

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-512824

(P2004-512824A)

(43) 公表日 平成16年4月30日(2004.4.30)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	4 B O 2 4
A 6 1 K 38/00	A 6 1 K 39/00	4 B O 6 3
A 6 1 K 39/00	A 6 1 K 39/395	4 B O 6 5
A 6 1 K 39/395	A 6 1 K 39/395	4 C O 8 4
A 6 1 K 45/00	A 6 1 K 45/00	4 C O 8 5
審査請求 未請求 予備審査請求 有 (全 395 頁) 最終頁に続く		

(21) 出願番号	特願2002-509377 (P2002-509377)	(71) 出願人	397069329
(86) (22) 出願日	平成13年7月10日 (2001.7.10)		コリクサ コーポレーション
(85) 翻訳文提出日	平成15年1月9日 (2003.1.9)		アメリカ合衆国 ワシントン 98104
(86) 国際出願番号	PCT/US2001/022058		, シアトル, コロンビア ストリート
(87) 国際公開番号	W02002/004514		1124, スイート 200
(87) 国際公開日	平成14年1月17日 (2002.1.17)	(74) 代理人	100078282
(31) 優先権主張番号	09/614, 124		弁理士 山本 秀策
(32) 優先日	平成12年7月11日 (2000.7.11)	(74) 代理人	100062409
(33) 優先権主張国	米国 (US)		弁理士 安村 高明
(31) 優先権主張番号	09/651, 563	(74) 代理人	100113413
(32) 優先日	平成12年8月29日 (2000.8.29)		弁理士 森下 夏樹
(33) 優先権主張国	米国 (US)	(72) 発明者	ワン, トントン
(31) 優先権主張番号	09/658, 824		アメリカ合衆国 ワシントン 98039
(32) 優先日	平成12年9月8日 (2000.9.8)		, メディナ, エヌ. イー. 28ティ
(33) 優先権主張国	米国 (US)		ーエイチ ストリート 8049
		最終頁に続く	

(54) 【発明の名称】 肺癌の治療および診断のための組成物および方法

(57) 【要約】

本発明によって、癌（特に、肺癌）の治療および診断のための組成物および方法が開示される。例示的な組成物は、1つ以上の肺癌ポリペプチド、その免疫原性部分、このようなポリペプチドをコードするポリヌクレオチド、このようなポリペプチドを発現する抗原提示細胞、およびこのようなポリペプチドを発現する細胞に特異的なT細胞を含む。この開示される組成物は、例えば、疾患（特に、肺癌）の診断、予防、および／または処置において有用である。

【特許請求の範囲】

【請求項1】

単離されたポリヌクレオチドであって、以下：

(a) 配列番号 878～1239、1240、1243、1247、1269、1272、1280、1283、1285、1286、1289、1300、1309、1318、1319、1327、1335、1339、1346、1359、1369、1370、1371、1393、1398、1405、1408、1413、1414、1417、1422、1429、1432、1435、1436、1438～1442、1447、1450、1453、1463、1467、1470、1473、1475、1482、1486、1491～1494、1501、1505、1506、1514～1517、1520、1522、1524、1535、1538、1542、1543、1547、1554、1557、1559、1561、1563～1664、1668、1669、1680～1706、1708～1732、1734、1736～1757、1759～1765、1767～1770、1772～1774、1776～1789、1824、1826～1829、1865～1868、1873、1875～1877、1879、1881、1883、1918、1922、1931、1938、1941 および 1974～2002 において提供される配列；

(b) 配列番号 878～1239、1240、1243、1247、1269、1272、1280、1283、1285、1286、1289、1300、1309、1318、1319、1327、1335、1339、1346、1359、1369、1370、1371、1393、1398、1405、1408、1413、1414、1417、1422、1429、1432、1435、1436、1438～1442、1447、1450、1453、1463、1467、1470、1473、1475、1482、1486、1491～1494、1501、1505、1506、1514～1517、1520、1522、1524、1535、1538、1542、1543、1547、1554、1557、1559、1561、1563～1664、1668、1669、1680～1706、1708～1732、1734、1736～1757、1759～1765、1767～1770、1772～1774、1776～1789、1824、1826～1829、1865～1868、1873、1875～1877、1879、1881、1883、1918、1922、1931、1938、1941 および 1974～2002 において提供される配列の相補体；

(c) 配列番号 878～1239、1240、1243、1247、1269、1272、1280、1283、1285、1286、1289、1300、1309、1318、1319、1327、1335、1339、1346、1359、1369、1370、1371、1393、1398、1405、1408、1413、1414、1417、1422、1429、1432、1435、1436、1438～1442、1447、1450、1453、1463、1467、1470、1473、1475、1482、1486、1491～1494、1501、1505、1506、1514～1517、1520、1522、1524、1535、1538、1542、1543、1547、1554、1557、1559、1561、1563～1664、1668、1669、1680～1706、1708～1732、1734、1736～1757、1759～1765、1767～1770、1772～1774、1776～1789、1824、1826～1829、1865～1868、1873、1875～1877、1879、1881、1883、1918、1922、1931、1938、1941 および 1974～2002 において提供される配列のうち、少なくとも 20 個連続した残基を含む配列；

(d) 配列番号 878～1239、1240、1243、1247、1269、1272、1280、1283、1285、1286、1289、1300、1309、1318、1319、1327、1335、1339、1346、1359、1369、1370、1371、1393、1398、1405、1408、1413、1414、1417

、1422、1429、1432、1435、1436、1438～1442、1447
 、1450、1453、1463、1467、1470、1473、1475、1482
 、1486、1491～1494、1501、1505、1506、1514～1517
 、1520、1522、1524、1535、1538、1542、1543、1547
 、1554、1557、1559、1561、1563～1664、1668、1669
 、1680～1706、1708～1732、1734、1736～1757、1759
 ～1765、1767～1770、1772～1774、1776～1789、1824
 、1826～1829、1865～1868、1873、1875～1877、1879
 、1881、1883、1918、1922、1931、1938、1941および19
 74～2002において提供される配列に、高度にストリンジェントな条件下でハイブリ
 10
 ダイズする配列；

(e) 配列番号878～1239、1240、1243、1247、1269、1272
 、1280、1283、1285、1286、1289、1300、1309、1318
 、1319、1327、1335、1339、1346、1359、1369、1370
 、1371、1393、1398、1405、1408、1413、1414、1417
 、1422、1429、1432、1435、1436、1438～1442、1447
 、1450、1453、1463、1467、1470、1473、1475、1482
 、1486、1491～1494、1501、1505、1506、1514～1517
 、1520、1522、1524、1535、1538、1542、1543、1547
 、1554、1557、1559、1561、1563～1664、1668、1669 20
 、1680～1706、1708～1732、1734、1736～1757、1759
 ～1765、1767～1770、1772～1774、1776～1789、1824
 、1826～1829、1865～1868、1873、1875～1877、1879
 、1881、1883、1918、1922、1931、1938、1941および19
 74～2002の配列に対して、少なくとも75%の同一性を有する配列；

(f) 配列番号878～1239、1240、1243、1247、1269、1272
 、1280、1283、1285、1286、1289、1300、1309、1318
 、1319、1327、1335、1339、1346、1359、1369、1370
 、1371、1393、1398、1405、1408、1413、1414、1417
 、1422、1429、1432、1435、1436、1438～1442、1447 30
 、1450、1453、1463、1467、1470、1473、1475、1482
 、1486、1491～1494、1501、1505、1506、1514～1517
 、1520、1522、1524、1535、1538、1542、1543、1547
 、1554、1557、1559、1561、1563～1664、1668、1669
 、1680～1706、1708～1732、1734、1736～1757、1759
 ～1765、1767～1770、1772～1774、1776～1789、1824
 、1826～1829、1865～1868、1873、1875～1877、1879
 、1881、1883、1918、1922、1931、1938、1941および19
 74～2002の配列に対して、少なくとも90%の同一性を有する配列；ならびに

(g) 配列番号878～1239、1240、1243、1247、1269、1272 40
 、1280、1283、1285、1286、1289、1300、1309、1318
 、1319、1327、1335、1339、1346、1359、1369、1370
 、1371、1393、1398、1405、1408、1413、1414、1417
 、1422、1429、1432、1435、1436、1438～1442、1447
 、1450、1453、1463、1467、1470、1473、1475、1482
 、1486、1491～1494、1501、1505、1506、1514～1517
 、1520、1522、1524、1535、1538、1542、1543、1547
 、1554、1557、1559、1561、1563～1664、1668、1669
 、1680～1706、1708～1732、1734、1736～1757、1759
 ～1765、1767～1770、1772～1774、1776～1789、1824 50

、1826～1829、1865～1868、1873、1875～1877、1879、1881、1883、1918、1922、1931、1938、1941および1974～2002において提供される配列の、縮重改変体、からなる群より選択される配列を含む、単離されたポリヌクレオチド。

【請求項2】

単離されたポリペプチドであって、以下：

(a) 配列番号1667、1670～1675、1678、1806～1822、1825、1830～1833、1834～1856、1863、1869～1872、1874、1876、1878、1880、1884～1890、1913、1917、1921、1927～1930、1932、1934、1937、1940および1942～1973； 10

(b) 配列番号1667、1670～1675、1678、1806～1822、1825、1830～1833、1834～1856、1863、1869～1872、1874、1876、1878、1880、1884～1890、1913、1917、1921、1927～1930、1932、1934、1937、1940および1942～1973において提供されるアミノ酸配列に対して、少なくとも70%の同一性を有する配列；

(c) 配列番号1667、1670～1675、1678、1806～1822、1825、1830～1833、1834～1856、1863、1869～1872、1874、1876、1878、1880、1884～1890、1913、1917、1921、1927～1930、1932、1934、1937、1940および1942～1973において提供されるアミノ酸配列に対して、少なくとも90%の同一性を有する配列； 20

(d) 請求項1に記載のポリヌクレオチドによってコードされる配列；

(e) 請求項1に記載のポリヌクレオチドによってコードされる配列に対して、少なくとも70%の同一性を有する配列；ならびに

(f) 請求項1に記載のポリヌクレオチドによってコードされる配列に対して、少なくとも90%の同一性を有する配列、

からなる群より選択されるアミノ酸配列を含む、単離されたポリペプチド。

【請求項3】

発現制御配列に作動可能に連結した請求項1に記載のポリヌクレオチドまたは請求項2に記載のポリペプチドをコードするポリヌクレオチドを含む、発現ベクター。 30

【請求項4】

請求項3に記載の発現ベクターで形質転換またはトランスフェクトした、宿主細胞。

【請求項5】

単離された抗体またはその抗原結合フラグメントであって、以下：

(a) 請求項2に記載のポリペプチド；ならびに

(b) 配列番号1948～1951、1959、1962～1968および1971において提供されるポリヌクレオチド配列またはその相補体によってコードされるアミノ酸配列を有する、ポリペプチド、 40

からなる群より選択されるポリペプチドに特異的に結合する、単離された抗体またはその抗原結合フラグメント。

【請求項6】

前記ポリペプチドが、配列番号1927～1929および1942～1973において提供される、請求項5に記載の単離された抗体または抗原結合フラグメント。

【請求項7】

患者における癌の存在を決定するための方法であって、該方法は、以下の工程：

(a) 該患者から、生物学的サンプルを得る工程；

(b) 該生物学的サンプルを、以下：

(i) 請求項2に記載のポリペプチド、 50

(i i) 配列番号 1677、1679、1806～1822、1925、1926 および 1934 において提供されるアミノ酸配列を有するポリペプチド、

(i i i) 配列番号 1677、1679、1806～1822、1925、1926 および 1934 において提供されるアミノ酸配列に対して、少なくとも 90% の同一性を有するポリペプチド、

(i v) 配列番号 1677、1679、1806～1822、1925、1926 および 1934 において提供されるアミノ酸配列に対して、少なくとも 95% の同一性を有するポリペプチド、

(v) 配列番号 1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199.、1301～1308、1310～1317、1320～1326、1328～1334、1336～1338、1340～1345、1347～1358、1360～1368、1372～1397、1399～1404、1406、1407、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924 および 1933 において提供されるポリヌクレオチドによってコードされるポリペプチド、ならびに

(v i) 配列番号 1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199.、1301～1308、1310～1317、1320～1326、1328～1334、1336～1338、1340～1345、1347～1358、1360～1368、1372～1397、1399～1404、1406、1407、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924 および 1933 において提供される配列に対して、少なくとも 90% の同一性を有するポリヌクレオチド配列によってコードされるポリペプチド、

からなる群より選択されるポリペプチドに結合する結合因子と接触させる工程；

(c) 該結合因子に結合するポリペプチドの量を、該サンプルにおいて検出する工程；ならびに

(d) 該ポリペプチドの量を、所定のカットオフ値と比較し、それにより該患者における癌の存在を決定する工程、を包含する、方法。

【請求項 8】

融合タンパク質であって、以下：

(i) 請求項 2 に記載のポリペプチド、

(i i) 配列番号 1677、1679、1806～1822、1925、1926 および

1934において提供されるアミノ酸配列を有するポリペプチド、

(iii) 配列番号1677、1679、1806～1822、1925、1926および1934において提供されるアミノ酸配列に対して、少なくとも90%の同一性を有するポリペプチド、

(iv) 配列番号1677、1679、1806～1822、1925、1926および1934において提供されるアミノ酸配列に対して、少なくとも95%の同一性を有するポリペプチド、

(v) 配列番号1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199.、1301～1308、1310～1317、1320～1326、1328～1334、1336、1338、1340～1345、1347～1358、1360～1368、1372～1397、1399～1404、1406、1407、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924および1933において提供されるポリヌクレオチドによってコードされるポリペプチド、ならびに

(vi) 配列番号1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199.、1301～1308、1310～1317、1320～1326、1328～1334、1336～1338、1340～1345、1347～1358、1360～1368、1372～1397、1399～1404、1406、1407、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924および1933において提供される配列に対して、少なくとも90%の同一性を有するポリヌクレオチドによってコードされるポリペプチド、
からなる群より選択される少なくとも1つのポリペプチドを含む、融合タンパク質。

【請求項9】

請求項8に記載の融合タンパク質であって、ここで、前記少なくとも1つのポリペプチドが、以下：

(a) 配列番号1876、1878、1913、1917、1921、1937、および1940；ならびに

(b) 配列番号1875、1877、1914、1918、1922、1938および1941において提供されるポリヌクレオチドによってコードされるポリペプチド、
において提供される、融合タンパク質。

【請求項10】

オリゴヌクレオチドであって、以下：

10

20

30

40

50

(a) 請求項 1 に記載の配列；ならびに

(b) 配列番号 1 2 4 1、1 2 4 2、1 2 4 4～1 2 4 6、1 2 4 8～1 2 6 8、1 2 7 0、1 2 7 1、1 2 7 3～1 2 7 9、1 2 8 1、1 2 8 2、1 2 8 4、1 2 8 7、1 2 8 8、1 2 9 0～1 9 9、1 3 0 1～1 3 0 8、1 3 1 0～1 3 1 7、1 3 2 0～1 3 2 6、1 3 2 8～1 3 3 4、1 3 3 6～1 3 3 8、1 3 4 0～1 3 4 5、1 3 4 7～1 3 5 8、1 3 6 0～1 3 6 8、1 3 7 2～1 3 9 7、1 3 9 9～1 4 0 4、1 4 0 6、1 4 0 7、1 4 0 9～1 4 1 2、1 4 1 5、1 4 1 6、1 4 1 8～1 4 2 1、1 4 2 3～1 4 2 8、1 4 3 0、1 4 3 1、1 4 3 3、1 4 3 4、1 4 3 7、1 4 4 3～1 4 4 6、1 4 4 8、1 4 4 9、1 4 5 1、1 4 5 2、1 4 5 4～1 4 6 2、1 4 6 4～1 4 6 6、1 4 6 8、1 4 6 9、1 4 7 1、1 4 7 2、1 4 7 4、1 4 7 6～1 4 8 1、1 4 8 3～1 4 8 5、1 4 8 7～1 4 9 0、1 4 9 5～1 5 0 0、1 5 0 2～1 5 0 4、1 5 0 7～1 5 1 3、1 5 1 8、1 5 1 9、1 5 2 1、1 5 2 3、1 5 2 5～1 5 3 4、1 5 3 6、1 5 3 7、1 5 3 9～1 5 4 1、1 5 4 4～1 5 4 6、1 5 4 8～1 5 5 3、1 5 5 5、1 5 5 6、1 5 5 8、1 5 6 0、1 5 6 2、1 6 7 6、1 7 0 7、1 7 3 3、1 7 3 5、1 7 5 8、1 7 6 6、1 7 7 1、1 7 7 5、1 7 9 0～1 8 0 5、1 8 2 3、1 9 1 0、1 9 2 3、1 9 2 4、および 1 9 3 3 において提供される配列、
 からなる群より選択される少なくとも 1 つの配列に、高度にストリンジェントな条件下でハイブリダイズする、オリゴヌクレオチド。

【請求項 1 1】

腫瘍タンパク質に特異的な T 細胞を刺激および／または拡大するための方法であって、該方法は、T 細胞を、以下：

(a) 請求項 2 に記載のポリペプチド；

(b) 配列番号 1 6 7 7、1 6 7 9、1 8 0 6～1 8 2 2、1 9 0 1～1 9 0 9、1 9 2 5、1 9 2 6、および 1 9 3 4 において提供されるアミノ酸配列を有するポリペプチド；

(c) 配列番号 1 6 7 7、1 6 7 9、1 8 0 6～1 8 2 2、1 9 0 1～1 9 0 9、1 9 2 5、1 9 2 6、および 1 9 3 4 に対して、少なくとも 9 0 % の同一性のアミノ酸配列を有するポリペプチド；

(d) 配列番号 1 6 7 7、1 6 7 9、1 8 0 6～1 8 2 2、1 9 0 1～1 9 0 9、1 9 2 5、1 9 2 6 および 1 9 3 4 に対して少なくとも 9 5 % の同一性のアミノ酸配列を有するポリペプチド；

(e) 請求項 1 に記載のポリヌクレオチド；

(f) (b)、(c) または (d) において提供されるアミノ酸配列を有するポリペプチドをコードするポリヌクレオチド；

(g) 配列番号 1 2 4 1、1 2 4 2、1 2 4 4～1 2 4 6、1 2 4 8～1 2 6 8、1 2 7 0、1 2 7 1、1 2 7 3～1 2 7 9、1 2 8 1、1 2 8 2、1 2 8 4、1 2 8 7、1 2 8 8、1 2 9 0～1 9 9、1 3 0 1～1 3 0 8、1 3 1 0～1 3 1 7、1 3 2 0～1 3 2 6、1 3 2 8～1 3 3 4、1 3 3 6～1 3 3 8、1 3 4 0～1 3 4 5、1 3 4 7～1 3 5 8、1 3 6 0～1 3 6 8、1 3 7 2～1 3 9 7、1 3 9 9～1 4 0 4、1 4 0 6、1 4 0 7、1 4 0 9～1 4 1 2、1 4 1 5、1 4 1 6、1 4 1 8～1 4 2 1、1 4 2 3～1 4 2 8、1 4 3 0、1 4 3 1、1 4 3 3、1 4 3 4、1 4 3 7、1 4 4 3～1 4 4 6、1 4 4 8、1 4 4 9、1 4 5 1、1 4 5 2、1 4 5 4～1 4 6 2、1 4 6 4～1 4 6 6、1 4 6 8、1 4 6 9、1 4 7 1、1 4 7 2、1 4 7 4、1 4 7 6～1 4 8 1、1 4 8 3～1 4 8 5、1 4 8 7～1 4 9 0、1 4 9 5～1 5 0 0、1 5 0 2～1 5 0 4、1 5 0 7～1 5 1 3、1 5 1 8、1 5 1 9、1 5 2 1、1 5 2 3、1 5 2 5～1 5 3 4、1 5 3 6、1 5 3 7、1 5 3 9～1 5 4 1、1 5 4 4～1 5 4 6、1 5 4 8～1 5 5 3、1 5 5 5、1 5 5 6、1 5 5 8、1 5 6 0、1 5 6 2、1 6 7 6、1 7 0 7、1 7 3 3、1 7 3 5、1 7 5 8、1 7 6 6、1 7 7 1、1 7 7 5、1 7 9 0～1 8 0 5、1 8 2 3、1 9 1 0、1 9 2 3、1 9 2 4、および 1 9 3 3 において提供される配列を有するポリヌクレオチド；

(h) 配列番号 1 2 4 1、1 2 4 2、1 2 4 4～1 2 4 6、1 2 4 8～1 2 6 8、1 2 7 0、1 2 7 1、1 2 7 3～1 2 7 9、1 2 8 1、1 2 8 2、1 2 8 4、1 2 8 7、1 2 8

8、1290～199.、1301～1308、1310～1317、1320～132
 6、1328～1334、1336～1338、1340～1345、1347～135
 8、1360～1368、1372～1397、1399～1404、1406、140
 7、1409～1412、1415、1416、1418～1421、1423～142
 8、1430、1431、1433、1434、1437、1443～1446、144
 8、1449、1451、1452、1454～1462、1464～1466、146
 8、1469、1471、1472、1474、1476～1481、1483～148
 5、1487～1490、1495～1500、1502～1504、1507～151
 3、1518、1519、1521、1523、1525～1534、1536、153
 7、1539～1541、1544～1546、1548～1553、1555、155 10
 6、1558、1560、1562、1676、1707、1733、1735、175
 8、1766、1771、1775、1790～1805、1823、1910、192
 3、1924、および1933において提供される配列の相補体；
 (i) 配列番号1241、1242、1244～1246、1248～1268、127
 0、1271、1273～1279、1281、1282、1284、1287、128
 8、1290～199.、1301～1308、1310～1317、1320～132
 6、1328～1334、1336～1338、1340～1345、1347～135
 8、1360～1368、1372～1397、1399～1404、1406、140
 7、1409～1412、1415、1416、1418～1421、1423～142
 8、1430、1431、1433、1434、1437、1443～1446、144 20
 8、1449、1451、1452、1454～1462、1464～1466、146
 8、1469、1471、1472、1474、1476～1481、1483～148
 5、1487～1490、1495～1500、1502～1504、1507～151
 3、1518、1519、1521、1523、1525～1534、1536、153
 7、1539～1541、1544～1546、1548～1553、1555、155
 6、1558、1560、1562、1676、1707、1733、1735、175
 8、1766、1771、1775、1790～1805、1823、1910、192
 3、1924、および1933において提供される配列に、高度にストリンジェントな条
 件下でハイブリダイズする配列；
 (j) 配列番号1241、1242、1244～1246、1248～1268、127 30
 0、1271、1273～1279、1281、1282、1284、1287、128
 8、1290～199.、1301～1308、1310～1317、1320～132
 6、1328～1334、1336～1338、1340～1345、1347～135
 8、1360～1368、1372～1397、1399～1404、1406、140
 7、1409～1412、1415、1416、1418～1421、1423～142
 8、1430、1431、1433、1434、1437、1443～1446、144
 8、1449、1451、1452、1454～1462、1464～1466、146
 8、1469、1471、1472、1474、1476～1481、1483～148
 5、1487～1490、1495～1500、1502～1504、1507～151
 3、1518、1519、1521、1523、1525～1534、1536、153 40
 7、1539～1541、1544～1546、1548～1553、1555、155
 6、1558、1560、1562、1676、1707、1733、1735、175
 8、1766、1771、1775、1790～1805、1823、1910、192
 3、1924、および1933の配列に対して、少なくとも90%の同一性を有する配列
 ；
 (k) 配列番号1241、1242、1244～1246、1248～1268、127
 0、1271、1273～1279、1281、1282、1284、1287、128
 8、1290～199.、1301～1308、1310～1317、1320～132
 6、1328～1334、1336～1338、1340～1345、1347～135
 8、1360～1368、1372～1397、1399～1404、1406、140 50

7、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924、および1933の配列に対して、少なくとも95%の同一性を有する配列 10

(1) 配列番号1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199.、1301～1308、1310～1317、1320～1326、1328～1334、1336～1338、1340～1345、1347～1358、1360～1368、1372～1397、1399～1404、1406、1407、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924、および1933において提供される配列の縮重改変体；ならびに 20

(m) (a)、(b)、(c)または(d)に記載のポリペプチドを発現する抗原提示細胞、
 からなる群より選択される少なくとも1つの成分と、T細胞の刺激および／または拡大を可能にするのに十分な条件下および時間で接触させる工程を包含する、方法。 30

【請求項12】

請求項11に記載の方法に従って調製されるT細胞を含む、単離されたT細胞集団。

【請求項13】

第1成分および第2成分を含む組成物であって、該第1成分は、生理学的に受容可能なキャリアおよび免疫刺激剤からなる群より選択され、そして第2の成分は、以下：

(a) 請求項2に記載のポリペプチド；
 (b) 配列番号1677、1679、1806～1822、1901～1909、1925、1926、および1934において提供されるアミノ酸配列を有するポリペプチド；
 (c) 配列番号1677、1679、1806～1822、1901～1909、1925、1926、および1934のうちの1つに対して、少なくとも90%の同一性を有するアミノ酸配列を有するポリペプチド； 40
 (d) 配列番号1677、1679、1806～1822、1901～1909、1925、1926および1934のうちの1つに対して、少なくとも95%の同一性を有するアミノ酸配列を有するポリペプチド；
 (e) 請求項1に記載のポリヌクレオチド；
 (f) (b)、(c)または(d)に提供されるアミノ酸配列を有するポリペプチドをコードするポリヌクレオチド；

(g) 配列番号1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199.、1301～1308、1310～1317、1320～132 50

6、1328～1334、1336～1338、1340～1345、1347～135
 8、1360～1368、1372～1397、1399～1404、1406、140
 7、1409～1412、1415、1416、1418～1421、1423～142
 8、1430、1431、1433、1434、1437、1443～1446、144
 8、1449、1451、1452、1454～1462、1464～1466、146
 8、1469、1471、1472、1474、1476～1481、1483～148
 5、1487～1490、1495～1500、1502～1504、1507～151
 3、1518、1519、1521、1523、1525～1534、1536、153
 7、1539～1541、1544～1546、1548～1553、1555、155
 6、1558、1560、1562、1676、1707、1733、1735、175 10
 8、1766、1771、1775、1790～1805、1823、1910、192
 3、1924、および1933において提供される配列を有するポリヌクレオチド；
 (h) 配列番号1241、1242、1244～1246、1248～1268、127
 0、1271、1273～1279、1281、1282、1284、1287、128
 8、1290～199.、1301～1308、1310～1317、1320～132
 6、1328～1334、1336～1338、1340～1345、1347～135
 8、1360～1368、1372～1397、1399～1404、1406、140
 7、1409～1412、1415、1416、1418～1421、1423～142
 8、1430、1431、1433、1434、1437、1443～1446、144
 8、1449、1451、1452、1454～1462、1464～1466、146 20
 8、1469、1471、1472、1474、1476～1481、1483～148
 5、1487～1490、1495～1500、1502～1504、1507～151
 3、1518、1519、1521、1523、1525～1534、1536、153
 7、1539～1541、1544～1546、1548～1553、1555、155
 6、1558、1560、1562、1676、1707、1733、1735、175
 8、1766、1771、1775、1790～1805、1823、1910、192
 3、1924、および1933において提供される配列の相補体；
 (i) 配列番号1241、1242、1244～1246、1248～1268、127
 0、1271、1273～1279、1281、1282、1284、1287、128
 8、1290～199.、1301～1308、1310～1317、1320～132 30
 6、1328～1334、1336～1338、1340～1345、1347～135
 8、1360～1368、1372～1397、1399～1404、1406、140
 7、1409～1412、1415、1416、1418～1421、1423～142
 8、1430、1431、1433、1434、1437、1443～1446、144
 8、1449、1451、1452、1454～1462、1464～1466、146
 8、1469、1471、1472、1474、1476～1481、1483～148
 5、1487～1490、1495～1500、1502～1504、1507～151
 3、1518、1519、1521、1523、1525～1534、1536、153
 7、1539～1541、1544～1546、1548～1553、1555、155
 6、1558、1560、1562、1676、1707、1733、1735、175 40
 8、1766、1771、1775、1790～1805、1823、1910、192
 3、1924、および1933において提供される配列に、高度にストリンジェントな条
 件下でハイブリダイズする配列；
 (j) 配列番号1241、1242、1244～1246、1248～1268、127
 0、1271、1273～1279、1281、1282、1284、1287、128
 8、1290～199.、1301～1308、1310～1317、1320～132
 6、1328～1334、1336～1338、1340～1345、1347～135
 8、1360～1368、1372～1397、1399～1404、1406、140
 7、1409～1412、1415、1416、1418～1421、1423～142
 8、1430、1431、1433、1434、1437、1443～1446、144 50

8、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924、および1933の配列に対して、少なくとも90%の同一性を有する配列；

(k) 配列番号1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199、1301～1308、1310～1317、1320～1326、1328～1334、1336～1338、1340～1345、1347～1358、1360～1368、1372～1397、1399～1404、1406、1407、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924、および1933の配列に対して、少なくとも95%の同一性を有する配列；

(l) 配列番号1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199、1301～1308、1310～1317、1320～1326、1328～1334、1336～1338、1340～1345、1347～1358、1360～1368、1372～1397、1399～1404、1406、1407、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924および1933において提供される配列の縮重改変体；

(m) 請求項5に記載の抗体；

(n) 請求項8に記載の融合タンパク質；

(o) 請求項12に記載のT細胞集団；ならびに

(p) (a)、(b)、(c)または(d)に記載のポリペプチドを発現する抗原提示細胞、

からなる群より選択される、組成物。

【請求項14】

患者における免疫応答を刺激するための方法であって、該方法は、請求項13に記載の組成物を該患者に投与する工程を包含する、方法。

【請求項15】

患者における肺癌を処置するための方法であって、該方法は、請求項 13 に記載の組成物を該患者に投与する工程を包含する、方法。

【請求項 16】

患者における癌の存在を決定するための方法であって、該方法は、以下の工程：

- (a) 該患者から生物学的サンプルを得る工程；
 - (b) 請求項 10 に記載のオリゴヌクレオチドと、該生物学的サンプルとを接触させる工程；
 - (c) 該オリゴヌクレオチドにハイブリダイズするポリヌクレオチドの量を、該サンプルにおいて検出する工程；ならびに
 - (d) 該オリゴヌクレオチドにハイブリダイズした該ポリヌクレオチドの量を、所定のカットオフ値と比較し、それによって該患者における該癌の存在を決定する工程、
- を包含する、方法。

【請求項 17】

請求項 10 に記載の少なくとも 1 つのオリゴヌクレオチドを含む、診断キット。

【請求項 18】

請求項 5 に記載の少なくとも 1 つの抗体および検出試薬を備える診断キットであって、ここで、該検出試薬は、レポーター基を含む、診断キット。

【請求項 19】

患者における肺癌の発症を阻害するための方法であって、該方法は、以下の工程：

- (a) 患者から単離した CD4+ および／または CD8+ T 細胞を、以下：
 - (i) 請求項 2 に記載のポリペプチド 2、
 - (ii) 配列番号 1677、1679、1806～1822、1901～1909、1925、1926、および 1934 において提供されるアミノ酸配列を有するポリペプチド、
 - (iii) 配列番号 1677、1679、1806～1822、1901～1909、1925、1926、および 1934 のうちの 1 つに対して、少なくとも 90% の同一性を有するアミノ酸配列を有するポリペプチド、
 - (iv) 配列番号 1677、1679、1806～1822、1901～1909、1925、1926、および 1934 のうちの 1 つに対して、少なくとも 95% の同一性を有するアミノ酸配列を有するポリペプチド、
 - (v) 請求項 1 に記載のポリヌクレオチド、
 - (vi) (i)、(ii)、(iii) または (iv) において提供されるアミノ酸配列を有するポリペプチドをコードするポリヌクレオチド、
 - (vii) 配列番号 1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199、1301～1308、1310～1317、1320～1326、1328～1334、1336～1338、1340～1345、1347～1358、1360～1368、1372～1397、1399～1404、1406、1407、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924、および 1933 において提供される配列を有するポリヌクレオチド、
 - (viii) 配列番号 1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、

1 2 8 8、1 2 9 0～1 9 9.、1 3 0 1～1 3 0 8、1 3 1 0～1 3 1 7、1 3 2 0～
 1 3 2 6、1 3 2 8～1 3 3 4、1 3 3 6～1 3 3 8、1 3 4 0～1 3 4 5、1 3 4 7～
 1 3 5 8、1 3 6 0～1 3 6 8、1 3 7 2～1 3 9 7、1 3 9 9～1 4 0 4、1 4 0 6、
 1 4 0 7、1 4 0 9～1 4 1 2、1 4 1 5、1 4 1 6、1 4 1 8～1 4 2 1、1 4 2 3～
 1 4 2 8、1 4 3 0、1 4 3 1、1 4 3 3、1 4 3 4、1 4 3 7、1 4 4 3～1 4 4 6、
 1 4 4 8、1 4 4 9、1 4 5 1、1 4 5 2、1 4 5 4～1 4 6 2、1 4 6 4～1 4 6 6、
 1 4 6 8、1 4 6 9、1 4 7 1、1 4 7 2、1 4 7 4、1 4 7 6～1 4 8 1、1 4 8 3～
 1 4 8 5、1 4 8 7～1 4 9 0、1 4 9 5～1 5 0 0、1 5 0 2～1 5 0 4、1 5 0 7～
 1 5 1 3、1 5 1 8、1 5 1 9、1 5 2 1、1 5 2 3、1 5 2 5～1 5 3 4、1 5 3 6、
 1 5 3 7、1 5 3 9～1 5 4 1、1 5 4 4～1 5 4 6、1 5 4 8～1 5 5 3、1 5 5 5、 10
 1 5 5 6、1 5 5 8、1 5 6 0、1 5 6 2、1 6 7 6、1 7 0 7、1 7 3 3、1 7 3 5、
 1 7 5 8、1 7 6 6、1 7 7 1、1 7 7 5、1 7 9 0～1 8 0 5、1 8 2 3、1 9 1 0、
 1 9 2 3、1 9 2 4、および1 9 3 3において提供される配列の相補体、
 (i x) 配列番号1 2 4 1、1 2 4 2、1 2 4 4～1 2 4 6、1 2 4 8～1 2 6 8、1 2
 7 0、1 2 7 1、1 2 7 3～1 2 7 9、1 2 8 1、1 2 8 2、1 2 8 4、1 2 8 7、1 2
 8 8、1 2 9 0～1 9 9.、1 3 0 1～1 3 0 8、1 3 1 0～1 3 1 7、1 3 2 0～1 3
 2 6、1 3 2 8～1 3 3 4、1 3 3 6～1 3 3 8、1 3 4 0～1 3 4 5、1 3 4 7～1 3
 5 8、1 3 6 0～1 3 6 8、1 3 7 2～1 3 9 7、1 3 9 9～1 4 0 4、1 4 0 6、1 4
 0 7、1 4 0 9～1 4 1 2、1 4 1 5、1 4 1 6、1 4 1 8～1 4 2 1、1 4 2 3～1 4
 2 8、1 4 3 0、1 4 3 1、1 4 3 3、1 4 3 4、1 4 3 7、1 4 4 3～1 4 4 6、1 4 20
 4 8、1 4 4 9、1 4 5 1、1 4 5 2、1 4 5 4～1 4 6 2、1 4 6 4～1 4 6 6、1 4
 6 8、1 4 6 9、1 4 7 1、1 4 7 2、1 4 7 4、1 4 7 6～1 4 8 1、1 4 8 3～1 4
 8 5、1 4 8 7～1 4 9 0、1 4 9 5～1 5 0 0、1 5 0 2～1 5 0 4、1 5 0 7～1 5
 1 3、1 5 1 8、1 5 1 9、1 5 2 1、1 5 2 3、1 5 2 5～1 5 3 4、1 5 3 6、1 5
 3 7、1 5 3 9～1 5 4 1、1 5 4 4～1 5 4 6、1 5 4 8～1 5 5 3、1 5 5 5、1 5
 5 6、1 5 5 8、1 5 6 0、1 5 6 2、1 6 7 6、1 7 0 7、1 7 3 3、1 7 3 5、1 7
 5 8、1 7 6 6、1 7 7 1、1 7 7 5、1 7 9 0～1 8 0 5、1 8 2 3、1 9 1 0、1 9
 2 3、1 9 2 4、および1 9 3 3において提供される配列に、高度にストリンジェントな
 条件下でハイブリダイズする配列、
 (x) 配列番号1 2 4 1、1 2 4 2、1 2 4 4～1 2 4 6、1 2 4 8～1 2 6 8、1 2 7 30
 0、1 2 7 1、1 2 7 3～1 2 7 9、1 2 8 1、1 2 8 2、1 2 8 4、1 2 8 7、1 2 8
 8、1 2 9 0～1 9 9.、1 3 0 1～1 3 0 8、1 3 1 0～1 3 1 7、1 3 2 0～1 3 2
 6、1 3 2 8～1 3 3 4、1 3 3 6～1 3 3 8、1 3 4 0～1 3 4 5、1 3 4 7～1 3 5
 8、1 3 6 0～1 3 6 8、1 3 7 2～1 3 9 7、1 3 9 9～1 4 0 4、1 4 0 6、1 4 0
 7、1 4 0 9～1 4 1 2、1 4 1 5、1 4 1 6、1 4 1 8～1 4 2 1、1 4 2 3～1 4 2
 8、1 4 3 0、1 4 3 1、1 4 3 3、1 4 3 4、1 4 3 7、1 4 4 3～1 4 4 6、1 4 4
 8、1 4 4 9、1 4 5 1、1 4 5 2、1 4 5 4～1 4 6 2、1 4 6 4～1 4 6 6、1 4 6
 8、1 4 6 9、1 4 7 1、1 4 7 2、1 4 7 4、1 4 7 6～1 4 8 1、1 4 8 3～1 4 8
 5、1 4 8 7～1 4 9 0、1 4 9 5～1 5 0 0、1 5 0 2～1 5 0 4、1 5 0 7～1 5 1
 3、1 5 1 8、1 5 1 9、1 5 2 1、1 5 2 3、1 5 2 5～1 5 3 4、1 5 3 6、1 5 3 40
 7、1 5 3 9～1 5 4 1、1 5 4 4～1 5 4 6、1 5 4 8～1 5 5 3、1 5 5 5、1 5 5
 6、1 5 5 8、1 5 6 0、1 5 6 2、1 6 7 6、1 7 0 7、1 7 3 3、1 7 3 5、1 7 5
 8、1 7 6 6、1 7 7 1、1 7 7 5、1 7 9 0～1 8 0 5、1 8 2 3、1 9 1 0、1 9 2
 3、1 9 2 4、および1 9 3 3の配列に対して、少なくとも90%の同一性を有する配列
 、
 (x i) 配列番号1 2 4 1、1 2 4 2、1 2 4 4～1 2 4 6、1 2 4 8～1 2 6 8、1 2
 7 0、1 2 7 1、1 2 7 3～1 2 7 9、1 2 8 1、1 2 8 2、1 2 8 4、1 2 8 7、1 2
 8 8、1 2 9 0～1 9 9.、1 3 0 1～1 3 0 8、1 3 1 0～1 3 1 7、1 3 2 0～1 3
 2 6、1 3 2 8～1 3 3 4、1 3 3 6～1 3 3 8、1 3 4 0～1 3 4 5、1 3 4 7～1 3
 5 8、1 3 6 0～1 3 6 8、1 3 7 2～1 3 9 7、1 3 9 9～1 4 0 4、1 4 0 6、1 4 50

07、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924、および1933の配列に対して、少なくとも95%の同一性を有する配列、

(x i i) 配列番号1241、1242、1244～1246、1248～1268、1270、1271、1273～1279、1281、1282、1284、1287、1288、1290～199、1301～1308、1310～1317、1320～1326、1328～1334、1336～1338、1340～1345、1347～1358、1360～1368、1372～1397、1399～1404、1406、1407、1409～1412、1415、1416、1418～1421、1423～1428、1430、1431、1433、1434、1437、1443～1446、1448、1449、1451、1452、1454～1462、1464～1466、1468、1469、1471、1472、1474、1476～1481、1483～1485、1487～1490、1495～1500、1502～1504、1507～1513、1518、1519、1521、1523、1525～1534、1536、1537、1539～1541、1544～1546、1548～1553、1555、1556、1558、1560、1562、1676、1707、1733、1735、1758、1766、1771、1775、1790～1805、1823、1910、1923、1924および1933において提供される配列の縮重改変体、ならびに

(x i i i) (i)、(i i)、(i i i)または(i v)に記載のポリペプチドを発現する抗原提示細胞、

からなる群より選択される少なくとも1つの成分と共に、T細胞が増殖するようにインキュベートする工程；

(b) 該増殖したT細胞の有効量を該患者に投与し、これによって該患者における癌の発症を阻害する工程、

を包含する、方法。

【発明の詳細な説明】

【0001】

(発明の技術分野)

本発明は、一般に、癌（例えば、肺癌）の治療および診断に関する。本発明は、より具体的には、肺腫瘍タンパク質の少なくとも一部を含むポリペプチド、およびこのようなポリペプチドをコードするポリヌクレオチドに関する。このようなポリペプチドおよびポリヌクレオチドは、薬学的組成物（例えば、ワクチン）および肺癌の診断および処置のための他の組成物において有用である。

【0002】

(参照によるCD-ROMの配列表の包含)

本出願に関連する配列表は、A I § 801 (a) に基づくペーパーコピーの代わりにCD-ROMで提供され、そしてこれは本明細書に参考として援用される。4枚のCD-ROMが提供され、これは配列表の同一のコピーを含む：CD-ROM No. 1は、「COPY 1-SEQUENCE LISTING PART」とラベルされ、2001年7月10日に作成された1.4MBの47802 pc. app. txtというファイルを含む；CD-ROM No. 2は、「COPY 2-SEQUENCE LISTING」とラベルされ、2001年7月10日に作成された1.4MBの47802 pc. app

. t x t というファイルを含む；C D - R O M N o . 3 は、「C O P Y 3 - S E Q U E N C E L I S T I N G P A R T」とラベルされ、2001年7月10日に作成され、1.4MBの47802 p c . a p p . t x t というファイルを含む；C D - R O M N o . 4 は、「C R F」とラベルされ、2001年7月10日に作成された1.4Mbの47802 p c . a p p . t x t というファイルを含む。

【0003】

(発明の背景)

癌は、世界中の重大な健康の問題である。癌の検出および治療において進歩はあったが、予防および処置のためのワクチンおよび他の普遍的に成功する方法は、現在利用可能でない。現在の治療（これは、一般に、化学療法または手術と放射線との組み合わせに基づく）は、多くの患者において不適切であるとの証明が続いている。 10

【0004】

肺癌は、米国における男性および女性の両方における癌での死亡の主な原因であり、1994年には推定172,000件の新たな症例が報告されている。全ての肺癌患者における5年間の生存率は、診断時における疾患の病期に関わらず、わずか13%である。これは、この疾患がまだ局在している間に検出された症例における46%の5年間生存率と対照的である。しかし、肺癌が広がる前に発見されるのはこの疾患の16%に過ぎない。

【0005】

早期発見は難しい。なぜなら、臨床症状は、この疾患が進行した病期に達するまでしばしばみられないからである。現在のところ、診断は、胸部X線、痰に含まれる細胞型の分析、および気管支道の光ファイバー試験の使用により補助される。処置レジメンは、癌の型および病期により決定され、そして手術、放射線治療および／または化学療法を含む。 20

【0006】

この癌および他の癌についての治療におけるかなりの研究にも関わらず、肺癌は、診断および効果的に処置することが困難なままである。従って、当該分野において、このような癌を検出および処置するための改良された方法の必要性が存在する。本発明は、これらの必要性を満たし、そしてさらに他の関連する利点を提供する。

【0007】

(発明の要旨)

1つの局面において、本発明は、以下： 30

(a) 配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002において提供される配列；

(b) 配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002において提供される配列の相補体； 40

(c) 配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002において提供 50

される配列のうち、少なくとも20、25、30、35、40、45、50、75および100個連続した残基を含む配列；

(d) 配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002において提供される配列に、中程度～高度にストリンジェントな条件下でハイブリダイズする配列；

(e) 配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002の配列に対して、少なくとも75%、80%、85%、90%、95%、96%、97%、98%、または99%の同一性を有する配列； 10

(f) 配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941、および1974～2002において提供される配列の縮重改変体、 20

からなる群より選択される配列を含むポリヌクレオチド組成物を提供する。

【0008】

1つの好ましい実施形態において、本発明のポリヌクレオチド組成物は、試験した肺腫瘍サンプルの少なくとも約20%、より好ましくは少なくとも約30%、そして最も好ましくは少なくとも50%において、正常な組織よりも少なくとも約2倍、好ましくは少なくとも約5倍、そして最も好ましくは少なくとも約10倍高いレベルで発現される。 30

【0009】

本発明は、別の局面において、上記のポリヌクレオチド配列によってコードされるアミノ酸配列を含むポリペプチド組成物を提供する。

【0010】

本発明は、配列番号324～340、783、786、787、789、791、793、795、797～799、805、806、809、827、1667、1670～1675、1677～1679、1806～1822、1825、1830～1833、1834～1856、1863、1864、1869～1872、1874、1876、1878、1880、1882、1884～1890、1901～1909、1913、1917、1921、1925～1930、1932、1934、1937、1940、および1942～1973に列挙される配列からなる群より選択されるアミノ酸配列を含む、ポリペプチド組成物をさらに提供する。 40

【0011】

特定の好ましい実施形態において、本発明のポリペプチドおよび／またはポリヌクレオチドは、免疫原性である（すなわち、これらは、免疫応答（特に、本明細書中にさらに記載されるような体液性免疫応答および／または細胞性免疫応答）を惹起し得る）。

【0012】

本発明は、開示されるポリペプチドおよび／またはポリヌクレオチド配列のフラグメント、改変体および／または誘導体をさらに提供し、ここで、このフラグメント、改変体および／または誘導体は、好ましくは、配列番号324～340、783、786、787、 50

789、791、793、795、797～799、805、806、809、827、1667、1670～1675、1677～1679、1806～1822、1825、1830～1833、1834～1856、1863、1864、1869～1872、1874、1876、1878、1880、1882、1884～1890、1901～1909、1913、1917、1921、1925～1930、1932、1934、1937、1940および1942～1973に示されるポリペプチド配列、または配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002に示されるポリヌクレオチド配列によってコードされるポリペプチド配列の免疫原性活性のうち、少なくとも約50%、好ましくは少なくとも約70%、そして最も好ましくは少なくとも90%の免疫原性活性レベルを有する。

10

【0013】

本発明は、上記のポリペプチドをコードするポリヌクレオチド、このようなポリヌクレオチドを含む発現ベクター、およびこのような発現ベクターで形質転換またはトランスフェクトされた宿主細胞をさらに提供する。

【0014】

他の局面において、本発明は、上記のポリペプチドまたはポリヌクレオチドおよび生理学的に受容可能なキャリアを含む薬学的組成物を提供する。

20

【0015】

本発明の関連する局面において、予防用途または治療用途のための薬学的組成物（例えば、ワクチン組成物）が提供される。このような組成物は、一般に、本発明の免疫原性ポリペプチドまたはポリヌクレオチド、および免疫刺激剤（例えば、アジュバント）を含む。

【0016】

本発明は、以下を含む薬学的組成物をさらに提供する：（a）本発明のポリペプチドまたはそのフラグメントに特異的に結合する抗体またはその抗原結合フラグメント；および（b）生理学的に受容可能なキャリア。

30

【0017】

さらなる局面において、本発明は、以下を含む薬学的組成物を提供する：（a）上記のポリペプチドを発現する抗原提示細胞、および（b）薬学的に受容可能なキャリアまたは賦形剤。例示的な抗原提示細胞としては、樹状細胞、マクロファージ、単球、線維芽細胞、およびB細胞が挙げられる。

【0018】

関連する局面において、以下を含む薬学的組成物が提供される：（a）上記のポリペプチドを発現する抗原提示細胞、および（b）免疫刺激剤。

【0019】

本発明は、他の局面において、代表的には生理学的に受容可能なキャリアおよび／または免疫刺激剤を含む薬学的組成物（例えば、ワクチン組成物）の形態で、少なくとも1つの上記ポリペプチドを含む融合タンパク質、ならびにこのような融合タンパク質をコードするポリヌクレオチドをさらに提供する。この融合タンパク質は、本明細書中に記載される複数の免疫原性ポリペプチドまたはその部分／改変体を含み得、そしてこのポリペプチドの発現、精製、および／または免疫原性を促進するための1つ以上のポリペプチドセグメントをさらに含み得る。

40

【0020】

さらなる局面において、本発明は、患者における免疫応答（好ましくは、ヒト患者におけるT細胞応答）を刺激するための方法を提供し、この方法は、本明細書中に記載の薬学的組成物を投与する工程を包含する。この患者は、肺癌に罹患している患者であり得る（こ

50

の場合、この方法はこの疾患の処置を提供する)か、またはこのような疾患の危険性があると考えられる患者は、予防的に処置され得る。

【0021】

さらなる局面において、本発明は、患者における癌の発症を阻害するための方法を提供し、この方法は、上記に列挙される薬学的組成物を患者に投与する工程を包含する。この患者は、肺癌に罹患している患者であり得る(この場合、この方法はこの疾患の処置を提供する)か、またはこのような疾患の危険性があると考えられる患者は、予防的に処置され得る。

【0022】

本発明は、他の局面において、生物学的サンプルから腫瘍細胞を除去するための方法をさらに提供し、この方法は、本発明のポリペプチドと特異的に反応するT細胞を、生物学的サンプルと接触させる工程を包含し、ここで、この接触させる工程は、このタンパク質を発現する細胞のこのサンプルからの除去を可能にするために十分な条件および時間で実施される。

10

【0023】

関連する局面において、患者における癌の発症を阻害するための方法が提供され、この方法は、上記のように処理した生物学的サンプルを患者に投与する工程を包含する。

【0024】

他の局面において、本発明のポリペプチドに特異的なT細胞を刺激および/または拡大するための方法がさらに提供され、この方法は、T細胞を、T細胞の刺激および/または拡大を可能にするのに十分な条件下および時間で、以下のうち1つ以上と接触させる工程を包含する:(i)上記のポリペプチド;(ii)このようなポリペプチドをコードするポリヌクレオチド;および/または(iii)このようなポリペプチドを発現する抗原提示細胞。上記のように調製されたT細胞を含む単離されたT細胞集団もまた、提供される。

20

【0025】

さらなる局面において、本発明は、患者における癌の発症を阻害するための方法を提供し、この方法は、上記のようなT細胞集団の有効量を患者に投与する工程を包含する。

【0026】

本発明は、患者における癌の発症を阻害するための方法をさらに提供し、この方法は、以下の工程を包含する:(a)以下:(i)本明細書中に開示されるポリペプチドの少なくとも免疫原性部分を含むポリペプチド;(ii)このようなポリペプチドをコードするポリヌクレオチド;および(iii)このようなポリペプチドを発現する抗原提示細胞、のうちの1つ以上と、患者から単離されたCD4⁺および/またはCD8⁺T細胞とをインキュベートする工程;ならびに(b)増殖したT細胞の有効量をこの患者に投与し、そしてこれによってこの患者における癌の発症を阻害する工程。増殖した細胞は、この患者への投与の前にクローニングされ得るが、その必要はない。

