/>	GOTERM_BP_FAT	transmembrane receptor protein tyrosine kinase signaling pathway	RT		13	4,2	1,3E-15	6,3E-13
<input type="checkbox"/>	GOTERM_BP_FAT	cell surface receptor linked signal transduction	RT		21	6,8	1,4E-13	4,3E-11
<input type="checkbox"/>	GOTERM_MF_FAT	protein tyrosine kinase activity	RT		11	3,5	1,5E-13	2,3E-11
<input type="checkbox"/>	GOTERM_MF_FAT	non-membrane spanning protein tyrosine kinase activity	RT		8	2,6	6,1E-13	4,6E-11
<input type="checkbox"/>	REACTOME_PATHWAY	REACT_604:Hemostasis	RT		15	4,8	1,1E-12	1,4E-11
<input type="checkbox"/>	KEGG_PATHWAY	Fc epsilon RI signaling pathway	RT		11	3,5	1,1E-12	6,8E-11
<input type="checkbox"/>	GOTERM_BP_FAT	peptidyl-tyrosine phosphorylation	RT		8	2,6	1,5E-12	3,5E-10
<input type="checkbox"/>	GOTERM_BP_FAT	peptidyl-tyrosine modification	RT		8	2,6	2,1E-12	3,9E-10
<input type="checkbox"/>	GOTERM_BP_FAT	phosphate metabolic process	RT		16	5,1	1,1E-11	1,7E-9
<input type="checkbox"/>	GOTERM_BP_FAT	phosphorus metabolic process	RT		16	5,1	1,1E-11	1,7E-9
<input type="checkbox"/>	REACTOME_PATHWAY	REACT_9417:Signaling by EGFR	RT		9	2,9	8,0E-11	5,2E-10
<input type="checkbox"/>	GOTERM_MF_FAT	protein domain specific binding	RT		11	3,5	1,4E-10	7,3E-9
<input type="checkbox"/>	GOTERM_BP_FAT	phosphorylation	RT		14	4,5	2,6E-10	3,5E-8
<input type="checkbox"/>	KEGG_PATHWAY	Glioma	RT		9	2,9	3,2E-10	1,0E-8
<input type="checkbox"/>	GOTERM_BP_FAT	intracellular signaling cascade	RT		16	5,1	4,0E-10	4,8E-8
<input type="checkbox"/>	GOTERM_BP_FAT	protein kinase cascade	RT		11	3,5	4,7E-10	4,9E-8
<input type="checkbox"/>	GOTERM_BP_FAT	protein amino acid phosphorylation	RT		13	4,2	5,4E-10	5,1E-8
<input type="checkbox"/>	KEGG_PATHWAY	Focal adhesion	RT		12	3,9	7,7E-10	1,6E-8
<input type="checkbox"/>	KEGG_PATHWAY	T cell receptor signaling pathway	RT		10	3,2	9,7E-10	1,5E-8
<input type="checkbox"/>	REACTOME_PATHWAY	REACT_16888:Signaling by PDGF	RT		9	2,9	1,1E-9	4,6E-9
<input type="checkbox"/>	GOTERM_BP_FAT	positive regulation of signal transduction	RT		10	3,2	1,5E-9	1,3E-7
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of peptidyl-tyrosine phosphorylation	RT		7	2,3	2,1E-9	1,7E-7
<input type="checkbox"/>	GOTERM_MF_FAT	insulin receptor substrate binding	RT		5	1,6	2,9E-9	1,1E-7
<input type="checkbox"/>	GOTERM_BP_FAT	positive regulation of cell communication	RT		10	3,2	3,9E-9	2,8E-7
<input type="checkbox"/>	KEGG_PATHWAY	ErbB signaling pathway	RT		9	2,9	4,5E-9	5,7E-8

Figura 26 - Layout de DAVID tools para funções presentes no cluster 1.

Para o cluster 1 foram encontrados 311 DAVID IDs. A opção a Funcional Annotation Chart encontrou um papel da rede DAA-Maq no mecanismo de transdução de sinais intracelulares destacando-se funções como fosforilação, receptor de superfícies celulares, receptores transmembranas, adesão celulares entre outras. O cluster 1 apresenta diversas DAVIDs IDs com valores de p-value e Benjamini ideais.

Para o cluster 2 foram encontrados 766 DAVID IDs. A opção a Funcional Annotation Chart revelou processos referente as classes das quinases como fosforilação, desfosforilação, entre outros, e também processos referentes ao processos de contato célula-biomaterial, como receptor transmembrana, regulação da proliferação, resposta a estímulos endógenos entre outros.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	transmembrane receptor protein tyrosine kinase signaling pathway	RT	31	4,0	1,4E-37	1,9E-34	
<input type="checkbox"/>	GOTERM_BP_FAT	enzyme linked receptor protein signaling pathway	RT	33	4,3	4,0E-35	2,7E-32	
<input type="checkbox"/>	GOTERM_MF_FAT	protein tyrosine kinase activity	RT	24	3,1	1,3E-29	2,3E-27	
<input type="checkbox"/>	GOTERM_BP_FAT	intracellular signaling cascade	RT	42	5,5	6,6E-28	3,0E-25	
<input type="checkbox"/>	GOTERM_BP_FAT	protein amino acid phosphorylation	RT	33	4,3	7,7E-26	2,6E-23	
<input type="checkbox"/>	GOTERM_MF_FAT	transmembrane receptor protein tyrosine kinase activity	RT	17	2,2	1,3E-24	1,2E-22	
<input type="checkbox"/>	GOTERM_BP_FAT	phosphorylation	RT	33	4,3	2,1E-23	5,8E-21	
<input type="checkbox"/>	GOTERM_MF_FAT	protein kinase activity	RT	28	3,7	1,3E-21	7,9E-20	
<input type="checkbox"/>	GOTERM_BP_FAT	phosphorus metabolic process	RT	33	4,3	8,4E-21	1,9E-18	
<input type="checkbox"/>	GOTERM_BP_FAT	phosphate metabolic process	RT	33	4,3	8,4E-21	1,9E-18	
<input type="checkbox"/>	GOTERM_BP_FAT	protein kinase cascade	RT	24	3,1	1,2E-20	2,4E-18	
<input type="checkbox"/>	GOTERM_BP_FAT	cell surface receptor linked signal transduction	RT	40	5,2	3,4E-19	5,8E-17	
<input type="checkbox"/>	KEGG_PATHWAY	ErbB signaling pathway	RT	18	2,3	5,2E-17	3,5E-15	
<input type="checkbox"/>	GOTERM_BP_FAT	peptidyl-tyrosine phosphorylation	RT	12	1,6	8,6E-17	1,7E-14	
<input type="checkbox"/>	GOTERM_BP_FAT	peptidyl-tyrosine modification	RT	12	1,6	1,4E-16	1,5E-14	
<input type="checkbox"/>	GOTERM_BP_FAT	positive regulation of cell proliferation	RT	21	2,7	9,9E-16	1,2E-13	
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of phosphorylation	RT	21	2,7	8,9E-15	1,0E-12	
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of phosphate metabolic process	RT	21	2,7	1,9E-14	2,0E-12	
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of phosphorus metabolic process	RT	21	2,7	1,9E-14	2,0E-12	
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of kinase activity	RT	18	2,3	2,5E-13	2,4E-11	
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of programmed cell death	RT	24	3,1	3,7E-13	3,4E-11	
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of cell death	RT	24	3,1	4,0E-13	3,4E-11	
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of transferase activity	RT	18	2,3	4,9E-13	3,9E-11	
<input type="checkbox"/>	GOTERM_BP_FAT	positive regulation of locomotion	RT	12	1,6	6,1E-13	4,6E-11	
<input type="checkbox"/>	KEGG_PATHWAY	Focal adhesion	RT	20	2,6	6,5E-13	2,2E-11	

Figura 27 – Layout de DAVID tools para funções biológicas obtidas com *Funcional Annotation Chart* do DAVID tools para o cluster 2.

4.7.2 Análise de redes proteicas para os grupos Ti-Maq.

OsteoBLAST retornou duas proteínas com fosforilação diferencial, comparado ao Maq, NCF1(S320) e PDPK1(Y373/376). Uma rede inicial foi construída com STRING utilizando os mesmos parâmetros que os grupos anteriores.

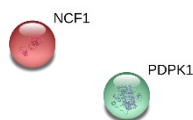


Figura 28 - As proteínas que apresentaram fosforilação diferencial para o grupo Ti-Maq não apresentam um comportamento de rede.

NCF1 e PDPK1 não possuem ligação descrita nos bancos de dados que STRING utiliza para predição (Figura 28). Uma rede expandida para estas proteínas foi então construídas com os parâmetros utilizados para a rede DAA-Maq (Figura 29).

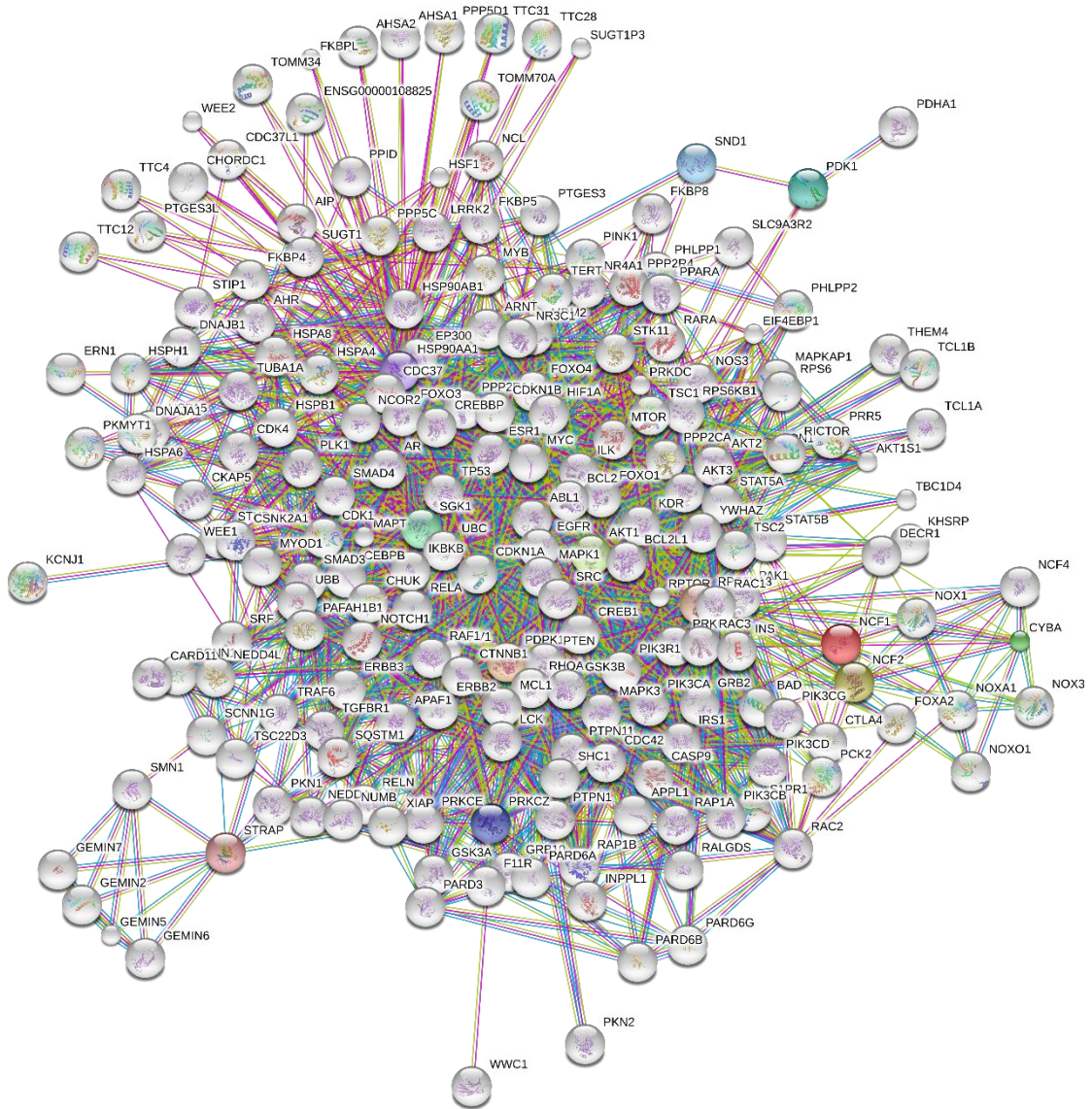


Figura 29 - Rede expandida para o grupo Ti-Maq. A rede foi construída a partir das proteínas NCF1 e PDPK1 que não possuíam ligações entre si.