30

【0027】

さらなる局面において、本発明は、患者における癌(好ましくは、肺癌)の存在または非存在を決定するための方法を提供し、この方法は、以下の工程を包含する:(a)上記のポリペプチドに結合する結合因子を、患者から得た生物学的サンプルに接触させる工程;(b)この結合因子に結合したポリペプチドの量を、このサンプルにおいて検出する工程;および(c)このポリペプチドの量と所定のカットオフ値とを比較して、それによってこの患者における癌の存在または非存在を決定する工程。好ましい実施形態において、この結合因子は、抗体であり、より好ましくはモノクローナル抗体である。

40

【0028】

本発明はまた、他の局面において、患者における癌の進行をモニタリングするための方法を提供する。このような方法は、以下の工程を包含する:(a)第1の時点で患者から得た生物学的サンプルを、上記ポリペプチドに結合する結合因子と接触させる工程;(b)この結合因子に結合するポリペプチドの量を、このサンプル中で検出する工程;(c)引き続き時点でこの患者から得た生物学的サンプルを使用して、工程(a)および(b)を

50

繰り返す工程；ならびに（d）工程（c）において検出したポリペプチドの量を、工程（b）において検出した量と比較し、それによってこの患者における癌の進行をモニタリングする工程。

【0029】

本発明は、他の局面において、患者における癌の存在または非存在を決定するための方法をさらに提供し、この方法は、以下の工程を包含する：（a）患者から得た生物学的サンプルを、本発明のポリペプチドをコードするポリヌクレオチドにハイブリダイズするオリゴヌクレオチドと接触させる工程；（b）このオリゴヌクレオチドにハイブリダイズするポリヌクレオチド（好ましくはmRNA）のレベルを、このサンプルにおいて検出する工程；および（c）このオリゴヌクレオチドにハイブリダイズするポリヌクレオチドのレベルを、所定のカットオフ値と比較し、それによってこの患者における癌の存在または非存在を決定する工程。特定の実施形態において、mRNAの量は、例えば、上記のようなポリペプチドをコードするポリヌクレオチド、またはこのようなポリヌクレオチドの相補体にハイブリダイズする少なくとも1つのオリゴヌクレオチドプライマーを使用するポリメラーゼ連鎖反応を介して、検出される。他の実施形態において、mRNAの量は、上記のポリペプチドをコードするポリヌクレオチド、またはこのようなポリヌクレオチドの相補体にハイブリダイズするオリゴヌクレオチドプローブを利用するハイブリダイゼーション技術を使用して、検出される。

10

【0030】

関連する局面において、患者における癌の進行をモニタリングするための方法が提供され、この方法は、以下の工程を包含する：（a）患者から得た生物学的サンプルを、本発明のポリペプチドをコードするポリヌクレオチドにハイブリダイズするオリゴヌクレオチドと接触させる工程；（b）このオリゴヌクレオチドにハイブリダイズしたポリヌクレオチドの量を、このサンプルにおいて検出する工程；（c）引き続く時点でこの患者から得た生物学的サンプルを使用して、工程（a）および（b）を繰り返す工程；ならびに（d）工程（c）において検出したポリヌクレオチドの量を、工程（b）において検出した量と比較し、それによってこの患者における癌の進行をモニタリングする工程。

20

【0031】

さらなる局面において、本発明は、上記のポリペプチドに結合する抗体（例えば、モノクローナル抗体）、ならびにこのような抗体を含む診断キットを提供する。上記の1つ以上のオリゴヌクレオチドプライマーまたはプローブを含む診断キットもまた、提供される。

30

【0032】

本発明のこれらの局面および他の局面は、以下の詳細な説明を参照して明らかとなる。本明細書中に開示される全ての参考文献は、その各々が別個に援用されるように、本明細書中でその全体において参考として援用される。

【0033】

（配列識別子）

配列番号1は、クローン#19038（L845Pともいわれる）について決定されたcDNA配列である。

配列番号2は、クローン#19036について決定されたcDNA配列である。

40

配列番号3は、クローン#19034について決定されたcDNA配列である。

配列番号4は、クローン#19033について決定されたcDNA配列である。

配列番号5は、クローン#19032について決定されたcDNA配列である。

配列番号6は、クローン#19030（L559Sともいわれる）について決定されたcDNA配列である。

配列番号7は、クローン#19029について決定されたcDNA配列である。

配列番号8は、クローン#19025について決定されたcDNA配列である。

配列番号9は、クローン#19023について決定されたcDNA配列である。

配列番号10は、クローン#18929について決定されたcDNA配列である。

配列番号11は、クローン#19010について決定されたcDNA配列である。

50

配列番号 12 は、クローン # 19009 について決定された cDNA 配列である。

配列番号 13 は、クローン # 19005、19007、19016、および 19017 について決定された cDNA 配列である。

配列番号 14 は、クローン # 19004 について決定された cDNA 配列である。

配列番号 15 は、クローン # 19002 および 18965 について決定された cDNA 配列である。

配列番号 16 は、クローン # 18998 について決定された cDNA 配列である。

配列番号 17 は、クローン # 18997 について決定された cDNA 配列である。

配列番号 18 は、クローン # 18996 について決定された cDNA 配列である。

配列番号 19 は、クローン # 18995 について決定された cDNA 配列である。 10

配列番号 20 は、クローン # 18994 (L846P としても知られている) について決定された cDNA 配列である。

配列番号 21 は、クローン # 18992 について決定された cDNA 配列である。

配列番号 22 は、クローン # 18991 について決定された cDNA 配列である。

配列番号 23 は、クローン # 18990 (クローン # 20111 ともいわれる) について決定された cDNA 配列である。

配列番号 24 は、クローン # 18987 について決定された cDNA 配列である。

配列番号 25 は、クローン # 18985 (L839P ともいわれる) について決定された cDNA 配列である。

配列番号 26 は、クローン # 18984 (L847P ともいわれる) について決定された cDNA 配列である。 20

配列番号 27 は、クローン # 18983 について決定された cDNA 配列である。

配列番号 28 は、クローン # 18976 および 18980 について決定された cDNA 配列である。

配列番号 29 は、クローン # 18975 について決定された cDNA 配列である。

配列番号 30 は、クローン # 18974 について決定された cDNA 配列である。

配列番号 31 は、クローン # 18973 について決定された cDNA 配列である。

配列番号 32 は、クローン # 18972 について決定された cDNA 配列である。

配列番号 33 は、クローン # 18971 (L801P ともいわれる) について決定された cDNA 配列である。 30

配列番号 34 は、クローン # 18970 について決定された cDNA 配列である。

配列番号 35 は、クローン # 18966 について決定された cDNA 配列である。

配列番号 36 は、クローン # 18964、18968 および 19039 について決定された cDNA 配列である。

配列番号 37 は、クローン # 18960 について決定された cDNA 配列である。

配列番号 38 は、クローン # 18959 について決定された cDNA 配列である。

配列番号 39 は、クローン # 18958 および 18982 について決定された cDNA 配列である。

配列番号 40 は、クローン # 18956 および 19015 について決定された cDNA 配列である。 40

配列番号 41 は、クローン # 18954 (L848P ともいわれる) について決定された cDNA 配列である。

配列番号 42 は、クローン # 18951 について決定された cDNA 配列である。

配列番号 43 は、クローン # 18950 について決定された cDNA 配列である。

配列番号 44 は、クローン # 18949 および 19024 (L844P ともいわれる) について決定された cDNA 配列である。

配列番号 45 は、クローン # 18948 について決定された cDNA 配列である。

配列番号 46 は、クローン # 18947 (L840P ともいわれる) について決定された cDNA 配列である。

配列番号 47 は、クローン # 18946、18953、18969 および 19027 につ 50

いて決定された c D N A 配列である。

配列番号 48 は、クローン # 18942 について決定された c D N A 配列である。

配列番号 49 は、クローン # 18940、18962、18963、19006、19008、19000 および 19031 について決定された c D N A 配列である。

配列番号 50 は、クローン # 18939 について決定された c D N A 配列である。

配列番号 51 は、クローン # 18938 および 18952 について決定された c D N A 配列である。

配列番号 52 は、クローン # 18938 について決定された c D N A 配列である。

配列番号 53 は、クローン # 18937 について決定された c D N A 配列である。

配列番号 54 は、クローン # 18934、18935、18993 および 19022 (L 548S ともしられる) について決定された c D N A 配列である。 10

配列番号 55 は、クローン # 18932 について決定された c D N A 配列である。

配列番号 56 は、クローン # 18931 および 18936 について決定された c D N A 配列である。

配列番号 57 は、クローン # 18930 について決定された c D N A 配列である。

配列番号 58 は、クローン # 19014 について決定された c D N A 配列である (この配列は、クローン L 773P (これはまた、1999 年 4 月 2 日出願の同時係属中の米国出願 09/285, 479 にもまた記載される) に対する相同性を有する)。

配列番号 59 は、クローン # 19127 について決定された c D N A 配列である。

配列番号 60 は、クローン # 19057 および 19064 について決定された c D N A 配列である。 20

配列番号 61 は、クローン # 19122 について決定された c D N A 配列である。

配列番号 62 は、クローン # 19120 および 18121 について決定された c D N A 配列である。

配列番号 63 は、クローン # 19118 について決定された c D N A 配列である。

配列番号 64 は、クローン # 19117 について決定された c D N A 配列である。

配列番号 65 は、クローン # 19116 について決定された c D N A 配列である。

配列番号 66 は、クローン # 19114 について決定された c D N A 配列である。

配列番号 67 は、クローン # 19112 (L 561S としても知られている) について決定された c D N A 配列である。 30

配列番号 68 は、クローン # 19110 について決定された c D N A 配列である。

配列番号 69 は、クローン # 19107 (L 552S ともしられる) について決定された c D N A 配列である。

配列番号 70 は、クローン # 19106 (L 547S ともしられる) について決定された c D N A 配列である。

配列番号 71 は、クローン # 19105 および 19111 について決定された c D N A 配列である。

配列番号 72 は、クローン # 19099 について決定された c D N A 配列である。

配列番号 73 は、クローン # 19095、19104 および 19125 (L 549S ともしられる) について決定された c D N A 配列である。 40

配列番号 74 は、クローン # 19094 について決定された c D N A 配列である。

配列番号 75 は、クローン # 19089 および 19101 について決定された c D N A 配列である。

配列番号 76 は、クローン # 19088 について決定された c D N A 配列である。

配列番号 77 は、クローン # 19087、19092、19096、19100 および 19119 について決定された c D N A 配列である。

配列番号 78 は、クローン # 19086 について決定された c D N A 配列である。

配列番号 79 は、クローン # 19085 (L 550S ともしられる) について決定された c D N A 配列である。

配列番号 80 は、クローン # 19084 (クローン # 19079 ともしられる) について 50

決定された c D N A 配列である。

配列番号 8 1 は、クローン # 1 9 0 8 2 について決定された c D N A 配列である。

配列番号 8 2 は、クローン # 1 9 0 8 0 について決定された c D N A 配列である。

配列番号 8 3 は、クローン # 1 9 0 7 7 について決定された c D N A 配列である。

配列番号 8 4 は、クローン # 1 9 0 7 6 (L 5 5 1 S と い わ れ る) について決定された c D N A 配列である。

配列番号 8 5 は、クローン # 1 9 0 7 4 (クローン # 2 0 1 0 2 と い わ れ る) について決定された c D N A 配列である。

配列番号 8 6 は、クローン # 1 9 0 7 3 (L 5 6 0 S と い わ れ る) について決定された c D N A 配列である。

10

配列番号 8 7 は、クローン # 1 9 0 7 2 および 1 9 1 1 5 について決定された c D N A 配列である。

配列番号 8 8 は、クローン # 1 9 0 7 1 について決定された c D N A 配列である。

配列番号 8 9 は、クローン # 1 9 0 7 0 について決定された c D N A 配列である。

配列番号 9 0 は、クローン # 1 9 0 6 9 について決定された c D N A 配列である。

配列番号 9 1 は、クローン # 1 9 0 6 8 (L 5 6 3 S と い わ れ る) について決定された c D N A 配列である。

配列番号 9 2 は、クローン # 1 9 0 6 6 について決定された c D N A 配列である。

配列番号 9 3 は、クローン # 1 9 0 6 5 について決定された c D N A 配列である。

配列番号 9 4 は、クローン # 1 9 0 6 3 について決定された c D N A 配列である。

20

配列番号 9 5 は、クローン # 1 9 0 6 1、1 9 0 8 1、1 9 1 0 8 および 1 9 1 0 9 について決定された c D N A 配列である。

配列番号 9 6 は、クローン # 1 9 0 6 0、1 9 0 6 7 および 1 9 0 8 3 (L 5 4 8 S と い わ れ る) について決定された c D N A 配列である。

配列番号 9 7 は、クローン # 1 9 0 5 9 および 1 9 0 6 2 について決定された c D N A 配列である。

配列番号 9 8 は、クローン # 1 9 0 5 8 について決定された c D N A 配列である。

配列番号 9 9 は、クローン # 1 9 1 2 4 について決定された c D N A 配列である。

配列番号 1 0 0 は、クローン # 1 8 9 2 9 について決定された c D N A 配列である。

配列番号 1 0 1 は、クローン # 1 8 4 2 2 について決定された c D N A 配列である。

30

配列番号 1 0 2 は、クローン # 1 8 4 2 5 について決定された c D N A 配列である。

配列番号 1 0 3 は、クローン # 1 8 4 3 1 について決定された c D N A 配列である。

配列番号 1 0 4 は、クローン # 1 8 4 3 3 について決定された c D N A 配列である。

配列番号 1 0 5 は、クローン # 1 8 4 4 4 について決定された c D N A 配列である。

配列番号 1 0 6 は、クローン # 1 8 4 4 9 について決定された c D N A 配列である。

配列番号 1 0 7 は、クローン # 1 8 4 5 1 について決定された c D N A 配列である。

配列番号 1 0 8 は、クローン # 1 8 4 5 2 について決定された c D N A 配列である。

配列番号 1 0 9 は、クローン # 1 8 4 5 5 について決定された c D N A 配列である。

配列番号 1 1 0 は、クローン # 1 8 4 5 7 について決定された c D N A 配列である。

配列番号 1 1 1 は、クローン # 1 8 4 6 6 について決定された c D N A 配列である。

40

配列番号 1 1 2 は、クローン # 1 8 4 6 8 について決定された c D N A 配列である。

配列番号 1 1 3 は、クローン # 1 8 4 7 1 について決定された c D N A 配列である。

配列番号 1 1 4 は、クローン # 1 8 4 7 5 について決定された c D N A 配列である。

配列番号 1 1 5 は、クローン # 1 8 4 7 6 について決定された c D N A 配列である。

配列番号 1 1 6 は、クローン # 1 8 4 7 7 について決定された c D N A 配列である。

配列番号 1 1 7 は、クローン # 2 0 6 3 1 について決定された c D N A 配列である。

配列番号 1 1 8 は、クローン # 2 0 6 3 4 について決定された c D N A 配列である。

配列番号 1 1 9 は、クローン # 2 0 6 3 5 について決定された c D N A 配列である。

配列番号 1 2 0 は、クローン # 2 0 6 3 7 について決定された c D N A 配列である。

配列番号 1 2 1 は、クローン # 2 0 6 3 8 について決定された c D N A 配列である。

50

[illegible]

配列番号 166 は、クローン # 19149. 1 について決定された cDNA 配列である。
配列番号 167 は、クローン # 19152 について決定された cDNA 配列である。
配列番号 168 は、クローン # 19153. 1 について決定された第 1 の cDNA 配列である。
配列番号 169 は、クローン # 19153. 2 について決定された第 2 の cDNA 配列である。
配列番号 170 は、クローン # 19155 について決定された cDNA 配列である。
配列番号 171 は、クローン # 19157 について決定された cDNA 配列である。
配列番号 172 は、クローン # 19159 について決定された cDNA 配列である。
配列番号 173 は、クローン # 19160 について決定された cDNA 配列である。 10
配列番号 174 は、クローン # 19161. 1 について決定された第 1 の cDNA 配列である。
配列番号 175 は、クローン # 19161. 2 について決定された第 2 の cDNA 配列である。
配列番号 176 は、クローン # 19162. 1 について決定された cDNA 配列である。
配列番号 177 は、クローン # 19166 について決定された cDNA 配列である。
配列番号 178 は、クローン # 19169 について決定された cDNA 配列である。
配列番号 179 は、クローン # 19171 について決定された cDNA 配列である。
配列番号 180 は、クローン # 19173. 1 について決定された第 1 の cDNA 配列である。 20
配列番号 181 は、クローン # 19173. 2 について決定された第 2 の cDNA 配列である。
配列番号 182 は、クローン # 19174. 1 について決定された cDNA 配列である。
配列番号 183 は、クローン # 19175 について決定された cDNA 配列である。
配列番号 184 は、クローン # 19177 について決定された cDNA 配列である。
配列番号 185 は、クローン # 19178 について決定された cDNA 配列である。
配列番号 186 は、クローン # 19179. 1 について決定された cDNA 配列である。
配列番号 187 は、クローン # 19179. 2 について決定された cDNA 配列である。
配列番号 188 は、クローン # 19180 について決定された cDNA 配列である。
配列番号 189 は、クローン # 19182. 1 について決定された第 1 の cDNA 配列である。 30
配列番号 190 は、クローン # 19182. 2 について決定された第 2 の cDNA 配列である。
配列番号 191 は、クローン # 19183. 1 について決定された cDNA 配列である。
配列番号 192 は、クローン # 19185. 1 について決定された cDNA 配列である。
配列番号 193 は、クローン # 19187 について決定された cDNA 配列である。
配列番号 194 は、クローン # 19188 について決定された cDNA 配列である。
配列番号 195 は、クローン # 19190 について決定された cDNA 配列である。
配列番号 196 は、クローン # 19191 について決定された cDNA 配列である。
配列番号 197 は、クローン # 19192 について決定された cDNA 配列である。 40
配列番号 198 は、クローン # 19193 について決定された cDNA 配列である。
配列番号 199 は、クローン # 19194. 1 について決定された第 1 の cDNA 配列である。
配列番号 200 は、クローン # 19194. 2 について決定された第 2 の cDNA 配列である。
配列番号 201 は、クローン # 19197 について決定された cDNA 配列である。
配列番号 202 は、クローン # 19200. 1 について決定された第 1 の cDNA 配列である。
配列番号 203 は、クローン # 19200. 2 について決定された第 2 の cDNA 配列である。 50

配列番号 204 は、クローン # 19202 について決定された cDNA 配列である。

配列番号 205 は、クローン # 19204. 1 について決定された第 1 の cDNA 配列である。

配列番号 206 は、クローン # 19204. 2 について決定された第 2 の cDNA 配列である。

配列番号 207 は、クローン # 19205 について決定された cDNA 配列である。

配列番号 208 は、クローン # 19206. 1 について決定された第 1 の cDNA 配列である。

配列番号 209 は、クローン # 19206. 2 について決定された第 2 の cDNA 配列である。

配列番号 210 は、クローン # 19207 について決定された cDNA 配列である。

配列番号 211 は、クローン # 19208 について決定された cDNA 配列である。

配列番号 212 は、クローン # 19211. 1 について決定された第 1 の cDNA 配列である。

配列番号 213 は、クローン # 19211. 2 について決定された第 2 の cDNA 配列である。

配列番号 214 は、クローン # 19214. 1 について決定された第 1 の cDNA 配列である。

配列番号 215 は、クローン # 19214. 2 について決定された第 2 の cDNA 配列である。

配列番号 216 は、クローン # 19215 について決定された cDNA 配列である。

配列番号 217 は、クローン # 19217. 2 について決定された第 1 の cDNA 配列である。

配列番号 218 は、クローン # 19217. 2 について決定された第 2 の cDNA 配列である。

配列番号 219 は、クローン # 19218. 1 について決定された第 1 の cDNA 配列である。

配列番号 220 は、クローン # 19218. 2 について決定された第 2 の cDNA 配列である。

配列番号 221 は、クローン # 19220. 1 について決定された第 1 の cDNA 配列である。

配列番号 222 は、クローン # 19220. 2 について決定された第 2 の cDNA 配列である。

配列番号 223 は、クローン # 22015 について決定された cDNA 配列である。

配列番号 224 は、クローン # 22017 について決定された cDNA 配列である。

配列番号 225 は、クローン # 22019 について決定された cDNA 配列である。

配列番号 226 は、クローン # 22020 について決定された cDNA 配列である。

配列番号 227 は、クローン # 22023 について決定された cDNA 配列である。

配列番号 228 は、クローン # 22026 について決定された cDNA 配列である。

配列番号 229 は、クローン # 22027 について決定された cDNA 配列である。

配列番号 230 は、クローン # 22028 について決定された cDNA 配列である。

配列番号 231 は、クローン # 22032 について決定された cDNA 配列である。

配列番号 232 は、クローン # 22037 について決定された cDNA 配列である。

配列番号 233 は、クローン # 22045 について決定された cDNA 配列である。

配列番号 234 は、クローン # 22048 について決定された cDNA 配列である。

配列番号 235 は、クローン # 22050 について決定された cDNA 配列である。

配列番号 236 は、クローン # 22052 について決定された cDNA 配列である。

配列番号 237 は、クローン # 22053 について決定された cDNA 配列である。

配列番号 238 は、クローン # 22057 について決定された cDNA 配列である。

配列番号 239 は、クローン # 22066 について決定された cDNA 配列である。

50

配列番号 290 は、クローン # 26964 について決定された cDNA 配列である。
配列番号 291 は、クローン # 26966 について決定された cDNA 配列である。
配列番号 292 は、クローン # 26968 について決定された cDNA 配列である。
配列番号 293 は、クローン # 26972 について決定された cDNA 配列である。
配列番号 294 は、クローン # 26973 について決定された cDNA 配列である。
配列番号 295 は、クローン # 26974 について決定された cDNA 配列である。
配列番号 296 は、クローン # 26976 について決定された cDNA 配列である。
配列番号 297 は、クローン # 26977 について決定された cDNA 配列である。
配列番号 298 は、クローン # 26979 について決定された cDNA 配列である。
配列番号 299 は、クローン # 26980 について決定された cDNA 配列である。 10
配列番号 300 は、クローン # 26981 について決定された cDNA 配列である。
配列番号 301 は、クローン # 26984 について決定された cDNA 配列である。
配列番号 302 は、クローン # 26985 について決定された cDNA 配列である。
配列番号 303 は、クローン # 26986 について決定された cDNA 配列である。
配列番号 304 は、クローン # 26993 について決定された cDNA 配列である。
配列番号 305 は、クローン # 26994 について決定された cDNA 配列である。
配列番号 306 は、クローン # 26995 について決定された cDNA 配列である。
配列番号 307 は、クローン # 27003 について決定された cDNA 配列である。
配列番号 308 は、クローン # 27005 について決定された cDNA 配列である。
配列番号 309 は、クローン # 27010 について決定された cDNA 配列である。 20
配列番号 310 は、クローン # 27011 について決定された cDNA 配列である。
配列番号 311 は、クローン # 27013 について決定された cDNA 配列である。
配列番号 312 は、クローン # 27016 について決定された cDNA 配列である。
配列番号 313 は、クローン # 27017 について決定された cDNA 配列である。
配列番号 314 は、クローン # 27019 について決定された cDNA 配列である。
配列番号 315 は、クローン # 27028 について決定された cDNA 配列である。
配列番号 316 は、クローン # 19060 についての全長 cDNA 配列である。
配列番号 317 は、クローン # 18964 についての全長 cDNA 配列である。
配列番号 318 は、クローン # 18929 についての全長 cDNA 配列である。
配列番号 319 は、クローン # 18991 についての全長 cDNA 配列である。 30
配列番号 320 は、クローン # 18996 についての全長 cDNA 配列である。
配列番号 321 は、クローン # 18966 についての全長 cDNA 配列である。
配列番号 322 は、クローン # 18951 についての全長 cDNA 配列である。
配列番号 323 は、クローン # 18973 (また L516S としても知られている) についての全長 cDNA 配列である。
配列番号 324 は、クローン # 19060 についてのアミノ酸配列である。
配列番号 325 は、クローン # 19063 についてのアミノ酸配列である。
配列番号 326 は、クローン # 19077 についてのアミノ酸配列である。
配列番号 327 は、クローン # 19110 についてのアミノ酸配列である。
配列番号 328 は、クローン # 19122 についてのアミノ酸配列である。 40
配列番号 329 は、クローン # 19118 についてのアミノ酸配列である。
配列番号 330 は、クローン # 19080 についてのアミノ酸配列である。
配列番号 331 は、クローン # 19127 についてのアミノ酸配列である。
配列番号 332 は、クローン # 19117 についてのアミノ酸配列である。
配列番号 333 は、クローン # 19095 (L549S ともいわれる) についてのアミノ酸配列である。
配列番号 334 は、クローン # 18964 についてのアミノ酸配列である。
配列番号 335 は、クローン # 18929 についてのアミノ酸配列である。
配列番号 336 は、クローン # 18991 についてのアミノ酸配列である。
配列番号 337 は、クローン # 18996 についてのアミノ酸配列である。 50

50

50

50

50

50

50

50

50

50

配列番号 788 は、L548S についての全長 cDNA 配列である。

配列番号 789 は、配列番号 788 によってコードされるアミノ酸配列である。

配列番号 790 は、L552S についての伸長 cDNA 配列である。

配列番号 791 は、配列番号 790 の cDNA 配列によってコードされる推定アミノ酸配列である。

配列番号 792 は、L552S のアイソフォームについて決定された cDNA 配列である。

配列番号 793 は、配列番号 792 によってコードされる推定アミノ酸配列である。

配列番号 794 は、L840P の伸長 cDNA 配列である。

配列番号 795 は、配列番号 794 によってコードされる推定アミノ酸配列である。

10

配列番号 796 は、L801P についての伸長 cDNA 配列である。

配列番号 797 は、配列番号 796 によってコードされる第 1 の推定アミノ酸配列である。

配列番号 798 は、配列番号 796 によってコードされる第 2 の推定アミノ酸配列である。

配列番号 799 は、配列番号 796 によってコードされる第 3 の推定アミノ酸配列である。

配列番号 800 は、L844P について決定された全長配列である。

配列番号 801 は、L551S についての 5' コンセンサス cDNA 配列である。

配列番号 802 は、L551S についての 3' コンセンサス cDNA 配列である。

20

配列番号 803 は、STY8 についての cDNA 配列である。

配列番号 804 は、L551S についての伸長 cDNA 配列である。

配列番号 805 は、STY8 についてのアミノ酸配列である。

配列番号 806 は、L551S についての伸長アミノ酸配列である。

配列番号 807 は、L773P について決定された全長 cDNA 配列である。

配列番号 808 は、L552S についての全長 cDNA 配列である。

配列番号 809 は、L552S についての全長アミノ酸配列である。

配列番号 810 は、クローン 50989 の決定された cDNA 配列である。

配列番号 811 は、クローン 50990 の決定された cDNA 配列である。

配列番号 812 は、クローン 50992 の決定された cDNA 配列である。

30

配列番号 813～824 は、肺腫瘍組織から単離されたクローンについて決定された cDNA 配列である。

配列番号 825 は、全長 L551S クローン 54305 について決定された cDNA 配列である。

配列番号 826 は、全長 L551S クローン 54298 について決定された cDNA 配列である。

配列番号 827 は、L551S についての全長アミノ酸配列である。

表 1～6 は、配列番号 828～1664 についての配列識別子を含む。

【0034】

【表 1】

40

表 1A

配列番号	クローン識別子	配列番号	クローン識別子
828	R0126:A02	869	R0126:D12
829	R0126:A03	870	R0126:E01
830	R0126:A05	871	R0126:E02
831	R0126:A06	872	R0126:E03
832	R0126:A08	873	R0126:E04
833	R0126:A09	874	R0126:E05
834	R0126:A10	875	R0126:E06
835	R0126:A11	876	R0126:E07
836	R0126:A12	877	R0126:E08
837	R0126:B01	878	R0126:E09
838	R0126:B03	879	R0126:E10
839	R0126:B04	880	R0126:E11
840	R0126:B05	881	R0126:E12
841	R0126:B06	882	R0126:F01
842	R0126:B07	883	R0126:F02
843	R0126:B08	884	R0126:F03
844	R0126:B09	885	R0126:F04
845	R0126:B11	886	R0126:F05
846	R0126:B12	887	R0126:F06
847	R0126:C01	888	R0126:F07
848	R0126:C02	889	R0126:F08
849	R0126:C03	890	R0126:F10
850	R0126:C05	891	R0126:F11
851	R0126:C06	892	R0126:F12
852	R0126:C07	893	R0126:G01
853	R0126:C08	894	R0126:G02
854	R0126:C09	895	R0126:G03
855	R0126:C10	896	R0126:G04
856	R0126:C11	897	R0126:G05
857	R0126:C12	898	R0126:G06
858	R0126:D01	899	R0126:G07
859	R0126:D02	900	R0126:G09
860	R0126:D03	901	R0126:G10
861	R0126:D04	902	R0126:G11
862	R0126:D05	903	R0126:G12
863	R0126:D06	904	R0126:H01
864	R0126:D07	905	R0126:H02

配列番号	クローン識別子	配列番号	クローン識別子
865	R0126:D08	906	R0126:H03
866	R0126:D09	907	R0126:H04
867	R0126:D10	908	R0126:H05
868	R0126:D11	909	R0126:H06

【 0 0 3 5 】

【 表 2 】

表 1B

配列番号	クローン識別子	配列番号	クローン識別子
910	R0126:H07	951	R0127:D10
911	R0126:H09	952	R0127:D11
912	R0126:H10	953	R0127:D12
913	R0126:H11	954	R0127:E02
914	R0127:A02	955	R0127:E03
915	R0127:A05	956	R0127:E04
916	R0127:A06	957	R0127:E05
917	R0127:A07	958	R0127:E06
918	R0127:A08	959	R0127:E07
919	R0127:A09	960	R0127:E08
920	R0127:A10	961	R0127:E09
921	R0127:A11	962	R0127:E10
922	R0127:A12	963	R0127:E11
923	R0127:B01	964	R0127:F01
924	R0127:B03	965	R0127:F02
925	R0127:B04	966	R0127:F03
926	R0127:B05	967	R0127:F04
927	R0127:B06	968	R0127:F05
928	R0127:B07	969	R0127:F06
929	R0127:B08	970	R0127:F07
930	R0127:B09	971	R0127:F08
931	R0127:B10	972	R0127:F10
932	R0127:B11	973	R0127:F11
933	R0127:B12	974	R0127:F12
934	R0127:C01	975	R0127:G01
935	R0127:C03	976	R0127:G02
936	R0127:C04	977	R0127:G03
937	R0127:C05	978	R0127:G04
938	R0127:C07	979	R0127:G05
939	R0127:C08	980	R0127:G06
940	R0127:C09	981	R0127:G07
941	R0127:C10	982	R0127:G08
942	R0127:C11	983	R0127:G09
943	R0127:D01	984	R0127:G10
944	R0127:D02	985	R0127:G11
945	R0127:D03	986	R0127:G12
946	R0127:D04	987	R0127:H01
947	R0127:D05	988	R0127:H02
948	R0127:D06	989	R0127:H03
949	R0127:D07	990	R0127:H04
950	R0127:D01	991	R0127:H05

10

20

30

40

【 0 0 3 6 】

【 表 3 】

表 1C

配列番号	クローン識別子	配列番号	クローン識別子
992	R0127:H06	1034	R0128:D11
993	R0127:H07	1035	R0128:D12
994	R0127:H08	1036	R0128:E01
995	R1027:H09	1037	R0128:E02
996	R1027:H10	1038	R0128:E03
997	R1027:H11	1039	R0128:E04
998	R1028:A02	1040	R0128:E05
999	R1028:A05	1041	R0128:E06
1000	R1028:A06	1042	R0128:E07
1001	R1028:A07	1043	R0128:E08
1002	R1028:A08	1044	R0128:E09
1003	R1028:A09	1045	R0128:E10
1004	R1028:A10	1046	R0128:E12
1005	R1028:B01	1047	R0128:F01
1006	R1028:B02	1048	R0128:F02
1007	R1028:B03	1049	R0128:F03
1008	R1028:B04	1050	R0128:F04
1009	R1028:B05	1051	R0128:F06
1010	R1028:B08	1052	R0128:F07
1011	R1028:B09	1053	R0128:F08
1012	R1028:B10	1054	R0128:F09
1013	R1028:B11	1055	R0128:F10
1014	R1028:B12	1056	R0128:F12
1015	R1028:C01	1057	R0128:G01
1016	R1028:C03	1058	R0128:G02
1017	R1028:C04	1059	R0128:G03
1018	R1028:C05	1060	R0128:G04
1019	R1028:C06	1061	R0128:G05
1020	R1028:C07	1062	R0128:G06
1021	R1028:C08	1063	R0128:G07
1022	R1028:C10	1064	R0128:G09
1023	R1028:C11	1065	R0128:G10
1024	R1028:C12	1066	R0128:G11
1025	R1028:D01	1067	R0128:G12
1026	R1028:D02	1068	R0128:H01
1027	R1028:D04	1069	R0128:H02
1028	R1028:D05	1070	R0128:H03
1029	R1028:D06	1071	R0128:H04
1030	R1028:D07	1072	R0128:H05
1031	R1028:D08	1073	R0128:H06
1032	R1028:D09	1074	R0128:H07
1033	R0128:D10	1075	R0128:H08

10

20

30

40

【 0 0 3 7 】

【 表 4 】

表 1D

配列番号	クローン識別子	配列番号	クローン識別子
1076	R0128:H09	1118	R0130:D12
1077	R0128:H10	1119	R0130:E01
1078	R0128:H11	1120	R0130:E02
1079	R0130:A02	1121	R0130:E03
1080	R0130:A05	1122	R0130:E04
1081	R0130:A06	1123	R0130:E05
1082	R0130:A08	1124	R0130:E06
1083	R0130:A09	1125	R0130:E07
1084	R0130:A10	1126	R0130:E08
1085	R0130:A11	1127	R0130:E09
1086	R0130:A12	1128	R0130:E10
1087	R0130:B01	1129	R0130:E11
1088	R0130:B02	1130	R0130:E12
1089	R0130:B03	1131	R0130:F02
1090	R0130:B04	1132	R0130:F03
1091	R0130:B05	1133	R0130:F05
1092	R0130:B06	1134	R0130:F06
1093	R0130:B08	1135	R0130:F07
1094	R0130:B09	1136	R0130:F08
1095	R0130:B10	1137	R0130:F09
1096	R0130:B11	1138	R0130:F10
1097	R0130:B12	1139	R0130:F11
1098	R0130:C02	1140	R0130:F12
1099	R0130:C03	1141	R0130:G01
1100	R0130:C04	1142	R0130:G02
1101	R0130:C05	1143	R0130:G03
1102	R0130:C06	1144	R0130:G04
1103	R0130:C07	1145	R0130:G05
1104	R0130:C08	1146	R0130:G06
1105	R0130:C09	1147	R0130:G07
1106	R0130:C10	1148	R0130:G08
1107	R0130:C11	1149	R0130:G09
1108	R0130:C12	1150	R0130:G10
1109	R0130:D02	1151	R0130:G11
1110	R0130:D03	1152	R0130:G12
1111	R0130:D04	1153	R0130:H01
1112	R0130:D05	1154	R0130:H02
1113	R0130:D06	1155	R0130:H04
1114	R0130:D07	1156	R0130:H05
1115	R0130:D09	1157	R0130:H06
1116	R0130:D10	1158	R0130:H07
1117	R0130:D11	1159	R0130:H08

10

20

30

40

【 0 0 3 8 】

【 表 5 】

表 1E

配列番号	クローン識別子	配列番号	クローン識別子
1160	R0130:H09	1200	R0131:E01
1161	R0130:H10	1201	R0131:E02
1162	R0130:H11	1202	R0131:E03
1163	R0131:A02	1203	R0131:E04
1164	R0131:A05	1204	R0131:E06
1165	R0131:A06	1205	R0131:E07
1166	R0131:A07	1206	R0131:E08
1167	R0131:A08	1207	R0131:E10
1168	R0131:A09	1208	R0131:E11
1169	R0131:A11	1209	R0131:E12
1170	R0131:A12	1210	R0131:F02
1171	R0131:B02	1211	R0131:F04
1172	R0131:B03	1212	R0131:F05
1173	R0131:B04	1213	R0131:F06
1174	R0131:B05	1214	R0131:F07
1175	R0131:B07	1215	R0131:F08
1176	R0131:B08	1216	R0131:F09
1177	R0131:B09	1217	R0131:F10
1178	R0131:B10	1218	R0131:F11
1179	R0131:B11	1219	R0131:F12
1180	R0131:C01	1220	R0131:G01
1181	R0131:C02	1221	R0131:G02
1182	R0131:C03	1222	R0131:G03
1183	R0131:C04	1223	R0131:G04
1184	R0131:C06	1224	R0131:G05
1185	R0131:C07	1225	R0131:G06
1186	R0131:C08	1226	R0131:G07
1187	R0131:C10	1227	R0131:G08
1188	R0131:C11	1228	R0131:G09
1189	R0131:C12	1229	R0131:G10
1190	R0131:D02	1230	R0131:G11
1191	R0131:D03	1231	R0131:G12
1192	R0131:D04	1232	R0131:H01
1193	R0131:D05	1233	R0131:H02
1194	R0131:D06	1234	R0131:H05
1195	R0131:D07	1235	R0131:H06
1196	R0131:D09	1236	R0131:H07
1197	R0131:D10	1237	R0131:H08
1198	R0131:D11	1238	R0131:H09
1199	R0131:D12	1239	R0131:H11

10

20

30

40

【 0 0 3 9 】

【 表 6 】

表 2:

NSCLC-SQL1 についてのクローン名および対応する配列番号

配列番号	クローン識別子
1240	コンティグ 54
1241	コンティグ 55
1242	コンティグ 57
1243	コンティグ 58
1244	コンティグ 60
1245	コンティグ 62
1246	コンティグ 63
1247	コンティグ 64
1248	コンティグ 65
1249	コンティグ 66
1250	コンティグ 67
1251	コンティグ 68
1252	コンティグ 69
1253	コンティグ 70
1254	コンティグ 71
1255	コンティグ 72
1256	コンティグ 73
1257	コンティグ 74
1258	コンティグ 75
1259	コンティグ 77
1260	コンティグ 78
1261	コンティグ 79
1262	コンティグ 80
1263	コンティグ 81
1264	コンティグ 83
1265	コンティグ 84
1266	コンティグ 86
1267	コンティグ 87
1268	コンティグ 88
1269	コンティグ 89
1270	コンティグ 90
1271	コンティグ 91
1272	コンティグ 92
1273	コンティグ 94
1274	コンティグ 95
1275	コンティグ 96
1276	コンティグ 97
1277	コンティグ 98
1278	コンティグ 99
1279	コンティグ100

10

20

30

40

【 0 0 4 0 】

【 表 7 】

表 3:

NSCLC-SCLI についてのクローン名および対応する配列番号

配列番号	クローン識別子
1280	コンティグ38
1281	コンティグ39
1282	コンティグ40
1283	コンティグ41
1284	コンティグ42
1285	コンティグ43
1286	コンティグ44
1287	コンティグ45
1288	コンティグ46
1289	コンティグ47
1290	コンティグ48
1291	コンティグ49
1292	コンティグ51
1293	コンティグ52
1294	コンティグ53
1295	コンティグ54
1296	コンティグ55
1297	コンティグ56
1298	コンティグ57
1299	コンティグ58
1300	コンティグ59
1301	コンティグ60
1302	コンティグ62
1303	コンティグ63
1304	コンティグ64
1305	コンティグ65
1306	コンティグ66
1307	コンティグ67
1308	コンティグ68
1309	コンティグ69
1310	コンティグ70
1311	コンティグ72
1312	コンティグ73
1313	コンティグ75
1314	コンティグ76
1315	コンティグ77
1316	コンティグ78
1317	コンティグ79
1318	コンティグ80
1319	コンティグ81
1320	コンティグ82

10

20

30

40

【0041】

【表8】

表 4A:
NSCLC-SCL3-SCL4 についてのクローン名および対応する配列番号

配列番号	クローン識別子	配列番号	クローン識別子
1321	コンティグ94	1363	コンティグ136
1322	コンティグ95	1364	コンティグ137
1323	コンティグ96	1365	コンティグ138
1324	コンティグ97	1366	コンティグ139
1325	コンティグ98	1367	コンティグ140
1326	コンティグ99	1368	コンティグ141
1327	コンティグ100	1369	コンティグ142
1328	コンティグ101	1370	コンティグ143
1329	コンティグ102	1371	コンティグ144
1330	コンティグ103	1372	コンティグ145
1331	コンティグ104	1373	コンティグ146
1332	コンティグ105	1374	コンティグ147
1333	コンティグ106	1375	コンティグ148
1334	コンティグ107	1376	コンティグ149
1335	コンティグ108	1377	コンティグ150
1336	コンティグ109	1378	コンティグ151
1337	コンティグ110	1379	コンティグ152
1338	コンティグ111	1380	コンティグ153
1339	コンティグ112	1381	コンティグ154
1340	コンティグ113	1382	コンティグ155
1341	コンティグ114	1383	コンティグ156
1342	コンティグ115	1384	コンティグ157
1343	コンティグ116	1385	コンティグ158
1344	コンティグ117	1386	コンティグ159
1345	コンティグ118	1387	コンティグ160
1346	コンティグ119	1388	コンティグ161
1347	コンティグ120	1389	コンティグ162
1348	コンティグ121	1390	コンティグ163
1349	コンティグ122	1391	コンティグ164
1350	コンティグ123	1392	コンティグ165
1351	コンティグ124	1393	コンティグ166
1352	コンティグ125	1394	コンティグ167
1353	コンティグ126	1395	コンティグ168
1354	コンティグ127	1396	コンティグ169
1355	コンティグ128	1397	コンティグ170
1356	コンティグ129	1398	コンティグ171
1357	コンティグ130	1399	コンティグ172
1358	コンティグ131	1400	コンティグ173
1359	コンティグ132	1401	コンティグ174
1360	コンティグ133	1402	コンティグ175
1361	コンティグ134	1403	コンティグ176

10

20

30

40

配列番号	クローン識別子	配列番号	クローン識別子
1362	コンティグ135		

【 0 0 4 2 】

【 表 9 】

50

表 4B:
NSCLC-SCL3-SCL4 についてのクローン名および対応する配列番号

配列番号	クローン識別子
1404	コンティグ 177
1405	コンティグ 178
1406	コンティグ 179
1407	コンティグ 180
1408	コンティグ 181
1409	コンティグ 182
1410	コンティグ 183
1411	コンティグ 184
1412	コンティグ 185
1413	コンティグ 186
1414	コンティグ 187

10

【 0 0 4 3 】

【 表 1 0 】

20

表 5:
SCLC-SQL1 についてのクローン名および対応する配列番号

配列番号	クローン識別子
1415	コンティグ 17
1416	コンティグ 18
1417	コンティグ 20
1418	コンティグ 23
1419	コンティグ 24
1420	コンティグ 25
1421	コンティグ 26
1422	コンティグ 27
1423	コンティグ 28
1424	コンティグ 29
1425	コンティグ 30
1426	コンティグ 31
1427	コンティグ 20
1428	コンティグ 21
1429	コンティグ 22
1430	コンティグ 23
1431	コンティグ 24
1432	コンティグ 25
1433	コンティグ 26
1434	コンティグ 27
1435	コンティグ 28
1436	コンティグ 29
1437	コンティグ 30
1438	コンティグ 31
1439	コンティグ 32
1440	コンティグ 33
1441	コンティグ 34
1442	コンティグ 35
1443	コンティグ 36
1444	コンティグ 37
1445	コンティグ 38

10

20

30

40

【 0 0 4 4 】

【 表 1 1 】

表 6A:

SCLC-SCL3-SCL4 についてのクローン名および対応する配列番号

配列番号	クローン識別子	配列番号	クローン識別子
1446	コンティグ 116	1488	コンティグ160
1447	コンティグ 117	1489	コンティグ161
1448	コンティグ 118	1490	コンティグ162
1449	コンティグ 119	1491	コンティグ163
1450	コンティグ 120	1492	コンティグ164
1451	コンティグ 122	1493	コンティグ165
1452	コンティグ 123	1494	コンティグ166
1453	コンティグ 124	1495	コンティグ167
1454	コンティグ125	1496	コンティグ168
1455	コンティグ126	1497	コンティグ169
1456	コンティグ127	1498	コンティグ170
1457	コンティグ128	1499	コンティグ171
1458	コンティグ129	1500	コンティグ172
1459	コンティグ130	1501	コンティグ173
1460	コンティグ131	1502	コンティグ174
1461	コンティグ132	1503	コンティグ175
1462	コンティグ133	1504	コンティグ176
1463	コンティグ135	1505	コンティグ177
1464	コンティグ136	1506	コンティグ178
1465	コンティグ137	1507	コンティグ179
1466	コンティグ138	1508	コンティグ181
1467	コンティグ139(L985P)	1509	コンティグ182
1468	コンティグ140	1510	コンティグ183
1469	コンティグ141	1511	コンティグ184
1470	コンティグ142	1512	コンティグ185
1471	コンティグ143	1513	コンティグ186
1472	コンティグ144	1514	コンティグ187
1473	コンティグ145	1515	コンティグ189
1474	コンティグ146	1516	コンティグ190
1475	コンティグ147	1517	コンティグ191
1476	コンティグ148	1518	コンティグ192
1477	コンティグ149	1519	コンティグ193
1478	コンティグ150	1520	コンティグ194
1479	コンティグ151	1521	コンティグ195
1480	コンティグ152	1522	コンティグ196
1481	コンティグ153	1523	コンティグ197
1482	コンティグ154	1524	コンティグ198
1483	コンティグ155	1525	コンティグ199
1484	コンティグ156	1526	コンティグ200
1485	コンティグ157	1527	コンティグ201
1486	コンティグ158	1528	コンティグ202

配列番号	クローン識別子	配列番号	クローン識別子
1487	コンティグ 159		

【 0 0 4 5 】

【 表 1 2 】

10

20

30

40

50

表 6B:

SCLC-SCL3-SCL4についてのクローン名および対応する配列番号

配列番号	クローン識別子
1529	コンティグ 203
1530	コンティグ 204
1531	コンティグ 205
1532	コンティグ 206
1533	コンティグ 207
1534	コンティグ 208
1535	コンティグ 209
1536	コンティグ 210
1537	コンティグ 211
1538	コンティグ 212
1539	コンティグ 213
1540	コンティグ 214
1541	コンティグ 215
1542	コンティグ 216
1543	コンティグ 217
1544	コンティグ 218
1545	コンティグ 219
1546	コンティグ 220
1547	コンティグ 221
1548	コンティグ 222
1549	コンティグ 223
1550	コンティグ 224
1551	コンティグ 225
1552	コンティグ 226
1553	コンティグ 227
1554	コンティグ 228
1555	コンティグ 229
1556	コンティグ 230
1557	コンティグ 231
1558	コンティグ 232
1559	コンティグ 233
1560	コンティグ 234
1561	コンティグ 235
1562	コンティグ 236
1563	コンティグ 237

10

20

30

40

【 0 0 4 6 】

【 表 1 3 】

表 7.