A análise da rede expandida foi realizada novamente com *Cytoscape*, através da ferramenta *NetworkAnalyzer*. HSP90AA1, UBC e AKT1 foram proteínas que apresentaram centralidades importantes. Src apareceu novamente com valores importantes para Betweenness (Figuras 30 e 31).

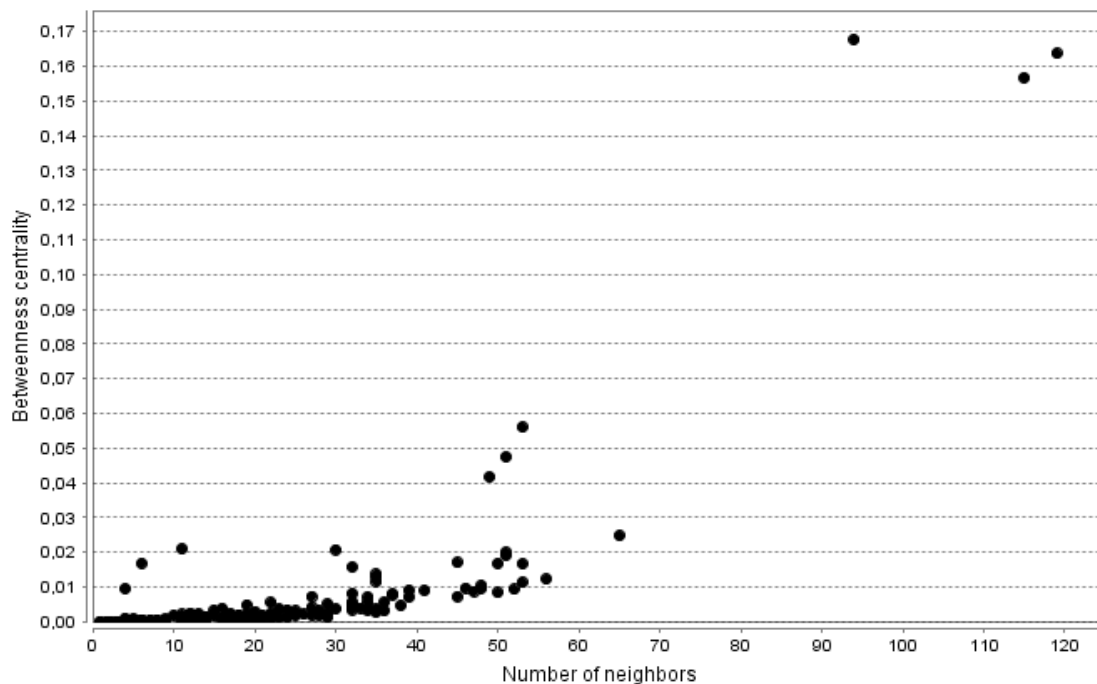


Figura 30 - Gráfico de Betweenness centrality para a rede de Ti-Maq. HSP90AA1(0.168), AKT1(0.164) e UBC(0.157) foram os nodos que apresentaram maiores valores para esta centralidade.

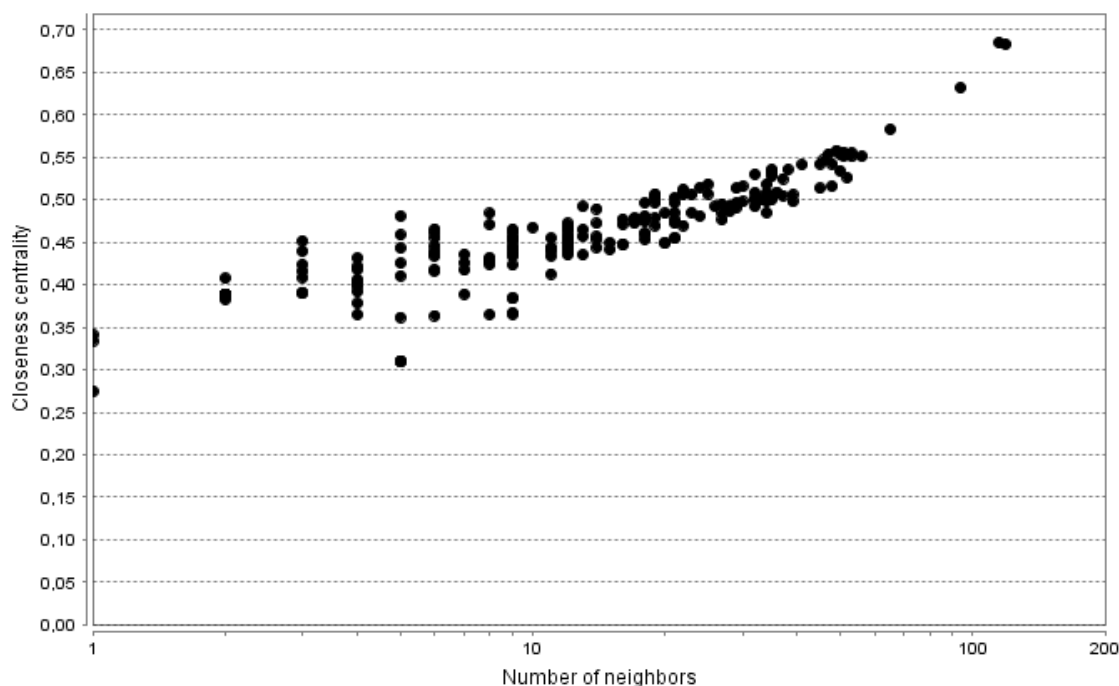


Figura 31 - Gráfico de Closeness centrality para a rede de Ti-Maq. UBC(0.685), AKT1(0.683), HSP90AA1(0.632) e Src(0.583) foram os nodos que apresentaram maiores valores para esta centralidade.

MCODE retornou seis clusters para a rede expandida de Ti-Maq. Os scores dos clusters foram de 12,333, 8,286, 6,717, 3,765, 3 e 3. Seguindo o critério utilizado nas

análise anteriores, selecionamos o cluster com score acima de 10 (Figura 32) para prosseguir as análises sobre as funções biológicas presentes na rede.

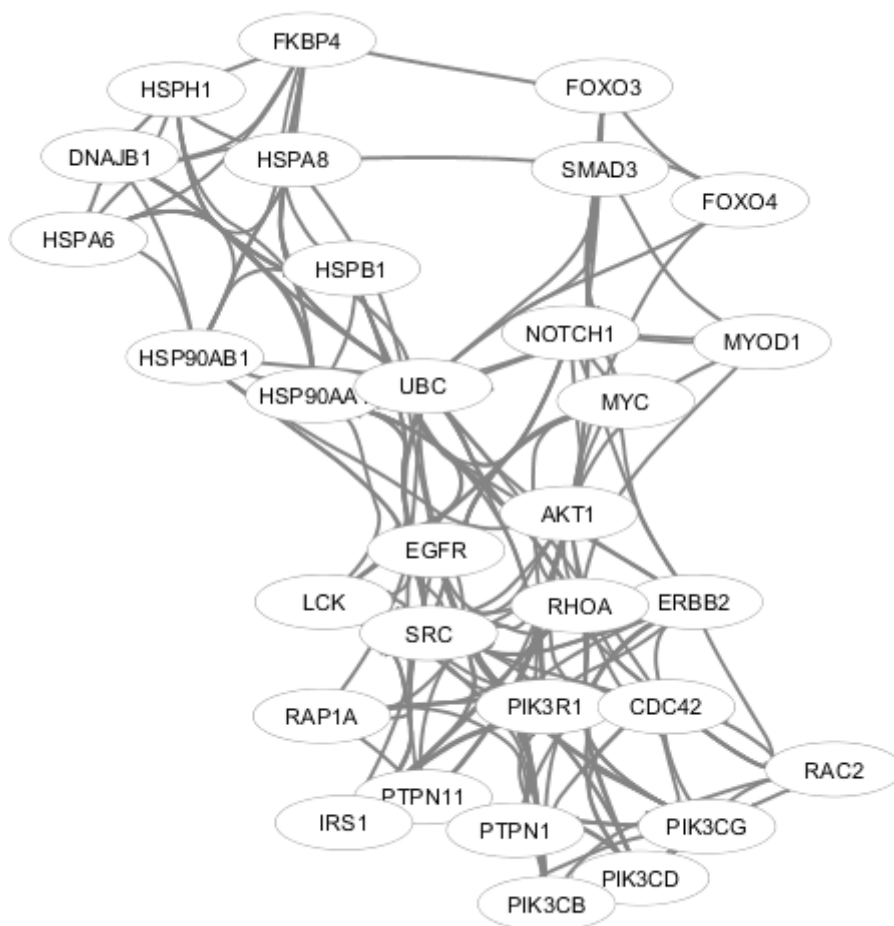


Figura 32 - Cluster para a rede de Ti-Maq com score de 12,333 e 31 nodos.

Devido as proteínas encontradas no processo de verificação de OsteoBLAST não apresentarem comportamento de rede não foi realizado a verificação de função biológica utilizando DAVID tools. O cluster encontrado para a rede expandida de Ti-Maq retornou 378 DAVID IDs. *Functional Annotation Chart* revelou funções biológicas relacionadas ao processo de adesão e migração celular. Além disso, vias de sinalização com VEGF e ErbB foram relatadas como presentes no cluster. Uma série de vias relatadas em diversos tipos de canceres foram relatadas, uma explicação para isso é o fato de células cancerígenas serem altamente proliferativas, possuírem uma alta motilidade. Tais características são desejáveis quando ocorre a interação de células com um biomaterial, podendo levar ao sucesso do implante.

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	response to organic substance	RT	18	4,8	1,2E-14	1,2E-11	
<input type="checkbox"/>	KEGG_PATHWAY	Pancreatic cancer	RT	10	2,6	2,3E-11	1,4E-9	
<input type="checkbox"/>	KEGG_PATHWAY	Endometrial cancer	RT	9	2,4	6,5E-11	2,0E-9	
<input type="checkbox"/>	KEGG_PATHWAY	Neurotrophin signaling pathway	RT	11	2,9	1,2E-10	2,6E-9	
<input type="checkbox"/>	KEGG_PATHWAY	Pathways in cancer	RT	14	3,7	6,0E-10	1,1E-8	
<input type="checkbox"/>	KEGG_PATHWAY	Focal adhesion	RT	12	3,2	7,7E-10	9,5E-9	
<input type="checkbox"/>	KEGG_PATHWAY	VEGF signaling pathway	RT	9	2,4	1,4E-9	1,4E-8	
<input type="checkbox"/>	GOTERM_CC_FAT	cytosol	RT	17	4,5	1,6E-9	1,8E-7	
<input type="checkbox"/>	KEGG_PATHWAY	Colorectal cancer	RT	9	2,4	3,4E-9	3,0E-8	
<input type="checkbox"/>	KEGG_PATHWAY	Non-small cell lung cancer	RT	8	2,1	4,2E-9	3,3E-8	
<input type="checkbox"/>	KEGG_PATHWAY	ErbB signaling pathway	RT	9	2,4	4,5E-9	3,1E-8	
<input type="checkbox"/>	KEGG_PATHWAY	Prostate cancer	RT	9	2,4	5,4E-9	3,4E-8	
<input type="checkbox"/>	GOTERM_BP_FAT	response to unfolded protein	RT	7	1,9	9,1E-9	4,7E-6	
<input type="checkbox"/>	GOTERM_BP_FAT	transmembrane receptor protein tyrosine kinase signaling pathway	RT	9	2,4	2,1E-8	7,4E-6	
<input type="checkbox"/>	KEGG_PATHWAY	Renal cell carcinoma	RT	8	2,1	2,7E-8	1,5E-7	
<input type="checkbox"/>	GOTERM_BP_FAT	enzyme linked receptor protein signaling pathway	RT	10	2,6	3,4E-8	8,8E-6	
<input type="checkbox"/>	KEGG_PATHWAY	Chronic myeloid leukemia	RT	8	2,1	4,4E-8	2,3E-7	
<input type="checkbox"/>	GOTERM_BP_FAT	negative regulation of apoptosis	RT	10	2,6	4,6E-8	9,5E-6	
<input type="checkbox"/>	KEGG_PATHWAY	Leukocyte transendothelial migration	RT	9	2,4	5,2E-8	2,5E-7	

Figura 33 - Layout de Functional Annotation Chart para o cluster de Ti-Maq.

4.7.3 A via de sinalização ErbB

As redes presentes apresentaram diversas funções biológicas e vias de sinalização envolvidas. Dentre estas, a via de sinalização de ErbB apresentou valores confiáveis para ambas as redes. Composta de quatro receptores de tirosina quinase localizados na membrana celular, ao receberem estímulos extracelulares estes vão disparar sinais responsáveis por regular diversas respostas biológicas, incluindo a proliferação, a diferenciação, a motilidade celular e sobrevivência (Figura 34).