配列番号	クローン識別子	配列番号	クローン識別子
1564	R0124:E05	1609	R0129:D09
1565	R0124:E06	1610	R0129:D10
1566	R0124:E08	1611	R0129:D11
1567	R0124:F07	1612	R0129:E02
1568	R0124:F08	1613	R0129:E03
1569	R0124:F09	1614	R0129:E04
1570	R0124:G04	1615	R0129:E05
1571	R0129:A02	1616	R0129:E06
1572	R0129:A03	1617	R0129:E07
1573	R0129:A06	1618	R0129:E08
1574	R0129:A07	1619	R0129:E09
1575	R0129:A08	1620	R0129:E11
1576	R0129:A09	1621	R0129:E12
1577	R0129:A10	1622	R0129:F01
1578	R0129:A11	1623	R0129:F02
1579	R0129:A12	1624	R0129:F03
1580	R0129:B02	1625	R0129:F04
1581	R0129:B03	1626	R0129:F06
1582	R0129:B04	1627	R0129:F07
1583	R0129:B05	1628	R0129:F08
1584	R0129:B06	1629	R0129:F09
1585	R0129:B07	1630	R0129:F10
1586	R0129:B08	1631	R0129:F11
1587	R0129:B09	1632	R0129:F12
1588	R0129:B10	1633	R0129:G01
1589	R0129:B11	1634	R0129:G02
1590	R0129:B12	1635	R0129:G03
1591	R0129:C01	1636	R0129:G04
1592	R0129:C02	1637	R0129:G05
1593	R0129:C03	1638	R0129:G06
1594	R0129:C04	1639	R0129:G07
1595	R0129:C06	1640	R0129:G08
1596	R0129:C07	1641	R0129:G09
1597	R0129:C08	1642	R0129:G10
1598	R0129:C09	1643	R0129:G11
1599	R0129:C10	1644	R0129:G12
1600	R0129:C11	1645	R0129:H01
1601	R0129:C12	1646	R0129:H02
1602	R0129:D01	1647	R0129:H03
1603	R0129:D03	1648	R0129:H04
1604	R0129:D04	1649	R0129:H05
1605	R0129:D05	1650	R0129:H08
1606	R0129:D06	1651	R0129:H09
1607	R0129:D07	1652	R0129:H10
1608	R0129:D08	1653	R0129:H11

10

20

30

40

【 0 0 4 7 】

【 表 1 4 】

表 8

配列番号	クローン識別子
1654	26484
1655	26496
1656	26517
1657	26531
1658	26022
1659	26026
1660	26810
1661	26815
1662	26869
1663	26883
1664	26902

10

配列番号 1665 および 1666 は、L548S のコード領域の増幅の際に使用されるプライマーである。

配列番号 1667 は、発現した組換え L7548S のタンパク質配列である。

配列番号 1668 は、発現した組換え L7548S の cDNA 配列である。

20

配列番号 1669 は、クローン #18971 (L801P) の伸長 cDNA 配列である。

配列番号 1670 は、配列番号 1669 によってコードされるオープンリーディングフレーム ORF4 のアミノ酸配列である。

配列番号 1671 は、配列番号 1669 によってコードされるオープンリーディングフレーム ORF5 のアミノ酸配列である。

配列番号 1672 は、配列番号 1669 によってコードされるオープンリーディングフレーム ORF6 のアミノ酸配列である。

配列番号 1673 は、配列番号 1669 によってコードされるオープンリーディングフレーム ORF7 のアミノ酸配列である。

配列番号 1674 は、配列番号 1669 によってコードされるオープンリーディングフレーム ORF8 のアミノ酸配列である。

30

配列番号 1675 は、配列番号 1669 によってコードされるオープンリーディングフレーム ORF9 のアミノ酸配列である。

配列番号 1676 は、コンティグ 139 (配列番号 1467) (L985P としても知られている) についての伸長 cDNA である。

配列番号 1677 は、配列番号 1676 によってコードされる L985P アミノ酸配列である。

配列番号 1678 は、配列番号 1669 のオープンリーディングフレーム ORF5X のアミノ酸配列である。

配列番号 1679 は、コンティグ 139 (配列番号 1467) についてのオープンリーディングフレームのアミノ酸配列である。

40

配列番号 1680～1788 は、表 9 に示され、肺チップ 5 上の S Q L 1、S C L 1、S C L 3 および S C L 4 ライブラリーのマイクロアレイ分析によって同定された cDNA クローンを表す。

【0048】

【表 15】

表 9

配列番号	クローン識別子
1680	58456
1681	58458
1682	58462
1683	58469
1684	58470
1685	58482
1686	58485
1687	58501
1688	58502
1689	58505
1690	58507
1691	58509
1692	58512
1693	58527
1694	58529
1695	58531
1696	58537
1697	58539
1698	58545
1699	59319
1700	59322
1701	59348
1702	59350
1703	59363
1704	59365
1705	59370
1706	59373
1707	59376
1708	61050
1709	61051
1710	61052
1711	61054
1712	61056
1713	61057
1714	61060
1715	61062
1716	61063
1717	61064
1718	61065
1719	61066
1720	61069

10

20

30

40

配列番号	クローン識別子
1721	61070
1722	61071
1723	61074
1724	61075
1725	61077
1726	61079
1727	61080
1728	61081
1729	61083
1730	61085
1731	61086
1732	61088
1733	61090
1734	61091
1735	61093
1736	61094
1737	61096
1738	61097
1739	61099
1740	61100
1741	61103
1742	61105
1743	61106
1744	61110
1745	61113
1746	61115
1747	61117
1748	61118
1749	61119
1750	61120
1751	61122
1752	61125
1753	61126
1754	61130
1755	61133
1756	61134
1757	61135
1758	61137
1759	61139
1760	61143
1761	61144
1762	61148
1763	61151

10

20

30

40

配列番号	クローン識別子
1764	61155
1765	61156
1766	61159
1767	61160
1768	61163
1769	61167
1770	61172
1771	61173
1772	61176
1773	61177
1774	61183
1775	61185
1776	61188
1777	61192
1778	61198
1779	61201
1780	61202
1781	61204
1782	61206
1783	61210
1784	61212
1785	61216
1786	61225
1787	61226
1788	61227

10

20

配列番号 1789 は、クローン # 47988 (L972P) の cDNA 配列である。 30

配列番号 1790 は、クローン # 48005 (L979P) の cDNA 配列である。

配列番号 1791 は、クローン # 48005 (L979P) についての伸長 cDNA 配列である。

配列番号 1792 は、クローン # 49826 (配列番号 1279 ; L980P) についての伸長 cDNA 配列である。

配列番号 1793 は、クローン # 20631 (配列番号 117 ; L973P) についての伸長 cDNA 配列である。

配列番号 1794 は、クローン # 20661 (配列番号 128 ; L974P) についての伸長 cDNA 配列である。

配列番号 1795 は、クローン # 50430 (配列番号 1442 ; L996P) についての伸長 cDNA 配列である。 40

配列番号 1796 は、クローン # 26961 (配列番号 288 ; L977P) についての伸長 cDNA 配列である。

配列番号 1797 は、クローン # 24928 (配列番号 1339 ; L978P) についての伸長 cDNA 配列である。

配列番号 1798 は、クローン # 50507 (配列番号 1446 ; L984P) についての伸長 cDNA 配列である。

配列番号 1799 は、クローン # 50645 (配列番号 1531 ; L988P) についての伸長 cDNA 配列である。

配列番号 1800 は、クローン # 50628 (配列番号 1533 ; L1423P) につい 50

ての伸長 c D N A 配列である。

配列番号 1 8 0 1 は、クローン # 5 0 5 6 0 (配列番号 1 5 2 7 ; L 9 8 7 P) についての伸長 c D N A 配列である。

配列番号 1 8 0 2 は、クローン # 2 7 6 9 9 (配列番号 4 6 8 ; L 9 9 8 P) についての伸長 c D N A 配列である。

配列番号 1 8 0 3 は、クローン # 5 9 3 0 3 (配列番号 9 4 9 ; L 1 4 2 5 P) についての伸長 c D N A 配列である。

配列番号 1 8 0 4 は、クローン # 5 9 3 1 4 (配列番号 1 1 5 6 ; L 1 4 2 6 P) についての伸長 c D N A 配列である。

配列番号 1 8 0 5 は、クローン # 5 9 2 9 8 (配列番号 9 2 1 ; L 1 4 2 7 P) についての伸長 c D N A 配列である。 10

配列番号 1 8 0 6 は、配列番号 1 7 9 1 によってコードされるアミノ酸配列である。

配列番号 1 8 0 7 は、配列番号 1 7 9 2 によってコードされるアミノ酸配列である。

配列番号 1 8 0 8 は、配列番号 1 7 9 3 によってコードされるアミノ酸配列である。

配列番号 1 8 0 9 は、配列番号 1 7 9 4 によってコードされるアミノ酸配列である。

配列番号 1 8 1 0 は、配列番号 1 7 9 5 によってコードされるアミノ酸配列である。

配列番号 1 8 1 1 は、配列番号 1 7 9 6 によってコードされるアミノ酸配列である。

配列番号 1 8 1 2 は、配列番号 1 7 9 7 によってコードされるアミノ酸配列である。

配列番号 1 8 1 3 は、配列番号 1 7 9 8 によってコードされるアミノ酸配列である。

配列番号 1 8 1 4 は、配列番号 1 7 9 9 によってコードされるアミノ酸配列である。 20

配列番号 1 8 1 5 は、配列番号 1 8 0 0 によってコードされるアミノ酸配列である。

配列番号 1 8 1 6 は、配列番号 1 5 2 7 (L 9 8 7 P) によってコードされるアミノ酸配列である。

配列番号 1 8 1 7 は、配列番号 1 8 2 3 によってコードされるアミノ酸配列である。

配列番号 1 8 1 8 は、配列番号 1 8 0 1 によってコードされるアミノ酸配列である。

配列番号 1 8 1 9 は、配列番号 1 8 0 2 によってコードされるアミノ酸配列である。

配列番号 1 8 2 0 は、配列番号 1 8 0 3 によってコードされるアミノ酸配列である。

配列番号 1 8 2 1 は、配列番号 1 8 0 4 によってコードされるアミノ酸配列である。

配列番号 1 8 2 2 は、配列番号 1 8 0 5 によってコードされるアミノ酸配列である。

配列番号 1 8 2 3 は、クローン # 5 0 5 6 0 (配列番号 1 5 2 7 ; L 9 8 7 P) についての伸長 c D N A 配列である。 30

配列番号 1 8 2 4 は、クローン L 8 7 2 P (配列番号 3 4) についての全長 c D N A 配列である。

配列番号 1 8 2 5 は、配列番号 1 8 2 4 によってコードされるアミノ酸配列である。

配列番号 1 8 2 6 は、L 5 5 2 S の N 末端部分をコードする c D N A 配列である。

配列番号 1 8 2 7 ~ 1 8 2 9 は、L 5 5 2 S の部分の c D N A 配列である。

配列番号 1 8 3 0 は、L 5 5 2 S の N 末端部分である。

配列番号 1 8 3 1 ~ 1 8 3 3 は、それぞれ、配列番号 1 8 2 7 ~ 1 8 2 9 によってコードされるアミノ酸配列である。

配列番号 1 8 3 4 ~ 1 8 5 6 は、L 5 4 8 S のペプチドのアミノ酸配列である。 40

配列番号 1 8 5 7 ~ 1 8 6 0 は、P C R プライマーである。

配列番号 1 8 6 1 は、P 8 0 1 P の R a 1 2 と O R F 4 との融合物について決定された D N A 配列である。

配列番号 1 8 6 2 は、P 8 0 1 P の R a 1 2 と O R F 5 との融合物について決定された D N A 配列である。

配列番号 1 8 6 3 は、P 8 0 1 P の R a 1 2 と O R F 4 との融合物のアミノ酸配列である。

配列番号 1 8 6 4 は、P 8 0 1 P の R a 1 2 と O R F 5 との融合物のアミノ酸配列である。

配列番号 1 8 6 5 は、クローン L 9 8 4 P _ (5 7 3 A) について決定された c D N A 配 50

列である。

配列番号 1866 は、クローン L984P__ (512A) について決定された cDNA 配列である。

配列番号 1867 は、クローン L984P__ (NCI-H128) について決定された cDNA 配列である。

配列番号 1868 は、クローン L984P__ (DMS-79) について決定された cDNA 配列である。

配列番号 1869 は、配列番号 1865 によってコードされるアミノ酸配列である。

配列番号 1870 は、配列番号 1866 によってコードされるアミノ酸配列である。

配列番号 1871 は、配列番号 1867 によってコードされるアミノ酸配列である。

配列番号 1872 は、配列番号 1868 によってコードされるアミノ酸配列である。

配列番号 1873 は、クローン L985P (配列番号 1467 において与えられる部分配列) についての全長 cDNA 配列である。

配列番号 1874 は、配列番号 1873 によってコードされる L985P についてのアミノ酸配列である。

配列番号 1875 は、Ra12 と L985P との融合物の推定および決定された cDNA 配列である。

配列番号 1876 は、Ra12 と配列番号 1875 によってコードされる L985P との融合物の推定アミノ酸配列である。

配列番号 1877 は、Ra12S と L985P との融合物の推定 cDNA 配列である。

配列番号 1878 は、Ra12S と配列番号 1877 によってコードされる L985P との融合物の推定アミノ酸配列である。

配列番号 1879 は、Ra12S と L985PE_x との融合物の推定 cDNA 配列である。

配列番号 1880 は、Ra12S と配列番号 1879 によってコードされる L985PE_x との融合物の推定アミノ酸配列である。

配列番号 1881 は、L985P の細胞外ループ 2 ペプチドについての推定 cDNA 配列である。

配列番号 1882 は、配列番号 1875 によってコードされる L985P の細胞外ループ 2 ペプチドについての推定アミノ酸配列である。

配列番号 1883 は、クローン #59316 (配列番号 1180 ; L1428P) についての伸長 cDNA 配列である。

配列番号 1884 は、配列番号 1883 によってコードされ、そして L1428P-ORF1 と命名された第 1 の推定アミノ酸配列である。

配列番号 1885 は、配列番号 1883 によってコードされ、そして L1428P-ORF2 と命名された第 2 の推定アミノ酸配列である。

配列番号 1886 は、配列番号 1883 によってコードされ、そして L1428P-ORF3 と命名された第 3 の推定アミノ酸配列である。

配列番号 1887 は、配列番号 1883 によってコードされ、そして L1428P-ORF4 と命名された第 4 の推定アミノ酸配列である。

配列番号 1888 は、配列番号 1883 によってコードされ、そして L1428P-ORF5 と命名された第 5 の推定アミノ酸配列である。

配列番号 1889 は、配列番号 1883 によってコードされ、そして L1428P-ORF6 と命名された第 6 の推定アミノ酸配列である。

配列番号 1890 は、配列番号 1883 によってコードされ、そして L1428P-ORF7 と命名された第 7 の推定アミノ酸配列である。

配列番号 1891 ~ 1900 は、表 17 に記載のデータベースヒットについてのヌクレオチド配列である。

配列番号 1901 ~ 1909 は、表 17 に記載のヌクレオチド配列によってコードされる推定アミノ酸配列である。

10

20

30

40

50

配列番号1910は、クローンL1437P（配列番号1896において与えられる部分配列）についての全長cDNAである。

配列番号1911は、クローンL548Sのコード領域についての順方向プライマーPDM-433である。

配列番号1912は、クローンL548Sのコード領域についての逆方向プライマーPDM-438である。

配列番号1913は、発現された組換えL548Sについてのアミノ酸配列である。

配列番号1914は、組換えL548SについてのDNAコード配列である。

配列番号1915は、クローンL551Sのコード領域についての順方向プライマーPDM-498である。

10

配列番号1916は、クローンL551Sのコード領域についての逆方向プライマーPDM-499である。

配列番号1917は、発現された組換えL551Sについてのアミノ酸配列である。

配列番号1918は、組換えL551SについてのDNAコード配列である。

配列番号1919は、クローンL552Sのコード領域についての順方向プライマーPDM-479である。

配列番号1920は、クローンL552Sのコード領域についての逆方向プライマーPDM-480である。

配列番号1921は、発現された組換えL552Sについてのアミノ酸配列である。

配列番号1922は、組換えL552SについてのDNAコード配列である。

20

配列番号1923は、クローン#19069（配列番号90において与えられる部分配列）についての推定全長cDNA配列である。

配列番号1924は、クローン#18965または#19002（配列番号15において与えられる部分配列）についての推定全長cDNA配列である。

配列番号1925は、配列番号1923によってコードされる推定アミノ酸配列である。

配列番号1926は、配列番号1924によってコードされる推定アミノ酸配列である。

配列番号1927は、第1のL552Sエピトープの決定されたアミノ酸配列である。

配列番号1928は、第2のL552Sエピトープの決定されたアミノ酸配列である。

配列番号1929は、第3のL552Sエピトープの決定されたアミノ酸配列である。

配列番号1930は、L985Pペプチド#3482についてのアミノ酸配列である。

30

配列番号1931は、クローン#61144（配列番号1761, L1439P）についての伸長cDNA配列である。

配列番号1932は、配列番号1931によってコードされる推定アミノ酸配列である。

配列番号1933は、NUF2R遺伝子の全長cDNAであり、これに対して配列番号1931はある程度の配列類似性を示す。

配列番号1934は、配列番号1933によってコードされる推定アミノ酸配列である。

配列番号1935は、クローンL552Sのコード領域についての順方向プライマーPDM-737である。

配列番号1936は、クローンL552Sのコード領域についての逆方向プライマーPDM-738である。

40

配列番号1937は、発現された組換えL552Sについてのアミノ酸配列である。

配列番号1938は、組換えL552SについてのDNAコード配列である。

配列番号1939は、クローンL552Sのコード領域についての別の順方向プライマーPDM-736である。

配列番号1940は、発現された第2の組換えL552Sについてのアミノ酸配列である。

配列番号1941は、第2の組換えL552SについてのDNAコード配列である。

配列番号1942は、第4のL552Sエピトープの決定されたアミノ酸配列である。

配列番号1943は、第1のXAGE-1エピトープの決定されたアミノ酸配列である。

配列番号1944は、第2のXAGE-1エピトープの決定されたアミノ酸配列である。

50

配列番号1945は、全長L552S（配列番号809）のアミノ酸残基1～20に対応する第1の20マーペプチドの決定されたアミノ酸配列である。

配列番号1946は、全長L552S（配列番号809）のアミノ酸残基6～25に対応する第2の20マーペプチドの決定されたアミノ酸配列である。

配列番号1947は、全長L552S（配列番号809）のアミノ酸残基11～30に対応する第3の20マーペプチドの決定されたアミノ酸配列である。

配列番号1948は、全長L552S（配列番号809）のアミノ酸残基16～35に対応する第4の20マーペプチドの決定されたアミノ酸配列である。

配列番号1949は、全長L552S（配列番号809）のアミノ酸残基21～40に対応する第5の20マーペプチドの決定されたアミノ酸配列である。

10

配列番号1950は、全長L552S（配列番号809）のアミノ酸残基26～45に対応する第6の20マーペプチドの決定されたアミノ酸配列である。

配列番号1951は、全長L552S（配列番号809）のアミノ酸残基31～50に対応する第7の20マーペプチドの決定されたアミノ酸配列である。

配列番号1952は、全長L552S（配列番号809）のアミノ酸残基36～55に対応する第8の20マーペプチドの決定されたアミノ酸配列である。

配列番号1953は、全長L552S（配列番号809）のアミノ酸残基41～60に対応する第9の20マーペプチドの決定されたアミノ酸配列である。

配列番号1954は、全長L552S（配列番号809）のアミノ酸残基46～65に対応する第10の20マーペプチドの決定されたアミノ酸配列である。

20

配列番号1955は、全長L552S（配列番号809）のアミノ酸残基51～70に対応する第11の20マーペプチドの決定されたアミノ酸配列である。

配列番号1955は、全長L552S（配列番号809）のアミノ酸残基56～75に対応する第12の20マーペプチドの決定されたアミノ酸配列である。

配列番号1956は、全長L552S（配列番号809）のアミノ酸残基61～80に対応する第13の20マーペプチドの決定されたアミノ酸配列である。

配列番号1957は、全長L552S（配列番号809）のアミノ酸残基66～85に対応する第14の20マーペプチドの決定されたアミノ酸配列である。

配列番号1958は、全長L552S（配列番号809）のアミノ酸残基71～90に対応する第15の20マーペプチドの決定されたアミノ酸配列である。

30

配列番号1959は、全長L552S（配列番号809）のアミノ酸残基76～95に対応する第16の20マーペプチドの決定されたアミノ酸配列である。

配列番号1961は、全長L552S（配列番号809）のアミノ酸残基81～100に対応する第17の20マーペプチドの決定されたアミノ酸配列である。

配列番号1962は、全長L552S（配列番号809）のアミノ酸残基86～105に対応する第18の20マーペプチドの決定されたアミノ酸配列である。

配列番号1963は、全長L552S（配列番号809）のアミノ酸残基91～110に対応する第19の20マーペプチドの決定されたアミノ酸配列である。

配列番号1964は、全長L552S（配列番号809）のアミノ酸残基96～115に対応する第20の20マーペプチドの決定されたアミノ酸配列である。

40

配列番号1965は、全長L552S（配列番号809）のアミノ酸残基101～120に対応する第21の20マーペプチドの決定されたアミノ酸配列である。

配列番号1966は、全長L552S（配列番号809）のアミノ酸残基106～125に対応する第22の20マーペプチドの決定されたアミノ酸配列である。

配列番号1967は、全長L552S（配列番号809）のアミノ酸残基111～130に対応する第23の20マーペプチドの決定されたアミノ酸配列である。

配列番号1968は、全長L552S（配列番号809）のアミノ酸残基116～135に対応する第24の20マーペプチドの決定されたアミノ酸配列である。

配列番号1969は、全長L552S（配列番号809）のアミノ酸残基121～140に対応する第25の20マーペプチドの決定されたアミノ酸配列である。

50

配列番号1970は、全長L552S（配列番号809）のアミノ酸残基126～145に対応する第26の20マーペプチドの決定されたアミノ酸配列である。

配列番号1971は、全長L552S（配列番号809）のアミノ酸残基131～150に対応する第27の20マーペプチドの決定されたアミノ酸配列である。

配列番号1972は、全長L552S（配列番号809）のアミノ酸残基136～155に対応する第28の20マーペプチドの決定されたアミノ酸配列である。

配列番号1973は、全長L552S（配列番号809）のアミノ酸残基141～160に対応する第29の20マーペプチドの決定されたアミノ酸配列である。

配列番号1974は、配列番号1945の20マーをコードするDNA配列である。

配列番号1975は、配列番号1946の20マーをコードするDNA配列である。

配列番号1976は、配列番号1947の20マーをコードするDNA配列である。

配列番号1977は、配列番号1948の20マーをコードするDNA配列である。

配列番号1978は、配列番号1949の20マーをコードするDNA配列である。

配列番号1979は、配列番号1950の20マーをコードするDNA配列である。

配列番号1980は、配列番号1951の20マーをコードするDNA配列である。

配列番号1981は、配列番号1952の20マーをコードするDNA配列である。

配列番号1982は、配列番号1953の20マーをコードするDNA配列である。

配列番号1983は、配列番号1954の20マーをコードするDNA配列である。

配列番号1984は、配列番号1955の20マーをコードするDNA配列である。

配列番号1985は、配列番号1956の20マーをコードするDNA配列である。

配列番号1986は、配列番号1957の20マーをコードするDNA配列である。

配列番号1987は、配列番号1958の20マーをコードするDNA配列である。

配列番号1988は、配列番号1959の20マーをコードするDNA配列である。

配列番号1989は、配列番号1960の20マーをコードするDNA配列である。

配列番号1990は、配列番号1961の20マーをコードするDNA配列である。

配列番号1991は、配列番号1962の20マーをコードするDNA配列である。

配列番号1992は、配列番号1963の20マーをコードするDNA配列である。

配列番号1993は、配列番号1964の20マーをコードするDNA配列である。

配列番号1994は、配列番号1965の20マーをコードするDNA配列である。

配列番号1995は、配列番号1966の20マーをコードするDNA配列である。

配列番号1996は、配列番号1967の20マーをコードするDNA配列である。

配列番号1997は、配列番号1968の20マーをコードするDNA配列である。

配列番号1998は、配列番号1969の20マーをコードするDNA配列である。

配列番号1999は、配列番号1970の20マーをコードするDNA配列である。

配列番号2000は、配列番号1971の20マーをコードするDNA配列である。

配列番号2001は、配列番号1972の20マーをコードするDNA配列である。

配列番号2002は、配列番号1973の20マーをコードするDNA配列である。

10

20

30

【0049】

（発明の詳細な説明）

本発明は、一般に、癌（特に、肺癌）の治療および診断における、組成物、ならびにその使用に関する。以下でさらに記載するように、本発明の例示的な組成物としては、以下を含むがこれらに限定されない：ポリペプチド、特に免疫原性ポリペプチド、このようなポリペプチドをコードするポリヌクレオチド、抗体および他の結合因子、抗原提示細胞（APC）ならびに免疫系細胞（例えば、T細胞）。

40

【0050】

本発明の実施は、特にそれと逆であることを示さない限り、当業者の範囲であるウイルス学、免疫学、細菌学、分子生物学、および組換えDNA技術の従来の方法（これらの多くは、例示目的で以下に記載される）を使用する。このような技術は、文献によって完全に説明されている。例えば、Sambrookら、Molecular Cloning: A Laboratory Manual（第2版、1989）；Maniatisら、

50

Molecular Cloning: A Laboratory Manual (1982); DNA Cloning: A Practical Approach、第I巻および第II巻 (D. Glover編); Oligonucleotide Synthesis (N. Gait編、1984); Nucleic Acid Hybridization (B. HamesおよびS. Higgins編、1985); Transcription and Translation (B. HamesおよびS. Higgins編、1984); Animal Cell Culture (R. Freshney編、1986); Perbal、A Practical Guide to Molecular Cloning (1984)を参照のこと。

【0051】

10

本明細書中で引用される、全ての刊行物、特許および特許出願は、上記または下記にかかわらず、全体が参考して本明細書によって援用される。

【0052】

本明細書中および添付の特許請求の範囲で使用される場合、単数形、「1つの(a、an)」、および「この、その(the)」は、文脈によって明らかに他の方法で示されていない限り、複数の言及を含む。

【0053】

(ポリペプチド組成物)

本明細書中で使用される場合、用語「ポリペプチド」は、その従来の意味、すなわちアミノ酸の配列として使用される。これらのポリペプチドは、特定の長さの生成物に限定されず; 従って、ペプチド、オリゴペプチド、およびタンパク質は、このポリペプチドの定義に含まれ、そしてこのような用語は、別段示さない限り、本明細書中で交換可能に使用され得る。この用語はまた、このポリペプチドの発現後修飾(例えば、グリコシル化、アセチル化、リン酸化など)ならびに天然および非天然の両方の、当該分野で公知の他の改変を称さないか、これらを除外する。ポリペプチドは、タンパク質全体でもあり、または部分配列でもあり得る。本発明の状況において特定の目的のポリペプチドは、エピトープ、すなわち、ポリペプチドの免疫原特性の実質的に原因となり、かつ免疫応答を誘起し得る抗原決定基を含むアミノ酸部分配列である。

20

【0054】

本発明の特に例示的なポリペプチドとしては、配列番号1~57、59~323、341~782、784~785、788、790、792、794、796、800~804、807、808、810~826、828~1664、1668、1669、1676、1680~1805、1824、1826~1829、1865~1868、1873、1875、1877、1879、1881、1883、1891~1900、1910、1914、1918、1922~1924、1931、1933、1938、1941および1974~2002のいずれか1つに示されたポリヌクレオチド配列、あるいは中度のストリンジェントな条件下または高度にストリンジェントな条件下で、配列番号1~57、59~323、341~782、784~785、788、790、792、794、796、800~804、807、808、810~826、828~1664、1668、1669、1676、1680~1805、1824、1826~1829、1865~1868、1873、1875、1877、1879、1881、1883、1891~1900、1910、1914、1918、1922~1924、1931、1933、1938、1941および1974~2002のいずれか1つに示されるポリヌクレオチド配列に対してハイブリダイズする配列によってコードされるポリペプチドを含む。特定の他の例示的な本発明のポリペプチドとしては、配列番号324~340、786、787、789、791、793、795、797~799、805、806、809、827、1667、1670~1675、1677~1679、1806~1822、1825、1830~1833、1834~1856、1863、1864、1869~1872、1874、1876、1878、1880、1882、1884~1890、1901~1909、1913、1917、1921、1925~1930、1932

30

40

50

、1934、1937、1940および1942～1973のいずれか1つに示されるアミノ酸配列が挙げられる。

【0055】

本発明のポリペプチドは、これらの同定が肺腫瘍サンプルにおけるこれらの発現の増大したレベルに少なくとも部分的に基づくことの現れとして、本明細書中で時々、肺腫瘍タンパク質または肺腫瘍ポリペプチドとして参照される。従って、「肺腫瘍ポリペプチド」または「肺腫瘍タンパク質」は、一般に本発明のポリペプチド配列、またはこのようなポリペプチドをコードするポリヌクレオチド配列を示し、これらは、例えば、本明細書中で提供された代表的なアッセイを使用して決定されるように、正常組織での発現レベルより、少なくとも2倍高いレベル、および好ましくは5倍高いレベルで、試験された肺腫瘍サン

10

【0056】

特定の好ましい実施形態において、本発明のポリペプチドは免疫原性であり、すなわちこれらは、肺癌を有する患者からの抗血清および／またはT細胞を用いる免疫アッセイ（例えば、ELISAまたはT細胞刺激アッセイ）で検出可能に反応する。免疫原性活性についてのスクリーニングは、当業者に周知の技術を使用して実施され得る。例えば、このよ

20

【0057】

当業者に認識されるように、本明細書中で開示されるポリペプチドの免疫原性部分もまた、本発明により包含される。本明細書中で用いられる場合、「免疫原性部分」は、それ自体が、本発明の免疫原性ポリペプチドを認識するB細胞および／またはT細胞表面抗原レセプターと免疫学的に反応性である（すなわち、特異的に結合する）、本発明の免疫原性ポリペプチドのフラグメントである。免疫原性部分は、一般に、周知の技術（例えば、Paul、Fundamental Immunology、第3版、243-247（Raven Press、1993）およびこの文献で引用される参考文献に要約されている技術）を用いて同定され得る。このような技術は、抗原特異的抗体、抗血清、および／またはT細胞株もしくはクローンと反応する能力について、ポリペプチドをスクリーニングすることを包含する。本明細書中で用いられる場合、抗血清および抗体は、これらが、抗原に特異的に結合する（すなわち、これらが、ELISAまたは他の免疫アッセイにおいてこのタンパク質と反応し、そして無関係のタンパク質とは検出可能に反応しない）場

30

40

【0058】

1つの好ましい実施形態において、本発明のポリペプチドの免疫原性部分は、（例えば、ELISAおよび／またはT細胞反応性アッセイにおいて）全長ポリペプチドの反応性よりも実質的に低くないレベルで、抗血清および／またはT細胞と反応する部分である。好ましくは、免疫原性部分の免疫原性活性のレベルは、全長ポリペプチドの免疫原性の少なくとも約50%、好ましくは少なくとも約70%、そして最も好ましくは90%より高い。いくつかの例において、対応する全長ポリペプチドよりも大きい免疫原性活性のレベルを有する（例えば、約100%または150%もしくはそれより大きい免疫原性活性を有

50

する) 好ましい免疫原性部分が同定される。

【0059】

特定の他の実施形態において、例示的な免疫原性部分は、N末端リーダー配列および／または膜貫通ドメインが欠失されたペプチドを含み得る。他の例示的な免疫原性部分は、成熟タンパク質に対して小さなN末端および／またはC末端の欠失（例えば、1～30アミノ酸、好ましくは5～15アミノ酸）を含む。

【0060】

別の実施形態において、本発明のポリペプチド組成物また、T細胞および／または本発明のポリペプチド（特に、本明細書中に開示されたアミノ酸配列を有するポリペプチドまたはそれらの免疫原性フラグメントもしくは改変体）に対して生成された抗体と免疫学的に反応性である1つ以上のポリペプチドを含み得る。 10

【0061】

本発明の別の実施形態において、本明細書中に記載される1つ以上のポリペプチドと免疫学的に反応性であるT細胞および／もしくは抗体を惹起し得る1つ以上のポリペプチド、または本明細書中に開示されるポリヌクレオチド配列に含まれる連続した核酸配列によってコードされる1つ以上のポリペプチド、またはそれらの免疫学的フラグメントもしくは改変体、あるいは中程度～高度なストリンジェンシーの条件下でこれらの配列の1つ以上にハイブリダイズする1つ以上に核酸配列、を含むポリペプチドが提供される。

【0062】

別の局面において、本発明は、本明細書中に示されるポリペプチド組成物（例えば、配列番号324～340、786、787、789、791、793、795、797～799、805、806、809、827、1667、1670～1675、1677～1679、1806～1822、1825、1830～1833、1834～1856、1863、1864、1869～1872、1874、1876、1878、1880、1882、1884～1890、1901～1909、1913、1917、1921、1925～1930、1932、1934、1937、1940および1942～1973に示されるポリペプチド、または配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002の配列に示されるポリヌクレオチド配列によりコードされるポリペプチド）の少なくとも約5、10、15、20、25、50、または100連続するアミノ酸またはそれ以上（全ての中間の長さを含む）を含むポリペプチドフラグメントを提供する。 20 30

【0063】

別の局面において、本発明は、本明細書中に記載されるポリペプチド組成物の改変体を提供する。一般的に本発明により包含されるポリペプチド改変体は、代表的に、本明細書中に示されるポリペプチド配列に対して、その長さに沿って、少なくとも約70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、もしくは99%またはそれ以上の同一性（以下に記載するように決定された）を示す。 40

【0064】

1つの好ましい実施形態において、本発明により提供されるポリペプチドフラグメントおよび改変体は、本明細書中に詳細に示される全長ポリペプチドと反応する抗体および／またはT細胞と免疫学的に反応性である。

【0065】

別の好ましい実施形態において、本発明により提供されるポリペプチドフラグメントおよび改変体は、本明細書中に詳細に示される全長ポリペプチド配列により示される免疫原性活性のレベルの少なくとも約50%、好ましくは少なくとも約70%、そして最も好まし 50

くは少なくとも約90%またはそれ以上の免疫原性活性のレベルを示す。

【0066】

ポリペプチド「改変体」は、本明細書中でこの用語が使用される場合、代表的に、本明細書中で詳細に開示されるポリペプチドと1つ以上の置換、欠失、付加、および／または挿入において異なるポリペプチドである。このような改変体は、天然に生じるか、または例えば、本発明の1つ以上の上記のポリペプチド配列を改変し、そして本明細書中に記載されるようなそれらの免疫学的活性を評価することにより、ならびに／または当該分野で周知の多くの技術のいずれかを用いて、合成的に作製され得る。

【0067】

例えば、本発明のポリペプチドの特定の例示的な改変体として、1つ以上の部分（例えば、N末端リーダー配列または膜貫通ドメイン）が除去された改変体が挙げられる。他の例示的な改変体として、小さな部分（例えば、1～30アミノ酸、好ましくは5～15アミノ酸）が成熟タンパク質のN末端および／またはC末端から除去された改変体が挙げられる。 10

【0068】

多くの例において、改変体は、保存的置換を含む。「保存的置換」は、アミノ酸が類似の特性を有する別のアミノ酸で置換され、その結果、ペプチド化学の当業者がポリペプチドの二次構造および疎水的性質が実質的に変化していないことを期待するような置換である。上記のように、改変は、本発明のポリヌクレオチドおよびポリペプチドの構造においてなされ得、そしてなお所望の特性（例えば、免疫原性特性）を有する改変体ポリペプチドまたは誘導体ポリペプチドをコードする機能的分子を獲得し得る。本発明のポリペプチドの等価物を（または本発明のポリペプチドの改善された免疫原性の改変体もしくは部分さえも）作製するようポリペプチドのアミノ酸配列が変更されることが所望される場合、当業者は、代表的に、表10に基づくコードDNA配列の1つ以上のコドンを変更する。 20

【0069】

例えば、特定のアミノ酸は、例えば、抗体の抗原結合領域または基質分子上の結合部位のような構造を有するタンパク質構造中の他のアミノ酸に、感知可能な相互作用的結合能の損失なしで、置換され得る。タンパク質の相互作用的な能力および性質がタンパク質の生物学的機能的活性を規定するので、特定のアミノ酸配列置換は、タンパク質配列、および当然ながら、その根底にあるDNAコード配列においてなされ得、そしてそれにも関わらず、同様の特性を有するタンパク質が入手される。従って、種々の変化が、ペプチドの生物学的有用性または活性の感知可能な損失を伴わずに、開示された組成物のペプチド配列またはそのペプチドをコードする対応するDNA配列においてなされ得ることが意図される。 30

【0070】

【表16】

表 10

アミノ酸			コドン						
アラニン	Ala	A	GCA	GCC	GCG	GCU			
システイン	Cys	C	UGC	UGU					
アスパラギン酸	Asp	D	GAC	GAU					
グルタミン酸	Glu	E	GAA	GAG					
フェニルアラニン	Phe	F	UUC	UUU					
グリシン	Gly	G	GGA	GGC	GGG	GGU			
ヒスチジン	His	H	CAC	CAU					
イソロイシン	Ile	I	AUA	AUC	AUU				
リジン	Lys	K	AAA	AAG					
ロイシン	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
メチオニン	Met	M	AUG						
アスパラギン	Asn	N	AAC	AAU					
プロリン	Pro	P	CCA	CCC	CCG	CCU			
グルタミン	Gln	Q	CAA	CAG					
アルギニン	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU	
セリン	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU	
スレオニン	Thr	T	ACA	ACC	ACG	ACU			
バリン	Val	V	GUA	GUC	GUG	GUU			
トリプトファン	Trp	W	UGG						
チロシン	Tyr	Y	UAC	UAU					

10

20

30

40

50

このような変化を作製する際に、アミノ酸の疎水性親水性指標が考慮され得る。タンパク質に相互作用的な生物学的機能を付与する際の疎水性親水性アミノ酸指標の重要性は、一般的に当該分野において理解されている（K y t e および D o o l i t t l e、1982、本明細書中に参考として援用される）。アミノ酸の相対的な疎水性親水性の性質は、得られるタンパク質の二次構造に寄与し、これが次に、他の分子（例えば、酵素、基質、レセプター、DNA、抗体、抗原など）とのそのタンパク質の相互作用を規定することが受け入れられている。各アミノ酸は、その疎水性および電荷特性に基づいて疎水性親水性指標を割り当てられている（K y t e および D o o l i t t l e、1982）。これらの値は以下である：イソロイシン（+4.5）；バリン（+4.2）；ロイシン（+3.8）；フェニルアラニン（+2.8）；システイン／シスチン（+2.5）；メチオニン（+1.9）；アラニン（+1.8）；グリシン（-0.4）；スレオニン（-0.7）；セリン（-0.8）；トリプトファン（-0.9）；チロシン（-1.3）；プロリン（-1.6）；ヒスチジン（-3.2）；グルタミン酸（-3.5）；グルタミン（-3.5）；アスパラギン酸（-3.5）；アスパラギン（-3.5）；リジン（-3.9）；およびアルギニン（-4.5）。

【0071】

特定のアミノ酸が類似の疎水性親水性指標またはスコアを有する他のアミノ酸によって置換され得、そしてなお類似の生物学的活性を有するタンパク質を生じる（すなわち、生物学的に機能的に等価なタンパク質をなお入手する）ことが、当該分野において公知である。このような変化を作製する際に、その疎水性親水性指標が±2以内であるアミノ酸の置換が好ましく、その疎水性親水性指標が±1以内であるアミノ酸の置換が特に好ましく、そしてその疎水性親水性指標が±0.5以内であるアミノ酸の置換がなおより特に好ましい。同様なアミノ酸の置換が、親水性に基づいて有効になされ得ることもまた、当該分野で理解されている。米国特許第4,554,101号（その全体が本明細書中に参考とし

て詳細に援用される)は、そのタンパク質の最大局所的平均親水性が、その隣接するアミノ酸の親水性によって支配されるので、そのタンパク質の生物学的特性と相関することを言及している。

【0072】

米国特許第4, 554, 101号に詳述されるように、以下の親水性値がアミノ酸残基に割り当てられた：アルギニン (+3.0)；リジン (+3.0)；アスパラギン酸 (+3.0 ± 1)；グルタミン酸 (+3.0 ± 1)；セリン (+0.3)；アスパラギン (+0.2)；グルタミン (+0.2)；グリシン (0)；スレオニン (-0.4)；プロリン (-0.5 ± 1)；アラニン (-0.5)；ヒスチジン (-0.5)；システイン (-1.0)；メチオニン (-1.3)；バリン (-1.5)；ロイシン (-1.8)；イソロイシン (-1.8)；チロシン (-2.3)；フェニルアラニン (-2.5)；トリプトファン (-3.4)。アミノ酸は、類似の親水性の値を有する別のアミノ酸に置換され得、そしてなお、生物学的な等価物（特に、免疫学的に等価なタンパク質）を得ることが理解される。このような変化において、その親水性値が±2以内であるアミノ酸の置換が好ましく、その親水性値が±1以内であるアミノ酸の置換が特に好ましく、そしてその親水性値が±0.5以内であるアミノ酸の置換がなおより特に好ましい。

10

【0073】

上記で概説したように、従って、アミノ酸置換は、一般的にアミノ酸側鎖置換基の相対的な類似性（例えば、その疎水性、親水性、電荷、大きさ、など）に基づく。前述の種々の特徴を考慮する典型的な置換は当業者に周知であり、そして以下を含む：アルギニンおよびリジン；グルタミン酸およびアスパラギン酸；セリンおよびスレオニン；グルタミンおよびアスパラギン；ならびにバリン、ロイシン、およびイソロイシン。

20

【0074】

さらに、任意のポリヌクレオチドが、インビボでの安定性を増加させるためにさらに改変され得る。可能性のある改変として以下が挙げられるがこれらに限定されない：5'末端および/または3'末端での隣接配列の付加；骨格におけるホスホジエステラーゼ結合ではなくホスホロチオエートまたは2'-O-メチルの使用；ならびに/あるいは従来とは異なる塩基（例えば、イノシン、キューオシン、およびワイプトシン、ならびにアデニン、シチジン、グアニン、チミン、およびウリジンのアセチル、メチル、チオ、および他の修飾形態）の含有。

30

【0075】

アミノ酸置換はさらに、残基の極性、電荷、溶解性、疎水性、親水性および/または両親媒性特性の類似性に基づいて作製され得る。例えば、負に荷電したアミノ酸としては、アスパラギン酸およびグルタミン酸；正に荷電したアミノ酸としては、リジンおよびアルギニン；そして類似の親水性値を有する非荷電性の極性ヘッド基を有するアミノ酸としては、ロイシン、イソロイシンおよびバリン；グリシンおよびアラニン；アスパラギンおよびグルタミン；ならびにセリン、トレオニン（スレオニン）、フェニルアラニンおよびチロシンが挙げられる。保存的变化を示し得るアミノ酸の他のグループとしては、(1) ala, pro, gly, glu, asp, gln, asn, ser, thr；(2) cys, ser, tyr, thr；(3) val, ile, leu, met, ala, phe；(4) lys, arg, his；および(5) phe, tyr, trp, hisが挙げられる。改変体はまた、またはあるいは、非保存的变化を含み得る。好ましい実施形態において、改変体ポリペプチドは、5アミノ酸またはそれより少ないアミノ酸の置換、欠失または付加によって、ネイティブの配列とは異なる。改変体はまた（またはあるいは）、例えば、ポリペプチドの免疫原性、二次構造および疎水性親水性特性に対する最小の影響を有するアミノ酸の欠失または付加によって、改変され得る。

40

【0076】

上記のように、ポリペプチドは、タンパク質のN末端にシグナル（または、リーダー）配列を含み得、これは、翻訳と同時に、または翻訳後に、そのタンパク質の転移を指向する。このポリペプチドはまた、このポリペプチドの合成、精製または同定を容易にするため

50

に、またはこのポリペプチドの固体支持体への結合を増強するために、リンカー配列または他の配列（例えば、ポリHis）に結合体化され得る。例えば、ポリペプチドは、免疫グロブリンFc領域に結合体化され得る。

【0077】

ポリペプチド配列が比較される場合、以下に記載されるように最大の一致で整列するとき、2つの配列中のアミノ酸の配列が同じである場合、2つの配列は、「同一」といわれる。2つの配列間の比較は、代表的には、比較ウィンドウによって配列を比較して、配列類似性の局所的領域を同定および比較することによって行われる。本明細書中で使用される「比較ウィンドウ」とは、少なくとも約20、通常は30～約75、40～約50の連続する位置のセグメントをいう。ここで、配列は、2つの配列が最適に整列された後、同じ数の連続する位置の参照配列と比較され得る。 10

【0078】

比較のための配列の最適な整列は、生命情報科学ソフトウェアのLasergeneスイート(suite)におけるMegalignプログラム(DNA STAR, Inc., Madison, WI)を使用して、デフォルトパラメータを用いて行われ得る。このプログラムは、以下の参考文献に記載のいくつかの整列スキームを統合する: Dayhoff, M. O. (編) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC 第5巻, 補遺3, 345～358頁におけるDayhoff, M. O. (1978) A model of evolutionary change in proteins—Matrices for detecting distant relationships.; Hein J. (1990) Unified Approach to Alignment and Phylogenesis 626～645頁 Methods in Enzymology 第183巻, Academic Press, Inc., San Diego, CA; Higgins, D. G. およびSharp, P. M. (1989) CABIOS 5:151-153; Myers, E. W. およびMuller W. (1988) CABIOS 4:11-17; Robinson, E. D. (1971) Comb. Theor 11:105; Saitou, N. Nei, M. (1987) Mol. Biol. Evol. 4:406-425; Sneath, P. H. A. およびSokal, R. R. (1973) Numerical Taxonomy—the Principles and Practice of Numerical Taxonomy, Freeman Press, San Francisco, CA; Wilbur, W. J. およびLipman, D. J. (1983) Proc. Natl. Acad., Sci. USA 80:726-730. 20

【0079】

あるいは、比較のための配列の最適な整列は、SmithおよびWaterman (1981) Add. Appl. Math 2:482の部分的同一性アルゴリズムによって、NeedlemanおよびWunsch (1970) J. Mol. Biol. 48:443の同一性整列アルゴリズムによって、PearsonおよびLipman (1988) Proc. Natl. Acad. Sci. USA 85:2444の類似性検索方法によって、これらのアルゴリズムのコンピュータ化した実行(Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI)におけるGAP、BESTFIT、BLAST、FASTA、およびTFASTA)によって、または検査によって行われ得る。 40

【0080】

配列同一性および配列類似性の割合を決定するために適切なアルゴリズムの1つの好ましい例は、BLASTおよびBLAST2.0アルゴリズムであり、これらはそれぞれAltschulら (1977) Nucl. Acids Res. 25:3389-3402 50

および Altschul ら (1990) J. Mol. Biol. 215: 403-410 に記載される。BLAST および BLAST 2.0 は、本発明のポリヌクレオチドおよびポリペプチドについての配列同一性の割合を決定するために、例えば、本明細書中に記載のパラメータを使用して使用され得る。BLAST 分析を行うためのソフトウェアは、National Center for Biotechnology Information を通して公に利用可能である。アミノ酸配列について、スコアリングマトリクスは、累積スコアを算出するために使用され得る。各指示におけるワードヒットの拡大は、以下の場合に停止する：累積整列スコアが、その最大到達値から量 X だけ低下した場合；累積スコアが、1 以上の負のスコアリングの残基整列の累積に起因してゼロ以下になった場合；またはいずれかの配列の末端に到達した場合。この BLAST アルゴリズムパラメーター W、T および X は、整列の感度および速度を決定する。 10

【0081】

好ましい 1 つのアプローチにおいて、「配列同一性のパーセンテージ」は、少なくとも 20 位置の比較ウィンドウにわたる 2 つの最適に整列した配列を比較することによって決定され、ここで比較ウィンドウ中のポリペプチド配列の部分は、参照配列（これは、付加または欠失を含まない）と比較して、2 つの配列の最適な整列について 20 パーセント以下、通常は 5～15 パーセント、または 10～12 パーセントの付加または欠失（すなわち、ギャップ）を含み得る。このパーセンテージは、両方の配列で同一のアミノ酸残基が生じる位置の数を決定して一致する位置の数を得、この一致する位置の数を参照配列における位置の総数（すなわち、ウィンドウサイズ）で除算し、そしてこの結果に 100 を掛けて配列同一性のパーセンテージを得ることによって算出される。 20

【0082】

他の例示的实施形態において、ポリペプチドは、本明細書中に記載の複数のポリペプチドを含むか、または本明細書に記載の少なくとも 1 つのポリペプチドおよび関連しない配列（例えば、公知の腫瘍タンパク質）を含む融合ポリペプチドであり得る。例えば、融合パートナーは、T ヘルパーエпитープ、好ましくはヒトによって認識される T ヘルパーエピトープを提供する際に補助し得るか（免疫学的融合パートナー）、またはネイティブの組換えタンパク質より高い収量でタンパク質を発現する際に補助し得る（発現エンハンサー）。特定の好ましい融合パートナーは、免疫学的融合パートナーおよび発現増強融合パートナーの両方である。他の融合パートナーは、ポリペプチドの溶解性を増加するように、またはポリペプチドが所望の細胞内コンパートメントに標的化されることを可能にするように選択され得る。なおさらなる融合パートナーとしては、親和性タグ（これは、ポリペプチドの精製を容易にする）が挙げられる。 30

【0083】

融合ポリペプチドは、一般に、標準的な技術（化学的結合体化を含む）を使用して調製され得る。好ましくは、融合ポリペプチドは、発現系において、組換えポリペプチドとして発現され、非融合ポリペプチドと比較して、増加したレベルの産生を可能にする。手短に言うと、このポリペプチド成分をコードする DNA 配列を、別々に組立て得、そして適切な発現ベクターに連結し得る。1 つのポリペプチド成分をコードする DNA 配列の 3' 末端は、ペプチドリinker を用いてまたは用いずに、第 2 のポリペプチド成分をコードする DNA 配列の 5' 末端に、これらの配列のリーディングフレームが同位相にあるように連結される。このことによって、両方の成分ポリペプチドの生物学的活性を保持する単一の融合ポリペプチドへの翻訳が可能になる。 40

【0084】

ペプチドリinker 配列は、各ポリペプチドがその二次構造および三次構造へと折り畳まれるのを保証するために十分な距離で第一および第二のポリペプチド成分を隔てるために用いられ得る。このようなペプチドリinker 配列は、当該分野で周知の標準的な技術を用いて融合ポリペプチド中に組み込まれる。適切なペプチドリinker 配列は、以下の因子に基づいて選択され得る：（1）伸長した可変（flexible）コンホメーションを取る能力；（2）第一および第二のポリペプチド上の機能的なエピトープと相互作用し得る二 50

次構造を取ることができないこと；ならびに（３）ポリペプチドの機能的なエピトープと反応し得る疎水性または荷電した残基の無いこと。好ましいペプチドリンカー配列は、G l y、A s nおよびS e r残基を含む。T h rおよびA l aのような中性に近い他のアミノ酸もまた、リンカー配列に用いられ得る。リンカーとして有用に用いられ得るアミノ酸配列としては、M a r a t e aら、G e n e 40：39-46、1985；M u r p h yら、P r o c . N a t l . A c a d . S c i . U S A 83：8258-8262、1986；米国特許第4,935,233号および米国特許第4,751,180号に開示されるアミノ酸配列が挙げられる。リンカー配列は、一般的に1から約50アミノ酸長であり得る。リンカー配列は、第一および第二のポリペプチドが、機能的ドメインを分離するため、および立体的な干渉を防ぐために用いられ得る非必須N末端アミノ酸領域を有する場合、必要とされない。

10

【0085】

連結されたDNA配列は、適切な転写または翻訳調節エレメントに作動可能に連結される。DNAの発現を担う調節エレメントは、第一のポリペプチドをコードするDNA配列の5'側にのみ位置する。同様に、翻訳および転写終結シグナルを終了するために必要とされる終止コドンは、第二のポリペプチドをコードするDNA配列の3'側にのみ存在する。

【0086】

融合ポリペプチドは、関連しない免疫原性タンパク質（例えば、リコール（r e c a l l）応答を惹起し得る免疫原性タンパク質）と共に、本明細書中に記載されるようなポリペプチドを含み得る。このようなタンパク質の例としては、破傷風タンパク質、結核タンパク質および肝炎タンパク質が挙げられる（例えば、S t o u t e rら、N e w E n g l . J . M e d . 、336：86-91（1997）を参照のこと）。

20

【0087】

1つの好ましい実施形態において、免疫学的融合パートナーは、M y c o b a c t e r i u m t u b e r c u l o s i s由来のR a 1 2フラグメントのようにM y c o b a c t e r i u m s p . 由来である。R a 1 2組成物、ならびに異種ポリヌクレオチド／ポリペプチド配列の発現および／または免疫原性を増強する際にR a 1 2を使用するための方法は、米国特許出願60／158,585（この開示は、本明細書中でその全体が参考として援用される）に記載されている。簡単には、R a 1 2は、M y c o b a c t e r i u m t u b e r c u l o s i s M T B 3 2 A核酸のサブ配列であるポリヌクレオチド領域をいう。M T B 3 2 Aは、M . t u b e r c u l o s i sの毒性菌株および無毒性菌株の遺伝子によってコードされる分子量32KDのセリンプロテアーゼである。M T B 3 2 Aのヌクレオチド配列およびアミノ酸配列が、記載されている（例えば、米国特許出願60／158,585；また、S k e i k yら、I n f e c t i o n a n d I m m u n . （1999）67：3998-4007も参照のこと。これらは、本明細書中で参考として援用される）。驚いたことに、M T B 3 2 Aコード配列の14KDのC末端フラグメントが、それぞれの上で高いレベルで発現し、精製プロセス全体にわたって可溶性ポリペプチドとして残存することが発見された。さらに、このフラグメントは、これが融合される異種抗原性ポリペプチドの免疫原性を増強し得る。この14KDのC末端フラグメントは、本明細書中ではR a 1 2といわれ、そしてM T B 3 2 Aのアミノ酸残基192-323のいくつかまたは全てを含むフラグメントを表す。

30

40

【0088】

他の好ましいR a 1 2ポリヌクレオチドは、一般的に、R a 1 2ポリペプチドの一部をコードする少なくとも約15個の連続するヌクレオチド、少なくとも約30個のヌクレオチド、少なくとも60個のヌクレオチド、少なくとも約100個のヌクレオチド、少なくとも約200個のヌクレオチド、または少なくとも約300個のヌクレオチドを含む。