Os quatros membros desta família são ErbB-1, ErbB-2, ErbB-3 e ErbB-4. ErbB-1 é conhecido como receptor do fator de crescimento epidermal, EGFR. Composta por 1210 resíduos de aminoácidos, EGFR possui 22 resíduos passíveis de fosforilação. OsteoBLAST detectou o resíduo Y1172. Localizado numa região da proteína chamada de cauda citosólica (HELIN; BEGUINOT, 1991), este resíduo é fosforilado pela própria EGFR, num processo conhecido como autofosforilação, e esta causa uma alteração conformacional responsável por desencadear diversas cascatas de sinalização. EGFR pode reagir a ligantes extracelulares como fator de crescimento epidermal (EGF) e fator de crescimento transformante alfa (TGF- α) principalmente. Ao se ligar a estes ocorrem processos de homo- ou hetero-dimerização de EGFR, neste caso com outros membros da família da ErbB, na superfície celular, estes são responsáveis

por uma mudança conformacional responsável pela autofosforilação dos resíduos presentes na cauda citosólica de EGFR, como o resíduo Y 1172 por exemplo (HERBST, 2004). A fosforilação dos resíduos presentes na cauda citosólica de EGFR esta presentes em uma série de eventos biológicos como processos cancerígenos (KIM et al., 2015), regulação positiva da cascata de MAPK (CROOK et al., 2016) esta relacionada à proliferação, sobrevivência e diferenciação celular.

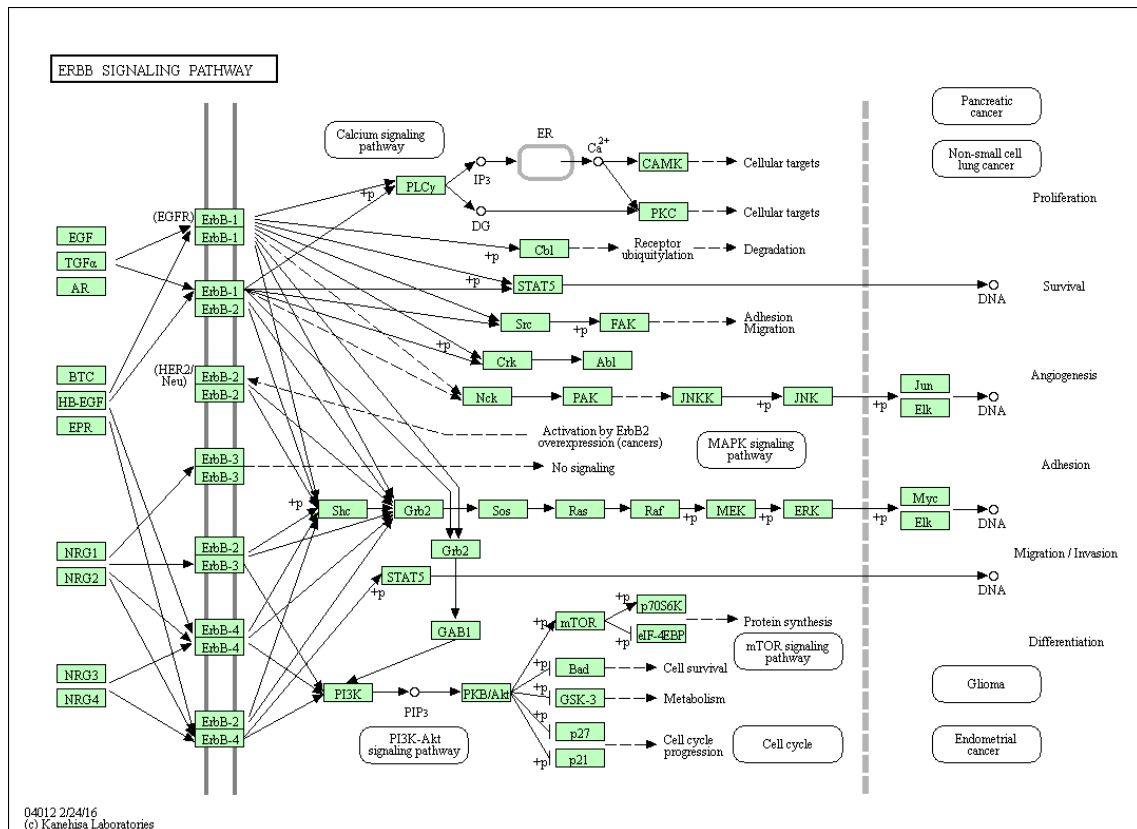


Figura 34 - Via de sinalização de ErbB. Fonte: KEGG pathway.

Similarmente ErbB-3 e ErbB-4 precisam se dimerizar para que haja fosforilação de sua cauda citosólica. ErbB-3 é estimulada por neuregulinas como NRG1 ou NRG2 e ErbB-4 pode ser estimulada por neuregulinas ou ligantes do tipo EGF (LI et al., 2007). A exceção na família das ErbB é a ErbB-2. Seu processo de dimerização não depende de nenhum ligante extracelular, para que ocorra sua dimerização é necessário que ela seja desligada de duas enzimas chaperonas, a HSP90 e CDC37 (BRAGIN et al., 2016; CITRI et al., 2004). Após estas chaperonas serem retiradas, ErbB-2 pode se dimerizar com outros membros da família ErbB, de modo homo ou heterodimérico, sinalizando cascatas como MAPK, AKT e CAMPK (Figura 35).

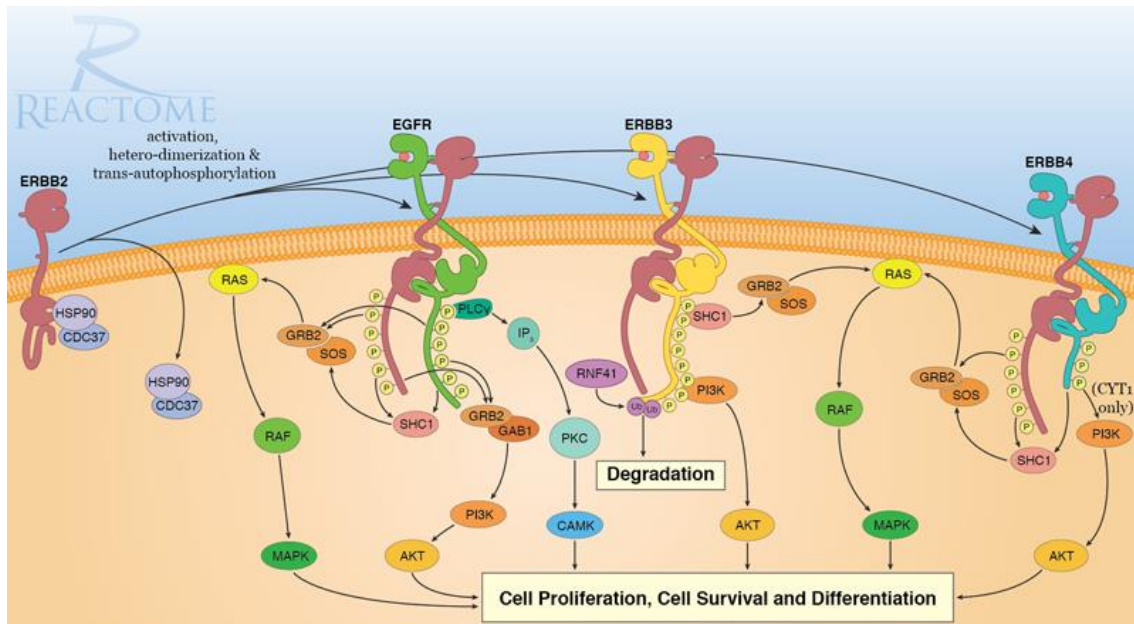


Figura 35 - Mecanismo de dimerização de membros da família ErbB é fundamental para a ativação da via. O processo de dimerização da ErbB-2 não depende de estímulos externos e depende do desligamento das chaperonas HSP90 e CDC37 de sua cauda citosólica. Ao ser desligada das chaperonas, ErbB-2 é capaz de se dimerizar com outros membros ErbB, e ao se dimerização fosforilar resíduos presentes na cauda citosólica que darão início ao mecanismo de sinalização.

De modo geral, a dimerização dos membros da família ErbB é necessária para a ativação da via. Esta dimerização provoca fosforilação de resíduos presentes na cauda citosólica que recrutará proteínas como GRB2 e SOS1, responsáveis pelo ajuste fino da regulação de vias de sinalização, especialmente MAPK (OOFT et al., 2015). GRB2 e HSP90 foram proteínas que apresentaram resultados de centralidades importantes, ao lado de Src e EGFR, em nosso estudo de redes. Este resultado ilustra a importância de uma abordagem sistêmica, uma vez que nossos resultados iniciais (PamChip® tirosina quinase) elencavam apenas EGFR como uma proteína importante na interação célula-biomaterial. Mas através desta abordagem de redes foi possível conhecer vias e outras proteínas que podem ser importantes neste evento fundamental para o sucesso de um biomaterial.

4.7.4 O ponto de adesão focal

Ponto de adesão focal é a região da célula responsável pelas junções célula-célula ou célula-matriz extracelular. Também conhecido por complexo de adesão focal, esta região é composta por proteínas fundamentais no processo de adesão celular, como as integrinas por exemplo. Integrinas são proteínas transmembranas que ligam o meio intra ao meio extracelular. A adesão por integrinas atua de maneira crucial em diversos

processos celulares (SUN et al., 2016). Filamentos de actina do citoesqueleto se ancoram às integrinas promovendo estabilidade mecânica a célula, além de participar de processos como motilidade e adesão celular (Figura 36).

As integrinas são responsáveis por captarem estímulos extracelulares como citocinas, hormônios ou mesmo adesão a superfícies (ZAMBUZZI et al., 2011a). Após a ativação das integrinas, FAK se autofosforila no resíduo Y397, esta fosforilação promove a interação com a proteína Src, que irá fosforilar FAK no resíduo Y576/577. Quando fosforilado no resíduo Y576/577 FAK assume uma conformação espacial que lhe permite estar ativada. Src possui dois sítios de fosforilação que merecem maior atenção. Quando fosforilada no resíduo Y416, Src é inativada, e quando a fosforilação ocorre no resíduo Y527 Src assume sua conformação ativa (ZAMBUZZI et al., 2011a; ZAMBUZZI; MILANI; TETI, 2010). O balanço da fosforilação destas duas proteínas é controlado por proteínas fosfatases como a PTP1B e LMW-PTP, que atuam desfosforilando-as (ARREGUI et al., 2013; FERNANDES et al., 2014). A mobilização de FAK e Src leva a uma cascata de sinalização que resultará num rearranjo do citoesqueleto. Esse processo de rearranjo do citoesqueleto, ou “spreading”, vem sendo utilizado como um método de análise da interação da célula a superfície do biomaterial, uma vez que em materiais que promovam uma melhor adesão em superfície este processo de “spreading” ocorra de modo mais acelerado.

Src foi uma proteína com centralidades importantes nas redes construídas com STRING. Src é uma proteína importantíssima em diversas vias de transdução de sinais participando de processos celulares como adesão, proliferação, diferenciação e sobrevivência, regulando positivamente a via MAPK. Src também participa de forma decisiva em processos cancerígenos (NTANASIS-STATHOPOULOS et al., 2016).