【0089】

R a 1 2ポリヌクレオチドは、ネイティブな配列（すなわち、R a 1 2ポリペプチドまたはその一部をコードする内因性配列）を含んでもよいし、このような配列の改変体を含ん

50

でもよい。R a 1 2 ポリヌクレオチド改変体は、1 以上の置換、付加、欠失および／または挿入を含み得、その結果、コードされた融合ポリペプチドの生物学的活性が、ネイティブな R a 1 2 ポリペプチドを含む融合ポリペプチドと比較して、実質的に減少していない。好ましくは、改変体は、ネイティブな R a 1 2 ポリペプチドまたはその一部をコードするポリヌクレオチド配列に、少なくとも約 7 0 % の同一性、より好ましくは少なくとも約 8 0 % の同一性、そして最も好ましくは少なくとも約 9 0 % の同一性を示す。

【0090】

好ましい他の実施形態において、免疫学的融合パートナーは、グラム陰性の細菌 *H a e m o p h i l u s i n f l u e n z a B* の表面タンパク質である、プロテイン D (W O 9 1 / 1 8 9 2 6) に由来する。好ましくは、プロテイン D 誘導体は、このタンパク質のほぼ 3 分の 1 (例えば、最初の N 末端 1 0 0 ~ 1 1 0 アミノ酸) を含み、そしてプロテイン D 誘導体は、脂質化 (l i p i d a t e d) され得る。特定の好ましい実施形態において、リボプロテイン D 融合パートナーの最初の 1 0 9 残基は、さらなる外因性 T 細胞エピトープを有するポリペプチドを提供するように、そして E . c o l i 中の発現レベルを増加するように、N 末端に含まれる (従って、発現エンハンサーとして機能する)。脂質テールは、抗原提示細胞への抗原の最適な提示を保証する。他の融合パートナーは、インフルエンザウイルス由来の非構造的タンパク質、N S 1 (血球凝集素) を含む。代表的に、N 末端の 8 1 アミノ酸が用いられるが、T ヘルパーエピトープを含む異なるフラグメントが使用され得る。

【0091】

別の実施形態において、免疫学的融合パートナーは、L Y T A として公知のタンパク質、またはその部分 (好ましくは C 末端部分) である。L Y T A は、アミダーゼ L Y T A (L y t A 遺伝子によりコードされる; G e n e 4 3 : 2 6 5 ~ 2 9 2 , 1 9 8 6) として公知の N - アセチル - L - アラニンアミダーゼを合成する *S t r e p t o c o c c u s p n e u m o n i a e* 由来である。L Y T A は、ペプチドグリカン骨格中の特定の結合を特異的に分解する自己溶解素である。L Y T A タンパク質の C 末端ドメインは、コリンまたはいくつかのコリンアナログ (例えば、D E A E) に対する親和性を担う。この性質は、融合タンパク質の発現のために有用な E . c o l i C - L Y T A 発現プラスミドの開発のために利用されてきた。アミノ末端で C - L Y T A フラグメントを含むハイブリッドタンパク質の精製が、記載されている (B i o t e c h n o l o g y 1 0 : 7 9 5 ~ 7 9 8 , 1 9 9 2 を参照のこと)。好ましい実施形態において、L Y T A の反復部分は、融合ポリペプチドに組み込まれ得る。反復部分は、残基 1 7 8 で開始する C 末端領域中に見出される。特に好ましい反復部分は、残基 1 8 8 ~ 3 0 5 を組み込む。

【0092】

なお別の例示的な実施形態は、融合ポリペプチド、およびそれらをコードするポリヌクレオチドを含み、ここでこの融合パートナーは、米国特許第 5 , 6 3 3 , 2 3 4 号に記載のようにポリペプチドをエンドソーム / リソソームコンパートメントへ指向させ得る標的シグナルを含む。本発明の免疫原性ポリペプチドは、この標的シグナルと融合する場合、より効率的に M H C クラス I I 分子と結合し、これによってこのポリペプチドに特異的な C D 4 ⁺ T 細胞のインビボでの増強された刺激が提供される。

【0093】

本発明のポリペプチドは、周知の種々の合成技術および／または組換え技術のいずれかを使用して調製される。これらの技術の後者を以下でさらに記載する。一般的に約 1 5 0 アミノ酸よりも少ないポリペプチド、部分および他の改変体は、当業者に周知の技術を使用して、合成手段により作製され得る。例示的な 1 つの例において、このようなポリペプチドは、市販の固相技術のいずれか (例えば、M e r r i f i e l d 固相合成方法) を使用して合成され、ここでアミノ酸は、成長しているアミノ酸鎖に連続的に付加される。M e r r i f i e l d , J . A m . C h e m . S o c . 8 5 : 2 1 4 9 - 2 1 4 6 , 1 9 6 3 を参照のこと。ポリペプチドの自動化された合成のための機器が、P e r k i n E l m e r / A p p l i e d B i o S y s t e m s D i v i s i o n (F o s t e r C i

10

20

30

40

50

t y, C A) などの供給業者から市販されており、そして製造者の指示に従って操作され得る。

【0094】

一般に、本発明のポリペプチド組成物（融合ポリペプチドを含む）が単離される。「単離された」ポリペプチドは、その元来の環境から取り出されたものである。例えば、天然に存在するタンパク質またはポリペプチドは、それが天然の系中で共存する物質のいくつかまたは全てから分離されている場合、単離されている。好ましくは、このようなポリペプチドはまた、精製され、例えば、少なくとも約90%純粋、より好ましくは少なくとも約95%純粋、そして最も好ましくは少なくとも約99%純粋である。

【0095】

10

（ポリヌクレオチド組成物）

他の局面において、本発明は、ポリヌクレオチド組成物を提供する。用語「DNA」および「ポリヌクレオチド」は、本明細書中で実質的に交換可能に使用され、特定の種の全ゲノムDNAを含まない単離されたDNA分子をいう。本明細書中で使用される場合、「単離された」は、ポリヌクレオチドが、他のコード配列から実質的に分離されており、そしてDNA分子が、無関係のコードDNAの大部分（例えば、大きな染色体フラグメントまたは他の機能的遺伝子またはポリペプチドコード領域）を含まないことを意味する。当然のことながら、これは、もともと単離されたDNA分子をいい、後で人工的にセグメントに付加された遺伝子またはコード領域を除外しない。

【0096】

20

当業者に理解されるように、本発明のポリヌクレオチド組成物は、タンパク質、ポリペプチド、ペプチドなどを発現するか、または発現し得るように適応されたゲノム配列、ゲノム外配列およびプラスミドにコードされる配列、ならびにより小さな操作された遺伝子セグメントを含み得る。このようなセグメントは、自然に単離され得るか、または人の手によって合成的に改変され得る。

【0097】

当業者に認識されるように、本発明のポリヌクレオチドは、一本鎖（コードもしくはアンチセンス）または二本鎖であり得、そしてDNA分子（ゲノム、cDNAもしくは合成）またはRNA分子であり得る。RNA分子は、HnRNA分子（これはイントロンを含み、そしてDNA分子に1対1の様式で対応する）、およびmRNA分子（これは、イントロンを含まない）を含み得る。さらなるコード配列または非コード配列は、本発明のポリヌクレオチド内に存在し得るがその必要はなく、そしてポリヌクレオチドは、他の分子および／または支持材料に連結され得るがその必要はない。

30

【0098】

ポリヌクレオチドは、ネイティブの配列（すなわち、本発明のポリペプチド／タンパク質、またはその部分をコードする内因性配列）を含み得るか、あるいはこのような配列の改変体または誘導体、そして好ましくは免疫原性改変体または誘導体をコードする配列を含み得る。

【0099】

従って、本発明の別の局面に従い、配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002のいずれか1つに示されるポリヌクレオチド配列、配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～

40

50

1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002のいずれか1つに示されるポリヌクレオチド配列の相補体、ならびに配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002のいずれか1つに示されるポリヌクレオチド配列の縮重改変体のうちいくつかまたは全てを含むポリヌクレオチド組成物が提供される。特定の好ましい実施形態において、本明細書中に示されるポリヌクレオチド配列は、上記のような免疫原性ポリペプチドをコードする。

【0100】

他の関連する実施形態において、本発明は、本明細書中で配列番号1～57、59～323、341～782、784～785、788、790、792、794、796、800～804、807、808、810～826、828～1664、1668、1669、1676、1680～1805、1824、1826～1829、1865～1868、1873、1875、1877、1879、1881、1883、1891～1900、1910、1914、1918、1922～1924、1931、1933、1938、1941および1974～2002に開示される配列に実質的な同一性を有するポリヌクレオチド改変体を提供する。例えば、これらは、本明細書中で記載の方法（例えば、以下に記載のような標準的なパラメーターを使用するBLAST分析）を使用して本発明のポリヌクレオチド配列と比較して、少なくとも70%の配列同一性、好ましくは少なくとも75%、80%、85%、90%、95%、96%、97%、98%、または99%より高い配列同一性を含む。当業者は、これらの値が、コドンの縮重、アミノ酸類似性、リーディングフレームの位置付けなどを考慮することによって、2つのヌクレオチド配列にコードされるタンパク質の対応する同一性を決定するように、適切に調整され得ることを理解する。

【0101】

代表的には、ポリヌクレオチド改変体は、1つ以上の置換、付加、欠失および／または挿入を含み、好ましくは、その結果、改変体ポリヌクレオチドによってコードされるポリペプチドの免疫原性は、本明細書中に具体的に示されるポリヌクレオチド配列によってコードされるポリペプチドに対して実質的に減少されない。用語「改変体」はまた、異種起源の相同遺伝子を包含することが理解されるべきである。

【0102】

さらなる実施形態において、本発明は、本明細書中に開示される配列の1つ以上に同一であるかまたは相補的な配列の種々の長さの連続したストレッチを含むポリヌクレオチドフラグメントを提供する。例えば、本明細書中に開示される配列の1つ以上の、少なくとも約10、15、20、30、40、50、75、100、150、200、300、400、500または1000以上連続したヌクレオチド、ならびにその間の全ての中間の長さの連続したヌクレオチドを含むポリヌクレオチドが、本発明によって提供される。この状況において、「中間の長さ」とは、示された値の間の任意の長さ（例えば、16、17、18、19など；21、22、23など；30、31、32など；50、51、52、53など；100、101、102、103など；150、151、152、153など；200～500；500～1,000などの間の全ての整数を含む）を意味することが容易に理解される。

【0103】

本発明の別の実施形態において、本明細書において提供されるポリヌクレオチド配列もしくはそのフラグメントまたはその相補的な配列に対して、中程度～高いストリンジェンシー条件下でハイブリダイズし得るポリヌクレオチド組成物が提供される。ハイブリダイゼ

ーション技術は、分子生物学の当該分野において周知である。例示の目的であるが、他のポリヌクレオチドと本発明のポリヌクレオチドとのハイブリダイゼーションを試験するために適切な中程度にストリンジェントな条件は、 $5 \times \text{SSC}$ 、 $0.5\% \text{ SDS}$ 、 1.0 mM EDTA ($\text{pH } 8.0$) の溶液中の事前洗浄； $50^\circ\text{C} \sim 60^\circ\text{C}$ 、 $5 \times \text{SSC}$ での、一晚のハイブリダイゼーション；続いて、 $0.1\% \text{ SDS}$ を含有する $2 \times$ 、 $0.5 \times$ および $0.2 \times \text{SSC}$ のそれぞれを用いた、 65°C で20分間の2回の洗浄を含む。当業者は、ハイブリダイゼーションのストリンジェンシーが、例えば、ハイブリダイゼーション溶液の塩含量および／またはハイブリダイゼーションが実施される温度を変更することによって、容易に操作され得ることを理解する。例えば、別の実施形態において、適切な高いストリンジェントのハイブリダイゼーション条件としては、ハイブリダイゼーションの温度が、例えば、 $60 \sim 65^\circ\text{C}$ または $65 \sim 70^\circ\text{C}$ に増大される点を除いて、上記の条件が挙げられる。

10

【0104】

特定の好ましい実施形態において、上記のポリヌクレオチド（例えば、ポリヌクレオチド改変体、フラグメント、およびハイブリダイズする配列）は、本明細書中に具体的に示されるポリペプチド配列と免疫学的に交差反応するポリペプチドをコードする。他の好ましい実施形態において、このようなポリヌクレオチドは、本明細書中に具体的に示されるポリペプチド配列の免疫原性活性の、少なくとも約 50% 、好ましくは少なくとも約 70% 、そしてより好ましくは少なくとも約 90% の免疫原性活性レベルを有するポリペプチドをコードする。

20

【0105】

本発明のポリヌクレオチド、またはそのフラグメントは、そのコード配列自体の長さに関わらず、他のDNA配列（例えば、プロモーター、ポリアデニル化シグナル、さらなる制限酵素部位、マルチクローニング部位、他のコードセグメントなど）と組み合わせ得、その結果、その全体の長さは、相当変化し得る。従って、ほとんどいずれの長さの核酸フラグメントをも使用し得ることが意図され、その全長は、好ましくは意図した組換えDNAプロトコルにおける調製および使用の容易さによって制限される。例えば、約10,000、約5000、約3000、約2,000、約1,000、約500、約200、約100、約50塩基対長など（全ての中間の長さを含む）の全長を有する例示的なポリヌクレオチドセグメントが、本発明の多くの実行において有用であることが意図される。

30

【0106】

ポリヌクレオチド配列を比較する場合、2つの配列におけるヌクレオチドの配列が、以下に記載されるように最大一致について整列されるのと同じ場合に、2つの配列が「同一」であるといわれる。2つの配列の間の比較は、代表的には、配列類似性の局部領域を同定および比較するために、比較ウィンドウにわたって配列を比較することによって行われる。本明細書中で使用される場合、「比較ウィンドウ」は、少なくとも約20、通常は30～約75、40～約50連続した位置のセグメントをいい、ここで、2つの配列が最適に整列された後に、配列は、連続した位置の同じ数の参照配列と比較され得る。

【0107】

比較のための配列の最適な整列は、バイオインフォマティクス（生命情報科学）ソフトウェアのLasergene suiteにおけるMegalignプログラム（DNASTAR, Inc., Madison, WI）を使用して、デフォルトパラメーターを用いて行われ得る。このプログラムは、以下の参考文献に記載のいくつかの整列スキームを統合する：Dayhoff, M. O.（編）Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, 補遺3, 345-358頁におけるDayhoff, M. O.（1978）A model of evolutionary change in proteins—Matrices for detecting distant relationships. ; Hein J.（1990）Unified Approach to Align

40

50

ment and Phylogenesis 626-645頁 Methods in Enzymology vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D. G. および Sharp, P. M. (1989) CABIOS 5:151-153; Myers, E. W. および Muller W. (1988) CABIOS 4:11-17; Robinson, E. D. (1971) Comb. Theor 11:105; Santou, N. Nes, M. (1987) Mol. Biol. Evol. 4:406-425; Sneath, P. H. A. および Sokal, R. R. (1973) Numerical Taxonomy—the Principles and Practice of Numerical Taxonomy, Freeman Press, San Francisco, CA; Wilbur, W. J. および Lipman, D. J. (1983) Proc. Natl. Acad. Sci. USA 80:726-730。 10

【0108】

あるいは、比較のための配列の最適な整列は、SmithおよびWaterman (1981) Add. Appl. Math 2:482の局部同定アルゴリズムによって、NeedlemanおよびWunsch (1970) J. Mol. Biol. 48:443の同一性整列アルゴリズムによって、PearsonおよびLipman (1988) Proc. Natl. Acad. Sci. USA 85:2444の類似性検索方法によって、これらのアルゴリズムのコンピュータ化した実行 (Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WIにおけるGAP、BESTFIT、BLAST、FASTA、およびTFASTA) によって、または 20
検査によって行われ得る。

【0109】

配列同一性および配列類似性の割合を決定するために適切なアルゴリズムの1つの好ましい例は、BLASTおよびBLAST2.0アルゴリズムであり、これらはそれぞれ、Altschulら (1977) Nucl. Acids Res. 25:3389-3402およびAltschulら (1990) J. Mol. Biol. 215:403-410に記載される。BLASTおよびBLAST2.0は、本発明のポリヌクレオチドについての配列同一性の割合を決定するために、例えば、本明細書中に記載のパラメータを用いて使用され得る。BLAST分析を行うためのソフトウェアは、National Center for Biotechnology Informationを通して公に 30
利用可能である。1つの例示的な例において、累積スコアは、ヌクレオチド配列について、パラメーターM (一致する残基の対についての報酬スコア (reward score) ; 常に >0) およびN (ミスマッチ残基についてのペナルティスコア ; 常に <0) を使用して算出され得る。各指示におけるワードヒットの拡大は、以下の場合に停止する : 累積整列スコアが、その最大到達値から量Xだけ低下した場合 ; 累積スコアが、1以上の負のスコアの残基整列の累積に起因してゼロ以下になった場合 ; またはいずれかの配列の末端に到達した場合。このBLASTアルゴリズムパラメーターW、TおよびXは、整列の感度および速度を決定する。BLASTNプログラム (ヌクレオチド配列について) は 40
、デフォルトとして、ワード長さ (W) 11、および期待値 (E) 10、そしてBLOSUM62スコアリングマトリックス (HenikoffおよびHenikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915を参照のこと) アラインメント、(B) 50、期待値 (E) 10、M=5、N=-4、および両方の鎖の比較を使用する。

【0110】

好ましくは、「配列同一性のパーセンテージ」は、少なくとも20の位置の比較ウィンドウにわたる2つの最適に整列した配列を比較することによって決定され、ここで比較ウィンドウ中のポリヌクレオチド配列の一部は、参照配列 (これは、付加または欠失を含まない) と比較して、2つの配列の最適な整列について20パーセント以下、通常は5~15 50

パーセント、または10～12パーセントの付加または欠失（すなわち、ギャップ）を含み得る。このパーセンテージは、両方の配列で同一の核酸塩基が生じる位置の数を決定して一致する位置の数を得、この一致する位置の数を参照配列中の位置の総数（すなわち、ウィンドウサイズ）で除算し、そしてこの結果に100を掛けて配列同一性のパーセンテージを得ることによって算出される。

【0111】

遺伝コードの縮重の結果として、本明細書中に記載されるようなポリペプチドをコードする多くのヌクレオチド配列が存在することが、当業者に理解される。これらのポリヌクレオチドのうちいくつかは、任意のネイティブな遺伝子のヌクレオチド配列に対して最小の相同性を有する。それにもかかわらず、コドンの用法における差異に起因して変化するポリヌクレオチドは、本発明によって具体的に意図される。さらに、本明細書中に提供されるポリヌクレオチド配列を含む遺伝子の対立遺伝子は、本発明の範囲内である。対立遺伝子は、1以上の変異（例えば、ヌクレオチドの欠失、付加および／または置換）の結果として変化する内因性遺伝子である。得られたmRNAおよびタンパク質は、変化した構造または機能を有し得るが、有する必要はない。対立遺伝子は、標準的な技術（例えば、ハイブリダイゼーション、増幅および／またはデータベース配列比較）を用いて、同定され得る。

10

【0112】

従って、本発明の別の実施形態において、変異誘発アプローチ（例えば、部位特異的変異誘発）は、本明細書中に記載のポリペプチドの免疫原性改変体および／または誘導体の調製のために使用される。このアプローチによって、ポリペプチド配列における特定の改変が、これらをコードする、基礎となるポリヌクレオチドの変異誘発を介してなされ得る。これらの技術は、例えば、ポリヌクレオチド中に1つ以上のヌクレオチド配列変化を導入することによって先の考慮の1つ以上を組み込む、配列改変体を調製および試験するための直接的なアプローチを提供する。

20

【0113】

部位特異的変異誘発は、所望の変異のDNA配列をコードする特定のオリゴヌクレオチド配列ならびに十分な数の隣接ヌクレオチドの使用を介して変異体の生成を可能にし、相対する欠失連結部の両側に安定な二重鎖を形成するために、十分なサイズおよび配列複雑性のプライマー配列を提供する。変異は、ポリヌクレオチド自体の特性を改善するか、変更するか、減少させるか、改変するか、さもなければ変化させるため、そして／またはコードされたポリペプチドの特性、活性、組成、安定性または一次配列を変更するために、選択されたポリヌクレオチド配列において使用され得る。

30

【0114】

本発明の特定の実施形態において、本発明者らは、コードされたポリペプチドの1つ以上の特性（例えば、ポリペプチドワクチンの免疫原性）を変更するために、開示されたポリヌクレオチド配列の変異誘発を意図する。部位特異的変異誘発の技術は、当該分野に周知であり、ポリペプチドおよびポリヌクレオチドの両方の改変体を作製するために広範に使用される。例えば、部位特異的変異誘発は、しばしば、DNA分子の特定部分を変更するために使用される。このような実施形態において、代表的に約14ヌクレオチド長～約25ヌクレオチド長程度の長さを含むプライマーが使用され、配列の連結部の両側にある、約5残基～約10残基が変更される。

40

【0115】

当業者によって理解されるように、部位特異的変異誘発技術は、しばしば、一本鎖形態および二本鎖形態の両方で存在するファージベクターを使用してきた。部位特異的変異誘発において有用である代表的なベクターとしては、M13ファージのようなベクターが挙げられる。これらのファージは、容易に商業的に入手可能であり、そしてそれらの使用は、一般的に当業者に周知である。二本鎖プラスミドもまた、目的の遺伝子をプラスミドからファージに転移する工程を除去する部位特異的変異誘発において慣用的に使用される。

【0116】

50

一般的に、本明細書に従う部位特異的変異誘発は、所望のペプチドをコードするDNA配列をその配列中に含む一本鎖ベクターを最初に入手する工程、またはこの配列を含む二本鎖ベクターの2本の鎖を融解して分ける工程によって実行される。所望の変異配列を保有するオリゴヌクレオチドプライマーは、一般的に、合成的に調製される。次いで、このプライマーは、一本鎖ベクターとともにアニーリングされ、そして変異を保有する鎖の合成を完了するために、DNA重合酵素（例えば、E. coli ポリメラーゼ I Klenowフラグメント）に供される。このようにヘテロ二重鎖が形成され、ここで一方の鎖が変異を有さないもとの配列をコードし、そして2つめの鎖は所望の変異を保有する。次いで、このヘテロ二重鎖ベクターは、適切な細胞（例えば、E. coli 細胞）を形質転換するために使用され、そして、変異した配列配置を保有する組換えベクターを含むクローンが選択される。 10

【0117】

部位特異的変異誘発を使用する、選択されたペプチドコードDNAセグメントの配列改変体の調製は、潜在的に有用な種を産生する手段を提供し、そしてこれは、限定を意味するものではない。なぜなら、ペプチドの配列改変体およびそれらをコードするDNA配列を入手し得る他の方法が存在するからである。例えば、所望のペプチド配列をコードする組換えベクターは、変異誘発剤（例えば、ヒドロキシルアミン）で処理されて、配列改変体を入手し得る。これらの方法およびプロトコルに関する具体的な詳細は、Malloyら、1994；Segal、1976；ProkopおよびBajpai、1991；Kuby、1994；およびManiatisら、1982（各々がその目的のために本明細書中に参考として援用される）の教示において見出される。 20

【0118】

本明細書中で使用される場合、用語「オリゴヌクレオチド特異的変異誘発手順」とは、テンプレート（鋳型）依存的プロセスおよびベクター媒介性増殖をいい、これは、その初期の濃度と比較して、特定の核酸分子の濃度の増加を生じるか、または検出可能なシグナルの濃度の増加（例えば、増幅）を生じる。本明細書中で使用される場合、用語「オリゴヌクレオチド特異的変異誘発手順」は、プライマー分子の鋳型依存的な伸長を含むプロセスをいうことが意図される。用語、鋳型依存的プロセスとは、新規に合成された核酸の鎖の配列が、相補的塩基対形成の周知の規則によって指示されるRNA分子またはDNA分子の核酸合成をいう（例えば、Watson、1987を参照のこと）。代表的には、ベクター媒介性 30の方法論は、DNAまたはRNAベクターへの核酸フラグメントの導入、ベクターのクローン性増幅、および増幅した核酸フラグメントの回収を包含する。このような方法論の例は、その全体が具体的に参考として本明細書中に援用される、米国特許第4,237,224号によって提供される。

【0119】

本発明のポリペプチド改変体の生成のための別のアプローチにおいて、米国特許第5,837,458号に記載のような、再帰的な配列組換えが使用され得る。このアプローチにおいて、組換えおよびスクリーニングまたは選択の反復性のサイクルが実施されて、例えば、増強された免疫原性活性を有する本発明の個々のポリヌクレオチド改変体を「展開させる」。 40

【0120】

本発明の他の実施形態において、本明細書中に提供されるポリヌクレオチド配列は、核酸ハイブリダイゼーションのためのプローブまたはプライマーとして、有利に使用され得る。このように、本明細書中に開示される15ヌクレオチド長の連続配列と同じ配列か、またはこれに相補的な配列を有する、少なくとも約15ヌクレオチド長の連続配列の配列領域を含む核酸セグメントが特定の有用性を見出すことが、意図される。例えば、約20、30、40、50、100、200、500、1000（全ての中間的な長さのものを含む）およびまさに全長配列までの、より長い連続した同一配列または相補的な配列がまた、特定の実施形態において使用される。

【0121】

このような核酸プローブが目的の配列に特異的にハイブリダイズする能力は、所定のサンプル中の相補的配列の存在を検出する際にこれらが使用されることを可能にする。しかし、他の用途（例えば、変異種プライマーまたは他の遺伝的構築物の調製の際の使用のためのプライマーを調製するための配列情報の使用）もまた意図される。

【0122】

10～14、15～20、30、50、または100～200ヌクレオチド程度（同様に中間的な長さを含む）の連続したヌクレオチドストレッチからなる配列領域を有し、本明細書中に開示されるポリヌクレオチド配列に同一または相補的なポリヌクレオチド分子は、例えばサザンブロッティングまたはノーザンブロッティングにおける使用のためのハイブリダイゼーションプローブとして特に意図される。これによって、遺伝子産物またはそのフラグメントの、多様な細胞型およびまた種々の細菌細胞の両方における分析が可能となる。フラグメントの全サイズ、ならびに相補的ストレッチのサイズは、究極的には、特定の核酸セグメントの意図される用途または適用に依存する。より小さなフラグメントは、一般にハイブリダイゼーション実施形態における用途を見出し、ここで連続した相補領域の長さが増加され得る（例えば、約15ヌクレオチドと約100ヌクレオチドとの間）が、検出を望む相補配列の長さに従って、より長い連続した相補ストレッチが使用され得る。

10

【0123】

約15～25ヌクレオチド長のハイブリダイゼーションプローブの使用は、安定かつ選択的な二重鎖分子の形成を可能にする。しかし、15塩基長より長いストレッチの連続した相補配列を有する分子が、ハイブリッドの安定性および選択性を増加するために一般に好ましく、それによって、得られる特異的ハイブリッド分子の質および程度が改善される。15～25の連続したヌクレオチド、または望まれる場合にはより長い遺伝子相補的ストレッチを有する核酸分子を設計するのが一般には好ましい。

20

【0124】

ハイブリダイゼーションプローブは、本明細書中に開示される配列のいずれかの、任意の部分から選択され得る。必要とされることの全ては、本明細書中に記載の配列、あるいはプローブまたはプライマーとしての利用を望む場合は、約15～25ヌクレオチド長から全長配列まで、および全長配列を含む配列の任意の連続した部分までを再検討することである。プローブおよびプライマー配列の選択は、種々の要因によって支配され得る。例えば、全配列の末端に向かってプライマーを使用することが望まれ得る。

30

【0125】

小さなポリヌクレオチドセグメントまたはフラグメントは、通常は自動オリゴヌクレオチド合成機を使用して行われるので、例えば化学的手段によってフラグメントを直接合成することによって、容易に調製され得る。また、フラグメントは、核酸複製技術（例えば、米国特許第4,683,202号（本明細書中に参考として援用される）のPCRTM技術）の適用によって、組換え産生のために選択配列を組換えベクターに導入することによって、および一般に分子生物学の分野の当業者に公知の他の組換えDNA技術によって、得られ得る。

【0126】

本発明のヌクレオチド配列は、目的の完全遺伝子または遺伝子フラグメントのいずれかの相補的ストレッチと、二重鎖分子を選択的に形成するその能力のために使用され得る。想定される適用に依存して、代表的には、標的配列に対するプローブの選択性の種々の程度を達成するために、ハイブリダイゼーションの種々の条件を使用することが望ましい。高い選択性を必要とする適用について、代表的には、ハイブリッドを形成するための比較的ストリンジェントな条件を使用することが望ましく、例えば、比較的低濃度の塩、および／または高温の条件（例えば、約50℃～約70℃の温度で、約0.02M～約0.15Mの塩の塩濃度によって提供されるような）を選択する。このような選択条件は、存在する場合、プローブとテンプレートまたは標的鎖との間のミスマッチにわずかに許容性であり、そして関連する配列の単離のために特に適切である。

40

50

【0127】

当然のことながら、いくつかの適用（例えば、潜在的なテンプレートにハイブリダイズする変異体プライマー鎖を使用する、変異体の調製が望ましい場合）について、低ストリンジェント（減少したストリンジェンシー）のハイブリダイゼーション条件が、ヘテロ二重鎖を形成するために代表的に必要とされる。これらの状況において、約20℃～約55℃の温度範囲で約0.15M～約0.9Mの塩の条件のような塩条件を使用することが望ましくあり得る。これによって、交差ハイブリダイズ種は、コントロールハイブリダイゼーションに関して、陽性ハイブリダイズシグナルとして容易に同定され得る。任意の場合において、ホルムアミド（これは、増加した温度と同じ様式で、ハイブリッド二重鎖を不安定化するように作用する）の増加した量の添加によって、条件がよりストリンジェントになり得ることが一般に理解される。従って、ハイブリダイゼーション条件は、容易に操作され得、従って、一般に、所望の結果に依存する最良の方法である。

10

【0128】

本発明の別の実施形態に従って、アンチセンスオリゴヌクレオチドを含むポリヌクレオチド組成物が、提供される。アンチセンスオリゴヌクレオチドは、タンパク質合成の効果的かつ標識化インヒビターであることが実証され、結果として、疾患に寄与するタンパク質の合成を阻害することによって疾患が処置され得る治療的なアプローチを提供する。タンパク質合成を阻害するためのアンチセンスオリゴヌクレオチドの効力は、よく確立されている。例えば、ポリガラクタウロナーゼ（polygalacturonase）およびムスカリン2型アセチルコリンレセプターの合成は、そのそれぞれのmRNA配列に対するアンチセンスオリゴヌクレオチドによって阻害される（米国特許第5,739,119号および米国特許第5,759,829号）。さらに、アンチセンス阻害の例が、核タンパク質サイクリン、多剤耐性遺伝子（MDG1）、ICAM-1、E-セレクチン、STK-1、線条体GABAAレセプターおよびヒトEGFを用いて示されている（Jaskulskiら、Science 1988 Jun 10; 240(4858): 1544-6; VasanthakumarおよびAhmed、Cancer Commun. 1989; 1(4): 225-32; Perisら、Brain Res Mol Brain Res. 1998 Jun 15; 57(2): 310-20; 米国特許第5,801,154号; 同第5,789,573号; 同第5,718,709号および同第5,610,288号）。種々の異常な細胞増殖（例えば、癌）を阻害しそして処置するために使用され得る、アンチセンス構築物もまた、記載されている（米国特許第5,747,470号; 同第5,591,317号および同第5,783,683号）。

20

30

【0129】

従って、特定の実施形態において、本発明は、本明細書中に記載されるポリヌクレオチド配列に特異的に結合し得る任意の配列もしくはその相補体のすべてまたは一部を含む、オリゴヌクレオチド配列を提供する。1つの実施形態において、そのアンチセンスオリゴヌクレオチドは、DNAまたはその誘導体を含む。別の実施形態において、そのオリゴヌクレオチドは、RNAまたはその誘導体を含む。第3の実施形態において、そのオリゴヌクレオチドは、ホスホロチオエート改変骨格を含む、改変DNAである。第4の実施形態において、そのオリゴヌクレオチド配列は、ペプチド核酸またはその誘導体を含む。各場合において、好ましい組成物は、本明細書中に開示されるポリヌクレオチドのうちの1つ以上の部分に相補的な、より好ましくは実質的に相補的な、そしてさらにより好ましくは完全に相補的な、配列領域を含む。所定の遺伝子配列に特異的なアンチセンス組成物の選択は、選択された標的配列の分析に基き、そして二次構造、T_m、結合エネルギー、および相対的安定性の決定に基く。アンチセンス組成物は、二量体、ヘアピン、または宿主細胞において標的mRNAへの特異的結合を減少または妨げる他の二次構造をそれらが形成できない相対的能力に基づいて選択され得る。mRNAの非常に好ましい標的領域は、AUG翻訳開始コドンの領域またはその付近の領域、およびmRNAの5'領域と実質的に相補的な配列である。これらの二次構造分析および標的部位選択の考慮は、例えば、OLIGOプライマー分析ソフトウェアのv. 4および/またはBLASTN 2.0.5アル

40

50

ゴリズムソフトウェア (A l t s c h u l ら、N u c l e i c A c i d s R e s . 1 9 9 7 , 2 5 (1 7) : 3 3 8 9 - 4 0 2) を使用して実施され得る。

【0130】

短いペプチドベクター (M P G (2 7 残基) と称する) を使用するアンチセンス送達法の使用もまた意図される。このM P G ペプチドは、H I V g p 4 1 の融合配列由来の疎水性ドメインと、S V 4 0 T 抗原の核局在化配列由来の親水性ドメインとを含む (M o r r i s ら、N u c l e i c A c i d s R e s . 1 9 9 7 J u l 1 5 ; 2 5 (1 4) : 2 7 3 0 - 6) 。このM P G ペプチドのいくつかの分子がアンチセンスオリゴヌクレオチドをコートし、そして比較的高効率 (9 0 %) で 1 時間未満で培養哺乳動物細胞へと送達され得ることが示された。さらに、M P G との相互作用は、ヌクレアーゼに対するそのオリゴヌクレオチドの安定性および形質膜を横切る能力の両方を強力に増加させる。 10

【0131】

本発明の別の実施形態に従って、本明細書中に記載されるポリヌクレオチド組成物は、腫瘍細胞における腫瘍ポリペプチドおよび本発明のタンパク質の発現を阻害するためのリボザイム分子の設計および調製に使用される。リボザイムは、部位特異的様式で核酸を切断するRNA-タンパク質複合体である。リボザイムは、エンドヌクレアーゼ活性を保有する特定の触媒ドメインを有する (K i m および C e c h 、P r o c N a t l A c a d S c i U S A . 1 9 8 7 D e c ; 8 4 (2 4) : 8 7 8 8 - 9 2 ; F o r s t e r および S y m o n s 、C e l l . 1 9 8 7 A p r 2 4 ; 4 9 (2) 2 1 1 - 2 0) 。例えば、多数のリボザイムが、高い程度の特異性でホスホエステル転移反応を促進し、しばしば、オリゴヌクレオチド基質中の数個のホスホエステルのうちの 1 つだけを切断する (C e c h ら、C e l l . 1 9 8 1 D e c ; 2 7 (3 P t 2) : 4 8 7 - 9 6 ; M i c h e l および W e s t h o f 、J M o l B i o l . 1 9 9 0 D e c 5 ; 2 1 6 (3) : 5 8 5 - 6 1 0 ; R e i n h o l d - H u r e k および S h u b 、N a t u r e . 1 9 9 2 M a y 1 4 ; 3 5 7 (6 3 7 4) : 1 7 3 - 6) 。この特異性は、この基質が、化学反応の前にリボザイムの内部ガイド配列 (「 I G S 」) との特異的塩基対形成相互作用を介して結合するという要件に帰せられている。 20

【0132】

天然に存在する酵素的RNAの6つの基本的変種が、現在公知である。各々が、生理学的条件下で、トランスで、RNAホスホエステル結合の加水分解を触媒し得る (従って、他のRNA分子を切断し得る) 。一般に、酵素的核酸は、まず、標的RNAへの結合によって作用する。このような結合は、標的RNAを切断するように作用する分子の酵素的部分に近接して保持される、酵素的核酸の標的結合部分を介して生じる。従って、その酵素的核酸はまず、標的RNAを認識し、次いで相補的塩基対形成を介してその標的RNAに結合し、そして一旦正確な部位に結合すると、その標的RNAを酵素的に切断するように作用する。このような標的RNAの戦略的切断は、コードされるタンパク質を直接合成する標的RNAの能力を破壊する。酵素的核酸がそのRNA標的に結合しそして切断した後、その酵素的核酸は、別の標的を探索するためにそのRNAから離れ、そして繰り返し、新しい標的に結合しそして切断し得る。 30

【0133】

リボザイムの酵素特性は、多くの技術 (例えば、アンチセンス技術 (ここで、核酸分子は、翻訳を阻害するために核酸標的に簡単に結合する)) に対して有利である。なぜなら、治療的処置に影響を与えるのに必要であるリボザイムの濃度は、アンチセンスオリゴヌクレオチの濃度よりも低いからである。この利点は、酵素学的に作用するリボザイムの能力を反映する。従って、単一のリボザイム分子は、標的RNAの多くの分子を切断し得る。さらに、このリボザイムは、高度に選択的なインヒビターであり、この阻害の特異性は、標的RNAに結合する塩基対形成の機構に依存するだけでなく、標的RNAの切断の機構にも依存する。切断部位付近の 1 個のミスマッチまたは塩基置換により、リボザイムの触媒活性が完全になくなり得る。アンチセンス分子内の同様のミスマッチでは、それらの作用は妨げられない (W o o l f ら、P r o c N a t l A c a d S c i U S A 40

． 1992 Aug 15 ; 89 (16) : 7305-9) 。従って、リボザイムの作用の特異性は、同じRNA部位に結合するアンチセンスオリゴヌクレオチドの作用の特異性よりも高い。

【0134】

この酵素学的核酸分子は、ハンマーヘッド (hammer head) モチーフ、ヘアピンモチーフ、 δ 型肝炎ウイルスモチーフ、I群イントロンモチーフまたはRNase P RNAモチーフ (RNA誘導配列に関連する) あるいはNeurospora VS RNAモチーフに形成され得る。ハンマーヘッドモチーフの例は、Rossiら, Nucleic Acids Res. 1992 Sep 11 ; 20 (17) : 4559-65に記載される。ヘアピンモチーフの例は、Hampelら (欧州特許出願公開番号EP 0 360257) 、HampelおよびTritz (Biochemistry 1989 Jun 13 ; 28 (12) : 4929-33) 、Hampelら, Nucleic Acids Res. 1990 Jan 25 ; 18 (2) : 299-304および米国特許第5, 631, 359号に記載される。 δ 型肝炎ウイルスモチーフの例は、PerottaおよびBeen, Biochemistry. 1992 Dec 1 ; 31 (47) : 11843-52に記載される ; RNase Pモチーフの例は、Guerrier-Takadara, Cell. 1983 Dec ; 35 (3 Pt 2) : 849-57に記載される ; Neurospora VS RNAリボザイムモチーフは、Collins (SavilleおよびCollins, Cell. 1990 May 18 ; 61 (4) : 685-96 ; SavilleおよびCollins, Proc Natl Acad Sci U S A. 1991 Oct 1 ; 88 (19) : 8826-30 ; CollinsおよびOlive, Biochemistry. 1993 Mar 23 ; 32 (11) : 2795-9) に記載される ; そしてI群イントロンの例は、米国特許第4, 987, 071号に記載される。本発明の酵素学的核酸分子において重要であることは、1以上の標的遺伝子RNA領域に相補的である特定の基質結合部位を有すること、およびこの分子にRNA切断活性を付与する、基質結合部位内のヌクレオチド配列またはこの基質結合部位の周囲のヌクレオチド配列を有することが全てである。従って、リボザイム構築物は、本明細書中に記載される特定のモチーフに限定される必要はない。

【0135】

リボザイムは、国際特許出願公開番号WO 93 / 23569および国際特許出願公開番号WO 94 / 02595 (各々は、本明細書中で参考として詳細に援用される) に記載されるように設計され得、そして記載されるようにインビトロおよびインビボにおいて試験するために合成され得る。このようなリボザイムはまた、送達するために最適化され得る。特定の例が提供されるが、当業者は、他の種において等価なRNA標的が、必要な場合に利用され得ることを理解する。

【0136】

リボザイムの活性は、リボザイムの結合アームの長さを変更することにより、または血清リボヌクレアーゼによる分解を防ぐ改変 (例えば、国際特許出願番号WO 92 / 07065 ; 国際特許公開番号WO 93 / 15187 ; 国際特許出願公開番号WO 91 / 03162 ; 欧州特許出願公開番号92110298. 4 ; 米国特許第5, 334, 711号、および国際特許出願公開番号WO 94 / 13688 (これには、酵素学的RNA分子の糖部分に対して成され得る、種々の化学的修飾が記載される) を参照のこと) 、細胞中でのリボザイムの有効性を増強する改変、ならびにRNAの合成時間を短縮および化学的必要性を低下させるための幹 (stem) II塩基の除去を用いて、リボザイムを化学的に合成することによって最適化され得る。

【0137】

Sullivanら (国際特許出願公開番号WO 94 / 02595) は、酵素学的RNA分子の送達のための一般的な方法を記載する。リボザイムは、当業者に公知の種々の方法によって細胞に投与され得、これらには、リボソームへのカプセル化、イオン泳動法、あるいは、ヒドロゲル、シクロデキストリン、生分解性ナノカプセルおよび生体接着性ミ

クロスフィアのような他のビヒクルへの取り込みによるもの、が挙げられるが、これらに限定されない。いくつかの適応症については、リボザイムは、エキソビボで細胞または組織に、上記のビヒクルを使用してか、または使用せずに直接送達され得る。あるいは、RNA／ビヒクルの組合せは、直接的な吸入、直接的な注射、またはカテーテル、注入ポンプまたはステントの使用により、局所的に送達され得る。他の送達経路としては、血管内注射、筋内注射、皮下注射または関節注射、エアロゾル吸入、経口送達（錠形態または丸薬形態）、局所送達、全身送達、眼送達、腹腔内送達および／またはくも膜下腔内送達（10）が挙げられるが、これらに限定されない。リボザイム送達および投与のより詳細な記載は、国際特許出願公開番号WO 94／02595および国際特許出願公開番号WO 93／23569（各々は、本明細書中で参考として詳細に援用される）に提供される。

【0138】

高濃度のリボザイムを細胞内に蓄積する別の手段は、リボザイムコード配列をDNA発現ベクターに組み込むことである。リボザイム配列の転写は、真核生物RNAポリメラーゼI（pol I）、RNAポリメラーゼII（pol II）、またはRNAポリメラーゼIII（pol III）のためのプロモーターにより駆動される。pol IIプロモーターまたはpol IIIプロモーターからの転写物は、全ての細胞内で高レベルで発現される；所定の細胞型における所定のpol IIプロモーターのレベルは、近くに存在する遺伝子調節配列（エンハンサー、サイレンサーなど）の性質に依存する。原核生物のRNAポリメラーゼプロモーターもまた使用され得るが、但し、原核生物RNAポリメラーゼ酵素は、適切な細胞において発現される。このようなプロモーターから発現する（20）リボザイムは、哺乳動物細胞において機能することが示されている。このような転写ユニットは、哺乳動物細胞内への導入のための種々のベクター（プラスミドDNAベクター、ウイルスDNAベクター（例えば、アデノウイルスベクターまたはアデノ随伴ベクター）、またはウイルスRNAベクター（例えば、レトロウイルスベクター、セムリキ森林ウイルスベクター、シンドビスウイルスベクター）が挙げられるがこれらに限定されない）に組み込まれ得る。

【0139】

本発明の別の実施形態において、ペプチド核酸（PNA）組成物が提供される。PNAは、DNA模倣物であり、核酸塩基は、偽ペプチド骨格に結合される（GoodおよびNielsen、Antisense Nucleic Acid Drug Dev. 1997 7（4）431～37）。PNAは、RNAまたはDNAを伝統的に使用した多くの方法において利用され得る。しばしば、PNA配列は、対応するRNA配列またはDNA配列よりも技術的により良く機能し、そしてRNAにもDNAにも固有でない有用性を有する。産生方法、特徴、および使用方法を含むPNAの総説は、Corey（Trends Biotechnol 1997 6月；15（6）：224～9）によって提供される。このように、特定の実施形態において、ACE mRNA配列の1以上の部分に相補的であるPNA配列を調製し得、そしてこのようなPNA組成物は、ACE-特異的mRNAの翻訳を調節、変更、減少、または低下させるために使用され得、これにより、（30）このようなPNA組成物が投与された宿主細胞におけるACE活性のレベルを変更する。

【0140】

PNAは、DNAの正常なホスホジエステル骨格を置換する2-アミノエチル-グリシン連結を有する（Nielsenら、Science 1991 Dec 6；254（5037）：1497～500；Hanveyら、Science. 1992 Nov 27；258（5087）：1481～5；HyrupおよびNielsen, Bioorg Med Chem. 1996 Jan；4（1）：5～23）。この化学は、3つの重要な結果を有する。第1に、DNAまたはホスホロチオエートオリゴヌクレオチドとは対照的に、PNAは、中性の分子であり；第2に、PNAは、アキラル（立体選択的な合成を開発する必要を避ける）であり；そして第3に、PNA合成は、標準的なBocプロトコルまたはFmocプロトコルを、固相ペプチド合成に使用する。一方、改良型Merrifield法を含む他の方法が使用されている。（40）

10

20

30

40

50

【0141】

PNAモノマーまたは既製のオリゴマーは、PerSeptive Biosystems (Framingham, MA) から市販されている。BocプロトコルまたはFmocプロトコルのいずれかによるPNAの合成は、手動プロトコルまたは自動プロトコルを使用して簡単に実施される (Nortonら、Bioorg Med Chem. 1995 Apr; 3 (4): 437~45)。この手動 (マニュアル) プロトコルは、それ自体で、化学的に修飾されたPNAの産生または密接に関係するPNAのファミリーの同時合成を導く。

【0142】

ペプチドの合成に関して、特定のPNA合成の成功は、選択した配列の特性に依存する。例えば、理論上、PNAは、ヌクレオチド塩基の任意の組合せを取り込み得るが、隣接するプリンの存在は、産物中に1以上の残基の欠失を導き得る。この困難性を見込みにおいて、隣接するプリンを有するPNAを産生する際に、付加されているようである残基の連結を非効率的に繰り返すべきであることが示唆される。この後、ペプチド合成の間に観測されるのと類似する産物の収率および純度を提供する逆相高圧液体クロマトグラフィーによってPNAを精製すべきである。

【0143】

所与の適用のためのPNAの修飾が、固相合成の間にアミノ酸を連結することによって、または露出されたN末端アミンにカルボン酸基を含む化合物を付加することによって達成され得る。あるいは、PNAは、導入されたリジンまたはシステインに連結することによって合成した後に修飾され得る。PNAが修飾され得る簡便さによって、より良い溶解度または特定の機能的必要条件の最適化が容易になる。一旦合成されると、PNAおよびそれらの誘導体の同定は、質量分析法によって確認され得る。いくつかの研究は、PNAの修飾物を作製および利用した (例えば、Nortonら、Bioorg Med Chem. 1995 Apr; 3 (4): 437~45; Petersenら、J Pept Sci. 1995 May-Jun; 1 (3): 175~83; Orumら、Biotechniques. 1995 Sep; 19 (3): 472~80; Footerら、Biochemistry. 1996 Aug 20; 35 (33): 10673~9; Griffithら、Nucleic Acid Res. 1995 Aug 11; 23 (15): 3003~8; Pardridgeら、Proc Natl Acad Sci USA 1995 Jun 6; 92 (12): 5592~6; Boffaら、Proc Natl Acad Sci USA 1995 Mar 14; 92 (6): 1901~5; Gambacorti-Passeriniら、Blood. 1996 Aug 15; 88 (4): 1411~7; Armitageら、Proc Natl Acad Sci USA. 1997 Nov 11; 94 (23): 12320~5; Seegerら、Biotechniques. 1997 Sep; 23 (3): 512~7)。米国特許第5,700,922号は、PNA-DNA-PNAキメラ分子および診断におけるその分子の使用、生物体におけるタンパク質の調節、ならびに治療に対して影響を受け易い状態の処置について議論する。

【0144】

PNAのアンチセンス結合特性を特徴付けする方法は、Rose (Anal Chem. 1993 Dec. 15; 65 (24): 3545~9) およびJensenら (Biochemistry. 1997 Apr 22; 36 (16): 5072~7) において議論される。Roseは、キャピラリーゲル電気泳動を使用して、相対的な結合速度論および化学量論を測定することで、PNAのそれらの相補的オリゴヌクレオチドに対する結合を決定する。同様のタイプの測定が、BIAcoreTM 技術を使用してJensenらによってなされた。

【0145】

記載され、そして当業者に明らかであるPNAの他の適用には、DNA鎖の侵入、アンチセンス阻害、変異分析、転写のエンハンサー、核酸の精製、転写活性遺伝子の単離、転写

10

20

30

40

50

因子結合の阻害、ゲノムの切断、バイオセンサー、インサイチュハイブリダイゼーションなどにおける使用が挙げられる。

【0146】

(ポリヌクレオチドの同定、特徴づけおよび発現)

本発明のポリヌクレオチド組成物は、種々の十分に確立された技術のいずれかを使用して、同定、調製および／または操作され得る（一般に、Sambrookら、Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989、および他の類似の参考文献を参照のこと）。例えば、ポリヌクレオチドは、cDNAのマイクロアレイを腫瘍に関連する発現（すなわち、本明細書中で提供される代表的なアッセイを使用して決定する場合、正常な組織における発現よりも、腫瘍での少なくとも2倍高い発現）についてスクリーニングすることによって、以下にさらに詳細に記載されるように、同定され得る。このようなスクリーニングは、例えば、製造者の指示に従って、Affymetrix, Inc. (Santa Clara, CA)のマイクロアレイ技術を使用して、（そして本質的にSchenarら、Proc. Natl. Acad. Sci. USA 93:10614~10619, 1996およびHellerら、Proc. Natl. Acad. Sci. USA 94:2150~2155, 1997に記載されるように）、行われ得る。あるいは、ポリヌクレオチドは、本明細書中に記載されるタンパク質を発現する細胞（例えば、腫瘍細胞）から調製されるcDNAから増幅され得る。