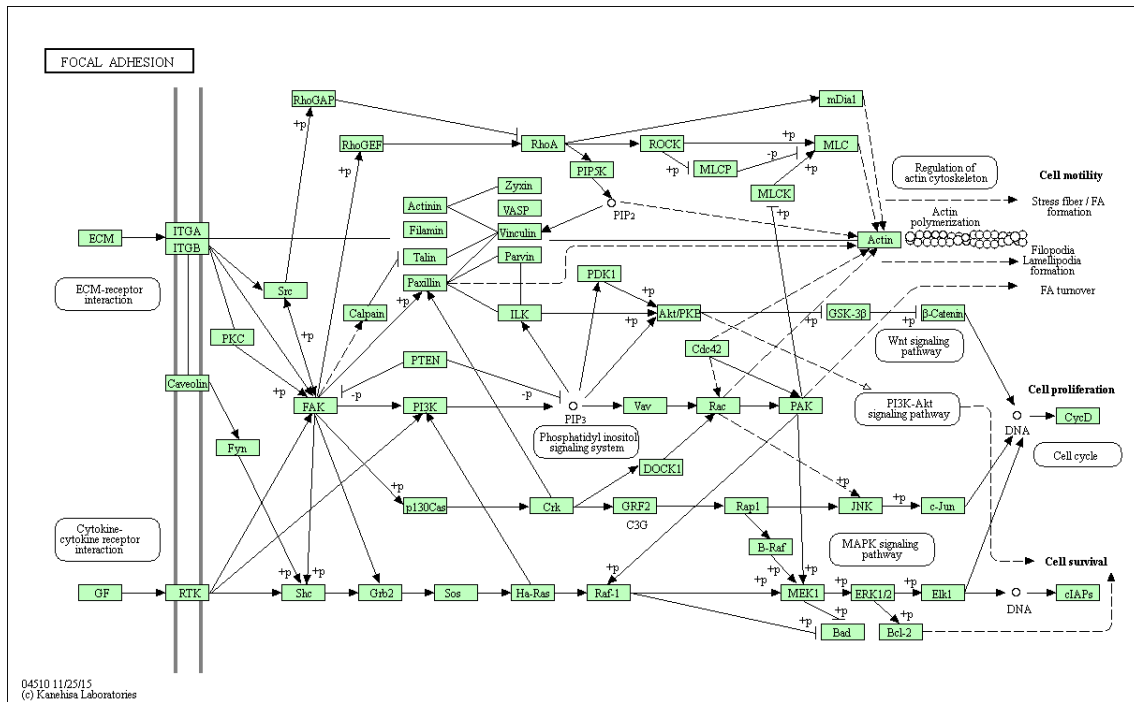


Figura 36 - Complexo de adesão focal. Fonte: KEGG pathway

4.7.5 Proteínas que não participam de redes

As proteínas KRTB6 e ENO2 foram detectadas como possuindo fosforilação diferencial por OsteoBLAST. Porém ao construirmos redes de interações com STRING estas não apresentaram comportamento de rede com as outras proteínas detectadas com OsteoBLAST. KRTB6, queratina tipo II citoesqueletal 6B, é uma proteína relacionada com a via de sinalização de Notch (ZHANG et al., 2015). ENO2, gama enolase, é uma proteína relacionada a gliconeogenese (Fonte Reactome). Existem poucos registros sobre as funções de ambos na literatura e em bancos de dados, sendo uma oportunidade para nosso grupo buscar o entendimento de suas funções no processo de interação célula-biomaterial.

5. Constatações

Os resultados obtidos durante a execução deste projeto nos permitiram constatar que:

- 1) PamChip® tirosina quinase surge como uma ferramenta poderosa na busca pelo quinoma;
- 2) O processo de análise dos resultados obtidos com PamChip® é passível de automatização;
- 3) A rotina computacional de OsteoBLAST é capaz de detectar a fosforilação diferencial;
- 4) A via de sinalização de ErbB é uma via importante no processo de adesão de células sobre a superfície de biomateriais;
- 5) A biologia sistêmica é uma ferramenta para inferências biológicas dos resultados de OsteoBLAST;
- 6) Embora os bancos de dados venham sendo atualizados constantemente, ainda faltam informações sobre alguns eventos biológicos;

6. Conclusão

Desta forma é possível concluir que, OsteoBLAST, aliado a uma abordagem de biologia sistêmica, pode auxiliar no desenvolvimento de biomateriais médico-odontológicos minimizando assim a experimentação animal na área.

7. Perspectivas futuras

Munidos dos resultados obtidos neste trabalho, pretendemos:

- ✓ Elaborar uma rotina totalmente automatizada de análise para OsteoBLAST;
- ✓ Criar um pacote na linguagem R para análise do quíoma diferencial;

8. Bibliografia

ARREGUI, C. O. et al. Protein tyrosine phosphatase PTP1B in cell adhesion and migration. **Cell adhesion & migration**, v. 7, n. 5, p. 418–423, 2013.

ARVIDSSON, A. K. et al. Tyr-716 in the platelet-derived growth factor beta-receptor kinase insert is involved in GRB2 binding and Ras activation. **Molecular and cellular biology**, v. 14, n. 10, p. 6715–6726, out. 1994.

ASHBURNER, M. et al. Gene Ontology: tool for the unification of biology. **Nat Genet**, v. 25, n. 1, p. 25–29, maio 2000.

BATEMAN, A. et al. UniProt: A hub for protein information. **Nucleic Acids Research**, v. 43, n. D1, p. D204–D212, 2015.

BAXTER, R. M. et al. Full activation of the platelet-derived growth factor beta-receptor kinase involves multiple events. **The Journal of biological chemistry**, v. 273, n. 27, p. 17050–17055, jul. 1998.

BLAKE, J. A. et al. Gene ontology consortium: Going forward. **Nucleic Acids Research**, v. 43, n. D1, p. D1049–D1056, 2015.

BRAGIN, P. E. et al. HER2 Transmembrane Domain Dimerization Coupled with Self-Association of Membrane-Embedded Cytoplasmic Juxtamembrane Regions. **Journal of Molecular Biology**, v. 428, n. 1, p. 52–61, 2016.

CARTER-SU, C.; SCHWARTZ, J.; ARGETSINGER, L. S. Growth hormone signaling pathways. **Growth hormone & IGF research: official journal of the Growth Hormone Research Society and the International IGF Research Society**, v. 28, p. 11–15, jun. 2016.

CAZARIN, K. C. C.; CORRÊA, C. L.; ZAMBRONE, F. A. D. Redução, refinamento e substituição do uso de animais em estudos toxicológicos: uma abordagem atual. **Revista Brasileira de Ciências Farmacêuticas**, v. 40, n. 3, p. 289–299, 2004.

CITRI, A. et al. Hsp90 restrains ErbB-2/HER2 signalling by limiting heterodimer formation. **EMBO reports**, v. 5, n. 12, p. 1165–1170, 2004.

CROOK, M. et al. EGFR Cell Survival Signaling Requires Phosphatidylcholine Biosynthesis. **G3 (Bethesda, Md.)**, set. 2016.

FERNANDES, G. V. O. et al. Osteoblast adhesion dynamics: a possible role for ROS and LMW-PTP. **Journal of cellular biochemistry**, v. 115, n. 6, p. 1063–9, 2014.

GEMINI-PIPERNI, S. et al. Cellular behavior as a dynamic field for exploring bone bioengineering: a closer look at cell-biomaterial interface. **Archives of biochemistry and biophysics**, v. 561, p. 88–98, nov. 2014.