10

20

【0147】

多数の鋳型（テンプレート）依存性プロセスが、サンプル中に存在する目的の標的配列を増幅するために利用可能である。最もよく知られた増幅方法の1つは、ポリメラーゼ連鎖反応（PCRTM）であり、これは、米国特許第4,683,195号、同第4,683,202号、および同第4,800,159号に詳細に記載され、これらの各々はその全体が本明細書中に参考として援用される。手短に言えば、PCRTMにおいては、標的配列の向かい合った相補鎖上の領域に対して相補的な2つのプライマー配列が調製される。過剰のデオキシヌクレオシド三リン酸を、DNAポリメラーゼ（例えば、Taqポリメラーゼ）とともに反応混合液に添加する。標的配列がサンプル中に存在する場合、プライマーはその標的に結合し、そしてポリメラーゼは、ヌクレオチドを付加することによって標的配列に沿ってプライマーを伸長させる。反応混合液の温度を上昇および下降させることによって、伸長したプライマーは、標的から解離して反応生成物を形成し、過剰のプライマーは標的および反応生成物に結合し、そしてこのプロセスが反復される。好ましくは、逆転写およびPCRTM増幅手順が、増幅されたmRNAの量を定量するために実行され得る。ポリメラーゼ連鎖反応方法論は、当該分野で周知である。

30

【0148】

多くの他のテンプレート依存性プロセスのいずれか（この多くは、PCRTM増幅技術のバリエーションである）は、当該分野で容易に知られており、利用可能である。例示として、いくつかのこのような方法としては、リガーゼ連鎖反応（LCRともいわれる）（例えば、欧州特許出願公開番号320,308号、米国特許第4,883,750号において記載される）；Q β レプリカーゼ（PCT国際特許出願公開番号PCT/US87/00880に記載される）；鎖置換増幅（Strand Displacement Amplification）（SDA）および修復鎖反応（Repair Chain Reaction）（RCR）が挙げられる。なお他の増幅法は、英国特許出願番号2202328号およびPCT国際特許出願公開番号PCT/US89/01025において記載される。他の核酸増幅手順には、転写に基づく増幅系（TAS）（PCT国際特許出願公開番号WO88/10315）が挙げられ、これには、核酸配列に基づく増幅（NASBA）および3SRが含まれる。欧州特許出願公開番号329,822は、一本鎖RNA（「ssRNA」）、ssDNA、および二本鎖DNA（dsDNA）をサイクル的に合成する工程を包含する核酸増殖プロセスを記載する。PCT国際特許出願公開番号

40

50

WO 89/06700は、プロモーター／プライマー配列の標的一本鎖DNA（「ssDNA」）へのハイブリダイゼーションに基づく核酸配列増幅スキーム、引き続くこの配列の多くのRNAコピーの転写を記載する。「RACE」（Frohman、1990）および「片側（one-sided）PCR」（Ohara、1989）のような他の増幅方法がまた、当業者に周知である。

【0149】

本発明のポリヌクレオチドの増幅した部分を使用して、適切なライブラリー（例えば、腫瘍cDNAライブラリー）から周知の技術を使用して全長遺伝子を単離し得る。このような技術において、増幅に適した1以上のポリヌクレオチドプローブまたはプライマーを使用して、ライブラリー（cDNAライブラリーまたはゲノムライブラリー）をスクリーニングする。好ましくは、ライブラリーはより大きな分子を含むようにサイズが選択される。ランダムプライムしたライブラリー（random primed library）もまた、遺伝子の5'領域および上流領域の同定ために好ましくあり得る。ゲノムライブラリーは、イントロンを入手することおよび5'配列を伸長させることについて好ましい。

10

【0150】

ハイブリダイゼーション技術に関して、部分配列は、周知の技術を使用して標識（例えば、ニックトランスレーションまたは³²Pでの末端標識）され得る。次いで、細菌ライブラリーまたはバクテリオファージライブラリーは、一般に、標識プローブと、変性した細菌コロニー（またはファージプラークを含む菌叢）を含むフィルターとをハイブリダイズすることによってスクリーニングされる（Sambrookら、Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989を参照のこと）。ハイブリダイズするコロニーまたはプラークを選択し、そして増殖させる。そしてそのDNAをさらなる分析のために単離する。cDNAクローンを、付加配列の量を決定するために、例えば、部分配列由来のプライマーおよびそのベクター由来のプライマーを使用するPCRによって分析し得る。制限酵素地図および部分配列を作成して、1以上の重複クローンを同定し得る。次いで、標準的な技術（これは、一連の欠失クローンを作製することを包含し得る）を使用して完全配列を決定し得る。次いで、得られた重複配列を1つの連続配列中に構築し得る。周知の技術を使用して、適切なフラグメントを連結することにより全長cDNA分子を生成し得る。

20

30

【0151】

あるいは、上記のような増幅技術は、部分的なcDNA配列から全長コード配列を得るために有用であり得る。このような増幅技術の1つは、逆（インバース）PCRである（Trigliaら、Nucl. Acids Res. 16:8186, 1988を参照のこと）。この技術は、制限酵素を使用して、その遺伝子の既知の領域内にフラグメントを生成する。次いで、このフラグメントを分子内連結により環化し、そして既知の領域に由来する多岐したプライマーを用いたPCRのためのテンプレート（鋳型）として使用する。代替のアプローチにおいて、部分配列に隣接した配列を、リンカー配列に対するプライマーおよび既知の領域に特異的なプライマーを使用する増幅により取り出し得る。この増幅した配列を、代表的に、同じリンカープライマーおよび既知の領域に特異的な第2のプライマーを使用する2回目の増幅に供する。既知の配列から反対方向に伸長を開始する2つのプライマーを使用するこの手順についての改変は、WO 96/38591に記載される。そのような別の技術は、「迅速なcDNA末端の増幅」またはRACEとして公知である。この技術は、公知配列の5'および3'である配列を同定するために、内部プライマーおよび外部プライマー（これらは、ポリA領域またはベクター配列にハイブリダイズする）の使用を含む。さらなる技術としては、キャプチャーPCR（Langerstromら、PCR Methods Applic. 1:111-19, 1991）およびウォーキングPCR（Parkerら、Nucl. Acids Res. 19:3055-60, 1991）が挙げられる。増幅を利用する他の方法もまた使用して、全長cDNA

40

50

A配列を入手し得る。

【0152】

特定の場合において、発現配列タグ（EST）データベース（例えば、GenBankより利用可能のもの）に提供される配列の分析により、全長cDNA配列を入手することが可能である。重複ESTの検索は、一般に、周知のプログラム（例えば、NCBI BLAST検索）を使用して行われ得、そしてこのようなESTを使用して連続した全長配列を生成し得る。全長DNA配列はまた、ゲノムフラグメントの分析により入手され得る。

【0153】

本発明の他の実施形態において、本発明のポリペプチドまたはその融合タンパク質もしくは機能的等価物をコードするポリヌクレオチド配列またはそのフラグメントは、適切な宿主細胞におけるポリペプチドの発現を指向するように組換えDNA分子において使用され得る。遺伝コードの固有の縮重に起因して、実質的に同じか、または機能的に等価なアミノ酸配列をコードする他のDNA配列が産生され得、そしてこれらの配列を使用して所定のポリペプチドをクローニングし、そして発現させ得る。

10

【0154】

当業者によって理解されるように、いくつかの場合において、天然に存在しないコドンをもつポリペプチドコードヌクレオチド配列を作製することが有利であり得る。例えば、特定の原核生物宿主または真核生物宿主に好まれるコドンは、タンパク質発現の速度を増加させるためか、または所望の特性（例えば、天然に存在する配列から生成される転写物の半減期よりも長い半減期）を有する組換えRNA転写物を産生するように選択され得る。

20

【0155】

さらに、本発明のポリヌクレオチド配列は、種々の理由（遺伝子産物のクローニング、プロセッシング、および／または発現を変更する改変が挙げられるが、それらに限定されない）のためにポリペプチドコード配列を変更するために、当該分野において一般的に公知の方法を使用して操作され得る。例えば、無作為断片化によるDNAシャッフリングならびに遺伝子フラグメントおよび合成オリゴヌクレオチドのPCR再アセンブリを使用して、ヌクレオチド配列を操作し得る。さらに、部位特異的変異誘発（site-directed mutagenesis）を使用して、新たな制限部位を挿入し得るか、グリコシル化パターンを改変し得るか、コドンの優先度（preference）を変化させ得るか、スプライス改変体を作製し得るか、または変異の導入などを行い得る。

30

【0156】

本発明の別の実施形態において、天然の核酸配列、改変された核酸配列、または組換え核酸配列は、融合タンパク質をコードする異種配列に連結され得る。例えば、ポリペプチド活性のインヒビターについてペプチドライブラリーをスクリーニングするために、市販の抗体により認識され得るキメラタンパク質をコードすることが有用であり得る。融合タンパク質はまた、ポリペプチドコード配列と異種タンパク質配列との間に位置する切断部位を含むように操作され得、その結果そのポリペプチドは、切断されて、そして異種部分から精製されて取り出され得る。

【0157】

所望のポリペプチドをコードする配列が、当該分野において周知の化学的方法を使用して、全体または部分的に合成され得る（Caruthers, M. H. ら、(1980) Nucleic Acids Res. Symp. Ser. 215~223、Horn, T. ら、(1980) Nucleic Acids Res. Symp. Ser. 225~232を参照のこと）。あるいは、タンパク質自体は、ポリペプチドのアミノ酸配列またはその一部を合成するための化学的方法を使用して産生され得る。例えば、ペプチド合成は、種々の固相技術を使用して実施され得（Roberge, J. Y. ら、(1995) Science 269:202~204）、そして自動化合成は、例えば、ABI 431A Peptide Synthesizer（Perkin Elmer, Palo Alto, CA）を使用して達成され得る。

40

50

【0158】

新たに合成されたペプチドは、分取高速液体クロマトグラフィー（例えば、Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N. Y.）または当該分野において利用可能な他の同等技術により実質的に精製され得る。合成ペプチドの組成は、アミノ酸分析または配列決定（例えば、エドマン分解手順）により確認され得る。さらに、ポリペプチドのアミノ酸配列またはその任意の部分は、直接合成の間改変されて、そして／または化学的方法を使用して、他のタンパク質もしくはその任意の部分に由来する配列と組み合わせられて、改変体ポリペプチドを産生し得る。

10

【0159】

所望のポリペプチドを発現させるために、ポリペプチドをコードするヌクレオチド配列または機能的等価物は、適切な発現ベクター（すなわち、挿入されたコード配列の転写および翻訳のために必要なエレメントを含むベクター）内に挿入され得る。当業者に周知である方法を使用して、目的のポリペプチドをコードする配列ならびに適切な転写制御エレメントおよび翻訳制御エレメントを含む発現ベクターを構築し得る。これらの方法としては、インビトロ組換えDNA技術、合成技術、およびインビボ遺伝子組換えが挙げられる。そのような技術は、例えば、Sambrook, J. ら、(1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N. Y., および Ausubel, F. M. ら、(1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N. Y. に記載される。

20

【0160】

種々の発現ベクター／宿主系が、ポリヌクレオチド配列を含み、そして発現させるように利用され得る。これらとしては、以下が挙げられるが、これらに限定されない：微生物（例えば、組換えバクテリオファージ、プラスミド、またはコスミドDNA発現ベクターで形質転換した細菌）；酵母発現ベクターで形質転換した酵母；ウイルス発現ベクター（例えば、バキュロウイルス）で感染させた昆虫細胞系；ウイルス発現ベクター（例えば、カリフラワーモザイクウイルス、CaMV；タバコモザイクウイルス、TMV）または細菌発現ベクター（例えば、TiもしくはpBR322プラスミド）で形質転換した植物細胞系；あるいは動物細胞系。

30

【0161】

発現ベクター内に存在する「制御エレメント」または「調節配列」は、転写および翻訳を実行するように宿主細胞タンパク質と相互作用する、それらのベクターの非翻訳領域（エンハンサー、プロモーター、5'および3'非翻訳領域）である。そのようなエレメントは、その長さおよび特異性を変更し得る。利用されるベクター系および宿主に依存して、多数の適切な転写エレメントおよび翻訳エレメント（構成的プロモーターおよび誘導性プロモーターを含む）が使用され得る。例えば、細菌系においてクローニングする場合、誘導性プロモーター（例えば、PBLUESCRIPTファージミド（Stratagene, La Jolla, Calif.）またはPSPORT1プラスミド（Gibco BRL, Gaithersburg, MD）などのハイブリッドlacZプロモーター）が使用され得る。哺乳動物細胞系において、哺乳動物遺伝子または哺乳動物ウイルスに由来するプロモーターが、一般的に好ましい。ポリペプチドをコードする配列の複数のコピーを含む細胞株を生成することが必要とされる場合、SV40またはEBVベースのベクターが、適切な選択マーカーと共に有利に使用され得る。

40

【0162】

細菌系において、多数の発現ベクターのいずれかが、発現されるポリペプチドに対して意図される使用に依存して選択され得る。例えば、大量に必要とされる場合（例えば、抗体の誘導のために）、容易に精製される融合タンパク質の高レベルの発現を指向するベクタ

50

ーが使用され得る。そのようなベクターとしては、以下が挙げられるがそれらに限定されない：多機能性 E. coli クローニングベクターおよび発現ベクター（例えば、BLUE SCRIPT (Stratagene)、ここでは目的のポリペプチドをコードする配列が、 β -ガラクトシダーゼのアミノ末端 Met およびそれに引続く 7 残基についての配列とインフレームでベクター内に連結され得、その結果、ハイブリッドタンパク質が産生される）；pINベクター（Van Heeke, G. および S. M. Schuster (1989) J. Biol. Chem. 264: 5503-5509) など。pGEXベクター（Promega, Madison, Wis.）もまた、グルタチオンSトランスフェラーゼ（GST）との融合タンパク質として外来ポリペプチドを発現させるために使用され得る。一般的に、そのような融合タンパク質は、可溶性であり、そしてグルタチオン-アガロースビーズに吸着させ、次に遊離のグルタチオンの存在下において溶出させることによって、溶解した細胞から容易に精製され得る。そのような系において作製されたタンパク質は、ヘパリン、トロニン、または第 XA 因子プロテアーゼ切断部位を含むように設計され得、その結果、クローニングされた目的のポリペプチドが、随意に GST 部分から放出され得る。

10

【0163】

酵母 (*Saccharomyces cerevisiae*) において、構成的プロモーターまたは誘導性プロモーター（例えば、 α 因子、アルコールオキシダーゼおよび PGH）を含む多数のベクターが使用され得る。概説については、Ausubel ら、(前出) および Grant ら、(1987) Methods Enzymol. 153: 516-544 を参照のこと。

20

【0164】

植物発現ベクターを使用する場合において、ポリペプチドをコードする配列の発現は、任意の多数のプロモーターにより駆動され得る。例えば、ウイルスプロモーター（例えば、CaMV の 35S プロモーターおよび 19S プロモーター）は、単独でか、または TMV に由来する ω リーダー配列と組み合わせて使用され得る (Takamatsu, N. (1987) EMBO J. 6: 307-311)。あるいは、植物プロモーター（例えば、RUBISCO の小サブユニットまたは熱ショックプロモーター）を使用し得る (Coruzzi, G. ら、(1984) EMBO J. 3: 1671-1680; Broglie, R. ら、(1984) Science 224: 838-843; および Winters, J. ら、(1991) Results Probl. Cell Differ. 17: 85-105)。これらの構築物は、直接的 DNA 形質転換または病原体媒介性トランスフェクションによって植物細胞内に導入され得る。そのような技術は、多数の一般的に入手可能な概説に記載されている（例えば、Hobbs, S. または Murry, L. E., McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N. Y.; 191-196 頁を参照のこと）。

30

【0165】

昆虫系はまた、目的のポリペプチドを発現させるために使用され得る。例えば、そのような 1 つの系において、オートグラファカリフォルニア核発汗病ウイルス (Autographa californica nuclear polyhedrosis virus) (AcNPV) は、Spodoptera frugiperda 細胞または Trichoplusia larvae において外来遺伝子を発現させるベクターとして使用される。ポリペプチドをコードする配列は、ウイルスの非必須領域内（例えば、ポリヘドリン (polyhedrin) 遺伝子) にクローニングされ得て、ポリヘドリンプロモーターの制御下に置かれ得る。ポリペプチドコード配列の首尾良い挿入は、ポリヘドリン遺伝子を不活性化し、そしてコートタンパク質を欠損している組換えウイルスを産生する。次いで、この組換えウイルスを使用して、例えば、目的のポリペプチドが発現され得る、S. frugiperda 細胞または Trichoplusia larvae に感染させ得る (Engelhard, E. K. ら、(1994) Proc. Natl. Aca

40

50

d. S c i . 9 1 : 3 2 2 4 - 3 2 2 7) 。

【0166】

哺乳動物宿主細胞において、多数のウイルスベースの発現系が一般的に利用可能である。例えば、アデノウイルスが発現ベクターとして使用される場合において、目的のポリペプチドをコードする配列は、後期プロモーターおよび3部からなるリーダー配列から構成されるアデノウイルス転写／翻訳複合体内に連結され得る。ウイルスゲノムの非必須E1またはE3領域における挿入を使用して、感染した宿主細胞においてポリペプチドを発現し得る、生存可能ウイルスを入手し得る (Logan, J. およびShenk, T. (1984) Proc. Natl. Acad. Sci. 81:3655-3659)。さらに、転写エンハンサー (例えば、ラウス肉腫ウイルス (RSV) エンハンサー) を使用して、哺乳動物宿主細胞における発現を増加させ得る。 10

【0167】

特定の開始シグナルはまた、目的のポリペプチドをコードする配列のより効率的な翻訳を達成するために使用され得る。そのようなシグナルとしては、ATG開始コドンおよび隣接配列が挙げられる。ポリペプチドをコードする配列、その開始コドンおよび上流配列が適切な発現ベクター内に挿入される場合において、さらなる転写制御シグナルまたは翻訳制御シグナルは必要とされなくとも良い。しかし、コード配列のみ、またはその一部のみが挿入される場合、ATG開始コドンを含む外因性翻訳制御シグナルが提供されるべきである。さらに、開始コドンは、インサート全体の翻訳を確実にするために、正確なリーディングフレーム内にあるべきである。外因性翻訳エレメントおよび開始コドンは、種々の起源 (天然および合成の両方) に由来し得る。発現の効率は、使用される特定の細胞系に適切なエンハンサー (例えば、文献 (Scharf, D. ら、(1994) Results Probl. Cell Differ. 20:125-162) に記載されるエンハンサー) の封入によって増大され得る。 20

【0168】

さらに、宿主細胞株は、挿入された配列の発現を調節するか、または所望の様式において発現されたタンパク質をプロセッシングするその能力について選択され得る。ポリペプチドのそのような改変としては、アセチル化、カルボキシル化、グリコシル化、リン酸化、リピデーション (lipidation) およびアシル化が挙げられるが、これらに限定されない。タンパク質の「プレプロ」形態を切断する翻訳後プロセッシングはまた、正確な挿入、折り畳みおよび／または機能を促進させるために使用され得る。異なる宿主細胞 (例えば、CHO、COS、HeLa、MDCK、HEK293およびWI38) (これらは、そのような翻訳後活性についての特定の細胞機構 (machinery) および特徴的機構 (mechanisms) を有する) は、正確な改変および外来タンパク質のプロセッシングを確実にするように選択され得る。 30

【0169】

長期間の組換えタンパク質の高収率の産生のために、安定な発現が一般的に好ましい。例えば、目的のポリヌクレオチドを安定に発現する細胞株が、ウイルス複製起点および／または内因性発現エレメントならびに同じもしくは別個のベクター上の選択マーカー遺伝子を含み得る、発現ベクターを使用して形質転換され得る。そのベクターの導入後、細胞は、それらが選択培地に切換えられる前に、富化 (enriched) 培地において1～2日間の増殖が可能とされ得る。選択マーカーの目的は、選択に対する耐性を与えることであり、そしてその存在は、導入された配列を首尾良く発現する細胞の増殖および回収を可能とする。安定に形質転換された細胞の耐性クローンは、その細胞型に適切な組織培養技術を使用して、増殖され得る。 40

【0170】

多数の選択系を使用して、形質転換細胞株を回収し得る。これらの選択系としては、それぞれ、tk. sup. 一細胞またはaprt. sup. 細胞において使用され得る、単純疱疹ウイルスチミジンキナーゼ (Wigler, M. ら、(1977) Cell 11:223-32) およびアデニンホスホリボシルトランスフェラーゼ (Lowy, I. ら、 50

(1990) *Cell* 22:817-23) 遺伝子が挙げられるが、これらに限定されない。また、代謝拮抗剤耐性、抗生物質耐性または除草剤耐性は、選択のための基礎として使用され得る；例えば、*dhfr* (これは、メトトレキセートに対する耐性を付与する (Wigler, M. ら、(1980) *Proc. Natl. Acad. Sci.* 77:3567-70))；*npt* (これは、アミノグリコシド、ネオマイシンおよびG-418に対する耐性を付与する (Colbere-Garapin, F. ら、(1981) *J. Mol. Biol.* 150:1-14))；ならびに*als*または*pat* (これらは、それぞれ、クロルスルフロンおよびホスフィノスリシンアセチルトランスフェラーゼに対する耐性を付与する (Murray, 前出))。さらなる選択遺伝子が記載されている。例えば、*trpB* (これは、細胞がトリプトファンの代わりにインドールを利用することを可能にする) または*hisD* (これは、細胞がヒスチジンの代わりにヒスチノールを利用することを可能にする) (Hartman, S. C. および R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51)。アントシアニン、 β -グルクロニダーゼおよびその基質のGUS、ならびにルシフェラーゼおよびその基質のルシフェリンのようなマーカーを用いる、可視マーカーの使用の popularity が高くなっており、形質転換体を同定するためのみならず、特定のベクター系に起因し得る一過性のタンパク質発現または安定なタンパク質発現の量を定量するためにもまた広範に使用されている (Rhodes, C. A. ら、(1995) *Methods Mol. Biol.* 55:121-131)。

10

20

30

40

50

【0171】

マーカー遺伝子発現の存在／非存在は、目的の遺伝子もまた存在するということを示唆しているが、その存在および発現は、確認されることを必要とし得る。例えば、ポリペプチドをコードする配列が、マーカー遺伝子配列内に挿入される場合、配列を含む組換え細胞は、マーカー遺伝子機能の非存在により同定され得る。あるいは、マーカー遺伝子は、1つのプロモーターの制御下にポリペプチドコード配列と直列に配置され得る。誘導または選択に応じてのマーカー遺伝子の発現は、通常、その直列遺伝子の発現も同様に示す。

【0172】

あるいは、所望のポリヌクレオチド配列を含み、かつ発現する宿主細胞は、当業者に公知の種々の手順によって同定され得る。これらの手順としては、以下が挙げられるが、これらに限定されない：核酸またはタンパク質の検出および／または定量化のための、例えば、膜ベースの技術、溶液ベースの技術、またはチップベースの技術を含む、DNA-DNAハイブリダイゼーション技術もしくはDNA-RNAハイブリダイゼーション技術およびタンパク質バイオアッセイ技術またはイムノアッセイ技術。

【0173】

ポリヌクレオチドコード化産物に特異的なポリクローナル抗体もしくはモノクローナル抗体のいずれかを使用して、ポリヌクレオチドコード化産物の発現を検出および測定するための種々のプロトコルは、当該分野において公知である。例としては、酵素結合免疫測定法 (ELISA)、ラジオイムノアッセイ (RIA) および蛍光活性化セルソーティング (FACS) が挙げられる。所定のポリペプチド上の2つの非干渉性エピトープに対して反応性であるモノクローナル抗体を利用する2つの部位の、モノクローナルベースのイムノアッセイが、いくつかの適用のために好まれ得るが、競合的結合アッセイもまた、利用され得る。これらのアッセイおよび他のアッセイは、とりわけ Hampton, R. ら、(1990; *Serological Methods, a Laboratory Manual*, APS Press, St Paul, Minn.) および Maddox, D. E. ら、(1983; *J. Exp. Med.* 158:1211-1216) に記載される。

【0174】

広範な種々の標識技術および結合技術が、当業者に公知であり、そして種々の核酸アッセイおよびアミノ酸アッセイにおいて使用され得る。ポリヌクレオチドに関連する配列を検出するための標識されたハイブリダイゼーションまたはPCRプローブを作製するための

手段としては、オリゴ標識 (oligolabeling)、ニックトランスレーション、末端標識または標識されたヌクレオチドを使用するPCR増幅が挙げられる。あるいは、mRNAプローブの産生のために、配列またはその任意の部分が、ベクター内にクローニングされ得る。そのようなベクターは、当該分野において公知であり、市販されており、そして適切なRNAポリメラーゼ (例えば、T7、T3、またはSP6) および標識されたヌクレオチドを添加することにより、RNAプローブをインビトロで合成するために使用され得る。これらの手順は、種々の市販キットを使用して実施され得る。使用され得る適切なレポーター分子または標識としては、放射性核種、酵素、蛍光、化学発光、または色素形成剤ならびに基質、コファクター (補因子)、インヒビター、磁気粒子などが挙げられる。

10

【0175】

目的のポリヌクレオチド配列で形質転換された宿主細胞は、タンパク質の発現および細胞培養物からの回収に適した条件下において培養され得る。組換え細胞により産生されたタンパク質は、使用された配列および／またはベクターに依存して、細胞内に分泌され得るかまたは含まれ得る。当業者によって理解されるように、本発明のポリヌクレオチドを含む発現ベクターは、原核細胞膜または真核細胞膜を介して、コードされたポリペプチドの分泌を指向するシグナル配列を含むように設計され得る。他の組換え構築物を使用して、目的のポリペプチドをコードする配列を、可溶性タンパク質の精製を促進するポリペプチドドメインをコードするヌクレオチド配列に結合させ得る。そのような精製促進ドメインとしては、以下が挙げられるが、これらに限定されない：固定された金属上での精製を可能にする金属キレート化ペプチド (例えば、ヒスチジントリプトファンモジュール)、固定された免疫グロブリン上での精製を可能にするプロテインAドメイン、およびFLAG伸長／アフィニティー精製システム (ImmuneX Corp., Seattle, Wash.) において利用されるドメイン。精製ドメインとコードされるポリペプチドとの間への、切断可能なリンカー配列 (例えば、第XA因子またはエンテロキナーゼに対して特異的な配列 (Invitrogen, San Diego, Calif.)) の封入を使用して、精製を促進させ得る。そのような発現ベクターの1つは、目的のポリペプチドを含む融合タンパク質およびチオレドキシンまたはエンテロキナーゼ切断部位の前に6つのヒスチジン残基をコードする核酸の発現を提供する。ヒスチジン残基は、Porath, J. ら、(1992, Prot. Exp. Purif. 3:263-281) に記載されるように、IMAC (固定化金属イオンアフィニティークロマトグラフィー) 上での精製を促進する一方で、エンテロキナーゼ切断部位は、融合タンパク質から所望のポリペプチドを精製するための手段を提供する。融合タンパク質を含むベクターの考察は、Kroll, D. J. ら、(1993; DNA Cell Boil. 12:441-453) において提供される。

20

30

【0176】

組換え産生方法に加えて、本発明のポリペプチドおよびそのフラグメントは、固相技術 (Merrifield J. (1963) J. Am. Chem. Soc. 85:2149-2154) を使用する、直接的ペプチド合成によって産生され得る。タンパク質合成は、手動技術を使用してか、または自動化によって実施され得る。自動化合成は、例えば、Applied Biosystems 431A ペプチド合成装置 (Perkin Elmer) を使用して、達成され得る。あるいは、種々のフラグメントは、別々に化学的に合成されて、そして全長分子を産生するために化学的方法を使用して組み合わせられ得る。

40

【0177】

(抗体組成物、そのフラグメントおよび他の結合因子)

別の局面によると、本発明はさらに、本明細書中で開示される腫瘍ポリペプチドに対して、またはその一部の改変体もしくは誘導体に対して免疫学的結合を示す結合因子 (例えば、抗体およびその抗原結合フラグメント) を提供する。抗体、またはその抗原結合フラグメントとは、本発明のポリペプチドに対して、「特異的に結合する」こと、「免疫学的に

50

結合する」ことを言い、そして／または、それが（例えば、E L I S A アッセイにおいて）検出可能なレベルでそのポリペプチドと反応し、かつ類似の条件下で非関連ポリペプチドと検出可能に反応しない場合、「免疫学的に反応性」であると言う。

【0178】

この文脈で使用される場合、免疫学的結合は、一般に、免疫グロブリン分子と、この免疫グロブリンが特異的な抗原との間で起こるタイプの非共有結合的な相互作用をいう。免疫学的結合相互作用の強度または親和性は、相互作用の解離定数 (K_d) の観点から表現され得、ここで、より小さい K_d は、より大きな親和性を示す。選択されたポリペプチドの免疫学的結合特性は、当該分野で周知の方法を使用して、定量され得る。1つのこのような方法は、抗原結合部位／抗原複合体形成および解離の速度を測定する工程を必要とし、ここで、これらの速度は、複合体パートナーの濃度、相互作用の親和性、および両方の方向における速度に等しく影響を与える幾何学的パラメータに依存する。従って、「会合速度定数 (on rate constant)」 (K_{on}) および「解離速度定数 (off rate constant)」 (K_{off}) の両方が、濃度ならびに会合および解離の実際の速度を計算することによって決定され得る。 K_{off} / K_{on} の比は、親和性に関係しない全てのパラメータの排除を可能にし、そして従って、解離定数 K_d に等しい。一般的には、Daviesら (1990) Annual Rev. Biochem. 59: 439-473を参照のこと。

【0179】

抗体の「抗原結合部位」または「結合部分」は、抗原結合に関係する免疫グロブリン分子の部分を用いる。この抗原結合部位は、重鎖 (「H」) および軽鎖 (「L」) のN末端可変 (「V」) 領域のアミノ酸残基によって形成される。重鎖および軽鎖のV領域内の3つの高度に分岐したストレッチが、「超可変領域」といわれ、これは、「フレームワーク領域」または「FR」として公知のより保存された隣接ストレッチの間に挿入される。従って、用語「FR」は、免疫グロブリンの超可変領域の間で、それに隣接して天然に見出されるアミノ酸配列をいう。抗体分子において、軽鎖の3つの超可変領域および重鎖の3つの超可変領域が、3次元空間において互いに相対的に配置されて、抗原結合表面を形成する。この抗原結合表面は、結合した抗原の3次元表面に相補的であり、そして重鎖および軽鎖の各々の3つの超可変領域は、「相補性決定領域」または「CDR」といわれる。

【0180】

結合因子は、本明細書中に提供される代表的なアッセイを使用して、癌（例えば、肺癌）を有する患者と有さない患者との間をさらに区別し得る。例えば、腫瘍タンパク質に結合する抗体または他の結合因子は、好ましくは、この疾患を有する患者の少なくとも約20%、より好ましくは患者の少なくとも約30%において、癌の存在を示すシグナルを生成する。あるいは、または加えて、この抗体は、癌を有さない個体の少なくとも約90%において、この疾患の非存在を示す陰性シグナルを生成する。結合因子が、この要求を満たすか否かを決定するために、癌（標準的な臨床試験によって決定されるような）を有する患者および有さない患者からの生物学的サンプル（例えば、血液、血清、痰、尿および／または腫瘍生検）が、この結合因子に結合するポリペプチドの存在について、本明細書に記載されるようにアッセイされ得る。好ましくは、疾患を有するサンプルおよび有さないサンプルの統計的に有意な数が、アッセイされる。各結合因子は、上記の基準を満足すべきである；しかし、当業者は、結合因子が、組合せて使用されて、感度を改善し得ることを認識する。

【0181】

上記の要求を満たす任意の薬剤は、結合因子であり得る。例えば、結合因子は、ペプチド成分を含むか含まないリボソーム、RNA分子またはポリペプチドであり得る。好ましい実施形態において、結合因子は、抗体またはその抗原結合フラグメントである。抗体は、当業者に公知の種々の技術のいずれかによって、調製され得る。例えば、HarlowおよびLane、Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988を参照のこと。一

10

20

30

40

50

般に、抗体は、細胞培養技術（本明細書に記載されるようなモノクローナル抗体の生成を含む）によって、または組換え抗体の生成を可能にするために、適切な細菌細胞宿主または哺乳動物細胞宿主への抗体遺伝子のトランスフェクションを介して、生成され得る。1つの技術において、このポリペプチドを含む免疫原が、最初に、広範な種々の哺乳動物（例えば、マウス、ラット、ウサギ、ヒツジまたはヤギ）のいずれかに注射される。この工程において、本発明のポリペプチドは、改変を有さない免疫原として機能し得る。あるいは、特に、比較的短いポリペプチドについて、このポリペプチドがキャリアタンパク質（例えば、ウシ血清アルブミンまたはキーホールリンペットヘモシアニン）に結合される場合に、優れた免疫応答が惹起され得る。この免疫原は、好ましくは、1回以上のブースト免疫を組み込む所定のスケジュールに従って、動物宿主に注射され、そしてこれらの動物が、定期的に採血される。このポリペプチドに特異的なポリクローナル抗体は、次いで、例えば、適切な固体支持体に結合されたポリペプチドを使用するアフィニティークロマトグラフィーによって、このような抗血清から精製され得る。

10

20

30

40

50

【0182】

目的の抗原性ポリペプチドに特異的なモノクローナル抗体が、例えば、K o h l l e r および M i l s t e i n, E u r. J. I m m u n o l. 6: 511-519, 1976の技術およびその改善物を使用して、調製され得る。簡潔には、これらの方法は、所望の特異性（すなわち、目的のポリペプチドとの反応性）を有する抗体を生成し得る不死細胞株の調製を含む。このような細胞株は、例えば、上記のように免疫された動物から得られる脾臓細胞から生成され得る。次いで、この脾臓細胞が、例えば、骨髓腫細胞融合パートナー、好ましくは、免疫された動物と同系であるものとの融合によって不死化される。種々の融合技術が使用され得る。例えば、脾臓細胞および骨髓腫細胞は、数分間、非イオン性界面活性剤と組合され、次いで、ハイブリッド細胞の増殖は支持するが、骨髓腫細胞の増殖は支持しない選択培地に低密度でプレートされ得る。好ましい選択技術は、H A T（ヒポキサンチン、アミノプテリン、チミジン）選択を使用する。十分な時間の後、通常、約1〜2週間の後、ハイブリッドのコロニーが観察される。単一コロニーを選択し、そしてそれらの培養上清を、このポリペプチドに対する結合活性について試験する。高い反応性および特異性を有するハイブリドーマが好ましい。

【0183】

モノクローナル抗体は、増殖しているハイブリドーマコロニーの上清から単離され得る。さらに、種々の技術が、収量を増加させるために使用され得、この技術は、例えば、適切な脊椎動物宿主（例えば、マウス）の腹腔へのハイブリドーマ細胞株の注入である。次いで、モノクローナル抗体が、腹水または血液から回収され得る。混入物が、従来の技術（例えば、クロマトグラフィー、ゲル濾過、沈澱、および抽出）によって抗体から除去され得る。本発明のポリペプチドは、例えば、アフィニティークロマトグラフィー工程における精製プロセスにおいて使用され得る。

【0184】

抗体分子の免疫学的結合特性を示し得る抗原結合部位を含む多くの治療的に有効な分子が、当該分野で公知である。このタンパク質分解性酵素パパインは、優先的に、I g G分子を切断して、いくつかのフラグメントを生じさせ、これら（「F（a b）フラグメント」）のうちの2つの各々は、インタクトな抗原結合部位を含む共有結合ヘテロダイマーを含む。酵素ペプシンは、I g G分子を切断して、いくつかのフラグメント（両方の抗原結合部位を含む「F（a b'）₂」フラグメントを含む）を提供し得る。「F v」フラグメントは、I g M、および稀な場合にはI g GまたはI g A免疫グロブリン分子の優先的なタンパク質分解的切断によって生成され得る。しかし、F vフラグメントは、当該分野で公知の組換え技術を使用して、より一般的に誘導される。F vフラグメントは、非共有結合的なV_H：：V_Lヘテロダイマーを含み、このヘテロダイマーは、ネイティブな抗体分子の抗原認識および結合能力のほとんどを保持する抗原結合部位を含む。I n b a r ら（1972）P r o c. N a t. A c a d. S c i. U S A 69: 2659-2662; H o c h m a n ら（1976）B i o c h e m 15: 2706-2710; および E h

r l i c h ら (1 9 8 0) B i o c h e m 1 9 : 4 0 9 1 - 4 0 9 6 .

【0185】

単鎖 F v (「s F v」) ポリペプチドは、共有結合された V_H : : V_L ヘテロダイマーであり、このヘテロダイマーは、ペプチドをコードするリンカーによって連結された V_H をコードする遺伝子および V_L をコードする遺伝子を含む遺伝子融合物から発現される。H u s t o n ら (1 9 8 8) P r o c . N a t . A c a d . S c i . U S A 8 5 (1 6) : 5 8 7 9 - 5 8 8 3 . 抗体 V 領域からの自然に凝集する (ただし、化学的に分離された) ポリペプチド軽鎖および重鎖を、s F v 分子に転換するための化学構造を識別するための多くの方法が記述され、この s F v は、抗原結合部位の構造に実質的に類似する 3 次元構造に折り畳まれる。例えば、米国特許第 5, 0 9 1, 5 1 3 号および同第 5, 1 3 2, 4 0 5 号 (H u s t o n ら) ; ならびに米国特許第 4, 9 4 6, 7 7 8 号 (L a d n e r ら) を参照のこと。

10

【0186】

上記分子の各々は、重鎖および軽鎖 C D R セットを含み、各々は、C D R S への支持を提供し、そして互いに対して C D R の空間的關係を規定する重鎖および軽鎖 F R セットの間に挿入される。本明細書中で使用される場合、用語「C D R セット」は、重鎖または軽鎖 V 領域の 3 つの超可変領域をいう。重鎖または軽鎖の N 末端から進んで、これらの領域は、各々、「C D R 1」、「C D R 2」および「C D R 3」と表記される。従って、抗原結合部位は、6 つの C D R を含み、この 6 つの C D R は、重鎖および軽鎖 V 領域の各々からの C D R セットを含む。単一 C D R (例えば、C D R 1、C D R 2、または C D R 3) を含むポリペプチドは、本明細書中で「分子認識ユニット」といわれる。多くの抗原-抗体複合体の結晶学的分析により、C D R のアミノ酸残基は、結合した抗原との広範な接触を形成することが立証され、ここで、最も広範な抗原接触は、重鎖 C D R 3 とである。従って、分子認識ユニットは、主に、抗原結合部位の特異性の原因である。

20

【0187】

本明細書中で使用される場合、用語「F R セット」は、重鎖または軽鎖の V 領域の C D R セットの C D R を構成する、4 つの隣接アミノ酸配列をいう。いくつかの F R 残基は、結合抗原と接触し得る ; しかし、F R は、主に、V 領域を抗原結合部位、特に C D R S に直接隣接する F R 残基に折り畳む原因である。F R において、特定のアミノ酸残基および特定の構造特徴が、非常に高度に保存される。この点において、全ての V 領域配列は、約 9 0 アミノ酸残基の内部ジスルフィドループを含む。V 領域が、結合部位に折り畳まれる場合、この C D R は、抗原結合表面を形成する突出したループモチーフとして表示される。正確な C D R アミノ酸配列に関わらず、特定の「正準」構造に折り畳まれた C D R ループの形状に影響する F R の保存された構造領域が存在することが一般的に認識される。さらに、特定の F R 残基は、抗体重鎖および軽鎖の相互作用を安定化する非共有結合的なドメイン間接触に関係することが知られている。

30

【0188】

非ヒト免疫グロブリン由来の抗原結合部位を含む、多くの「ヒト化」抗体分子が記載され、この抗体分子には、げっ歯類 V 領域およびヒト定常領域に融合されたそれらの会合 C D R (W i n t e r ら (1 9 9 1) N a t u r e 3 4 9 : 2 9 3 - 2 9 9 ; L o b u g l i o ら (1 9 8 9) P r o c . N a t . A c a d . S c i . U S A 8 6 : 4 2 2 0 - 4 2 2 4 ; S h a w ら (1 9 8 7) J I m m u n o l . 1 3 8 : 4 5 3 4 - 4 5 3 8 ; および B r o w n ら (1 9 8 7) C a n c e r R e s . 4 7 : 3 5 7 7 - 3 5 8 3) 、適切なヒト抗体定常ドメインとの融合の前に、ヒト支持 F R にグラフトされたげっ歯類 C D R (R i e c h m a n n ら (1 9 8 8) N a t u r e 3 3 2 : 3 2 3 - 3 2 7 ; V e r h o e y e n ら (1 9 8 8) S c i e n c e 2 3 9 : 1 5 3 4 - 1 5 3 6 ; および J o n e s ら (1 9 8 6) N a t u r e 3 2 1 : 5 2 2 - 5 2 5) 、ならびに組換えベニアリングされた (r e c o m b i n a n t l y v e n e e r e d) げっ歯類 F R によって支持されるげっ歯類 C D R (欧州特許公開番号 5 1 9, 5 9 6 (1 9 9 2 年 1 2 月 2 3 日 公開)) を含むキメラ抗体が挙げられる。これらの「ヒト化」分子は、ヒトレシピエントに

40

50

おけるこれらの部分の治療的適用の持続期間および有効性を制限するげっ歯類抗ヒト抗体分子に対する所望されない免疫学的応答を最小化するように設計される。

【0189】

本明細書中で使用される場合、用語「ベニアリングされたFR」および「組換えベニアリングされたFR」は、ネイティブのFRポリペプチド折り畳み構造の実質的に全てを保持する抗原結合部位を含む異種分子を提供するために、例えば、げっ歯類重鎖または軽鎖のV領域からのFR残基のヒトFR残基での選択的置換をいう。ベニアリング技術は、抗原結合部位のリガンド結合特徴が、主に抗原結合表面内の重鎖および軽鎖のCDRセットの構造および相対的な配置によって決定されるという理解に基づく。Daviesら(1990) Ann. Rev. Biochem. 59: 439-473。従って、抗原結合特異性は、CDR構造、互いとのそれらの相互作用、およびV領域ドメインの残りとのそれらの相互作用が注意深く維持されるヒト化抗体のみにおいて保存され得る。ベニアリング技術を使用することによって、外部（例えば、溶媒がアクセス可能な）FR残基（これは、免疫系に容易に遭遇する）は、ヒト残基と選択的に置換されて、弱い免疫原性ベニアリング表面または実質的に非免疫原性のベニアリング表面のいずれかを含むハイブリッド分子を提供する。

10

【0190】

ベニアリングのプロセスは、Sequences of Proteins of Immunological Interest, 第4版(U. S. Dept. of Health and Human Services, U. S. Government Printing Office, 1987)において、Kabataらによって編集されたヒト抗体可変ドメインについての利用可能な配列データを使用し、Kabataデータベース、ならびに他の米国および海外のアクセス可能なデータベース（核酸およびタンパク質の両方）に更新される。V領域アミノ酸の溶媒アクセス可能性は、ヒトおよびマウス抗体フラグメントについての既知の3次元構造から推定され得る。マウス抗原結合部位をベニアリングする際に2つの一般的な工程が存在する。最初に、目的の抗体分子の可変ドメインのFRが、上記の供給源から得られたヒト可変ドメインの対応するFR配列と比較される。次いで、最も相同的なヒトV領域が、対応するマウスアミノ酸と、残基ごとに比較される。ヒト対応物とは異なるマウスFRにおける残基は、当該分野において周知の組換え技術を使用して、ヒト部分に存在する残基によって置換される。残基スイッチングは、少なくとも部分的に露出した（溶媒アクセス可能である）部分を用いて実施されるのみであり、そしてV領域ドメインの三次構造に対する有意な効果を有し得るアミノ酸残基（例えば、プロリン、グリシンおよび荷電したアミノ酸）の置換において、注意がなされる。

20

30

【0191】

この様式において、得られた「ベニアリングされた」マウス抗原結合部位は、従って、マウスCDR残基、CDRに実質的に隣接する残基、埋没したかほとんど埋没した（溶媒のアクセスが不可能）として同定された残基、重鎖ドメインと軽鎖ドメインとの間の非共有結合的な（例えば、静電的および疎水的）接触に関係すると考えられる残基、およびCDRループの「正準」三次構造に影響を与えると考えられるFRの保存された構造領域からの残基を保持するように設計される。次いで、これらの設計基準は、ヒト様FRへのマウス抗原結合部位の軽鎖および重鎖の両方のCDRを組合せる組換えヌクレオチド配列を調製するために使用され、このヒト様FRは、マウス抗体分子の抗原特異性を示す組換えヒト抗体の発現のために哺乳動物細胞をトランスフェクトするために使用され得る。

40

【0192】

本発明の別の実施形態において、本発明のモノクローナル抗体は、1つ以上の治療剤に結合され得る。この点に関して、適切な薬剤としては、放射性核種、分化インデューサー、薬物、毒素、およびその誘導体が挙げられる。好ましい放射性核種としては、 ^{90}Y 、 ^{123}I 、 ^{125}I 、 ^{131}I 、 ^{186}Re 、 ^{188}Re 、 ^{211}At および ^{212}Bi が挙げられる。好ましい薬物としては、メトトレキサート、ならびにピリミジンおよびプリンアナログが挙げられる。好ましい分化インデューサーとしては、ホルボールエステル

50

および酪酸が挙げられる。好ましい毒素としては、リシン、アブリン、ジフテリア毒素、コレラ毒素、ゲロニン、Pseudomonas 体外毒素、Shigella 毒素、およびアメリカマゴボウ抗ウイルスタンパク質が挙げられる。

【0193】

治療剤は、適切なモノクローナル抗体に直接的または間接的に（例えば、リンカー基を介して）のいずれかで結合（例えば、共有結合）され得る。薬剤と抗体との間の直接的な反応は、各々が他のものと反応し得る置換基を有する場合に可能である。例えば、一方の上の求核基（例えば、アミノ基またはスルフヒドリル基）は、他方のカルボニル含有基（例えば、酸無水物または酸ハロゲン化物）または良好な脱離基（例えば、ハロゲン化物）を含むアルキル基と反応し得る。

10

【0194】

あるいは、リンカー基を介して治療剤と抗体とを結合させることが所望され得る。リンカー基は、結合の可能性を妨げることを回避するために、抗体を薬剤から隔てるためのスペーサーとして機能し得る。リンカー基はまた、薬剤または抗体上の置換基の化学的反応性を増加させるように働き得、従って結合効率を増大させる。化学的反応性の増大はまた、薬剤または薬剤上の官能基の使用（さもなければ、可能ではない）を促進し得る。

【0195】

種々の二官能性または多官能性試薬、ホモ官能性とヘテロ官能性との両方（例えば、Pierce Chemical Co., Rockford, IL のカタログ中に記載されるもの）が、リンカー基として使用され得ることが当業者には明らかである。結合は、例えば、アミノ基、カルボキシル基、スルフヒドリル基、または酸化された炭水化物残基を介してもたらされ得る。このような方法論を記載する多数の参考文献（例えば、Rodwell らに対する米国特許第 4, 671, 958 号）が存在する。

20

【0196】

本発明の免疫結合体の抗体部分がないときに治療剤がより強力である場合、細胞中へのインターナリゼーションの間に、またはその際に切断可能なリンカー基を使用することが望ましくあり得る。多数の異なる切断可能なリンカー基が記載されている。これらのリンカー基からの薬剤の細胞内放出についての機構は、ジスルフィド結合の還元（例えば、Spittler への米国特許第 4, 489, 710 号）、感光性結合の照射（例えば、Senter への米国特許第 4, 625, 014 号）、誘導体化されたアミノ酸側鎖の加水分解（例えば、Kohn への米国特許第 4, 638, 045 号）、血清補体媒介性加水分解（例えば、Rodwell への米国特許第 4, 671, 958 号）、および酸触媒加水分解（例えば、Blattler への米国特許第 4, 569, 789 号）による切断を含む。

30

【0197】

1 つより多い薬剤を抗体に結合させることが望ましくあり得る。1 つの実施形態において、薬剤の複数の分子が 1 つの抗体分子に結合される。別の実施形態において、1 つより多い型の薬剤が 1 つの抗体に結合され得る。特定の実施形態に関わらず、1 つより多い薬剤を有する免疫結合体は、種々の方法で調製され得る。例えば、1 つより多い薬剤が、抗体分子に直接的に結合され得るか、または付着のための複数の部位を提供するリンカーが使用され得る。あるいは、キャリアが使用され得る。

40

【0198】

キャリアは、種々の方法（直接的にかまたはリンカー基を介するかのいずれかの共有結合を含む）で薬剤を保有し得る。適切なキャリアとしては、アルブミンのようなタンパク質（例えば、Kato への米国特許第 4, 507, 234 号）、ペプチド、およびアミノデキストランのような多糖類（例えば、Shih への米国特許第 4, 699, 784 号）が挙げられる。キャリアはまた、例えばリポソーム小胞内に、非共有結合によってかまたはカプセル化によって、薬剤を保有し得る（例えば、米国特許第 4, 429, 008 号および同第 4, 873, 088 号）。放射性核種薬剤に特異的なキャリアは、放射性ハロゲン化低分子およびキレート化合物を含む。例えば、米国特許第 4, 735, 792 号は

50

、代表的な放射性ハロゲン化低分子およびそれらの合成を開示する。放射性核種キレートは、金属、または金属酸化物、放射性核種を結合するためのドナー原子として窒素原子および硫黄原子を含むものを含む、キレート化合物から形成され得る。例えば、Davisonらへの米国特許第4,673,562号は、代表的なキレート化合物およびそれらの合成を開示する。

【0199】

(T細胞組成物)

別の局面において、本発明は、本明細書中に開示される腫瘍ポリペプチドに特異的なT細胞、または本明細書中に開示される腫瘍ポリペプチドの改変体もしくは誘導体に特異的なT細胞を提供する。このような細胞は、一般的に標準的手順を使用して、インビトロまたはエキソビボで調製され得る。例えば、T細胞は、市販の細胞分離システム（例えば、IsoplexTM システム (Nexell Therapeutics, Inc. (Irvine, CA; 米国特許第5,240,856号; 米国特許第5,215,926号; WO89/06280; WO91/16116およびWO92/07243もまた参照のこと) から入手可能) を使用して、患者の骨髄、末梢血あるいは骨髄または末梢血の画分中から単離され得る。あるいは、T細胞は、関連または無関連のヒト、非ヒト動物、細胞株または培養物から誘導され得る。

10

【0200】

T細胞は、ポリペプチド、ポリペプチドをコードするポリヌクレオチドおよび／またはそのようなポリペプチドを発現する抗原提示細胞 (APC) を用いて刺激され得る。このような刺激は、目的のポリペプチドに特異的であるT細胞の生成を可能にする条件下および十分な時間、行われる。好ましくは、本発明の腫瘍ポリペプチドまたはポリヌクレオチドは、送達ビヒクル（例えば、ミクロスフェア）中に存在して、特定のT細胞の生成を容易にする。