HELIN, K.; BEGUINOT, L. Internalization and down-regulation of the human epidermal growth factor receptor are regulated by the carboxyl-terminal tyrosines. **The Journal of biological chemistry**, v. 266, n. 13, p. 8363–8368, maio 1991.

HENKEL, J. et al. Bone Regeneration Based on Tissue Engineering Conceptions – A 21st Century Perspective. **Bone Research**, v. 1, n. 3, p. 216–248, 2013.

HERBST, R. S. Review of epidermal growth factor receptor biology. **International Journal of Radiation Oncology Biology Physics**, v. 59, n. 2 SUPPL., p. 21–26, 2004.

HUANG, D. W.; SHERMAN, B. T.; LEMPICKI, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. **Nature protocols**, v. 4, n. 1, p. 44–57, 2009.

KARR, J. R. et al. A Whole-Cell Computational Model Predicts Phenotype from Genotype. **Cell**, v. 150, n. 2, p. 389–401, 2012.

KIM, Y. et al. Differential Effects of Tyrosine Kinase Inhibitors on Normal and Oncogenic EGFR Signaling and Downstream Effectors. **Molecular cancer research : MCR**, v. 13, n. 4, p. 765–774, abr. 2015.

KUHN, M. et al. STITCH 4: Integration of protein-chemical interactions with user data. **Nucleic Acids Research**, v. 42, n. D1, p. 401–407, 2014.

LI, Z. et al. Neuregulin-1 only induces trans-phosphorylation between ErbB receptor heterodimer partners. **Cellular Signalling**, v. 19, n. 3, p. 466–471, 2007.

LONGO, G. et al. Improving Osteoblast Response In Vitro by a Nanostructured Thin Film with Titanium Carbide and Titanium Oxides Clustered around Graphitic Carbon. **PloS one**, v. 11, n. 3, p. e0152566, 2016.

MANISCALCO, P.; CAFORIO, M. The importance of early rehabilitation in proximal humeral fracture: A clinical trial of efficacy and safety of a new endomedullary nail.

Journal of back and musculoskeletal rehabilitation, jul. 2016.

MANNING, G. et al. Evolution of protein kinase signaling from yeast to man. **Trends in biochemical sciences**, v. 27, n. 10, p. 514–520, out. 2002.

MILANI, R. **Desenvolvimento de um pipeline para análise em larga escala de um chip de proteínas quinases**. [s.l.] Universidade estadual de Campinas, 2010.

NANDAGOPAL, N.; ELOWITZ, M. B. Synthetic biology: integrated gene circuits. **Science (New York, N.Y.)**, v. 333, n. 6047, p. 1244–1248, set. 2011.

NTANASIS-STATHOPOULOS, I. et al. The Emerging Role of Tyrosine Kinase Inhibitors in Ovarian Cancer Treatment: A Systematic Review. **Cancer investigation**, v. 34, n. 7, p. 313–339, ago. 2016.

OKEGAWA, T. et al. The role of cell adhesion molecule in cancer progression and its application in cancer therapy. **Acta biochimica Polonica**, v. 51, n. 2, p. 445–457, 2004.

OOFT, M. L. et al. Prognostic significance of the EGFR pathway in nasopharyngeal carcinoma: a systematic review and meta-analysis. **Biomarkers in medicine**, v. 9, n. 10, p. 997–1010, 2015.

PAUKKU, K. et al. Platelet-derived growth factor (PDGF)-induced activation of signal transducer and activator of transcription (Stat) 5 is mediated by PDGF beta-receptor and is not dependent on c-src, fyn, jak1 or jak2 kinases. **The Biochemical journal**, v. 345 Pt 3, p. 759–766, fev. 2000.

ROSENBERGER, A. F. N. et al. Protein Kinase Activity Decreases with Higher Braak Stages of Alzheimer's Disease Pathology. **Journal of Alzheimer's disease : JAD**, out. 2015.

RUSSELL, W. M. S.; BURCH, R. L. **The Principles of Humane Experimental Technique**. [s.l.: s.n.].

SHANNON, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome research**, v. 13, n. 11, p. 2498–2504, nov. 2003.

SOHRABI, K. et al. How successful are small-diameter implants? A literature review. 2012.

SUN, Z. et al. Kank2 activates talin, reduces force transduction across integrins and induces central adhesion formation. **Nat Cell Biol**, v. 18, n. 9, p. 941–953, set. 2016.

SZKLARCZYK, D. et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. **Nucleic Acids Research** , v. 43, n. D1, p. D447–D452, 28 jan. 2015.

VARDAKAS, K. Z. et al. An analysis of factors contributing to PubMed’s growth. **Journal of Informetrics**, v. 9, n. 3, p. 592–617, 2015.

WARDEGA, P.; HELDIN, C. H.; LENNARTSSON, J. Mutation of tyrosine residue 857 in the PDGF α -receptor affects cell proliferation but not migration. **Cellular Signalling**, v. 22, n. 9, p. 1363–1368, 2010.

WEE, H. et al. Finite Element-Derived Surrogate Models of Locked Plate Fracture Fixation Biomechanics. **Annals of Biomedical Engineering**, 2016.

ZAMBUZZI, W. F. et al. On the road to understanding of the osteoblast adhesion: Cytoskeleton organization is rearranged by distinct signaling pathways. **Journal of Cellular Biochemistry**, v. 108, n. 1, p. 134–144, 2009.

ZAMBUZZI, W. F. et al. Intracellular Signal Transduction as a Factor in the Development of “ Smart ” Biomaterials for Bone Tissue Engineering. v. 108, n. 6, p. 1246–1250, 2011a.

ZAMBUZZI, W. F. et al. Intracellular signal transduction as a factor in the development of “smart” biomaterials for bone tissue engineering. **Biotechnology and bioengineering**, v. 108, n. 6, p. 1246–1250, jun. 2011b.

ZAMBUZZI, W. F. et al. Nanometer scale titanium surface texturing are detected by signaling pathways involving transient FAK and Src activations. **PloS one**, v. 9, n. 7, p. e95662, 2014.

ZAMBUZZI, W. F.; MILANI, R.; TETI, A. Expanding the role of Src and protein-tyrosine phosphatases balance in modulating osteoblast metabolism: Lessons from mice. **Biochimie**, v. 92, n. 4, p. 327–332, 2010.

ZHANG, H. et al. KRT6B, a key mediator of notch signaling in honokiol-induced human hepatoma cell apoptosis. **International journal of clinical and experimental medicine**, v. 8, n. 9, p. 16880–16889, 2015.