20

【0201】

T細胞は、このT細胞が、特異的に増殖するか、サイトカインを分泌するか、または本発明のポリペプチドで被覆されるか、もしくはこのポリペプチドをコードする遺伝子が発現する標的細胞を殺傷する場合に、本発明のポリペプチドに特異的であるとみなされる。T細胞特異性は、種々の標準的技術のいずれかを使用して評価され得る。例えば、クロム放出アッセイまたは増殖アッセイにおいて、ネガティブコントロールと比較して、溶解および／または増殖における2倍を超える増加の刺激指標は、T細胞特異性を示す。このようなアッセイは、例えば、Chenら、Cancer Res. 54:1065-1070, 1994に記載されるように、実行され得る。あるいは、T細胞の増殖の検出は、種々の公知の技術によって達成され得る。例えば、T細胞増殖は、DNA合成の速度の増加を測定することによって検出され得る（例えば、トリチウム化チミジンでT細胞の培養物をパルス標識し、そしてDNAに取り込まれたトリチウム化チミジンの量を測定することによって）。3～7日間の腫瘍ポリペプチド (100 ng/ml～100 μg/ml、好ましくは、200 ng/ml～25 μg/ml) との接触は、代表的に、T細胞の増殖において少なくとも2倍の増加を生じる。2～3時間の上記のような接触は、標準的なサイトカインアッセイを使用して測定されるように（ここで、サイトカイン（例えば、TNFまたはIFN-γ）放出のレベルの2倍の増加が、T細胞の活性化を示す）、T細胞の活性化を生じるはずである (Coliganら、Current Protocols in Immunology, 第1巻、Wiley Interscience (Green 1998) を参照のこと)。腫瘍ポリペプチド、ポリヌクレオチドまたはポリペプチド発現APCに応答して活性化されたT細胞は、CD4⁺ および／またはCD8⁺ であり得る。腫瘍ポリペプチド特異的T細胞は、標準的な技術を使用して拡大され得る。好ましい実施形態において、T細胞は、患者、関連するドナーまたは無関連のドナーに由来し、そして刺激および拡大後にその患者に投与される。

30

40

【0202】

治療目的で、腫瘍ポリペプチド、ポリヌクレオチドまたはAPCに応答して増殖するCD

50

4^+ T細胞または $CD8^+$ T細胞は、インビトロまたはインビボのいずれかで大量に拡大され得る。このようなT細胞のインビトロでの増殖は、種々の方法で達成され得る。例えば、T細胞は、T細胞増殖因子（例えば、インターロイキン-2）の添加を伴うか、または伴わずに、腫瘍ポリペプチド、またはこのようなポリペプチドの免疫原性部分に対応する短いペプチドに対して再曝露され得、そして／または腫瘍ポリペプチドを合成する刺激細胞に対して再曝露され得る。あるいは、腫瘍ポリペプチドの存在下で増殖する1つ以上のT細胞は、クローニングによって数の上で拡大され得る。細胞をクローニングするための方法は、当該分野で周知であり、そしてこれらとしては、限界希釈が挙げられる。

【0203】

（T細胞レセプター組成物）

T細胞レセプター（TCR）は、ジスルフィド結合によって連結されている、2つの異なる、非常に可変性のポリペプチド鎖（T細胞レセプターの α 鎖および β 鎖と呼ばれる）からなる（Janeway, Travers, Walport. Immunobiology. 第4版、148～159. Elsevier Science Ltd/Garland Publishing. 1999）。この α/β ヘテロダイマー（異種二量体）は、細胞膜で不変のCD3鎖と複合体化する。この複合体は、MHC分子に結合した特定の抗原性ペプチドを認識する。TCR特異性の膨大な多様性は、体細胞遺伝子再配列により、まるで免疫グロブリンの多様性のように生成される。この β 鎖遺伝子は、50を超える可変性領域（V）、2つの多様性領域（D）、10を超える連結セグメント（J）、および2つの定常領域セグメント（C）を含む。 α 鎖遺伝子は、70を超えるVセグメント、および60を超えるJセグメントを含むが、Dセグメントを含まず、そして1つのCセグメントを含む。胸腺におけるT細胞発達の間、 β 鎖のD～J遺伝子再配列が生じ、続いて、V遺伝子セグメントのDJへの再配列が生じる。この機能的VDJ β エキソンは、転写され、そしてスプライシングされてC β に連結される。 α 鎖に関しては、V α 遺伝子セグメントは、J α 遺伝子セグメントに再配列して、機能的エキソンを形成し、これが次に転写され、C α にスプライシングされてる。多様性は、さらに β 鎖のVセグメントと、Dセグメントと、Jセグメントとの間、そして α 鎖のVセグメントと、Jセグメントとの間のPおよびN-ヌクレオチドの無作為な付加によって、組み換えプロセスの間にさらに増大する（Janeway, Travers, Walport. Immunobiology. 第4版、98および150. Elsevier Science Ltd/Garland Publishing. 1999）。 10 20 30

【0204】

本発明は、別の局面において、本明細書に開示のポリペプチド、またはその改変体もしくは誘導体に特異的なTCRを提供する。本発明に従って、本明細書に記載の腫瘍ポリペプチドを認識する、T細胞レセプターの α 鎖および β 鎖について、V-JもしくはV-D-J接合領域、またはその部分について、ポリヌクレオチドおよびアミノ酸の配列が提供される。一般に、本発明のこの局面は、MHCの文脈において提示される腫瘍ポリペプチドを認識するか、またはそれに結合するT細胞レセプターに関する。好ましい実施形態において、T細胞レセプターによって認識される腫瘍抗原は、本発明のポリペプチドを含む。例えば、腫瘍ペプチドに特異的なTCRをコードするcDNAは、標準的な分子生物学技術および組み換えDNA技術を用いて、腫瘍ポリペプチドに特異的なT細胞から単離され得る。 40

【0205】

本発明はさらに、腫瘍ポリペプチドを認識するか、またはそれに結合する、本発明のT細胞レセプターと実質的に同じ機能または活性を有する、T細胞レセプターまたはそのアナログを包含する。このようなレセプターとしては、本明細書に提供されるT細胞レセプターの、レセプターフラグメント、または本明細書に提供されるT細胞レセプターの、置換、付加、もしくは欠失の変異体が挙げられるがこれらに限定されない。本発明はまた、本明細書に提供されるT細胞レセプターに実質的に相同であるか、または実質的に同じ活性を保持する、ポリペプチドまたはペプチドを包含する。用語「アナログ」とは、本明細書 50

に提供されるT細胞レセプターと実質的に同一のアミノ酸残基配列（ここで、1つ以上の残基、好ましくは5残基以下、より好ましくは、25残基以下が、機能的に類似の残基で保存的に置換されている）を有し、そして本明細書に記載のようなT細胞レセプターの機能的局面を示す、任意のタンパク質またはポリペプチドを含む。

【0206】

本発明はさらに、適切な哺乳動物宿主細胞であって、本明細書に記載のポリペプチドに特異的なTCRをコードするポリヌクレオチドでトランスフェクトされている（これによって、この宿主細胞はこのポリペプチドに特異的にされる）、適切な哺乳動物宿主細胞、例えば、非特異的T細胞を提供する。TCRの α 鎖および β 鎖は、別々の発現ベクターに含まれ得るか、または単一の発現ベクター（これはまた、リボソーム内侵入部位（IRES）の下流の遺伝子のcap依存性翻訳のためのIRESを含む）に含まれ得る。このポリペプチドに特異的なTCRを発現するこの宿主細胞は、例えば、以下にさらに考察されるような肺癌の養子免疫療法のために、用いられ得る。

10

【0207】

本発明のさらなる局面において、本明細書に挙げたポリペプチドに特異的なクローニングされたTCRは、肺癌の診断のためのキットに用いられ得る。例えば、腫瘍特異的TCRの核酸配列またはその一部は、生物学的サンプル中で、特定のTCRをコードする再配列された遺伝子の発現を検出するためのプローブまたはプライマーとして用いられ得る。従って、本発明はさらに、ポリペプチドに特異的なTCRをコードするメッセンジャーRNAまたはDNAを検出するためのアッセイを提供する。

20

【0208】

（薬学的組成物）

さらなる実施形態において、本発明は、細胞または動物への投与のための、単独でかまたは治療の1つ以上の他の様相と組合わせてかのいずれかでの、薬学的に受容可能なキャリア中の本明細書中に開示される1つ以上のポリヌクレオチド、ポリペプチド、T細胞および／または抗体組成物の処方物に関する。

【0209】

所望される場合、本明細書中に開示される組成物が、他の薬剤（例えば、他のタンパク質もしくはポリペプチドまたは種々の薬学的に活性な薬剤など）と組合わせても同様に投与され得ることが、理解される。事実、さらなる薬剤が標的細胞または宿主細胞と接触した際に有意な有害な影響をもたらさないとすれば、さらに含まれ得る他の成分に実質的に制限はない。従って、この組成物は、特定の例において必要とされる種々の他の薬剤と共に送達され得る。このような組成物は、宿主細胞または他の生物学的供給源から精製され得るか、あるいは、本明細書中に記載されるように化学的合成され得る。同様に、このような組成物はさらに、置換または誘導体化されたRNA組成物またはDNA組成物を含み得る。

30

【0210】

従って、本発明の別の局面において、生理学的に受容可能なキャリアと組合わせた、本明細書中に記載される1つ以上のポリヌクレオチド、ポリペプチド、抗体、および／またはT細胞組成物を含む薬学的組成物が、提供される。特定の好ましい実施形態において、本発明の薬学的組成物は、予防ワクチン適用および治療ワクチン適用における使用のための、本発明の免疫原性ポリヌクレオチド組成物および／または免疫原性ポリペプチド組成物を含む。ワクチン調製物は、一般的に、例えば、M. F. PowellおよびM. J. Newman編、「Vaccine Design (the subunit and adjuvant approach)」、Plenum Press (NY, 1995)に記載される。一般的にこのような組成物は、1つ以上の免疫刺激剤と組合わせた、本発明のポリヌクレオチド組成物および／またはポリペプチド組成物の1つ以上を含む。

40

【0211】

本明細書中に記載される任意の薬学的組成物が、本発明のポリヌクレオチドおよびポリペプチドの薬学的に受容可能な塩を含み得ることは、明らかである。このような塩は、例え

50

ば、薬学的に受容可能な非毒性の塩基（有機塩基（例えば、一級アミン、二級アミンおよび三級アミンならびに塩基性アミノ酸の塩）および無機塩基（例えば、ナトリウム塩、カリウム塩、リチウム塩、アンモニウム塩、カルシウム塩およびマグネシウム塩）を含む）から調製され得る。

【0212】

別の実施形態において、本発明の例示的な免疫原性組成物（例えば、ワクチン組成物）は、上記のようなポリペプチドの1つ以上をコードするDNAを含み、その結果このポリペプチドは、インサイチュで生成される。上記のように、ポリヌクレオチドは、当業者に公知の種々の送達系の内のいずれかで投与され得る。実際、多数の遺伝子送達技術（例えば、Rolland, Crit. Rev. Therap. Drug Carrier Systems 15:143-198, 1998およびこの中に引用される参考文献に記載される遺伝子送達技術）が、当該分野で周知である。当然、適切なポリヌクレオチド発現系は、患者における発現のための必要な調節性のDNA調節配列（例えば、適切なプロモーターおよび終結シグナル）を含む。あるいは、細菌送達系は、その細胞表面上にポリペプチドの免疫原性部分を発現するか、またはエプトープなどを分泌する細菌（例えば、Bacillus-Calmette-Guérin）の投与を含み得る。

10

【0213】

従って、特定の実施形態において、本明細書中に記載される免疫原性ポリペプチドをコードするポリヌクレオチドは、多数の公知のウイルスに基づく系のうちのいずれかを用いて、発現のための適切な哺乳動物宿主細胞に導入される。1つの例示的な実施形態において、レトロウイルスが、遺伝子送達系のための簡便かつ有効な基盤を提供する。本発明のポリペプチドをコードする選択されたヌクレオチド配列は、当該分野で公知の技術を用いて、ベクターに挿入され得、そしてレトロウイルス粒子にパッケージングされ得る。次いで、組換えウイルスが、単離され、そして被験体に送達され得る。多数の例示的なレトロウイルス系が、記載されている（例えば、米国特許第5,219,740号；MillerおよびRosman (1989) BioTechniques 7:980-990；Miller, A. D. (1990) Human Gene Therapy 1:5-14；Scarpaら (1991) Virology 180:849-852；Burnsら (1993) Proc. Natl. Acad. Sci. USA 90:8033-8037；ならびにBoris-LawrieおよびTemin (1993) Cur. Opin. Genet. Develop. 3:102-109）。

20

30

【0214】

さらに、多数の例示的なアデノウイルスに基づく系もまた、記載されている。宿主ゲノムに組込むレトロウイルスとは異なり、アデノウイルスは、染色体外に残り、このようにして、挿入変異誘発に関連する危険性を最小化する（Haj-AhmadおよびGraham (1986) J. Virol. 57:267-274；Bettrら (1993) J. Virol. 67:5911-5921；Mitterederら (1994) Human Gene Therapy 5:717-729；Sethら (1994) J. Virol. 68:933-940；Barrら (1994) Gene Therapy 1:51-58；Berkner, K. L. (1988) BioTechniques 6:616-629；ならびにRichら (1993) Human Gene Therapy 4:461-476）。

40

【0215】

種々のアデノ随伴ウイルス（AAV）ベクター系もまた、ポリヌクレオチド送達に関して開発されている。AAVベクターは、当該分野で周知の技術を用いて容易に構築され得る。例えば、米国特許第5,173,414号および同第5,139,941号；国際公開番号WO 92/01070およびWO 93/03769；Lebkowskiら (1988) Molec. Cell. Biol. 8:3988-3996；Vincentら (1990) Vaccines 90 (Cold Spring Harbor Laboratory Press)；Carter, B. J. (1992) Current

50

Opinion in Biotechnology 3:533-539; Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129; Kotin, R. M. (1994) Human Gene Therapy 5:793-801; ShellingおよびSmith (1994) Gene Therapy 1:165-169; ならびに Zhouら (1994) J. Exp. Med. 179:1867-1875を参照のこと。

【0216】

遺伝子移入による、本発明のポリペプチドをコードするポリヌクレオチドを送達するのに有用なさらなるウイルスベクターとしては、ポックスウイルスのファミリー（例えば、ワクシニアウイルスおよび鳥類ポックスウイルス）から誘導されるものが挙げられる。例として、新規分子を発現するワクシニアウイルスの組換え体は、以下のように構築され得る。ポリペプチドをコードするDNAを最初に、ワクシニアプロモーターおよび隣接ワクシニアDNA配列（例えば、チミジンキナーゼ（TK）をコードする配列）に隣接するように、適切なベクターに挿入する。次いで、このベクターを使用して、ワクシニアで同時に感染させる細胞にトランスフェクトする。相同組換えによって、ワクシニアプロモーターおよび目的のポリペプチドをコードする遺伝子を、ウイルスゲノムに挿入する。生じるTK^{sup}（-）組換え体を、5-ブロモデオキシウリジンの存在下でこの細胞を培養し、そして5-ブロモデオキシウリジンに耐性のウイルスプラークをピックアップすることによって、選択し得る。

10

【0217】

ワクシニアに基づく感染／トランスフェクション系は、生物の宿主細胞における、本明細書中に記載される1つ以上のポリペプチドの誘導可能な一過性発現または同時発現を提供するために、好都合に使用され得る。この特定の系において、細胞を最初に、インビトロで、バクテリオファージT7 RNAポリメラーゼをコードするワクシニアウイルスの組換え体で感染させる。このポリメラーゼは、それがT7プロモーターを保有する鋳型のみを転写する点で、鋭敏な特異性を示す。感染後、細胞を、T7プロモーターによって駆動される目的のポリヌクレオチドでトランスフェクトさせる。ワクシニアウイルスの組換え体から細胞質中で発現されるポリメラーゼは、トランスフェクトされたDNAをRNAに転写し、このRNAは次いで、宿主翻訳機構によってポリペプチドに翻訳される。この方法は、大量のRNAおよびその翻訳産物の、高レベルで一過性の細胞質産生を提供する。例えば、Elroy-SteinおよびMoss, Proc. Natl. Acad. Sci. USA (1990) 87:6743-6747; Fuerstら、Proc. Natl. Acad. Sci. USA (1986) 83:8122-8126を参照のこと。

20

30

【0218】

あるいは、アビポックスウイルス（avipoxvirus）（例えば、鶏痘ウイルスおよびカナリア痘ウイルス）もまた使用されて、目的のコード配列を送達し得る。哺乳動物病原体由来の免疫原を発現する組換えアビポックスウイルスは、非鳥類種に投与された場合に、防御免疫を与えることが公知である。アビポックスベクターの使用は、ヒトおよび他の哺乳動物種において特に望ましい。なぜなら、アビポックス属のメンバーは、感受性の鳥類種においてのみ生産的に複製し得、故に、哺乳動物細胞において感染性ではないからである。組換えアビポックスウイルスを産生するための方法は、当該分野で公知であり、そしてワクシニアウイルスの産生に関して上記に記載されるように、遺伝子組換えを使用する。例えば、WO 91/12882; WO 89/03429; およびWO 92/03545を参照のこと。

40

【0219】

多数のアルファウイルスベクターのいずれかもまた、本発明のポリヌクレオチド組成物の送達のために使用され得、これらとしては、米国特許第5,843,723号; 同第6,015,686号; 同第6,008,035号および同第6,015,694号に記載されるベクターが挙げられる。ベネズエラウマ脳脊髄炎（VEE）に基づく特定のベクターもまた、使用され得、この例示的な例は、米国特許第5,505,947号および同第5

50

、643、576号に見出され得る。

【0220】

さらに、分子結合体化ベクター（例えば、Michaelら、J. Biol. Chem. (1993) 268:6866~6869およびWagnerら、Proc. Natl. Acad. Sci. USA (1992) 89:6099~6103に記載のアデノウイルススキメラベクター）がまた、本発明における遺伝子送達に用いられ得る。

【0221】

これらおよび他の公知のウイルスベースの送達系に対するさらなる例示的情報は、例えば、以下に見出され得る：Fisher-Hochら、Proc. Natl. Acad. Sci. USA 86:317~321、1989；Flexnerら、Ann. N. Y. Acad. Sci. 569:86~103、1989；Flexnerら、Vaccine 8:17~21、1990；米国特許第4,603,112号、同第4,769,330号および同第5,017,487号；WO89/01973；米国特許第4,777,127号；GB 2,200,651:EP 0,345,242号；WO 91/02805；Berkner、Biotechniques 6:616~627、1988；Rosenfeldら、Science 252:431~434、1991；Kollsら、Proc. Natl. Acad. Sci. USA 91:215~219、1994；Kass-Eislerら、Proc. Natl. Acad. Sci. USA 90:11498~11502、1993；Guzmanら、Circulation 88:2838~2848、1993；ならびにGuzmanら、Cir. Res. 73:1202~1207、1993。

【0222】

特定の実施形態において、ポリヌクレオチドを標的細胞のゲノムに組み込み得る。この組み込みは、相同組み換え（遺伝子置換）を介して特定の位置および方向であり得るか、または無作為な非特異的位置に組み込まれ得る（遺伝子増大）。なおさらなる実施形態において、ポリヌクレオチドはDNAの別のエピソームセグメントとして細胞において安定に維持され得る。このようなポリヌクレオチドセグメントすなわち「エピソーム」は、宿主細胞の周期に独立してまたは同調して維持および複製を可能にするのに十分な配列をコードする。発現構築物が細胞に送達され、そしてこの細胞中にこのポリヌクレオチドが残る様式は、使用される発現構築物のタイプに依存する。

【0223】

本発明の別の実施形態において、例えば、Ulmerら、Science 259:1745~1749、1993に記載され、そしてCohen、Science 259:1691~1692、1993によって概説されるように、ポリヌクレオチドは、「裸の」DNAとして投与/送達される。裸のDNAの取り込みは、生分解性ビーズ（これは、細胞に効率的に輸送される）上にそのDNAをコーティングすることによって増加され得る。

【0224】

なお別の実施形態において、本発明の組成物は、微粒子銃アプローチ（その多くが記載されている）を介して送達され得る。1つの代表的な例において、ガス駆動粒子加速（gas-driven acceleration）は、Powderject Pharmaceuticals PLC（Oxford, UK）およびPowderject Vaccines Inc.（Madison, WI）によって製造されたデバイスのようなデバイスで達成され得る。そのいくつかの例は、米国特許第5,846,796号；同第6,010,478号；同第5,865,796号；同第5,584,807号；および欧州特許番号第0500799号に記載されている。このアプローチは、注射針のない（無注射針）（ニードルフリー）送達アプローチを提供する。ここでは、微視的な粒子（例えば、ポリヌクレオチド粒子、またはポリペプチド粒子）の乾燥粉末処方物を、手持ちデバイスによって生成されたヘリウムガスジェット内で高速に加速し、目的の標的組織内へ粒子を噴射する。

【0225】

関連の実施形態において、本発明の組成物のガス駆動注射針なし注入に有用であり得る他のデバイスおよび方法は、Bioject, Inc. (Portland, OR)によって提供されるものが挙げられ、そのいくつかの例は、米国特許第4,790,824号；同第5,064,413号；同第5,312,335号；同第5,383,851号；同第5,399,163号；同第5,520,639号；および同第5,993,412号に記載されている。

【0226】

別の実施形態に従って、本明細書に記載される薬学的組成物は、本発明の免疫原性ポリヌクレオチド、ポリペプチド、抗体、T細胞および／またはAPC組成物に加えて、1つ以上の免疫賦活剤（免疫刺激因子）を含む。免疫賦活剤とは、本質的に、外因性抗原に対する免疫応答（抗原および／または細胞媒介）を増大または増強する任意の物質をいう。免疫賦活剤の1つの好ましい型は、アジュバントを含む。多くのアジュバントは、迅速な異化作用から抗原を保護するように設計された物質（例えば、水酸化アルミニウムまたは鉱油）および免疫応答の刺激物質（例えば、脂質A、Bordetella pertussisまたはMycobacterium tuberculosis誘導化タンパク質）を含む。特定のアジュバントは、例えば、フロイント不完全アジュバントおよびフロイント完全アジュバント（Difco Laboratories, Detroit, MI）；Merck Adjuvant 65（Merck and Company, Inc., Rahway, NJ）；AS-2（SmithKline Beecham, Philadelphia, PA）；アルミニウム塩（例えば、水酸化アルミニウムゲル（ミョウバン）またはリン酸アルミニウム）；カルシウム、鉄または亜鉛の塩；アシル化チロシンの不溶性懸濁液；アセチル化糖；カチオンの誘導化多糖またはアニオンの誘導化多糖；ポリフォスファゼン（polyphosphazene）；生分解性ミクロスフェア：モノホスホリル脂質Aおよびモノホスホリルクイル（quil）Aとして、市販されている。サイトカイン（例えば、GM-CSF、インターロイキン-2、インターロイキン-7、インターロイキン-12および成長因子のような他のもの）もまた、アジュバントとして使用され得る。

【0227】

本発明の特定の実施形態において、アジュバント組成物は、Th1型の免疫応答を優勢に誘導するものが好ましい。高レベルのTh1型サイトカイン（例えば、IFN- γ 、TNF α 、IL-2およびIL-12）は、投与される抗原に対する細胞媒介性応答の誘導を支持する傾向がある。対照的に、高レベルのTh2型サイトカイン（例えば、IL-4、IL-5、IL-6およびIL-10）は、体液性免疫応答の誘導を支持する傾向がある。本明細書中に提供されるようなワクチンの適用の後、患者は、Th1型応答およびTh2型応答を含む免疫応答を支持する。応答が優勢にTh1型である好ましい実施形態において、Th1型サイトカインのレベルは、Th2型サイトカインのレベルより高い程度まで増加する。これらのサイトカインのレベルは、標準的アッセイを使用して容易に評価され得る。サイトカインのファミリーの概説については、MosmannおよびCoffman、Ann. Rev. Immunol. 7:145~173、1989を参照のこと。

【0228】

Th1型優勢の応答を誘発するための特定の好ましいアジュバントは、例えば、モノホスホリルリピドA、好ましくは3-de-O-アシル化モノホスホリルリピドAとアルミニウム塩との組み合わせを含む。MPL（登録商標）アジュバントは、Corixa Corporation（Seattle, WA；例えば、米国特許第4,436,727号；同第4,877,611号；同第4,866,034号および同第4,912,094号を参照のこと）から入手可能である。CpG含有オリゴヌクレオチド（ここで、CpGジヌクレオチドはメチル化されていない）はまた、Th1優勢の応答を誘導する。このようなオリゴヌクレオチドは周知であり、そして例えばWO96/02555、WO99/33488ならびに米国特許第6,008,200号および同第5,856,462号に

記載される。免疫賦活剤DNA配列がまた、例えば、S a t oら、S c i e n c e 273:352, 1996によって記載される。別の好ましいアジュバントは、サポニン（例えば、Q u i l A）、またはその誘導体（Q S 21およびQ S 7（A q u i l a B i o p h a r m a c e u t i c a l s I n c., F r a m i n g h a m, M A）を含む）；E s c i n；D i g i t o n i n（ジキトニン）；またはG y p s o p h i l aもしくはC h e n o p o d i u m q u i n o aサポニンを含む。他の好ましい処方物としては、本発明のアジュバントの組み合わせ、例えば、Q S 21、Q S 7、Q u i l A、 β -エスシン、またはジキトニンを含む以下の群のうち少なくとも2つの組み合わせ、において1つより多いサポニンを含む。

【0229】

あるいは、このサポニン処方物は、キトサン、または他のポリカチオン性ポリマーからなるワクチンビヒクル、ポリラクチドおよびポリラクチド- α -グリコリド粒子、ポリ-N-アセチルグルコサミンベースのポリマーマトリックス、ポリサッカライドまたは化学的に改変されたポリサッカライドからなる粒子、リポソームおよび脂質ベースの粒子、グリセロールモノエステルからなる粒子、などと組み合わせられ得る。サポニンはまた、コレステロールの存在下で処方されて、リポソームまたはI S C O Mのような粒子構造を形成し得る。さらに、サポニンは、非粒子の溶液もしくは懸濁液で、または小数層リポソームもしくはI S C O Mのような粒子状構造で、ポリオキシエチレンエーテルまたはエステルと一緒に処方され得る。サポニンはまた、C a r b o p o l（登録商標）のような賦形剤と一緒に処方されて、粘度を増大され得るか、またはラクトースのような粉末賦形剤とともに乾燥粉末形態に処方され得る。

【0230】

1つの好ましい実施形態において、アジュバント系は、モノホスホリルリピドAとサポニン誘導体との組み合わせ（例えば、W O 94/00153に記載されるような、Q S 21と3 D-M P L（登録商標）アジュバントとの組み合わせ、またはW O 96/33739に記載されるような、Q S 21がコレステロールでクエンチ（q u e n c h）される、あまり反応発生的（r e a c t o g e n i c）でない組成物）を含む。他の好ましい処方物は、水中油型エマルジョンおよびトコフェロールを含む。水中油型エマルジョン中のQ S 21、3 D-M P L（登録商標）アジュバントおよびトコフェロールを使用する、別の特に好ましいアジュバント処方物は、W O 95/17210に記載されている。

【0231】

別の強化されたアジュバント系は、C p G含有オリゴヌクレオチドとサポニン誘導体の組み合わせを含み、特にC p GとQ S 21の組み合わせがW O 00/09159に開示されている。好ましくは、この処方物は、さらに、水中油型エマルジョンおよびトコフェロールを含む。

【0232】

本発明の薬学的組成物における使用のためのさらなる代表的アジュバントとしては、M o n t a n i d e I S A 720（S e p p i c, F r a n c e）、S A F（C h i r o n, C a l i f o r n i a, U n i t e d S t a t e s）、I S C O M S（C S L）、M F-59（C h i r o n）、アジュバントのS B A Sシリーズ（例えば、S B A S-2またはS B A S-4（S m i t h K l i n e B e e c h a m, R i x e n s a r t, B e l g i u mから入手可能）、D e t o x（E n h a n z y n（登録商標））（C o r i x a, H a m i l t o n, M T）、R C-529（C o r i x a, H a m i l t o n, M T）および他のアミノアルキルグルコサミニド4-ホスフェート（A G P）（例えば、係属中の米国出願登録番号08/853, 826および09/074, 720（これらの開示は、その全体が本明細書中に参考として援用される）に記載されるアジュバント）、ならびにW O 99/52549 A1に記載のポリオキシエチレンエーテルアジュバントのようなアジュバントが挙げられる。

【0233】

他の好ましいアジュバントは、一般式（I）： $\text{HO}(\text{CH}_2\text{CH}_2\text{O})_n\text{-A-R}$ のアジ

10

20

30

40

50

ュバント分子を含み、

ここで、 n は、1～50であり、 A は、結合または $-C(O)-$ であり、 R は、 $C_1 \sim 50$ アルキルまたはフェニル $C_1 \sim 50$ アルキルである。

【0234】

本発明の1つの実施形態は、一般式(I)のポリオキシエチレンエーテルを含むワクチン処方物からなり、ここで、 n は、1と50との間、好ましくは4～24、最も好ましくは9であり、この R 成分は、 $C_1 \sim 50$ 、好ましくは $C_4 \sim C_{20}$ アルキル、そして最も好ましくは C_{12} アルキルであり、そして A は、結合である。ポリオキシエチレンエーテルの濃度は、0.1～20%の範囲、好ましくは0.1～10%、そして最も好ましくは0.1～1%の範囲にあるべきである。好ましいポリオキシエチレンエーテルは、以下の群から選択される：ポリオキシエチレン-9-ラウロイルエーテル、ポリオキシエチレン-9-ステアシル(*stearyl*)エーテル、ポリオキシエチレン-8-ステアシル(*stearyl*)エーテル、ポリオキシエチレン-4-ラウリルエーテル、ポリオキシエチレン-35-ラウリルエーテルおよびポリオキシエチレン-23-ラウリルエーテル。ポリオキシエチレンラウリルエーテルのようなポリオキシエチレンエーテルは、*Merck Index* (第12版、7717項目)に記載される。これらのアジュバント分子は、WO99/52549に記載される。

【0235】

上記の一般式(I)に従うポリオキシエチレンエーテルは、所望の場合、別のアジュバントと組み合わせられ得る。例えば、好ましいアジュバントの組み合わせは、係属中の英国特許出願第GB9820956.2号に記載されるようなCpGとの組み合わせである。

【0236】

本発明の別の実施形態に従って、本明細書に記載される免疫原性組成物は、抗原提示細胞(APC)(例えば、樹状細胞、マクロファージ、B細胞、単球、および有効なAPCであるように操作され得る他の細胞)を介して宿主に送達される。このような細胞は、抗原を提示する能力を増大するように、T細胞応答の活性化および/または維持を改良するように、それ自体で抗腫瘍効果を有するように、そして/あるいは受け手と免疫学的に適合性(すなわち、一致するHLAハプロタイプ)であるように遺伝学的に改変されてもよいが、改変される必要はない。APCは、一般に、種々の生物学的な流体および器官(腫瘍および腫瘍周辺組織を含む)のいずれかから単離されてもよいし、そして自己細胞、同種異系細胞、同系細胞、または異種細胞であってもよい。

【0237】

本発明の特定の好ましい実施形態は、抗原提示細胞として、樹状細胞またはその前駆細胞を使用する。樹状細胞は、高度に強力なAPCであり(*Banchereau*および*Steinman, Nature* 392:245-251, 1998)、そして予防的または治療的な抗腫瘍免疫性を誘発するための生理学的アジュバントとして有効であることが示されてきた(*Timmerman*および*Levy, Ann. Rev. Med.* 50:507-529, 1999を参照のこと)。一般に、樹状細胞は、それらの代表的な形状(インサイチュでは星状、インビトロでは目に見える顕著な細胞質プロセス(樹状突起)を有する)、高い効率で抗原を取り込み、プロセッシングし、そして提示するそれらの能力、および未処理のT細胞応答を活性化するそれらの能力に基づいて同定され得る。もちろん樹状細胞は、インビボまたはエキソビボで樹状細胞上に通常見出されない特定の細胞表面レセプターまたはリガンドを発現するように操作され得、このような改変樹状細胞は本発明によって意図される。樹状細胞の代替として、分泌小胞抗原装荷樹状細胞(*secreted vesicles antigen-loaded dendritic cell*)(エキソソーム(*exosome*)と呼ばれる)がワクチン内で使用され得る(*Zitvogel*ら、*Nature Med.* 4:594-600, 1998を参照のこと)。

【0238】

樹状細胞および前駆細胞は、末梢血、骨髓、腫瘍浸潤細胞、腫瘍周辺組織浸潤細胞、リンパ節、脾臓、皮膚、臍帯血、または任意の他の適切な組織もしくは流体から得られ得る。例えば、樹状細胞は、末梢血から収集された単球の培養物に、GM-CSF、IL-4、IL-13および／またはTNF α のようなサイトカインの組み合わせを添加することによってエキソビボで分化され得る。あるいは、末梢血、臍帯血または骨髓から収集されたCD34陽性細胞は、培養培地にGM-CSF、IL-3、TNF α 、CD40リガンド、LPS、flt3リガンドおよび／または樹状細胞の分化、成熟、および増殖を誘導する他の化合物の組み合わせを添加することによって、樹状細胞に分化され得る。

【0239】

樹状細胞は、「未熟」細胞および「成熟」細胞として都合良く分類され、このことは、2つの充分に特徴付けられた表現型の間を単純な方法で区別することを可能にする。しかしこの命名は、あらゆる可能な分化の中間段階を排除すると解釈されるべきではない。未熟な樹状細胞は、抗原の取り込みおよびプロセッシングの高い能力を有するAPCとして特徴付けられ、この能力は、Fc γ レセプターおよびマンノースレセプターの高度な発現と相関する。成熟表現型は、代表的に、クラスIおよびクラスII MHC、接着分子（例えば、CD54およびCD11）ならびに同時刺激性分子（例えば、CD40、CD80、CD86および4-1BB）のようなT細胞活性化を担う細胞表面分子の高度な発現ではなく、これらのマーカーのより低い発現によって特徴付けられる。

【0240】

APCは、一般に、本発明のポリヌクレオチド（またはその部分もしくは他の改変体）を用いてトランスフェクトされ得、その結果、このコードされたポリペプチドまたはその免疫原性部分が細胞表面上に発現される。このようなトランスフェクションはエキソビボで生じ得、次いでこのようなトランスフェクトされた細胞を含む薬学的組成物は、本明細書中に記載されるように、治療目的のために使用され得る。あるいは、樹状細胞または他の抗原提示細胞を標的とする遺伝子送達ビヒクルが、患者に投与され得、インビボで起こるトランスフェクションを生じる。樹状細胞のインビボおよびエキソビボでのトランスフェクションは、例えば、WO97/24447に記載される方法、またはMahvira、Immunology and Cell Biology 75:456-460、1997によって記載される遺伝子銃アプローチのような当該分野で公知の任意の方法を使用して一般に実施され得る。樹状細胞の抗原ローディングは、樹状細胞または前駆細胞を、腫瘍ポリペプチド、DNA（裸のもしくはプラスミドベクター中の）またはRNA；あるいは抗原発現性組換え細菌またはウイルス（例えば、牛痘、鶏痘、アデノウイルスまたはレンチウイルスのベクター）とインキュベートすることによって達成され得る。ローディングの前に、ポリペプチドは、T細胞補助を提供する免疫学的パートナー（例えば、キャリア分子）に共有結合され得る。あるいは、樹状細胞は、別々にまたはポリペプチドの存在下で、結合していない免疫学的パートナーでパルスされ得る。

【0241】

当業者に公知の任意の適切なキャリアが、本発明の薬学的組成物において使用され得るが、このタイプのキャリアは、代表的に投与の様式に依存して変化する。本発明の組成物は、投与の任意の適切な様式、例えば、局所投与、経口投与、経鼻投与、粘膜投与、静脈投与、頭蓋内投与、腹腔内投与、皮下投与、および筋肉内投与を含む様式のために処方され得る。

【0242】

このような薬学的組成物内での使用のためのキャリアは、生体適合性であり、そしてまた生分解性であり得る。特定の実施形態において、好ましくは、この処方物は比較的一定レベルの活性成分の放出を提供する。しかし、他の実施形態において、投与直後のより迅速な放出速度が所望され得る。このような組成物の処方十分に、公知の技術を使用する当業者のレベル内である。この点に関して有用な例示的なキャリアとしては、ポリ（ラクチド-c o -グリコリド）、ポリアクリレート、ラテックス、デンプン、セルロース、デキストランなどの微粒子が挙げられる。他の例示的な徐放性キャリアとしては、非液性（n

on-liquid) 親水性コア (例えば、架橋ポリサッカリドまたはオリゴサッカリド) を含む超分子バイオベクター、ならびに必要なに応じて、両親媒性化合物を含む外部層 (例えば、リン脂質) (例えば、米国特許第 5, 151, 254 号および PCT 出願 WO 94/20078、WO/94/23701 および WO 96/06638 を参照のこと) を含む超分子バイオベクターが挙げられる。徐放性処方物内に含まれる活性な化合物の量は、移植の部位、放出の速度および予期される持続期間、ならびに処置または予防されるべき状態の性質に依存する。

【0243】

別の例示的な実施形態において、生分解性ミクロスフェア (例えば、ポリ乳酸ポリグリコレート) は、本発明の組成物のためのキャリアとして使用される。適切な生分解性ミクロスフェアは、例えば、米国特許第 4, 897, 268 号; 同第 5, 075, 109 号; 同第 5, 928, 647 号; 同第 5, 811, 128 号; 同第 5, 820, 883 号; 同第 5, 853, 763 号; 同第 5, 814, 344 号; 同第 5, 407, 609 号および同第 5, 942, 252 号に開示される。改変 B 型肝炎コアタンパク質キャリア系 (例えば、WO/99 40934 およびそこで引用される参考文献に記載されるような系) もまた、多くの用途に有用である。別の例示的なキャリア送達系 (システム) は、キャリア含有粒子-タンパク質複合体 (例えば、米国特許第 5, 928, 647 号に記載される複合体) を使用し、これは、宿主においてクラス I 拘束細胞傷害性 T リンパ球応答を誘導し得る。

10

【0244】

本発明の薬学的組成物はしばしば、1 つ以上の、緩衝剤 (例えば、中性の緩衝化生理食塩水またはリン酸緩衝化生理食塩水); 糖質 (例えば、グルコース、マンノース、スクロースまたはデキストラン); マンニトール; タンパク質; ポリペプチドまたはアミノ酸 (例えば、グリシン); 抗酸化剤; 静菌剤; キレート剤 (例えば、EDTA またはグルタチオン); アジュバント (例えば、水酸化アルミニウム); 処方物をレシipient の血液に対して等張性、低張性または弱く高張性にする溶質; 懸濁剤; 濃化剤および/または保存剤をさらに含む。あるいは、本発明の組成物は、凍結乾燥物として処方され得る。

20

【0245】

本明細書中に記載される薬学的組成物は、単回用量容器または多用量容器 (例えば、密閉アンプルまたはバイアル) 中に存在し得る。このような容器は、代表的に、使用まで処方物の無菌性および安定性を維持するような様式で、密封される。一般に、処方物は、油性ビヒクルまたは水性ビヒクル中の、懸濁液、溶液またはエマルジョンとして保存され得る。あるいは、薬学的組成物は、使用の直前に無菌水性キャリアの添加のみを必要とする、凍結乾燥状態で保存され得る。

30

【0246】

様々な処置レジメンにおいて本明細書中で記載される特定の組成物を使用するための適切な投薬レジメンおよび処置レジメン (例えば、経口、非経口、静脈内、鼻腔内、および筋肉内の投与および処方を含む) の開発は、当該分野で周知であり、これらのいくつかは、一般的な例示目的のために以下で簡単に考察される。

【0247】

特定の用途において、本明細書中で開示される薬学的組成物は、経口投与によって動物に送達され得る。このように、これらの組成物は、不活性希釈剤と共にかまたは同化可能食用キャリアと共に処方され得るか、あるいはこれらの組成物は、硬質または軟質ゼラチンカプセルに入れられ得るか、あるいはこれらの組成物は錠剤に圧縮され得るか、あるいはこれらの組成物は食餌の食品に直接組み込まれ得る。

40

【0248】

活性化合物は、さらに賦形剤と共に組み込まれ得、そして経口摂取錠剤、経頬粘膜錠剤、トローチ、カプセル剤、エリキシル、懸濁液、シロップ、ウェハ剤などの形態で使用される (例えば、Mathiowitz ら、Nature 1997 Mar 27; 386 (6623): 410-4; Hwang ら、Crit Rev Ther Drug C

50

carrier Syst 1998; 15 (3): 243-84; 米国特許第5, 641, 515号; 米国特許第5, 580, 579号および米国特許第5, 792, 451号を参照のこと)。錠剤、トローチ、丸剤、カプセル剤などはまた、任意の種々のさらなる成分を含み得、例えば、結合因子（例えば、トラガカントゴム、アラビアゴム、コーンスターチまたはゼラチン）；賦形剤（例えば、リン酸二カルシウム）；崩壊剤（例えば、コーンスターチ、ポテトデンプン、アルギン酸など）；潤沢剤（例えば、ステアリン酸マグネシウム）；および甘味料（例えば、スクロース、ラクトースまたはサッカリン）または香料（例えば、ペパーミント、ウインターグリーン油、またはチェリー香料）が添加され得る。投薬量単位形態がカプセル剤である場合、これは上記のタイプの材料に加えて液体キャリアを含み得る。様々な他の材料が、コーティングとして存在し得るか、またはそうでなければ投薬単位の物理的形態を改変し得る。例えば、錠剤、丸剤またはカプセル剤は、セラック、糖またはその両方でコーティングされ得る。もちろん、任意の投薬単位形態を調製する際に使用される任意の材料は、薬学的に純粋でありかつ用いられる量で実質的に非毒性でなければならない。さらに、活性化合物が、持続性放出調製物および処方物に組み込まれ得る。

【0249】

代表的に、これらの処方物は、少なくとも約0.1%またはそれよりも多い活性化合物を含み得るが、活性成分の割合は、もちろん、変化し得、そして好都合には、全処方物の重量または体積の約1または2%と、約60%または70%以上のとの間であり得る。当然、治療的に有用な組成物の各々の中の活性化合物の量は、適切な投薬量が、化合物の任意の所定の単位用量において得られるような様式で、調製され得る。溶解度、バイオアベイラビリティ、生物学的半減期、投与の経路、製品の有効期限のような因子、および他の薬理学的考慮が、このような薬学的処方物を調製する分野の当業者によって意図され、そしてそのようなものとして、種々の投薬量および処置レジメンが所望され得る。

【0250】

あるいは、経口投与について、本発明の組成物は、うがい薬、歯みがき剤、バツカル錠、経口スプレー、または舌下経口投与処方物の形態で、1つ以上の賦形剤と混合され得る。あるいは、活性成分は、経口溶液（例えば、ホウ酸ナトリウム、グリセリンおよび炭酸水素カリウムを含む溶液）に組み込まれ得るか、または歯みがき剤に分散され得るか、または治療の有効量で、水、結合因子、研磨剤、香料、発泡剤、および湿潤剤を含み得る組成物に添加され得る。あるいは、これらの組成物は、舌下に置かれ得るか、そうでなければ口の中で溶解され得る錠剤形態または溶液形態に成形され得る。

【0251】

特定の状況において、本明細書中で開示される薬学的組成物を、非経口送達、静脈内送達、筋肉内送達またはさらに腹腔内送達することが所望される。このようなアプローチは、当業者に周知であり、これらのいくつかは例えば、例えば、米国特許第5, 543, 158号; 米国特許第5, 641, 515号および米国特許第5, 399, 363号にさらに記載される。特定の実施形態において、遊離塩基または薬学的に受容可能な塩としての活性化合物の溶液は、界面活性剤（例えば、ヒドロキシプロピルセルロース）と適切に混合された水中で調製され得る。分散物もまた、グリセロール、液体ポリエチレングリコール、およびそれらの混合物中で、ならびにオイル中で調製され得る。保存および使用の通常の条件下で、これらの調製物は一般に、微生物の増殖を防止するために、防腐剤を含む。

【0252】

注入用途のために適切な例示的な薬学的形態として、滅菌水性液剤または分散物および滅菌注射用液剤または分散物の即時調製のための滅菌散剤が挙げられる（例えば、米国特許第5, 466, 468号を参照のこと）。全ての場合において、その形態は、滅菌でなければならない、そして容易に注射することができる程度に、流動性でなければならない。それは、製造および保存の条件下で安定でなければならない、そして微生物（例えば、細菌および真菌）の汚染作用に対して保存されなければならない。キャリアは、例えば、水、エタノール、ポリオール（例えば、グリセロール、プロピレングリコール、および液体ポリ

エチレングリコールなど)、それらの適切な混合物、および／もしくは植物油を含む溶媒または分散媒体であり得る。適切な流動性は、例えば、コーティング(例えば、レシチン)の使用によって、分散物の場合には必要とされる粒子サイズの維持によって、そして／または界面活性剤の使用によって、維持され得る。微生物の作用の防止は、種々の抗菌剤および抗真菌剤(例えば、パラベン(paraben)、クロロブタノール、フェノール、ソルビン酸、チメロサルなど)によって促進され得る。多くの場合において、等張剤(例えば、糖または塩化ナトリウム)を含めることが、望ましい。注射可能な組成物の長期にわたる吸収は、組成物中での吸収を遅延させる薬剤(例えば、モノステアリン酸アルミニウムおよびゼラチン)の使用によって、もたらされ得る。

【0253】

一実施形態において、水溶液の非経口投与のために、必要ならばその溶液は、適切に緩衝化されるべきであり、そして液体希釈剤が、最初に、十分な生理食塩水またはグルコースで等張にされる。これらの特定の水溶液は、特に、静脈内投与、筋肉内投与、皮下投与、および腹腔内投与に適切である。これに関連して、使用され得る滅菌水性媒体は、本開示の観点から当業者に公知である。例えば、一投与量は、1mlの等張性NaCl溶液に溶解され得、そして1000mlの皮下注入流体に添加され得るか、または注入予定部位で注入され得るかのいずれかである(例えば、「Remington's Pharmaceutical Sciences」、第15版、1035~1038頁および1570~1580頁を参照のこと)。投薬量におけるいくつかの変化が、処置される被験体の状態に依存して必然的に生じる。さらに、ヒト投与について、調製物は、もちろん、FDA Office of Biological Standardsによって要求されるような、無菌性、発熱性、ならびに一般的な安全性標準および純度標準を好ましく満たす。

【0254】

本発明の別の実施形態において、本明細書中で開示される組成物は、中性形態または塩形態で処方され得る。例示的な薬学的に受容可能な塩としては、酸付加塩(タンパク質の遊離アミノ基とともに形成される)が挙げられ、そしてこれらの塩は、無機酸(例えば、塩酸またはリン酸など)、または酢酸、シュウ酸、酒石酸、マンデル酸などの有機酸とともに形成される。遊離カルボキシル基とともに形成される塩もまた、無機塩基(例えば、水酸化ナトリウム、水酸化カリウム、水酸化アンモニウム、水酸化カルシウム、または水酸化鉄)、およびイソプロピルアミン、トリメチルアミン、ヒスチジン、プロカインなどの有機塩基から誘導され得る。処方の際に、溶液は、投薬処方物と適合する様式でかつ治療的に有効な量で投与される。

【0255】

キャリアは、任意および全ての溶媒、分散媒体、ビヒクル、コーティング、希釈剤、抗菌剤および抗真菌剤、等張剤および吸収を遅延させる薬剤、緩衝液、キャリア溶液、懸濁液、コロイドなどをさらに含み得る。薬学的活性物質のためのこのような媒体および薬剤の使用は、当該分野で周知である。任意の従来の媒体または薬剤が、活性成分と不適合である場合を除いては、治療組成物におけるその使用が、企図される。補助的活性成分もまた、これらの組成物に組み込まれ得る。句「薬学的に受容可能な」とは、ヒトに投与される場合に、アレルギー反応または類似の厄介な反応を生じない分子実体および組成物をいう。

【0256】

特定の実施形態において、薬学的組成物は、鼻腔内スプレー、吸入、および／または他のエアロゾル送達ビヒクルによって送達され得る。遺伝子、核酸およびペプチド組成物を、経鼻エアロゾルスプレーを介して肺に直接送達するための方法は、例えば、米国特許第5,756,353号および米国特許第5,804,212号に記載されている。同様に、鼻腔内微粒子樹脂(Takenagaら、J Controlled Release 1998 Mar 2; 52(1-2): 81-7)およびリゾホスファチジルグリセロール化合物(米国特許第5,725,871号)を使用する薬物の送達はまた、薬学分野で周知である。同様に、ポリテトラフルオロエチレン支持マトリックスの形態での例示

10

20

30

40

50

的な経粘膜薬物送達は、米国特許第5,780,045号に記載される。

【0257】

特定の実施形態において、リポソーム、ナノカプセル、微粒子、脂質粒子、小胞などを、適切な宿主細胞／生物への本発明の組成物の導入のために使用する。特に、本発明の組成物は、脂質粒子、リポソーム、小胞、ナノスフィア、またはナノ粒子などのいずれかにカプセル化して送達するために処方され得る。あるいは、本発明の組成物は、このようなキャリアビヒクルの表面に、共有結合的または非共有結合的のいずれかで結合され得る。

【0258】

潜在的な薬物キャリアとしてのリポソームおよびリポソーム様調製物の形成および使用は一般に、当業者に公知である（例えば、Lasica, Trends Biotechnol 1998 Jul; 16(7): 307-21; Takakura, Nippon Rinsho 1998 Mar; 56(3): 691-5; Chandranら、Indian J Exp Biol. 1997 Aug; 35(8): 801-9; Margalit, Crit Rev Ther Drug Carrier Syst. 1995; 12(2-3): 233-61; 米国特許第5,567,434号; 米国特許第5,552,157号; 米国特許第5,565,213号; 米国特許第5,738,868号および米国特許第5,795,587号（これらの各々は、その全体が本明細書中で参考として詳細に援用される）を参照のこと）。

【0259】

リポソームは、T細胞懸濁液、初代肝細胞培養物およびPC12細胞を含む他の手順によるトランスフェクションが通常困難である多くの細胞型とともに、首尾よく使用されている（Renneisenら、J Biol Chem. 1990 Sep 25; 265(27): 16337-42; Mullerら、DNA Cell Biol. 1990 Apr; 9(3): 221-9）。さらに、リポソームは、ウイルスベースの送達系に典型的なDNA長の制限がない。リポソームは、遺伝子、種々の薬物、放射線治療剤、酵素、ウイルス、転写因子およびアロステリックエフェクターなどを、種々の培養細胞株および動物に導入するために効果的に使用されている。さらに、リポソームの使用は、全身送達後の、自己免疫応答にも受容不可能な毒性にも関連しないようである。

【0260】

特定の実施形態において、リポソームは、水性媒体中に分散したリン脂質から形成され、そして多層同心性二重層小胞（多層小胞（MLV）とも呼ばれる）を自発的に形成する。

【0261】

あるいは、他の実施形態において、本発明は、本発明の組成物の薬学的に受容可能なナノカプセル処方物を提供する。ナノカプセルは一般に、化合物を安定かつ再現可能な様式で捕獲し得る（例えば、Quintanar-Guerreroら、Drug Dev Ind Pharm. 1998 Dec; 24(12): 1113-28を参照のこと）。細胞内ポリマー過負荷に起因する副作用を回避するために、このような超微粒子（約0.1 μ mのサイズ）は、インビボで分解され得るポリマーを使用して設計され得る。このような粒子は、例えば、Couvreurら、Crit Rev Ther Drug Carrier Syst. 1988; 5(1): 1-20; zur Muhlenら、Eur J Pharm Biopharm. 1998 Mar; 45(2): 149-55; Zambauxら、J Controlled Release. 1998 Jan 2; 50(1-3): 31-40; および米国特許第5,145,684号に記載されるようにして作製され得る。

【0262】

（癌治療法）

癌治療への免疫学的アプローチは、癌細胞が、異常な細胞および分子または外来の細胞および分子に対する人体の防御をしばしば回避し得、そしてこれらの防御が、失った基盤（ground）を取り戻すために免疫的に刺激され得るという認識に基づく（例えば、Klein, Immunology (Wiley-Interscience, New York) 1998; 70(1): 1-10)。

ork、1982)、623～648頁)。種々の免疫エフェクターが腫瘍の増殖を直接的または間接的に阻害し得るという多数の最近の知見は、癌治療に対するアプローチにおける更新された関心を導いた(例えば、Jagerら、Oncology 2001; 60(1): 1-7; Rennerら、Ann Hematol 2000 Dec; 79(12): 651-9)。

【0263】

抗腫瘍細胞の免疫および人体からの腫瘍細胞の除去に関連する機能を有する4つの基本的な細胞型は、以下の通りである: i) 非自己侵入細胞を認識および標識するために免疫グロブリンを血漿に分泌するBリンパ球; ii) 免疫グロブリンでコーティングされた標的侵入細胞の溶解および処理を担う補体タンパク質を分泌する単球; iii) 腫瘍細胞の破壊、抗体依存性細胞傷害およびナチュラルキリング(natural killing)のための2種類の機構を有するナチュラルキラーリンパ球; ならびにiv) 抗原特異的レセプターを保持し、そして相補的マーカー分子を保有する腫瘍細胞を認識する能力を有するTリンパ球(Schreiber, H., 1989, Fundamental Immunology (編)、W. E. Paul, 923～955頁)。

10

【0264】

癌の免疫治療は、一般に、体液性免疫応答、細胞性免疫応答、またはそれらの両方を誘導することに焦点が当てられている。さらに、CD4⁺Tヘルパー細胞の誘導は、抗体または細胞傷害性CD8⁺T細胞のいずれかを二次的に誘導するために必要であることが十分に確立されている。癌細胞、特に、肺癌細胞に対して選択的であるか、または理想的には特異的であるポリペプチド抗原は、肺癌に対する免疫応答を誘導するための強力なアプローチを提供し、そして本発明の重要な局面である。

20

【0265】

従って、本発明のさらなる局面において、本明細書に記載される薬学的組成物は、癌を処置するため、特に、肺癌の免疫治療のために使用され得る。このような方法において、本明細書中に記載される薬学的組成物は、患者、代表的には、温血動物、好ましくはヒトに投与される。患者は、癌に罹患していてもしていなくてもよい。従って、上記薬学的組成物は、癌の発症を予防するため、または癌に罹患した患者を処置するために使用され得る。薬学的組成物およびワクチンは、原発性腫瘍の外科的除去および/または放射線治療剤もしくは従来の化学療法剤の投与のような処置の前、あるいはその後のいずれかに投与され得る。上記のように、薬学的組成物の投与は、任意の適切な方法(静脈内、腹腔内、筋肉内、皮下、鼻腔内、皮内、肛門、膣、局所的小および経口経路による投与を含む)によってであり得る。

30

【0266】

特定の実施形態において、免疫療法は、能動的免疫療法であり得、この療法にける処置は、免疫応答改変剤(例えば、本明細書中で提供されるポリペプチドおよびポリヌクレオチド)の投与を用いる、腫瘍に対して反応する内因性宿主免疫系のインビボ刺激に依存する。

【0267】

他の実施形態において、免疫療法は、受動的免疫療法であり得、この療法における処置は、確立された腫瘍免疫反応性を有する因子(例えば、エフェクター細胞または抗体)(これらは、抗腫瘍効果を直接的または間接的に媒介し得、そして必ずしもインタクトな宿主免疫系に依存しない)の送達を含む。エフェクター細胞の例としては、上記のようなT細胞、本明細書中に提供されるポリペプチドを発現するTリンパ球(例えば、CD8⁺細胞傷害性Tリンパ球およびCD4⁺Tヘルパー腫瘍浸潤性リンパ球)、キラー細胞(例えば、ナチュラルキラー細胞およびリンホカイン活性化キラー細胞)、B細胞および抗原提示細胞(例えば、樹状細胞およびマクロファージ)が挙げられる。本明細書中に列挙されるポリペプチドに特異的なT細胞レセプターおよび抗体レセプターは、養子免疫療法のために他のベクターまたはエフェクター細胞中にクローニングされ、発現され、そして移入され得る。本明細書に提供されるポリペプチドはまた、受動免疫療法のための抗体または抗

40

50

イディオタイプ抗体（上記および米国特許第4,918,164号に記載される）を生成するために用いられ得る。

【0268】

モノクローナル抗体は、検出、診断アッセイまたは治療適用において、所望される選択的な利用のために、任意の種々の標識で標識され得る（米国特許第6,090,365号；同第6,015,542号；同第5,843,398号；同第5,595,721号；および同第4,708,930号（各々が個々に援用されるかのように、本明細書中でこれらの全体が参考として援用されている）に記載されている）。各々の場合において、抗原の決定部位への標識されたモノクローナル抗体の結合は、特定の治療剤の検出または異常な細胞上の抗原決定部位への特定の治療剤の送達を示す。本発明のさらなる目的は、このようなモノクローナル抗体の所望される選択的な利用を達成するために適切に標識される、特定のモノクローナル抗体を提供することである。 10

【0269】

エフェクター細胞は、一般に、本明細書中に記載されるように、インビトロでの増殖により養子免疫治療のために十分な量で得られ得る。単一の抗原特異的エフェクター細胞を、インビボでの抗原認識を保持しながら数十億まで増殖させるための培養条件は当該分野で周知である。このようなインビトロの培養条件は代表的に、しばしばサイトカイン（例えば、IL-2）および分裂しない支持細胞の存在下で、抗原での間欠刺激を用いる。上で述べたように、本明細書中で提供される免疫反応性ポリペプチドは、抗原特異的T細胞培養物を急速に増殖するために用いられ、免疫治療に十分な数の細胞を生成し得る。詳細には、抗原提示細胞（例えば、樹状細胞、マクロファージ、単球、繊維芽細胞、および／またはB細胞）は、当該分野で周知の標準的技術を用いて、免疫反応性ポリペプチドでパルスされ得るか、または1つ以上のポリヌクレオチドでトランスフェクトされ得る。例えば、抗原提示細胞は、組換えウイルスまたは他の発現系における発現を増大するのに適切なプロモーターを有するポリヌクレオチドでトランスフェクトされ得る。治療において使用するための培養されたエフェクター細胞は、インビボで増殖されかつ広範に分散され得、そして長期間生存し得なければならない。培養されたエフェクター細胞が、インビボで増殖し、そしてIL-2を補充された抗原での反復刺激によって、長期間、多数生存するように誘導され得ることが研究で示されている（例えば、Cheeverら、Immunological Reviews 157:177、1997を参照のこと）。 20 30

【0270】

あるいは、本明細書において列挙されるポリペプチドを発現するベクターは、患者から得られた抗原提示細胞に導入され得、そして同じ患者に戻す移植のためにエキソビボでクローニ的に増殖され得る。トランスフェクトされた細胞は、当該分野で公知の任意の手段（好ましくは、静脈投与、腔内投与、腹腔内投与、または腫瘍内投与による滅菌形態）を用いて患者に再導入され得る。

【0271】

本明細書中に記載される治療的組成物の投与の経路および頻度、ならびに投薬量は、個々人で異なり、そして標準的技術を用いて容易に確立され得る。概して、薬学的組成物およびワクチンは、注射（例えば、皮内、筋肉内、静脈内、または皮下）により、経鼻的に（例えば、吸引により）または経口的に、投与され得る。好ましくは、52週間にわたって1~10用量の間が投与され得る。好ましくは、1ヶ月の間隔で6用量が投与され、そしてブースター（追加）ワクチン接種がその後定期的に与えられ得る。交互のプロトコルが個々の患者に適切であり得る。適切な用量は、上記のように投与された場合、抗腫瘍免疫応答を促進し得、そして基底（すなわち、未処置）レベルより少なくとも10~50%上である、化合物の量である。このような応答は、患者内の抗腫瘍抗体を測定することによってか、または患者の腫瘍細胞をインビトロで殺傷し得る細胞溶解性エフェクター細胞のワクチン依存性の生成によってモニターされ得る。このようなワクチンはまた、ワクチン接種されていない患者と比較すると、ワクチン接種された患者において、改善された臨床的結果（例えば、より頻繁な寛解、完全もしくは部分的に疾患を有さないか、またはより 40 50

長く疾患を有さない生存)を導く免疫応答を生じ得るはずである。一般に、1つ以上のポリペプチドを含む薬学的組成物およびワクチンについて、用量中に存在する各ポリペプチドの量は、宿主の体重(kg)あたり、約25 µg~5mgの範囲である。適切な用量サイズは、患者の大きさで変化するが、代表的には約0.1mL~約5mLの範囲である。

【0272】

一般に、適切な投薬量および処置レジメンは、治療的および／または予防的利点を提供するのに十分な量の活性化合物を提供する。このような応答は、処置されていない患者と比較して、処置された患者において、改善された臨床的結果(例えば、より頻繁な寛解、完全なまたは部分的な、あるいはより長い疾患なしでの生存)を確立することによってモニターされ得る。腫瘍タンパク質に対する既存の免疫応答における増加は、一般的に、改善された臨床的結果と関連する。このような免疫応答は、一般的に、標準的な増殖アッセイ、細胞障害性アッセイまたはサイトカインアッセイを使用して評価され得、これは、処置の前または後に患者から得られるサンプルを使用して行われ得る。

10

【0273】

(癌の検出および診断組成物、方法およびキット)

一般的に、癌は、患者から得られた生物学的サンプル(例えば、血液、血清、痰、尿、および／または腫瘍生検)における1つ以上の肺腫瘍タンパク質および／またはこのようなタンパク質をコードするポリヌクレオチドの存在に基づいてこの患者において検出され得る。言い換えると、このようなタンパク質は、肺癌のような癌の存在または非存在を示すためのマーカーとして使用され得る。さらに、このようなタンパク質は、他の癌の検出のために有用であり得る。本明細書中で提供される結合因子は、一般に、生物学的サンプル中で、この薬剤に結合する抗原のレベルの検出を可能にする。

20

【0274】

ポリヌクレオチドプライマーおよびプローブは、癌の存在または非存在もまた示す、腫瘍タンパク質をコードするmRNAのレベルを検出するために使用され得る。一般に、腫瘍の配列は、腫瘍が発生するのと同じ型の正常な組織よりも、腫瘍組織において、少なくとも2倍、好ましくは3倍、そしてより好ましくは5倍以上のレベルで存在するはずである。この腫瘍が発生するのとは異なる細胞型の特定の腫瘍配列の発現レベルは、特定の診断実施形態において無関係である。なぜならば、この腫瘍細胞の存在は、同一の型の正常組織における発現レベルに対する、腫瘍組織において予め決定された異なる発現レベル(例えば、2倍、5倍など)の観測により確認され得るからである。

30

【0275】

他の異なる発現パターンは、診断目的のために有利に利用され得る。例えば、本発明の1つの局面において、同一の型だが他の正常組織の型ではない、腫瘍組織および正常組織(例えば、PBMC)における腫瘍配列の過剰発現は、診断的に開発され得る。この場合において、例えば、循環組織部位または腫瘍が発生する組織部位とは異なるいくつかの組織部位から採取されたサンプル中における転移性腫瘍細胞の存在は、例えば、RT-PCR分析を使用して、このサンプル中の腫瘍配列の発現を検出することによって同定および／または確認され得る。多くの場合において、目的のサンプル中の腫瘍細胞(例えば、PBMC)を、細胞捕獲(cell capture)技術または他の類似の技術を使用して、富化することが所望される。

40

【0276】

サンプル中のポリペプチドマーカーを検出するために結合因子を使用するための、当業者に公知の種々のアッセイ型式が存在する。例えば、HarlowおよびLane、Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, 1988を参照のこと。一般的に、患者における癌の存在または非存在は、(a)患者から得られた生物学的サンプルを結合因子と接触させる工程；(b)この結合因子に結合するポリペプチドのレベルをサンプルにおいて検出する工程；および(c)ポリペプチドのレベルと所定のカットオフ値とを比較する工程によって決定され得る。

50

【0277】

好ましい実施形態において、このアッセイは、ポリペプチドに結合するためおよびサンプルの残りからポリペプチドを除くために固体支持体上に固定された結合因子の使用を含む。次いで、結合されたポリペプチドは、レポーター基を含み、結合因子／ポリペプチド複合体に特異的に結合する検出試薬を使用して検出され得る。このような検出試薬は、例えば、ポリペプチドまたは抗体に特異的に結合する結合因子あるいは結合因子に特異的に結合する他の薬剤（例えば、抗免疫グロブリン、プロテインG、プロテインAまたはレクチン）を含み得る。あるいは、競合アッセイが、使用され得、ここで、ポリペプチドは、レポーター基で標識され、そしてサンプルと結合因子のインキュベーション後にその固定された結合因子に結合される。サンプルの成分が、標識ポリペプチドの結合因子への結合を阻害する程度は、サンプルの固定された結合因子との反応性を示す。このようなアッセイにおける使用に適切なポリペプチドは、上記のような、全長肺腫瘍タンパク質および結合因子が結合するそのポリペプチド部分を含む。

10

【0278】

固体支持体は、腫瘍タンパク質が付着され得る当業者に公知の任意の材料であり得る。例えば、固体支持体は、マイクロタイタープレートにおける試験ウェルあるいはニトロセルロースまたは他の適切な膜であり得る。あるいは、その支持体は、ビーズまたはディスク（例えば、ガラス）、ファイバークラス、ラテックス、またはプラスチック物質（例えば、ポリスチレン、またはポリ塩化ビニル）であり得る。その支持体はまた、磁気粒子または光ファイバーセンサー（例えば、米国特許第5,359,681号に記載のような）であり得る。この結合因子は、当業者に公知の種々の技術を使用して固体支持体上に固定され得、これらは特許および科学文献に十分に記載されている。本発明の状況において、用語「固定化」とは、非共有結合的な会合（例えば、吸着）および共有結合的な付着（これは、薬剤と支持体上の官能基との間で直接結合され得るかまたは架橋剤を用いる結合であり得る）の両方をいう。マイクロタイタープレートのウェル、または膜への吸着による固定化が好ましい。このような場合、吸着は、適切な緩衝液中で固体支持体を用いて適切な時間量で結合因子に接触させることによって達成され得る。接触時間は、温度によって変化するが、代表的には、約1時間から約1日の間である。一般的には、約10ng～約10μg、そして好ましくは約100ng～約1μgの範囲の量の結合因子とプラスチックマイクロタイタープレート（例えば、ポリスチレンまたはポリ塩化ビニル）のウェルを接触させることは、適切な量の結合因子を固定化するのに十分である。

20

30

【0279】

固体支持体への結合因子の共有結合的付着は、一般に、支持体および結合因子上の官能基（例えば、水酸基またはアミノ基）の両方と反応する二官能性試薬と支持体を最初に反応させることによって達成され得る。例えば、この結合因子は、ベンゾキノンを用いるかまたは結合パートナー上のアミンおよび活性水素と支持体上のアルデヒド基との縮合によって、適切なポリマーコーティングを有する支持体に、共有結合的に付着され得る（例えば、Pierce Immunotechnology Catalog and Handbook, 1991, A12-A13を参照のこと）。

【0280】

特定の実施形態において、このアッセイは、2抗体サンドイッチアッセイである。本アッセイは、最初に、固体支持体（一般に、マイクロタイタープレートのウェル）上で固定されている抗体をサンプルに接触させて、その結果、サンプル内のポリペプチドを固定された抗体に結合させることによって実施され得る。次いで、非結合サンプルは固定されたポリペプチド-抗体複合体から除去され、そして検出試薬（好ましくは、そのポリペプチド上の異なる部位に結合し得る第2の抗体（レポーター基を含む））が添加される。次いで、固体支持体に結合したままである検出試薬の量が、特定のレポーター基に関して適切な方法を用いて決定される。

40

【0281】

より詳細には、一旦抗体が上記のように支持体上に固定化されると、支持体上の残りのタ

50

ンパク質結合部位は、典型的にはブロックされる。任意の適切なブロック剤（例えば、ウシ血清アルブミンまたは Tween 20TM（Sigma Chemical Co., St. Louis, MO））は、当業者に公知である。固定された抗体は次いで、サンプルとインキュベートされ、そしてポリペプチドをこの抗体に結合させる。インキュベーションの前に、このサンプルは適切な希釈液（例えば、リン酸緩衝化生理食塩水（PBS））で希釈され得る。概して、適切な接触時間（すなわち、インキュベーション時間）は、肺癌を有する個体から得られたサンプル内のポリペプチドの存在を検出するのに十分な時間である。好ましくは、この接触時間は、結合ポリペプチドと非結合ポリペプチドとの間の平衡で達成される結合レベルの少なくとも約 95% である結合レベルを達成するのに十分な時間である。当業者は、ある時間にわたって起こる結合レベルをアッセイすることによって、平衡に達するまでに必要な時間が容易に決定され得ることを認識する。室温では、一般に、約 30 分間のインキュベーション時間で十分である。

10

【0282】

次いで、非結合サンプルが、適切な緩衝液（例えば、0.1% Tween 20TM を含む PBS）を用いて固体支持体を洗浄することによって除去される。次いで、レポーター基を含む第 2 の抗体が固体支持体に添加され得る。好ましいレポーター基は、上記の基を含む。

【0283】

次いで、検出試薬が、結合されたポリペプチドを検出するのに十分な量の時間、固定された抗体-ポリペプチド複合体とインキュベートされる。適切な量の時間は、一般に、ある時間にわたって起こる結合のレベルをアッセイすることによって決定され得る。次いで、非結合の検出試薬は除去され、そして結合した検出試薬は、レポーター基を用いて検出される。レポーター基を検出するために使用される方法は、レポーター基の性質に依存する。放射性基について、一般的には、シンチレーション計数法またはオートラジオグラフィ法が適切である。分光法は、色素、発光基および蛍光基を検出するために使用され得る。ビチオンは、異なるレポーター基（一般に、放射性もしくは蛍光基または酵素）に結合されたアビジンを使用して検出され得る。酵素レポーター基は、一般に、基質の添加（一般には、特定の時間の間）、続いて反応産物の分光分析または他の分析により検出され得る。

20

【0284】

癌（例えば、肺癌）の存在または非存在を決定するために、固体支持体に結合したままのレポーター基から検出されるシグナルが、一般に、所定のカットオフ値と対応するシグナルと比較される。1つの好ましい実施形態において、癌の検出のためのカットオフ値は、固定された抗体を、癌を有さない患者由来のサンプルとインキュベートした際に得られる平均シグナル値である。概して、所定のカットオフ値を 3 標準偏差上回るシグナルを生じるサンプルが、癌に対して陽性とみなされる。代替の好ましい実施形態において、このカットオフ値は、Sackett ら、Clinical Epidemiology: A Basic Science for Clinical Medicine, Little Brown and Co., 1985, 106~7 頁の方法に従って、レシーバーオペレーターカーブ（Receiver Operator Curve）を使用して決定される。簡単に言うと、本実施形態において、このカットオフ値は、診断試験結果についての各可能なカットオフ値に対応する真の陽性割合（すなわち、感度）および偽陽性割合（100% - 特異性）の対のプロットから決定され得る。プロット上の上方左手の角に最も近いカットオフ値（すなわち、最大領域を囲む値）が、最も正確なカットオフ値であり、そして本方法によって決定されたカットオフ値より高いシグナルを生ずるサンプルが陽性に見なされ得る。あるいは、カットオフ値は、偽陽性割合を最小にするためにプロットに沿って左へシフトされ得るか、または偽陰性割合を最小にするために右へシフトされ得る。概して、本方法によって決定されたカットオフ値より高いシグナルを生ずるサンプルが、癌に対して陽性に見なされる。

30

40

【0285】

50

関連の実施形態において、このアッセイは、フロースルー試験形式またはストリップ試験形式で実行される（ここで、結合因子は、ニトロセルロースのような膜上で固定化される）。フロースルー試験では、サンプル内のポリペプチドは、サンプルが膜を通過するにつれて固定された結合因子に結合する。次いで、第2の標識化された結合因子が、この第2の結合因子を含む溶液がその膜を介して流れるにつれて、結合因子-ポリペプチド複合体と結合する。次いで、結合した第2の結合因子の検出は、上記のように実行され得る。ストリップ試験形式では、結合因子が結合される膜の一端を、サンプルを含む溶液中に浸す。このサンプルは、膜に沿って、第2の結合因子を含む領域を通して、そして固定された結合因子の領域まで移動する。固定された抗体の領域での第2の結合因子の濃度が、癌の存在を示す。代表的には、その部位での第2の結合因子の濃度は、視覚的に読みとられ得るパターン（例えば、線）を生成する。このようなパターンを示さないことは陰性の結果を示す。概して、この膜上に固定化される結合因子の量は、生物学的サンプルが、上記の形式において、2抗体サンドイッチアッセイにおいて陽性シグナルを生じるのに十分であるレベルのポリペプチドを含む場合、視覚的に識別可能なパターンを生じるように選択される。このようなアッセイにおける使用に好ましい結合因子は、抗体およびその抗原結合フラグメントである。好ましくは、膜上に固定化される抗体の量は、約25 ng～約1 μgの範囲であり、そしてより好ましくは、約50 ng～約500 ngの範囲である。このような試験は、代表的には、非常に少ない量の生物学的サンプルを用いて実行され得る。

10

【0286】

もちろん、本発明の腫瘍タンパク質または結合因子との使用に適する多数の他のアッセイプロトコルが存在する。上記の記載は、例示であることを意図するにすぎない。例えば、上記プロトコルが、腫瘍ポリペプチドを使用するために容易に改変され得て、生物学的サンプル内でこのようなポリペプチドに結合する抗体を検出し得ることが当業者に明かである。このような腫瘍タンパク質特異的抗体の検出は、癌の存在と相関し得る。

20

【0287】

癌もまた、あるいは癌は、生物学的サンプル中の腫瘍タンパク質と特異的に反応するT細胞の存在に基づいて検出され得る。特定の方法では、患者から単離されたCD4⁺T細胞および／またはCD8⁺T細胞を含む生物学的サンプルは、腫瘍ポリペプチド、このようなポリペプチドをコードするポリヌクレオチドおよび／またはこのようなポリペプチドの少なくとも免疫原性部分を発現するAPCとともにインキュベートされ、そしてT細胞の特異的活性化の存在または非存在が検出される。適切な生物学的サンプルとしては、単離されたT細胞が挙げられるがこれらに限定されない。例えば、T細胞は、慣用技術によって（例えば、末梢血リンパ球のFicoll/Hypaque密度勾配遠心分離法によって）患者から単離され得る。T細胞は、2～9日間（代表的に4日間）、37℃にてポリペプチド（例えば、5～25 μg/ml）とともにインビトロでインキュベートされ得る。T細胞サンプルの別のアリコート、コントロールとして役立てるために、腫瘍ポリペプチドの非存在下でインキュベートすることが所望され得る。CD4⁺T細胞に関して、活性化は好ましくは、T細胞の増殖を評価することによって検出される。CD8⁺T細胞に関しては、活性化は好ましくは、細胞溶解活性を評価することによって検出される。疾患のない患者におけるよりも少なくとも2倍高い増殖レベルおよび／または少なくとも20%高い細胞溶解活性レベルは、患者における癌の存在を示す。

30

40

【0288】

上記のように、癌もまた、あるいは癌は、生物学的サンプル中の腫瘍タンパク質をコードするmRNAのレベルに基づいて検出され得る。例えば、少なくとも2つのオリゴヌクレオチドプライマーをポリメラーゼ連鎖反応（PCR）に基づくアッセイにおいて用いて、生物学的サンプルに由来する腫瘍cDNAの一部を増幅し得、ここで、オリゴヌクレオチドプライマーのうちの少なくとも1つは、腫瘍タンパク質をコードするポリヌクレオチドに特異的である（すなわち、ハイブリダイズする）。次いで、増幅されたcDNAが、当該分野で周知の技術（例えば、ゲル電気泳動）を用いて分離され、そして検出される。

【0289】

50

同様に、腫瘍タンパク質をコードするポリヌクレオチドに特異的にハイブリダイズするオリゴヌクレオチドプローブをハイブリダイゼーションアッセイにおいて用いて、生物学的サンプル中でこの腫瘍タンパク質をコードするポリヌクレオチドの存在を検出し得る。

【0290】

アッセイ条件下でのハイブリダイゼーションを可能にするために、オリゴヌクレオチドのプライマーおよびプローブは、少なくとも10ヌクレオチド、そして好ましくは少なくとも20ヌクレオチドの長さの、本発明の腫瘍タンパク質をコードするポリヌクレオチドの一部に対して、少なくとも約60%、好ましくは少なくとも約75%、そしてより好ましくは少なくとも約90%の同一性を有するオリゴヌクレオチド配列を含むべきである。好ましくは、オリゴヌクレオチドプライマーおよび／またはプローブは、上記に規定されるような、中程度にストリンジェントな条件下で、本明細書中に記載されるポリペプチドをコードするポリヌクレオチドにハイブリダイズする。本明細書中に記載される診断方法において有用に用いられ得るオリゴヌクレオチドのプライマーおよび／またはプローブは、好ましくは、少なくとも10～40ヌクレオチドの長さである。好ましい実施形態において、オリゴヌクレオチドプライマーは、本明細書中に開示される配列を有するDNA分子の少なくとも10の連続するヌクレオチド、より好ましくは少なくとも15の連続するヌクレオチドを含む。PCRに基づくアッセイおよびハイブリダイゼーションアッセイの両方についての技術は、当該分野で周知である（例えば、Mullisら, Cold Spring Harbor Symp. Quant. Biol., 51:263, 1987; Erlich編, PCR Technology, Stockton Press, N Y, 1989を参照のこと）。 10 20

【0291】

1つの好ましいアッセイは、RT-PCRを用い、RT-PCRでは、PCRは、逆転写と組み合わせて適用される。代表的に、RNAは生物学的サンプル（例えば、生検組織）から抽出され、そして逆転写されてcDNA分子を生成する。少なくとも1つの特異的プライマーを用いるPCR増幅は、cDNA分子を生成し、このcDNA分子は、例えば、ゲル電気泳動を用いて分離および可視化され得る。増幅は、試験患者および癌に罹患していない個体から採取された生物学的サンプルについて行われ得る。増幅反応は、2桁の大きさに及ぶいくつかのcDNA希釈物について行われ得る。試験患者サンプルのいくつかの希釈物における発現の増加が、癌のないサンプルの同一希釈の物と比較して2倍以上である場合、これは、代表的に、陽性とみなされる。 30

【0292】

本発明の別の局面において、細胞捕獲技術は、例えば、リアルタイムPCRと組み合わせて使用されて、肺腫瘍抗原を発現する転移細胞の検出のためのより感度の高いツールを提供し得る。生物学的サンプル（例えば、骨髄サンプル、末梢血、および小針吸引サンプル）における肺癌細胞の検出は、肺癌患者の診断および予後判定に望ましい。

【0293】

細胞表面マーカーに対する特異的モノクローナル抗体または4量体抗体複合体でコーティングされた免疫磁性ビーズを用いて、サンプル中の癌細胞を最初に富化し得るか、またはサンプル中の癌細胞をポジティブに選択し得る。種々の市販のキットが用いられ得る。これらのキットとしては、Dynabeads（登録商標）Epithelial Enrich（Dynal Biotech, Oslo, Norway）、StemSepTM（StemCell Technologies, Inc., Vancouver, BC）、およびRosetteSep（StemCell Technologies）が挙げられる。当業者は、他の方法論およびキットが用いられて、所望の細胞集団が富化またはポジティブに選択され得ることを認識する。Dynabeads（登録商標）Epithelial Enrichは、正常上皮組織および新生物上皮組織上で発現された2つの糖タンパク質膜抗原に特異的なmAbでコーティングされた磁性ビーズを含む。このコーティングされたビーズをサンプルに添加し得、次いで、このサンプルを磁石に接触させ得、それによりこのビーズに結合した細胞を捕獲し得る。所望でない細胞は洗い流され、 40 50

磁石により単離された細胞は、ビーズから溶出され、さらなる分析において用いられる。

【0294】

R o s e t t e S e p は、血液サンプルから直接細胞を富化するために用いられ得、そして種々の所望でない細胞を標的とし、サンプル中に存在する赤血球（R B C）上のグリコホリンAにそれらを架橋して、ロゼットを形成する4量体の抗体のカクテルからなる。F i c o l l 上で遠心分離すると、標的とした細胞は、遊離のR B Cとともにペレットを形成する。枯渇カクテルにおける抗体の組み合わせは、どの細胞が除去され、そして結果的にどの細胞が回収されるかを決定する。利用可能な抗体としては、以下が挙げられるが、これらに限定されない：C D 2、C D 3、C D 4、C D 5、C D 8、C D 10、C D 11b、C D 14、C D 15、C D 16、C D 19、C D 20、C D 24、C D 25、C D 29、C D 33、C D 34、C D 36、C D 38、C D 41、C D 45、C D 45 R A、C D 45 R O、C D 56、C D 66 B、C D 66 e、H L A - D R、I g E、およびT C R $\alpha \beta$ 。 10

【0295】

さらに、本発明において、肺腫瘍抗原に特異的なm A bが生成され得、そして同様な様式で用いられ得ることが意図される。例えば、腫瘍特異的細胞表面抗原に結合するm A bは、磁性ビーズに結合体化され得るか、または4量体抗体複合体に処方され得、そしてサンプルから転移性の肺腫瘍細胞を富化するか、またはポジティブに選択するために用いられ得る。一旦、サンプルが富化されるか、またはポジティブに選択されると、細胞は溶解され得、そしてR N Aが単離され得る。次いで、R N Aは、本明細書中に記載のようにリアルタイムP C Rアッセイにおいて肺腫瘍特異的プライマーを用いてR T - P C R分析に供され得る。当業者は、富化されたかまたは選択された集団の細胞が他の方法（例えば、インサイチュハイブリダイゼーションまたはフローサイトメトリー）によって分析され得ることを認識する。 20

【0296】

別の実施形態において、本明細書中に記載された組成物は、癌の進行についてのマーカーとして使用され得る。この実施形態において、癌の診断について上記に記載されるようなアッセイは、経時的に実行され得、そして反応性ポリペプチドまたはポリヌクレオチドのレベルの変化を評価し得る。例えば、このアッセイは、6ヶ月～1年の期間、24～72時間毎に実行され得、そしてその後、必要に応じて実行され得る。一般に、癌は、検出されるこのポリペプチドまたはポリヌクレオチドのレベルが経時的に増大する患者において進行している。対照的に、癌は、反応性ポリペプチドまたはポリヌクレオチドのレベルが一定のままであるか、または時間とともに減少するかのいずれかである場合、進行していない。 30

【0297】

特定のインビボ診断アッセイは、腫瘍上で直接実施され得る。1つのこのようなアッセイは、腫瘍細胞を結合因子と接触させる工程を包含する。次いで、結合された結合因子は、レポーター基によって直接的または間接的に検出され得る。このような結合因子はまた、組織学的な用途において使用され得る。あるいは、ポリヌクレオチドプローブは、このような用途において使用され得る。 40

【0298】

上記のように、感度を改善するために、複数の腫瘍タンパク質マーカーは、所定のサンプル内でアッセイされ得る。本明細書中に提供される異なるタンパク質に特異的な結合因子が単一のアッセイにおいて組み合わせられ得ることは明かである。さらに、複数のプライマーまたはプローブが同時に用いられ得る。腫瘍タンパク質マーカーの選択は、最適な感度をもたらす組み合わせを決定する慣用実験に基づき得る。さらに、または代替として、本明細書中に提供される腫瘍タンパク質のアッセイは、他の公知の腫瘍抗原に対するアッセイと組み合わせられ得る。

【0299】

本発明はさらに、上記の診断方法のうちのいずれかにおいて使用するためのキットを提供 50

する。このようなキットは代表的に、診断アッセイを行うのに必要な2以上の構成要素を備える。構成要素は、化合物、試薬、容器および／または器具であり得る。例えば、キット内の1つの容器は、腫瘍タンパク質に特異的に結合するモノクローナル抗体またはそのフラグメントを含み得る。このような抗体またはフラグメントは、上記のように支持体材料に付着されて提供され得る。1以上のさらなる容器は、アッセイにおいて使用される要素（例えば、試薬または緩衝液）を封入し得る。このようなキットもまた、あるいはこのようなキットは、抗体結合の直接的または間接的な検出に適切なレポーター基を含む上記のような検出試薬を備え得る。

【0300】

あるいは、キットは、生物学的サンプル中の腫瘍タンパク質をコードするmRNAレベルを検出するように設計され得る。このようなキットは一般に、腫瘍タンパク質をコードするポリヌクレオチドにハイブリダイズする、上記のような、少なくとも1つのオリゴヌクレオチドのプロブまたはプライマーを備える。このようなオリゴヌクレオチドは、例えば、PCRまたはハイブリダイゼーションアッセイにおいて用いられ得る。このようなキット内に存在し得るさらなる構成要素としては、腫瘍タンパク質をコードするポリヌクレオチドの検出を容易にする、第2のオリゴヌクレオチドおよび／または診断試薬もしくは容器が挙げられる。

【0301】

以下の実施例は、例示のために提供され、そして限定のためではない。

【0302】

(実施例1)

(肺腫瘍cDNAの同定および特徴付け)

本実施例は、肺腫瘍タンパク質をコードするcDNA分子の同定を説明する。

【0303】

(A. 従来のcDNAライブラリーサブトラクションを使用する、肺腺癌ライブラリーからのcDNA配列の単離)

ヒト肺腺癌cDNA発現ライブラリーを、Superscript Plasmid System for cDNA Synthesis and Plasmid Cloningキット(BRL Life Technologies, Gaithersburg, MD)を使用して製造業者のプロトコルに従って、患者組織(#40031486)由来のポリA⁺RNAから構築した。詳細には、肺癌腫組織を、ポリトロン(polyttron; Kinematica, Switzerland)で均質化し、そして全RNAを、Trizol試薬(BRL Life Technologies)を使用して製造業者の指示に従って抽出した。次いで、ポリA⁺RNAをSambrookら、Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989に記載されているように、オリゴdTセルロースカラムを使用して精製した。1本鎖のcDNAを、NotI/Oligo-dT18プライマーを使用して合成した。2本鎖のcDNAを合成し、BstXI/EcoRIアダプター(Invitrogen, San Diego, CA)と連結し、そしてNotIで消化した。cDNAサイズ分画カラム(BRL Life Technologies)でのサイズ分画後、cDNAを、pcDNA3.1(Invitrogen)のBstXI/NotI部位に連結し、そしてエレクトロポレーションによってElectroMax E.coli DH10B細胞(BRL Life Technologies)に形質転換した。全部で3×10⁶個の独立コロニーを生成させた。

【0304】

同じ手順を使用して、正常なヒトcDNA発現ライブラリーを、肺、肝臓、脾臓、皮膚、腎臓、脳、および休止PBMCを含む正常な組織標本のパネルから調製した。

【0305】

cDNAライブラリーサブトラクションを、いくつかの改変を用いてHaraら(Bl o 50

od、84:189-199、1994)に記載されているように、肺腺癌および正常組織のcDNAライブラリーを使用して行った。詳細には、肺腺癌特異的サブトラクトcDNAライブラリーを、以下のように作成した。正常組織のcDNAライブラリー(80μg)を、BamHIおよびXhoIで消化し、続いてDNAポリメラーゼクレノウフラグメントでのフィルイン反応を行った。フェノールクロロホルム抽出、およびエタノール沈殿後、DNAを133μlのH₂O中に溶解させ、熱変性させ、そして133μl(133μg)のPhotoprobeビオチン(Vector Laboratories, Burlingame, CA)と混合した。製造業者によって推奨されているように、得られた混合物を270Wの太陽灯で20分間、氷上で照射した。さらにPhotoprobeビオチン(67μl)を添加し、そしてビオチニル化反応を繰り返した。5倍のブタノールでの抽出後、DNAをエタノール沈殿させ、そして23μlのH₂O中に溶解させた。得られたDNA、および以前の肺サブトラクションにおいて頻繁に回収された他の非常に豊富なcDNAクローンによって、ドライバードNAを形成した。

【0306】

トレーサードNAを形成するために、10μgの肺腺癌cDNAライブラリーを、NotIおよびSpeIで消化し、フェノールクロロホルムで抽出し、そしてChromaspin-400カラム(Clontech, Palo Alto, CA)を通過させた。典型的には、5μgのcDNAをカラムのサイジング後に回収した。エタノール沈殿後、トレーサードNAを5μlのH₂O中に溶解させた。トレーサードNAを、15μlのドライバードNAおよび20μlの2×ハイブリダイゼーション緩衝液(1.5MのNaCl/10mMのEDTA/50mMのHEPES pH7.5/0.2%のドデシル硫酸ナトリウム)と混合し、ミネラルオイルで重層し、そして完全に熱変性させた。サンプルを、すぐに68℃の水浴に移し、そして20時間インキュベートした(長いハイブリダイゼーション[LH])。次いで、反応混合物をストレプトアビジン処理、続いてフェノール/クロロホルム抽出に供した。このプロセスをさらに3回繰り返した。サブトラクトDNAを沈殿させ、12μlのH₂O中に溶解させ、8μlのドライバードNAおよび20μlの2×ハイブリダイゼーション緩衝液と混合し、そして68℃で2時間の間、68℃でハイブリダイゼーションに供した(短いハイブリダイゼーション[SH])。ビオチニル化した2本鎖DNAの除去後、サブトラクトcDNAを、クロラムフェニコール耐性pBCSK⁺(Stratagene, La Jolla, CA)のNotI/SpeI部位に連結し、そしてエレクトロポレーションによってElectroMax E.coli DH10B細胞中に形質転換して、LAT-S1と呼ぶ肺腺癌特異的サブトラクトcDNAライブラリーを作成した。同様に、LAT-S2を、さらなるドライバードとして、トレーサード中に過剰発現される23個の遺伝子を含むことによって、作成した。

【0307】

第2のヒト肺腺癌cDNA発現ライブラリーを、第2の患者(#86-66)由来の腺癌組織を使用して構築し、そして正常組織の同じパネルおよびLAT-S1中で過剰発現されるさらなる遺伝子を使用して、上記のように、第2の肺腺癌特異的サブトラクトcDNAライブラリー(LAT2-S2と呼ぶ)を調製するために使用した。

【0308】

第3のヒト転移性肺腺癌ライブラリーを、肺および胃の腺癌に起源する2つの肺胸膜滲出液のプールから構築した。サブトラクトcDNAライブラリー(Mets-sub2と呼ぶ)を、正常組織の同じパネルを使用して上記のように作成した。Mets-sub3サブトラクトライブラリーを、ドライバードとして51個のさらなる遺伝子を含むことによって構築した。これらの51個の遺伝子を、テスター中に過剰発現されるハウスキーピング(housekeeping)遺伝子を提示するMets-sub2中に回収した。

【0309】

上記の、LAT-S1から単離した全部で16個のcDNAフラグメント、LAT-S2から単離した585個のcDNAフラグメント、LAT2-S2に由来する568個のcDNAクローン、Mets-sub2に由来する15個のcDNAクローン、およびMe

10

20

30

40

50

t s - s u b 3 に由来する 3 4 3 個の c D N A クローンを、コロニー P C R で増幅し、そして肺腫瘍、正常な肺、ならびに種々の他の正常組織および腫瘍組織中でのそれらの m R N A の発現レベルを、マイクロアレイ技術 (I n c y t e , P a l o A l t o , C A) を使用して決定した。簡潔には、P C R 増幅産物を、それぞれの産物がアレイ中の特有の位置を占有するアレイ形式でスライド上に点在させた。m R N A を、試験する組織サンプルから抽出し、逆転写し、そして蛍光標識した c D N A プローブを生成した。マイクロアレイを標識した c D N A プローブでプローブし、スライドをスキャンし、そして蛍光強度を測定した。この強度は、ハイブリダイゼーション強度と相関する。7 3 個の豊富ではない c D N A クローン (そのうちの 4 2 が固有であることを見出した) は、肺腫瘍中で過剰発現を示し、試験した正常組織 (肺、皮膚、リンパ節、結腸、肝臓、膵臓、乳房、心臓、骨髄、大腸、腎臓、胃、脳、小腸、膀胱、および唾液腺) 中での発現は検出不可能であるか、または肺腺癌腫瘍と比較して有意に低いレベルのいずれかであった。これらのクローンをさらに、P e r k i n E l m e r / A p p l i e d B i o s y s t e m s D i v i s i o n A u t o m a t e d S e q u e n c e r M o d e l 3 7 3 A および / または M o d e l 3 7 7 (F o s t e r C i t y , C A) を用いる D N A 配列決定によって特徴付けた。

10

【0310】

配列を、E M B L G e n B a n k データベース (r e l e a s e 9 6) を使用して遺伝子バンク中の既知の配列と比較した。配列番号 6 7 に提供する配列に対しては有意な相 20
同性は見られず、以前に同定された発現された配列タグ (E S T) に対する明らかな相 20
同性はなかった。配列番号 6 0、6 2、6 5、6 6、6 9 ~ 7 1、7 4、7 6、7 9、8 0、8 4、8 6、8 9 ~ 9 2、9 5、9 7、および 9 8 の配列が、以前に同定された発現された配列タグ (E S T) に対していくらかの相同性を示すことを見出した。配列番号 5 9、6 1、6 3、6 4、6 7、6 8、7 2、7 3、7 5、7 7、7 8、8 1 ~ 8 3、8 5、8 7、8 8、9 3、9 4、9 6、9 9、および 1 0 0 の c D N A 配列は、以前に同定された遺伝子に対して相同性を示した。配列番号 9 6 および 1 0 0 のクローンの全長の c D N A 配列を、それぞれ配列番号 3 1 6 および 3 1 8 に提供する。配列番号 5 9、6 1、6 3、6 4、6 8、7 3、8 2、8 3、9 4、9 6、および 1 0 0 のクローンのアミノ酸配列を、それぞれ、配列番号 3 3 1、3 2 8、3 2 9、3 3 2、3 2 7、3 3 3、3 3 0、3 2 6、3 2 5、3 2 4、および 3 3 5 に提供する。配列番号 6 9 の配列 (L 5 5 2 S と呼 30
ぶ) によってコードされるアミノ酸配列を、配列番号 7 8 6 に提供する。

30

【0311】

さらなる研究によって、L 5 5 2 S (配列番号 7 9 0) の延長した c D N A 配列およびオープンリーディングフレームの単離を導いた。配列番号 7 9 0 の c D N A 配列によってコードされるアミノ酸配列を、配列番号 7 9 1 に提供する。続く研究によって、L 5 5 2 S の全長の c D N A 配列の単離を導いた (配列番号 8 0 8) 。L 5 5 2 S の全長の c D N A 配列は、4 8 0 塩基対のオープンリーディングフレーム (配列番号 7 9 0) を有し、そして 1 6 0 個のアミノ酸の推定のポリペプチド (配列番号 8 0 9) をコードする。

【0312】

最初のデータベース検索は、データベース中のタンパク質と相同である任意の配列を検出 40
することができず、L 5 5 2 S が未知の機能の新規のタンパク質をコードすることを示唆している。最近、精巣癌抗原 X A G E - 1 が、ユーイング肉腫 (E w i n g ' s s a r c o m a) 中で過剰発現されることが見出された (L i u ら、2 0 0 0 C a n c e r R e s . 6 0 : 4 7 5 2 - 4 7 5 5) 。X A G E - 1 の決定された c D N A 配列を配列番号 7 9 2 に提供し、対応するアミノ酸配列を配列番号 7 9 3 に提供する。L 5 5 2 S と X A G E - 1 との配列比較は、顕著な同一性ならびに差異を明らかにする。C 末端配列の大部分は互いに同一である。L 5 5 2 S と X A G E - 1 から推定されるポリペプチドは、異なる N 末端配列を有する。L 5 5 2 S アミノ酸の親水性分析は、推定される膜貫通ドメインを有さない極めて親水性であるタンパク質を示唆した。L 5 5 2 S の P S O R T 分析は 40
同じ結果を明らかにした。L 5 5 2 S もまた X 番染色体中に配置されるので、これはおそ 50

50

らく、精巣癌抗原 X A G E - 1 の新しいイソ型である。遺伝子配列分析は、両方の遺伝子が X 番染色体の同じ領域に配置され、そして両方ともが 4 個のエキソンを有することを明らかにした。後ろの 3 個のエキソンは、L 5 5 2 S および X A G E - 1 について同一である。しかし、X A G E - 1 についての第 1 のエキソンは、L 5 5 2 S についての第 1 のエキシソンの上流に存在し、そしてこれによって、L 5 5 2 S および X A G E - 1 の異なる 5 ' ヌクレオチドおよびアミノ酸配列を生じる。従って、L 5 5 2 S および X A G E - 1 は代替的にスプライシングされるアイソタイプである。

【0313】

L 5 5 2 S のこの N 末端部分のアミノ酸配列を配列番号 1 8 3 0 に提供し、対応する c D N A 配列を配列番号 1 8 2 6 に提供する。配列番号 1 8 2 7 ~ 1 8 2 9 に提供する c D N A 配列は、配列番号 8 0 8 の b p 3 9 4 ~ 6 8 1、配列番号 8 0 8 の b p 3 9 4 ~ 5 3 4、および配列番号 7 9 2 の b p 2 1 4 ~ 3 9 4 をそれぞれ示す。対応するアミノ酸配列を、それぞれ、配列番号 1 8 3 1 ~ 1 8 3 3 に提供する。 10

【0314】

L 5 5 2 S について全長をクローニングする試みによって、転移性肺癌腫ライブラリーからの 3 個のさらなる c D N A 配列（それぞれ、配列番号 8 1 0 ~ 8 1 2 ; クローン 5 0 9 8 9、5 0 9 9 0、および 5 0 9 9 2 と呼ぶ）の単離を導いた。配列番号 8 1 0 の配列が、以前に同定されたヒト D N A 配列のいくつかの相同性を示すことを見出した。配列番号 8 1 1 の配列が、以前に同定された D N A 配列といくつかの相同性を示すことを見出した。配列番号 8 1 2 の配列が、以前に同定された E S T といくつかの相同性を示すことを見出した。 20

【0315】

配列番号 8 4 (L 5 5 1 S と呼ぶ) の遺伝子が、9 個の原発性腺癌のうちの 2 個において過剰発現され、そして 2 個の転移性腺癌のうちの 2 個および 2 個の扁平上皮細胞癌のうちの 1 個において低いレベルで発現されることを、リアルタイム R T - P C T 分析によって決定した。正常な胃中での非常に低い発現を除いて、正常な組織中では発現は観察されなかった。L 5 5 1 S についてのさらなる研究によって、配列番号 8 0 1 および 8 0 2 にそれぞれ提供する、5 ' および 3 ' c D N A コンセンサス配列の単離を導いた。L 5 5 1 S の 5 ' 配列が、以前に同定された遺伝子 S T Y 8 (配列番号 8 0 3 に提供する c D N A 配列; 対応するアミノ酸配列を配列番号 8 0 5 に提供する) (これは、有糸分裂促進因子活性化プロテインキナーゼホスファターゼである) に対していくつかの相同性を示すことを見出した。しかし、L 5 5 1 S の 3 ' 配列に対しては、有意な相同性は見出せなかった。続いて、L 5 5 1 S についての延長した c D N A 配列を単離した (配列番号 8 0 4)。対応するアミノ酸配列を、配列番号 8 0 6 に提供する。さらなる研究によって、L 5 5 1 S の 2 つの独立した全長クローンの単離を導いた (5 4 2 9 8 および 5 4 3 0 5 と呼ぶ)。これらの 2 つのクローンは、S T Y 8 D N A 配列と比較して 5 個のヌクレオチドの差異を有する。これらの差異のうちの 2 つは、単一のヌクレオチド多形性である。これはコードされるアミノ酸配列には影響を与えない。他の 3 個のヌクレオチドの差異は、2 つの L 5 5 1 S クローンの間で一致するが、S T Y 8 タンパク質配列とは異なるコードされるアミノ酸配列を導く。L 5 5 1 S の全長クローン 5 4 3 0 5 および 5 4 2 9 8 について決定した c D N A 配列を、それぞれ配列番号 8 2 5 および 8 2 6 に提供し、L 5 5 1 S のアミノ酸配列を配列番号 8 2 7 に提供する。 30 40

【0316】

(B. P C R に基づく c D N A ライブラリーサブトラクションを使用する、肺腺癌ライブラリーからの c D N A 配列の単離)

皮膚、結腸、肺、食道、脳、腎臓、脾臓、膵臓、および肝臓を含む 9 個の正常なヒト組織の c D N A のプールに対してサブトラクトした 2 つのヒト肺原発性腺癌のプール由来の c D N A を含有している、サブトラクトライブラリー由来の c D N A クローン (C l o n t e c h, P a l o A l t o, C A) を、初回の P C R 増幅に供した。このライブラリー (A L T - 1 と呼ぶ) を、製造業者のプロトコールに従って 2 回目の P C R 増幅に供した 50

。肺腫瘍、正常な肺、ならびに種々の他の正常組織および腫瘍組織中の760個のcDNAクローンの発現レベルを、上記のようなマイクロアレイ技術を使用して試験した。全部で118個のクローン（そのうちの55個が固有であった）が、肺腫瘍組織中で過剰発現され、試験した正常組織（肺、皮膚、リンパ節、結腸、肝臓、脾臓、乳房、心臓、骨髄、大腸、腎臓、胃、脳、小腸、膀胱、および唾液腺）中での発現は検出不可能であるかまたは有意に低いレベルで発現されるかのいずれかであることを見出した。配列を、EMBLおよびGenBankデータベース（release 96）を使用して遺伝子バンク中の既知の配列と比較した。配列番号44に提供する配列に対する有意な相同性（ESTを含む）は見出せなかった。配列番号1、11、13、15、20、23～27、29、30、33、34、39、41、43、45、46、51、および57の配列が、以前に同定された発現された配列タグ（EST）に対していくらかの相同性を示すことを見出した。配列番号2～10、12、14、16～19、21、22、28、31、32、35～38、40、42、44、47～50、52～56、および58のcDNA配列は、以前に同定された遺伝子に対する相同性を示した。配列番号18、22、31、35、36、および42のクローンの全長のcDNA配列を、それぞれ、配列番号320、319、323、321、317、321、および322に提供し、対応するアミノ酸配列を、それぞれ、配列番号337、336、340、338、334、および339に提供する。

【0317】

さらなる研究によって、配列番号33のクローンについての延長したcDNA配列（L801Pと呼ぶ）の単離を導いた。この延長したcDNA配列（配列番号796に提供する）が、3個のオープンリーディングフレーム（ORF）を含む可能性のあることを見出した。これらの3個のORFによってコードされる推定のアミノ酸配列を、それぞれ配列番号797～799に提供する。全長をクローニングするさらなる努力によって、配列番号1669に示す、L801Pのさらに延長されたcDNA配列、延長されたcDNA配列によってコードされるさらなる5個の可能性のあるオープンリーディングフレーム（ORF 4～9と呼ぶ；それぞれ、配列番号1670～1675）を導いた。L801Pを、染色体領域20p13にマップし、そしてORF4に対応するこの遺伝子領域による137アミノ酸のORFを同定した（配列番号1670）。このことは、これがおそらくL801Pの確実なORFであることを示唆している。

【0318】

マイクロアレイ分析によって、L801Pが、正常組織と比較して、肺腫瘍組織中で2倍またはそれ以上過剰発現されることを見出した。リアルタイムPCR分析によると、肺腺癌の50%より多く、および試験した肺扁平上皮細胞性癌腫瘍サンプルの30%より多くが、正常な肺組織と比較して上昇したL801Pの発現を有した。上昇したL801Pを示したこれらのうちでは、発現レベルは、正常な肺組織サンプル中よりも10倍以上大きかった。さらに、L801Pの低い発現または発現がないことを、正常組織のRNAの大量のパネルにおいて検出した。

【0319】

L801Pの発現をまた、乳房、前立腺、卵巣、および結腸の腫瘍を含む多数の他の腫瘍型において検出し、従って、L801Pの発現は、これらの型の癌における診断および／または治療上の有用性を有し得る。

【0320】

続く研究において、配列番号44のクローン（L844Pと呼ぶ）の全長のcDNA配列を単離した（配列番号800に提供する）。この配列の公開されているデータベース中の配列との比較によって、配列の5'末端の470塩基が、既知の遺伝子ジヒドロジオールデヒドロゲナーゼに対して相同性を示すことを明らかにした。従って、このことは、L884Pが以前に同定された3'非翻訳領域の2007塩基対を有しているジヒドロジオールデヒドロゲナーゼファミリーの新規の転写物であることを示している。

【0321】

配列番号46の配列（L840Pと呼ぶ）によってコードされる推定されるアミノ酸配列

を、配列番号 787 に提供する。L840P についての延長した cDNA 配列（これは、オープンリーディングフレームを含むと決定した）を配列番号 794 に提供する。配列番号 794 の cDNA 配列によってコードされる推定されたアミノ酸配列を、配列番号 795 に提供する。配列番号 54 のクローン（L548S と呼ぶ）についての全長の cDNA 配列を、配列番号 788 に提供し、対応するアミノ酸配列を配列番号 789 に提供する。

【0322】

配列番号 25 および 46 の遺伝子（それぞれ、L839P および L840P と呼ぶ）のノーザンブロット分析は顕著に類似していた。両方の遺伝子は、2 個の肺腺癌のうちの 1 個において、3.6 kb および 1.6 kb の 2 つのバンドとして発現された。L839P の発現は、正常な肺または気管では観察されなかった。L840P の発現は、正常な骨髄、休止 PBMC または活性化 PBMC、食道、あるいは正常な肺では観察されなかった。同様の発現パターンを仮定すると、L839P および L840P は同じ遺伝子に由来し得る。

10

【0323】

さらなる肺腺癌 cDNA クローンを、以下のように単離した。cDNA ライブラリーを、2 つの肺腺癌のプールから調製し、そして肺、脳、肝臓、腎臓、脾臓、皮膚、心臓、および脾臓を含む正常組織のパネル由来の cDNA に対してサブトラクトした。サブトラクションを、より大きなフラグメントを作成するように改変した PCR に基づくプロトコール（Clontech）を使用して行った。このプロトコールにおいては、テスターおよびドライバーの 2 本鎖 cDNA を、6 ヌクレオチドの制限部位（MluI、MscI、PvuII、SalI、および StuI）を認識する 5 個の制限酵素で別々に消化した。この消化によって、Clontech プロトコールに従う RsaI での消化によって生じる平均 300 bp の大きさの cDNA ではなく、平均 600 bp の cDNA の大きさを生じた。制限消化したテスター cDNA の末端をフィルインしてアダプターの連結のための平滑末端を作成した。この改変は、サブトラクションの効率には影響を与えなかった。次いで、2 つのテスターの集団を異なるアダプターを用いて作成し、そしてドライバーのライブラリーはアダプターを有さないままにした。次いで、テスターおよびドライバーのライブラリーを過剰なドライバー cDNA を使用してハイブリダイズさせた。最初のハイブリダイゼーション工程においては、ドライバーを、2 つのテスター cDNA の集団のそれぞれと別々にハイブリダイズさせた。これによって、（a）ハイブリダイズしていないテスター cDNA、（b）他のテスター cDNA にハイブリダイズしたテスター cDNA、（c）ドライバー cDNA にハイブリダイズしたテスター cDNA、および（d）ハイブリダイズしていないドライバー cDNA の集団を生じた。次いで、2 つの別々のハイブリダイゼーション反応物を混合し、そしてさらなる変性させたドライバー cDNA の存在下で再度ハイブリダイズさせた。この 2 回目のハイブリダイゼーション後、集団（a）から（d）に加えて、1 つのアダプターを有するテスター cDNA が第 2 のアダプターを有するテスター cDNA にハイブリダイズした第 5 の集団（e）を作成した。従って、第 2 のハイブリダイゼーション工程によって、アダプター特異的プライマーを用いた PCR 増幅のための鋳型として使用され得る、異なって発現される配列の富化を生じた。

20

30

【0324】

次いで、末端をフィルインし、そして PCR 増幅を、アダプター特異的プライマーを使用して行った。ドライバー cDNA にハイブリダイズしていないテスター cDNA を含む集団（e）だけが、指数関数的に増幅した。次いで、第 2 回目の PCR 増幅工程を行って、バックグラウンドを減少させ、そして異なって発現された配列をさらに富化させた。

40

【0325】

57 個の cDNA クローンをサブトラクトライブラリー（LAP1 と呼ぶ）から単離し、そして配列決定した。これらのクローンのうちの 16 個について決定した cDNA 配列を、配列番号 101～116 に提供する。配列番号 101 および 114 の配列は、以前に同定された配列に対して有意な相同性は示さなかった。配列番号 102～109 および 112 の配列は、以前に同定された配列に対していくらかの類似性を示し、一方、配列番号 1

50

13、115、および116の配列は、以前に同定されたESTに対していくらかの類似性を示した。

【0326】

LAP1ライブラリーから分析したさらなる502個のクローンを配列決定し、そして決定したcDNA配列を、配列番号828～1239および1564～1653に示す。

【0327】

(C. PCRに基づくcDNAライブラリーサブトラクションを使用する、小細胞性肺癌腫ライブラリーからのcDNA配列の単離)

小細胞性肺癌腫のサブトラクトcDNAライブラリー(SCL1と呼ぶ)を、本質的には、上記の改変したPCRに基づくサブトラクションプロセスを使用して調製した。小細胞性肺癌腫由来のcDNAを、正常な肺、脳、腎臓、肝臓、脾臓、皮膚、心臓、リンパ節、および脾臓を含む正常組織のパネル由来のcDNAに対してサブトラクトした。テスターおよびドライバーの両方のポリA+RNAを、最初に、SMART PCR cDNA合成キット(Clontech, Palo Alto, CA)を使用して増幅した。テスターおよびドライバーの2本鎖cDNAを5個の制限酵素(DraI、MscI、PvuII、SmaI、およびStuI)で別々に消化した。これらの制限酵素は平滑な末端切断を作成し、そして消化によって600bpの平均の挿入物の大きさを生じた。このセットの制限酵素での消化によって、cDNA末端をフィルインすることによる平滑末端を作成するために必要とされる工程を排除した。これらの改変は、サブトラクションの効率には影響を与えなかった。

【0328】

85個のクローンを単離し、そして配列決定した。これらのクローンのうちの31個についての決定したcDNA配列を、配列番号117～147に提供する。配列番号122、124、126、127、130、131、133、136、139、および147の配列は、以前に同定された配列に対して有意な相同性は示さなかった。配列番号120、129、135、137、140、142、144、および145の配列は、以前に同定された遺伝子配列に対していくらかの類似性を示し、一方、配列番号114、118、119、121、123、125、128、132、134、138、141、143、および147の配列は、以前に単離されたESTに対していくらかの類似性を示した。

【0329】

さらなる研究において、3個のさらなるcDNAライブラリーを、9個の正常組織(肺、脳、腎臓、肝臓、脾臓、皮膚、心臓、唾液腺、および脾臓)由来のポリA+RNAのプールに対してサブトラクトした単一の小細胞性肺腺癌サンプル由来のポリA+RNAより作成した。第1のライブラリー(SCL2と呼ぶ)については、サブトラクションを、テスターおよびドライバーをPvuII、StuI、MscIおよびDraIで消化したことを除いて、本質的にはLAP1ライブラリーについての上記のように行った。使用したテスターcDNAとドライバーcDNAとの比は、Clontechによって推奨されている通りであった。第2のライブラリー(SCL3と呼ぶ)については、サブトラクションを、SCL2ライブラリーから同定した非常に豊富なクローンのcDNAをドライバーcDNA中に含んだことを除いて、本質的にはSCL2と同様に行った。SCL4ライブラリーの構築を、テスターに対するドライバーのより高い比を使用したことを除いて、本質的にはSCL3ライブラリーについて記載したように行った。

【0330】

それぞれのライブラリーを、DNA配列決定およびデータベース分析によって特徴付けた。SCL2ライブラリーから単離した35個のクローンについて決定したcDNA配列を、配列番号245～279に提供し、SCL3ライブラリーから単離した21個のクローンについて決定したcDNA配列、およびSCL4ライブラリーから単離した15個のクローンについて決定したcDNA配列を、それぞれ、配列番号280～300および301～315に提供する。配列番号246、254、261、262、304、309、および311の配列は、以前に同定された配列に対して有意な相同性を示さなかった。配列

番号245、248、255、266、270、275、280、282、283、288～290、292、295、301、および303の配列は、以前に単離されたESTに対しいくらかの相同性を示し、一方、配列番号247、249～253、256～260、263～265、267～269、271～274、276～279、281、284～287、291、293、294、296～300、302、305～308、310、および312～315の配列は、以前に同定された遺伝子配列に対していくらかの相同性を示した。

【0331】

本明細書中に開示する配列を、マイクロアレイ分析によって特異的腫瘍組織中での過剰発現についてさらに(f i r t j e r)評価した。このアプローチを使用して、cDNA配列をPCR増幅し、そして腫瘍および正常組織中でのそれらのmRNA発現プロファイルを、本質的には記載されているようにcDNAマイクロアレイ技術を使用して試験した(Shena, M. ら、1995 Science 270:476～70)。簡潔には、クローンを複数のレプリカとして、ガラススライド上に並べた。それぞれの位置は固有のcDNAクローンに対応する(5500個程度の多くのクローンを1つのスライドまたはチップ上に並べることができる)。それぞれのチップを、Cy3およびCy5で蛍光標識したcDNAプローブの対と、それぞれハイブリダイズさせる。典型的には、1μgのポリA⁺RNAを、それぞれのcDNAプローブを作成するために使用した。ハイブリダイゼーション後、チップをスキャンし、そしてCy3およびCy5チャンネルの両方についての蛍光強度を記録した。複数の内部品質管理(b u i l t - i n q u a l i t y c o n t r o l)工程が存在した。第1に、プローブの品質を、偏在して発現される遺伝子のパネルを使用してモニターした。第2に、対照プレートもまた、酵母DNAフラグメントを含み得る。酵母DNAフラグメントの相補性DNAは、プローブの品質および分析の感度を測定するためのプローブ合成をスパイクし得る。一般に、この技術は、mRNAの100,000コピー分の1の感度を提供する。最後に、この技術の再現性は、種々の位置に2連の対照cDNA成分を含むことによって確実にされ得る。

【0332】

3個のPCRに基づくサブトラクトcDNAライブラリー由来の3264個のcDNAクローンを、上記のcDNAマイクロアレイ技術によって分析した。cDNAクローンを、Lung Chip 5上に並べた。これらのcDNAクローンのうち、960個のクローンはS QL 1ライブラリー由来であり、768個のクローンはS CL 1ライブラリー由来であり、そして1536個のクローンはS CL 3およびS CL 4ライブラリー由来であった。蛍光標識したcDNAプローブの35個の対を、マイクロアレイ分析に使用した。それぞれのプローブの対は、正常組織プローブと対にした肺腫瘍プローブを含んだ。発現データを分析した。498個のcDNAクローンが、正常組織のプローブ群と比較して、小細胞性および／または非小細胞性肺腫瘍プローブ群において2倍またはそれよりも大きく過剰発現されることを見出した。また、正常組織中でのこれらのクローンの平均の発現値は、0.1未満であった(発現の範囲は、0.001から10まで)。配列番号1240～1563に開示したcDNA配列は、324個の豊富ではないクローンを示す。

【0333】

以下の配列は、GenBankおよびGeneSeqを含むデータベース分析に基づく新規であった：配列番号1240、1243、1247、1269、1272、1280、1283、1285、1286、1289、1300、1309、1318、1319、1327、1335、1339、1346、1359、1369、1370、1371、1393、1398、1405、1408、1413、1414、1417、1422、1429、1432、1435、1436、1438～1442、1447、1450、1453、1463、1467、1470、1473、1475、1482、1486、1491～1494、1501、1505、1506、1514～1517、1520、1522、1524、1535、1538、1542、1543、1547、1554、1557、1559、1561、および1563。

【0334】

コンティグ (c o n t i g) 139 (配列番号1467)の一部の配列の延長したcDNA配列 (L985Pとしてもまた公知)を、問い合わせ配列として配列番号1467を使用して公開されているデータベースを検索することによって推定した。このアプローチによって、配列番号1467が、細胞表面免疫調節因子-2 (C S I M M - 2)をコードするcDNA配列 (配列番号1676)に対して相同性を有することを見出した。配列番号1676のcDNA配列は、配列番号1677に示す配列を有しているタンパク質をコードする。

【0335】

マイクロアレイ分析によると、L985Pは、正常組織のプローブ群と比較して、肺腫瘍プローブ群において2倍またはそれより大きく過剰発現された。さらに、正常組織中でのL985Pの平均の発現値は、0.2未満であった (発現の範囲は、0.01から10であった)。リアルタイムPCR分析によると、試験した小細胞性肺癌腫肺腫瘍サンプルの40%以上が、正常な肺組織と比較して増大したL985Pの発現を有した。増大したL985Pを示したこれらについて、発現レベルは、正常な肺組織サンプル中よりも3倍以上高かった。L985Pの低い発現または発現がないことを、正常組織のRNAの大量のパネル中で検出した。L985Pについてのこれらの知見は、患者における肺癌の存在を検出するための診断マーカーとして、および/または肺癌の処置のための免疫治療用の標的としての両方のその用途を支持する。

10

【0336】

(D. PCRに基づくcDNAライブラリーサブトラクションを使用する、神経内分泌ライブラリーからのcDNA配列の単離)

改変したPCRに基づくサブトラクションプロセスを使用して、本質的には、LAP1サブトラクトライブラリーについて上記のように、サブトラクトcDNAライブラリー (MLN1と呼ぶ)を、正常な肺、脳、腎臓、肝臓、脾臓、皮膚、心臓、リンパ節、および脾臓を含む9個の正常組織のパネルを用いるサブトラクションによって、下方気管竜骨リンパ節 (s u b c a r i n a l l y m p h n o d e)に転移した肺神経内分泌癌腫から誘導した。

20

【0337】

91個の独立クローンを単離し、そして配列決定した。これらのクローンの58個について決定したcDNA配列を、配列番号147~222に提供する。配列番号150、151、154、157、158、159、160、163、174、175、178、186~190、192、193、195~200、208~210、212~215、および220の配列は、以前に同定された配列に対して有意な相同性は示さなかった。配列番号152、155、156、161、165、166、176、179、182、184、185、191、194、221、および222の配列は、以前に同定された遺伝子配列に対していくつかの類似性を示し、一方、配列番号148、149、153、164、167~173、177、180、181、183、201~207、211、および216~219の配列は、以前に単離されたESTに対していくつかの類似性を示した。

30

【0338】

MLN1ライブラリーから単離したさらなる442個のクローンの決定したcDNA配列を、配列番号341~782に提供し、MLN1ライブラリーから単離したさらなる11個のクローンの決定したcDNA配列を、配列番号1654~1664に提供する。

40

【0339】

(E. PCRに基づくcDNAライブラリーサブトラクションを使用する、扁平上皮細胞性肺癌腫ライブラリーからのcDNA配列の単離)

扁平上皮細胞性肺癌腫についてのサブトラクトcDNAライブラリー (SQL1と呼ぶ)を、テスターおよびドライバの2本鎖cDNAを4個の制限酵素 (D r a I、M s c I、P v u I I、およびS t u I)で別々に消化したことを除いて、本質的には上記の改変したPCRに基づくサブトラクションプロセスを使用して調製した。2つの扁平上皮細胞

50

性肺癌腫のプールに由来する c D N A を、正常な肺、脳、腎臓、肝臓、膵臓、皮膚、心臓、脾臓、食道、および気管を含む 10 個の正常組織のプール由来の c D N A に対してサブトラクトした。

【0340】

74 個のクローンを単離し、そして配列決定した。これらのクローンの 22 個について決定した c D N A 配列を、配列番号 223 ~ 244 に提供する。配列番号 241 の配列は、以前に同定された配列に対して有意な相同性は示さなかった。配列番号 223、225、232、233、235、238、239、242、および 243 の配列は、以前に同定された遺伝子配列に対していくらかの類似性を示し、一方、配列番号 224、226 ~ 231、234、236、237、240、241、および 244 の配列は、以前に単離された E S T に対していくらかの類似性を示した。 10

【0341】

肺腫瘍組織から調製した c D N A ライブラリーの特徴付けの間に単離したさらなる 12 個のクローンの配列を、配列番号 813 ~ 824 に提供する。これらの配列の G e n B a n k データベースおよび G e n e S e q D N A データベースとの比較は、以前に同定された配列に対する有意な相同性は明らかにしなかった。

【0342】

(実施例 2)

(ポリペプチドの合成)

ポリペプチドを、H P T U (O-ベンゾトリアゾール-N, N, N', N'-テトラメチルウロニンヘキサフルオロホスフェート) 活性化を用いて F M O C 化学を使用して、P e r k i n E l m e r / A p p l i e d B i o s y s t e m s D i v i s i o n 430A ペプチド合成装置上で合成し得る。G l y - C y s - G l y 配列を、固定化表面への共役、結合、またはペプチドの標識の方法を提供するために、ペプチドのアミノ末端に付着させ得る。固体支持体からのペプチドの切断は、以下の切断混合物を使用して行うことができる：トリフルオロ酢酸：エタンジチオール：チオアニソール：水：フェノール (40 : 1 : 2 : 2 : 3)。2 時間の切断後、ペプチドを冷却したメチル-tert-ブチルエーテル中で沈殿させ得る。次いで、ペプチドのペレットを、0.1% のトリフルオロ酢酸 (T F A) を含有している水中に溶解させ得、そして C 18 逆相 H P L C による精製の前に凍結乾燥させ得る。水中 (0.1% の T F A を含有している) の 0% ~ 60% のアセトニトリル (0.1% の T F A を含有している) の勾配を、ペプチドを溶出させるために使用し得る。純粋な画分の凍結乾燥後、ペプチドを、電気スプレー (e l e c t r o s p r a y) または他のタイプの質量スペクトル分析を使用し、そしてアミノ酸分析によって特徴付け得る。 20 30

【0343】

(実施例 3)

(L548S H i s タグ融合タンパク質の E. c o l i 中での発現)

L548S コード領域を以下のプライマーを用いて P C R 増幅した：

アミノ酸 2 で開始する順方向プライマー：

P D M - 433 : 5' g c t a a a g g t g a c c c c a a g a a a c c a a a g 3' Tm 60 °C (配列番号 1665) 40

終結コドンの後に X h o I 部位を作成する逆方向プライマー：

P D M - 438 : 5' c t a t t a a c t c g a g g g a g a c a g a t a a a c a g t t t c t t t a 3' Tm 61 °C (配列番号 1666)。

次いで、P C R 産物を、X h o I 制限酵素で消化し、ゲル精製し、次いで p P D M H i s 中 (E c o 72 I および X h o I 制限酵素で消化した、インフレイムで H i s タグを有する改変型 p E T 28 ベクター) にクローニングした。正確な構築物を、D N A 配列分析によって確認し、次いで B L 21 (D E 3) p L y s S および B L 21 (D E 3) C o d o n P l u s R I L 発現宿主中に形質転換した。発現した組換え L548S のタンパク質配列を、配列番号 1667 に示し、そして発現した組換え L7548S の D N A 配列 50

を、配列番号 1 6 6 8 に示す。

【0 3 4 4】

(実施例 4)

(肺チップ 5、S Q L 1、S C L 1、S C L 3、および S C L 4 ライブラリーのさらなる分析)

さらなる分析を、正常組織に対して、腫瘍プローブ群中での 2 倍以上の過剰発現および 0.2 以下の正常組織中での平均の発現という判断基準を使用して、肺チップ 5 について行った。これによって、肺癌腫中で過剰発現される 1 0 9 個の豊富ではないクローンの同定を得た。以下の表 1 1 にまとめるように、1 9 個の c D N A クローンを肺扁平上皮癌サブトラクトライブラリー S Q L 1 から回収し、9 個の c D N A クローンを小細胞性肺癌腫ライブラリー S C L 1 から回収し、そして 8 1 個の c D N A クローンを、小細胞性肺癌腫ライブラリー S C L 3 および S C L 4 から回収した。

10

【0 3 4 5】

【表 1 7】

表 11

配列番号	Seq. Ref.	エレメント (384)	エレメント (96)	比	平均 シグナル 1	平均 シグナル 2	ライブラ リー
1680	58456	p0003r03c13	R0001 E7	3.09	0.424	0.137	SQL1
1681	58458	p0003r03c10	R0001 F5	2.31	0.408	0.176	SQL1
1682	58462	p0003r04c16	R0001 H8	2.22	0.257	0.116	SQL1
1683	58469	p0003r07c12	R0002 F6	2.1	0.289	0.138	SQL1
1684	58470	p0003r09c21	R0003 A11	2.55	0.493	0.194	SQL1
1685	58482	p0003r12c19	R0003 G10	2.16	0.36	0.167	SQL1
1686	58485	p0003r12c10	R0003 H5	2.48	0.273	0.11	SQL1
1687	58501	p0004r04c23	R0005 G12	2.04	0.26	0.128	SQL1
1688	58502	p0004r04c03	R0005 G2	2.17	0.289	0.133	SQL1
1689	58505	p0004r05c23	R0006 A12	3.08	0.454	0.148	SQL1
1690	58507	p0004r06c11	R0006 C6	3.22	0.49	0.152	SQL1
1691	58509	p0004r07c15	R0006 E8	3.26	0.421	0.129	SQL1
1692	58512	p0004r09c03	R0007 A2	3.16	0.559	0.177	SQL1
1693	58527	p0004r12c22	R0007 H11	2.03	0.278	0.137	SQL1
1694	58529	p0004r14c09	R0008 C5	2.26	0.45	0.199	SQL1
1695	58531	p0004r16c01	R0008 G1	2.84	0.387	0.138	SQL1
1696	58537	p0005r02c08	R0009 D4	2.03	0.355	0.175	SQL1
1697	58539	p0005r03c08	R0009 F4	2.34	0.42	0.18	SQL1
1698	58545	p0005r07c21	R0010 E11	2.96	0.361	0.122	SQL1
1699	59319	p0005r10c04	R0011 D2	3.1	0.478	0.154	SCL1
1700	59322	p0005r12c01	R0011 G1	2.16	0.255	0.118	SCL1
1701	59348	p0006r11c12	R0015 F6	2.33	0.269	0.116	SCL1
1702	59350	p0006r14c13	R0016 C7	2.41	0.447	0.185	SCL1
1703	59363	p0007r02c16	R0017 D8	2.12	0.421	0.199	SCL1
1704	59365	p0007r03c20	R0017 F10	3.07	0.584	0.19	SCL1
1705	59370	p0007r04c10	R0017 H5	2.06	0.284	0.138	SCL1
1706	59373	p0007r05c23	R0018 A12	2.95	0.472	0.16	SCL1
1707	59376	p0007r06c02	R0018 D1	2.13	0.246	0.116	SCL1
1708	61050	p0011r02c10	R0033 D5	2.23	0.306	0.137	SCL3/4
1709	61051	p0011r03c23	R0033 E12	2.9	0.298	0.103	SCL3/4
1710	61052	p0011r03c08	R0033 F4	2.18	0.265	0.122	SCL3/4
1711	61054	p0011r03c16	R0033 F8	2.11	0.415	0.197	SCL3/4
1712	61056	p0011r04c13	R0033 G7	2.73	0.314	0.115	SCL3/4
1713	61057	p0011r04c10	R0033 H5	2.45	0.463	0.189	SCL3/4
1714	61060	p0011r05c11	R0034 A6	3.28	0.536	0.164	SCL3/4
1715	61062	p0011r06c21	R0034 C11	2.73	0.526	0.192	SCL3/4
1716	61063	p0011r06c05	R0034 C3	3.61	0.513	0.142	SCL3/4
1717	61064	p0011r06c04	R0034 D2	2.58	0.477	0.185	SCL3/4
1718	61065	p0011r06c14	R0034 D7	4.91	0.55	0.112	SCL3/4
1719	61066	p0011r06c18	R0034 D9	2.38	0.285	0.12	SCL3/4
1720	61069	p0011r07c16	R0034 F8	2.25	0.426	0.189	SCL3/4
1721	61070	p0011r08c21	R0034 G11	2	0.234	0.117	SCL3/4
1722	61071	p0011r08c03	R0034 G2	2.76	0.321	0.116	SCL3/4

10

20

30

40

(表 1 1 続き)

配列番号	Seq. Ref.	エレメント (384)	エレメント (96)	比	平均 シグナル 1	平均 シグナル 2	ライブ ラリー
1723	61074	p0011r08c16	R0034 H8	3.02	0.399	0.132	SCL3/4
1724	61075	p0011r09c05	R0035 A3	3.83	0.496	0.13	SCL3/4
1725	61077	p0011r10c21	R0035 C11	2.12	0.306	0.144	SCL3/4
1726	61079	p0011r11c23	R0035 E12	2.04	0.22	0.108	SCL3/4
1727	61080	p0011r11c15	R0035 E8	2.76	0.299	0.108	SCL3/4
1728	61081	p0011r11c14	R0035 F7	2.37	0.303	0.128	SCL3/4
1729	61083	p0011r12c15	R0035 G8	2.29	0.351	0.153	SCL3/4
1730	61085	p0011r13c05	R0036 A3	2.62	0.43	0.164	SCL3/4
1731	61086	p0011r13c09	R0036 A5	2.53	0.398	0.157	SCL3/4
1732	61088	p0011r14c05	R0036 C3	4.26	0.702	0.165	SCL3/4
1733	61090	p0011r15c07	R0036 E4	3.16	0.429	0.136	SCL3/4
1734	61091	p0011r16c16	R0036 H8	3.54	0.634	0.179	SCL3/4
1735	61093	p0012r02c03	R0037 C2	2.2	0.265	0.121	SCL3/4
1736	61094	p0012r02c11	R0037 C6	15.17	1.79	0.118	SCL3/4
1737	61096	p0012r02c08	R0037 D4	2.44	0.27	0.111	SCL3/4
1738	61097	p0012r02c10	R0037 D5	4.52	0.81	0.179	SCL3/4
1739	61099	p0012r03c02	R0037 F1	3.34	0.39	0.117	SCL3/4
1740	61100	p0012r03c06	R0037 F3	2.03	0.233	0.114	SCL3/4
1741	61103	p0012r04c17	R0037 G9	2.48	0.413	0.167	SCL3/4
1742	61105	p0012r05c11	R0038 A6	3.26	0.501	0.154	SCL3/4
1743	61106	p0012r05c08	R0038 B4	2.46	0.354	0.144	SCL3/4
1744	61110	p0012r06c15	R0038 C8	2.18	0.41	0.188	SCL3/4
1745	61113	p0012r07c09	R0038 E5	2.47	0.376	0.152	SCL3/4
1746	61115	p0012r07c13	R0038 E7	2.57	0.483	0.188	SCL3/4
1747	61117	p0012r07c24	R0038 F12	2.18	0.235	0.108	SCL3/4
1748	61118	p0012r07c18	R0038 F9	4.44	0.605	0.136	SCL3/4
1749	61119	p0012r08c03	R0038 G2	2.97	0.35	0.118	SCL3/4
1750	61120	p0012r08c07	R0038 G4	2.23	0.323	0.144	SCL3/4
1751	61122	p0012r08c18	R0038 H9	2.23	0.373	0.168	SCL3/4
1752	61125	p0012r10c17	R0039 C9	2.1	0.22	0.105	SCL3/4
1753	61126	p0012r10c16	R0039 D8	2.47	0.345	0.14	SCL3/4
1754	61130	p0012r12c12	R0039 H6	2.66	0.282	0.106	SCL3/4
1755	61133	p0012r13c24	R0040 B12	2.25	0.27	0.12	SCL3/4
1756	61134	p0012r14c23	R0040 C12	2.23	0.228	0.102	SCL3/4
1757	61135	p0012r14c03	R0040 C2	2.05	0.298	0.146	SCL3/4
1758	61137	p0012r14c02	R0040 D1	8.63	1.463	0.17	SCL3/4
1759	61139	p0012r14c14	R0040 D7	2.69	0.3	0.111	SCL3/4
1760	61143	p0012r16c02	R0040 H1	2.55	0.318	0.125	SCL3/4
1761	61144	p0012r16c18	R0040 H9	2.85	0.318	0.112	SCL3/4
1762	61148	p0013r02c19	R0041 C10	2.33	0.463	0.199	SCL3/4
1763	61151	p0013r02c03	R0041 C2	2.25	0.336	0.149	SCL3/4
1764	61155	p0013r04c07	R0041 G4	2.13	0.366	0.171	SCL3/4
1765	61156	p0013r05c05	R0042 A3	2.73	0.38	0.139	SCL3/4
1766	61159	p0013r06c24	R0042 D12	4.57	0.831	0.182	SCL3/4
1767	61160	p0013r07c19	R0042 E10	8.6	1.191	0.138	SCL3/4
1768	61163	p0013r07c18	R0042 F9	2.18	0.278	0.128	SCL3/4
1769	61167	p0013r10c12	R0043 D6	3.13	0.39	0.124	SCL3/4
1770	61172	p0013r12c03	R0043 G2	2	0.396	0.198	SCL3/4

10

20

30

40

(表 1 1 続き)

配列番号	Seq. Ref.	エレメント (384)	エレメント (96)	比	平均 シグナル 1	平均 シグナル 2	ライブ ラリー
1771	61173	p0013r12c07	R0043 G4	3.73	0.72	0.193	SCL3/4
1772	61176	p0013r13c04	R0044 B2	2.34	0.446	0.19	SCL3/4
1773	61177	p0013r14c01	R0044 C1	3.9	0.539	0.138	SCL3/4
1774	61183	p0013r15c14	R0044 F7	5.49	0.959	0.175	SCL3/4
1775	61185	p0013r16c24	R0044 H12	2.25	0.409	0.182	SCL3/4
1776	61188	p0014r01c07	R0045 A4	2.14	0.271	0.127	SCL3/4
1777	61192	p0014r02c19	R0045 C10	2.33	0.321	0.138	SCL3/4
1778	61198	p0014r04c24	R0045 H12	2.3	0.321	0.14	SCL3/4
1779	61201	p0014r06c22	R0046 D11	2.43	0.269	0.111	SCL3/4
1780	61202	p0014r06c08	R0046 D4	2.57	0.346	0.135	SCL3/4
1781	61204	p0014r07c07	R0046 E4	4.27	0.516	0.121	SCL3/4
1782	61206	p0014r07c12	R0046 F6	2.18	0.364	0.167	SCL3/4
1783	61210	p0015r09c02	R0051 B1	2.43	0.463	0.19	SCL3/4
1784	61212	p0015r10c15	R0051 C8	2.64	0.406	0.154	SCL3/4
1785	61216	p0015r11c16	R0051 F8	2.28	0.278	0.122	SCL3/4
1786	61225	p0015r14c12	R0052 D6	2.25	0.25	0.111	SCL3/4
1787	61226	p0015r14c14	R0052 D7	2.54	0.3	0.118	SCL3/4
1788	61227	p0015r16c18	R0052 H9	2.06	0.312	0.151	SCL3/4

10

20

上記の表中のシグナル 2 に対するシグナル 1 の比は、正常組織に対して腫瘍中で同定された配列の発現レベルの測定値を提供する。例えば、配列番号 1 6 6 9 については、腫瘍特異的シグナルは、試験した正常組織についてのシグナルの 3. 0 9 倍であり；配列番号 1 6 7 0 については、腫瘍特異的シグナルは、正常組織についてのシグナルの 2. 3 1 倍であったなど。

【 0 3 4 6 】

(実施例 5)

(肺腫瘍配列のリアルタイム P C R 分析)

リアルタイム P C R を、種々の腫瘍および正常組織中でのそれらの発現プロファイルをさらに評価するために、本明細書中で開示した肺腫瘍配列のサブセットに対して行った。簡潔には、P C R 産物の定量は、数回のサイクルによる。ここでは、D N A の量は、バックグラウンドをかるうじて上回るところからプラトーにまで対数的に増幅する。連続的な蛍光モニタリングを使用して、閾値サイクル数（ここでは、D N A が対数的に増幅する）をそれぞれの P C R 反応において容易に決定する。2 つの蛍光検出システムが存在する。1 つは、2 本鎖 D N A 特異的結合色素 S Y B R Green I 色素に基づく。他方は、R e p o r t e r 色素を 5 ' 末端 (F A M) に、そして Q u e n c h e r 色素を 3 ' 末端 (T A M R A) に含有している T a q M a n プローブ (P e r k i n E l m e r / A p p l i e d B i o s y s t e m s D i v i s i o n, F o s t e r C i t y, C A) を使用する。標的特異的 P C R 増幅によって、A m p l i T a q G o l d ^{T M} (P e r k i n E l m e r / A p p l i e d B i o s y s t e m s D i v i s i o n, F o s t e r C i t y, C A) のヌクレアーゼ活性により、Q u e n c h e r 含有プローブからの R e p o r t e r 色素の切断および放出を生じる。このように、放出されたレポーター色素から生じる蛍光シグナルは、P C R 産物の量に比例する。両方の検出方法によって、比較可能な結果を生じることが見出されている。複数の組織サンプル中での遺伝子発現の相対的なレベルを比較するために、組織および／または細胞株由来の R N A を使用して c D N A のパネルを構築し、そしてリアルタイム P C R を、それぞれの c D N A サンプル中のコピー数を定量するために遺伝子特異的プライマーを使用して行う。それぞれの c D N A サンプルについて、一般的には、2 連で行い、そしてそれぞれの反応を 2 連のプレートで繰り返した。最終的なリアルタイム P C R の結果を、典型的には、それぞれの c D N

30

40

50

Aサンプル中の内部アクチン数に対して基準化した目的の遺伝子のコピー数の平均として報告する。リアルタイムPCR反応を、SYBR Green I色素を使用するGeneAmp 5700 DetectorまたはTaqManプローブ(Perkin Elmer/Applied Biosystems Division, Foster City, CA)を使用するABI PRISM 7700 Detector上で行い得る。

【0347】

本明細書中で開示する多数の肺腫瘍特異的配列のリアルタイムPCR分析によって得た結果を、以下の表にまとめる。さらに、これらのクローンの多くについての延長したcDNA配列を、公開されている配列データベースを検索することによって得た。延長した配列、およびこれらの配列によってコードされるタンパク質を、以下の表12に配列番号によって同定する。

10

【0348】

【表18】

表 12

ライブラリー	クローン名	クローン番号	配列番号	リアルタイム PCR の結果	延長した cDNA 配列	コードされる ポリペプチド 配列
SQL1	L972P	47988	1789	1/7の肺鱗状腫瘍、1/3のHN鱗状腫瘍において過剰発現された。正常組織中での発現は低いかまたは存在しなかった。		
SQL1	L979P	48005	1790	2/7の鱗状肺腫瘍、1/3のHN鱗状腫瘍、1/2のアデノ(adenocarcinoma)肺腫瘍において過剰発現された。正常組織中での発現は低いかまたは存在しなかった。	1791	1806
SQL1	L970P	49853	1269	1/7の肺鱗状腫瘍、および1/3のHN鱗状腫瘍において高く過剰発現された。正常組織中での発現は低いかまたは存在しなかった。		

20

(表 12 続き)

ライブラリー	クローン名	クローン番号	配列番号	リアルタイム PCR の結果	延長した cDNA 配列	コードされるポリペプチド配列
SQL1	L981P	49865	1272	1/6 の鱗状肺腫瘍、および 1/3 の H N 鱗状腫瘍において過剰発現された。正常組織中での発現は低いかまたは存在しなかった。		
SQL1	L980P	49826	1279	3/7 の鱗状肺腫瘍、1/3 の H N 鱗状腫瘍、1/2 アデノ肺腫瘍において過剰発現された。正常組織中での発現は低いかまたは存在しなかった。	1792	1807
SCL1	L973P	20631	117	異型類癌腫 MET および腺癌において過剰発現された。いくつかの正常組織中で発現された。	1793	1808
SCL1	L974P	20661	128	初代小細胞、鱗状、および腺癌において過剰発現された。いくつかの正常組織中での発現が観察された。	1794	1809
SCL1	L996P	50430	1442	2/2 の初代小細胞、6/6 の小細胞性細胞株、1/1 の異型類癌腫 MET、1/1 のアデノ、1/1 の鱗状において過剰発現された。正常組織中での発現は非常に低いかまたは存在しなかった。	1795	1810
SCL3	L977P	26961	268	1/2 の初代小細胞、1/6 の小細胞性細胞株、および 1/1 の類癌腫 MET において過剰発現された。正常組織中での発現は非常に低いかまたは存在しなかった。	1796	1811
SCL2	L978P	24928	1339	2/2 の初代小細胞、3/6 の小細胞性細胞株、1/1 の類癌腫 MET、アデノおよび鱗状腫瘍のプールにおいて過剰発現された。正常組織中での発現は非常に低いかまたは存在しなかった。	1797	1812
SCL3/4	L984P	50507	1446	1/2 の初代小細胞性腫瘍、および 4/6 の小細胞性腫瘍細胞株において高く発現された。正常組織中での発現は低いかまたは存在しなかった。	1798	1813
SCL3/4	L580S	50536	1449	選択された小細胞性腫瘍および鱗状腫瘍において過剰発現された。正常な脳、細気管支、軟口蓋、および気管においてはいくつかの発現が観察された。		
SCL3/4	L988P	50645	1531	1/2 の初代小細胞、1/2 の初代小細胞、6/6 の小細胞性細胞株、1/1 の類癌腫 MET、アデノ、および鱗状腫瘍のプールにおいて過剰発現された。いくつかの正常組織（脳、副腎、唾液腺、気管、胸腺）中で発現された。	1799	1814
SCL3/4	L1423P	50625	1533	1/2 の初代小細胞、5/6 の小細胞性細胞株、1/1 の類癌腫 MET において過剰発現された。正常な脳および下垂体中でもまた発現された。	1800	1815
SCL3/4	L986P	50483	1490	1/2 の初代小細胞、5/6 の小細胞性細胞株、1/1 の類癌腫 MET、アデノおよび鱗状腫瘍のプールにおいて過剰発現された。正常な下垂体、および脊髄中で発現された。		

10

20

30

(表 12 続き)

ライブラリー	クローン名	クローン番号	配列番号	リアルタイム PCR の結果	延長した cDNA 配列	コードされるポリペプチド配列
SCL3/4	L987P	50560	1527	1/2 の初代小細胞、6/6 の小細胞性細胞株、1/1 の類癌腫 MET において過剰発現された。正常な下垂体および副腎中で発現された。	1801	1816-1818
SCL3/4	L1424P	50639	1547	1/2 の初代小細胞および 1/1 の類癌腫 MET において過剰発現された。正常組織中での発現は低いかまたは存在しなかった。		
MLN1	L997P	26749	730	1/1 の異型類癌腫 MET において過剰発現された。正常組織中では発現されなかった。		
MLN1	L999P	26752	733	2/2 の初代小細胞、6/6 の小細胞性細胞株、1/1 の類癌腫 MET において過剰発現された。いくつかの正常組織中で発現された。		
MLN1	L1400P	26529	405	2/6 の小細胞性細胞株、および 2/2 の初代小細胞において過剰発現された。いくつかの正常組織中で中程度から低く発現された。		
MLN1	L998P	27699	468	1/1 の異型類癌腫 MET において過剰発現された。正常組織中では低く発現された。	1802	1819
LAP1	L1425P	59303	949	4/7 の鱗状腫瘍、1/2 の腺癌腫瘍および 6 個の小細胞性肺癌腫のプールにおいて過剰発現された。正常な脳、腎臓、および骨格筋においては、中程度から高い発現が観察された。	1803	1820
LAP1	L1426P	59314	1156	1 個の肺鱗状腫瘍、および 1 個の H N 鱗状腫瘍において高く発現された。正常組織中において非常に低い発現または発現しないことが観察された。	1804	1821
LAP1	L1427P	59298	921	3/12 の腺癌腫瘍において高く過剰発現された。正常組織中での発現は非常に低いかまたは存在しなかった。	1805	1822
LAP1	L1428P	59316	1180	4/12 の腺癌腫瘍において過剰発現され、そしていくつかの他の腺癌腫瘍において低いレベルで発現された。正常組織中での発現は非常に低いかまたは存在しなかった。		

10

20

30

(実施例 6)

(肺腫瘍抗原に由来する C D 4 免疫原性 T 細胞エピトープの同定)

抗原 L 5 4 8 S (配列番号 7 8 9) に特異的な C D 4 T 細胞株を、以下のように作製した。

【0349】

L 5 4 8 S 配列全体にまたがる一連の重複する 20 マーのペプチドを合成した (それぞれ、配列番号 1 8 3 4 ~ 1 8 5 6)。初回免疫のために、ペプチドを 4 ~ 5 個のペプチドのプールに合わせて、そして 96 ウェル U 底プレート中で、樹状細胞および精製した C D 4 + T 細胞とともに $2 \mu\text{g}/\text{ml}$ で培養した。100 個の培養物を、それぞれのペプチドのプールについて作製した。培養物を、ペプチドのプールで充填した新しい樹状細胞で毎週再刺激した。全部で 3 回の刺激サイクルの後、細胞をさらに 1 週間休ませ、そしてインターフェロン γ E L I S A および増殖アッセイを使用して、ペプチドプールでパルスした抗原提示細胞 (A P C) に対する特異性について試験した。これらのアッセイのために、関連ペプチドのプール、組換え L 5 4 8 S、または無関係なペプチドのいずれかで充填した接着性の単球を、抗原提示細胞 (A P C) として使用した。以下の表 13 に示すように、多数の C D 4 T 細胞株が、初回免疫ペプチドならびに組換え L 5 4 8 S タンパク質

40

50

との反応性を示した。これらの株を、プール由来の個々のペプチドの認識について、ならびに組換えL548Sの認識について試験するために、さらに増殖させた。

【0350】

これらの株の優位な反応性は、ペプチド21（配列番号1854）（これは、L548Sのアミノ酸161～180に対応する）とであるようである。

【0351】

【表19】

表13

刺激		増殖 (CPM)		刺激指数	
ペプチド	陽性細胞株	抗原なし	ペプチド	ペプチド	L548S
p1-5	B1	700	14891	8791	21
	B2	1135	48724	53944	42
	G2	1227	7609	3193	6
p6-10					
p11-15	E2	8315	33723	13391	4
	E4	22097	100040	44171	5
p16-19					
p20-23	E3	3937	45367	15524	11
	F4	2648	130947	12927	49

10

（実施例7）

（患者の血清中の肺腫瘍抗原に対する抗体の検出）

肺腫瘍抗原L548S（配列番号789）およびL552S（配列番号809）に特異的な抗体が、肺癌患者の滲出液または血清中に存在するが、正常なドナーには存在しないことを示した。より詳細には、肺癌患者から得た滲出液中および正常なドナーに由来する血清中のL548S、L551S（配列番号827）およびL552Sに対する抗体の存在を、組換えタンパク質およびHRP結合体化抗ヒトIgを使用してELISAによって試験した。簡潔には、それぞれのタンパク質（100ng）を、pH9.5で96ウェルプレート中に被覆した。平行して、BSA（ウシ血清アルブミン）をもまた対照タンパク質として被覆した。BSA（[N]）に対するシグナル（[S]、405nmで測定した吸光度）を決定した。これらの研究の結果を、表14に示す。ここでは、－は[S] / [N] < 2を示し；+／－は[S] / [N] > 2を示し；++は[S] / [N] > 3を示し；そして+++は[S] / [N] > 5を示す。

20

【0352】

【表20】

30

表 14

肺腫瘍抗原に対する抗体の検出

	L548S	L551S	L552S
滲出液			
#1	+/-	-	-
#2	-	-	-
#3	-	-	+++
#4	-	-	-
#5	+/-	-	-
#7	-	-	-
#8	-	-	+/-
#10	-	-	+/-
#11	-	-	-
#12	=	-	+/-
#13	-	-	-
#14	+/-	-	+/-
#15	-	-	-
#17	-	-	++
#18	-	-	-
#19	-	-	++
#20	-	-	-
正常な血清			
#21	-	-	-
#22	-	-	-
#23	-	-	-
#24	-	-	-
#25	-	-	-

10

20

30

ウェスタンブロット分析を使用して、L 5 5 2 S に対する抗体が、肺癌患者に由来する滲出液の 4 個のサンプルのうちの 1 個に存在し、L 5 5 2 S 抗体は、試験した正常な血清の 3 個のサンプル中では検出されないことを見出した。

【0353】

(実施例 8)

(肺腫瘍抗原の融合タンパク質)

全長の R a 1 2 の、L 8 0 1 P O R F 4 (配列番号 1 6 7 0) または L 8 0 1 P O R F 5 (配列番号 1 6 7 1) のいずれかとの融合タンパク質を調製し、そして以下のように E . c o l i 中で 1 つの組換えタンパク質として発現させた。

40

【0354】

L 8 0 1 P の O R F 4 の c D N A を、全長の L 8 0 1 P の c D N A、ならびに配列番号 1 8 5 7 および 1 8 5 8 のプライマーを用いて P C R によって得た。L 8 0 1 P の O R F 5 の c D N A を、全長の L 8 0 1 P の c D N A、ならびに配列番号 1 8 5 9 および 1 8 6 0 のプライマーを用いて P C R によって得た。予想された大きさを有する P C R 産物をアガロースゲルから回収し、制限酵素 E c o R I および X h o I で消化し、そして続く E . c o l i 中での発現のために、発現ベクター p C R X 1 中の対応する部位にクローニングした。R a 1 2 と O R F 4 との融合体について、最良の発現を、2 × Y S 培地中の H M S 1 7 4 (D E 3) p L y s S 中で、約 3 時間 3 7 °C で I P T G を使用して誘導した組換えタンパク質を用いて得た。R a 1 2 の O R F 5 との融合体については、最良の発現を、2 ×

50

Y S 培地中の H M S 1 7 4 (D E 3) p L y s S 中で、再び、約 3 時間 3 7 °C で I P T G を使用して誘導した組換えタンパク質を用いて得た。融合タンパク質の産生のために使用したプラスミドを、D N A 配列決定によって確認した。O F R 4 および O R F 5 の融合体について決定した c D N A 配列を、配列番号 1 8 6 1 および 1 8 6 2 にそれぞれ提供し、対応するアミノ酸配列を、配列番号 1 8 6 3 および 1 8 6 4 にそれぞれ提供する。

【0355】

(実施例 9)

(全長の L 9 8 4 P をコードする c D N A のクローニング)

本実施例は、4 個の異なる c D N A 供給源からの、P C R 増幅による L 9 8 4 P をコードする c D N A 配列の単離を説明する。簡潔には、クローン L 9 8 4 P の先に単離した c D N A 配列を、ヒト a c h a e t e - s c u t e ホモログ 1 (A S H 1) をコードする D N A 配列に対して相同性を有しているとして同定した。遺伝子特異的プライマーを、A S H 1 について、公開されているドメイン中に存在する配列情報 (g e n b a n k 登録番号 N M - 0 0 4 3 1 6) を使用して作製した。これらの遺伝子特異的プライマーを P C R 増幅において使用して、L 9 8 4 P を、4 個の異なる c D N A 供給源からクローニングした。4 個の c D N A 供給源は、小細胞性肺癌腫初代腫瘍サンプル (R N A I d . 5 7 3 A) 、M E T 神経内分泌異型類癌腫サンプル (R N A I d . 5 1 2 A) 、および 2 つの小細胞性肺癌腫細胞株 (細胞株 I d . N C I H 1 2 8 および D M S 7 9) であった。これらの 4 個のクローンについて決定した c D N A 配列を、それぞれ、配列番号 1 8 6 5 ~ 1 8 6 8 に提供する。c D N A のコード領域は、ヒトゲノム D N A 中で多型を示すトリプレット C A G の繰り返しを含む。この多型は、4 個の異なる供給源から、ならびに A S H 1 の c D N A からクローニングされた L 9 8 4 P において観察することができ、そしてヒト第 4 染色体に由来する。g e n b a n k データベースに寄託された A S H 1 の c D N A (g e n b a n k 登録番号 N M - 0 0 6 6 8 8) は、1 4 コピーのトリプレット C A G を含むが、ヒト第 4 染色体の配列に由来する c D N A 配列 (g e n b a n k 登録番号 X M _ 0 0 6 6 8 8) は、1 2 コピーのトリプレットを含む。5 7 3 A および 5 1 2 A の両方からクローニングした c D N A は、1 2 コピーの C A G トリプレットを含む。小細胞性肺癌腫細胞株 D M S 7 9 からクローニングした c D N A は、1 3 コピーのトリプレット C A G を含み、一方、小細胞性肺癌腫細胞株 N C I H 1 2 8 からクローニングした c D N A は、わずか 1 0 コピーのトリプレット C A G を含む。多型はコード領域中に存在するので、これによって同様にタンパク質配列において多型を生じる (配列番号 1 8 6 9 ~ 1 8 7 2 を参照のこと)。

【0356】

(実施例 10)

(全長の L 9 8 5 P をコードする c D N A のクローニング)

実施例 1 C に先に開示したように、コンティグ 1 3 9 (配列番号 1 4 6 7、L 9 8 5 P としてもまた既知である) の配列を問い合わせとして使用して公開データベースの検索を行い、そしてこの配列が、細胞表面免疫調節因子-2 (C S I M M - 2、G e n e s e q データベースから得られた配列) をコードする c D N A (配列番号 1 6 7 6) に対して相同性を有することを示した。このクローンの全長の配列を、元々クローニングされた配列 (配列番号 1 4 6 7) の放射標識プローブを用いて、小細胞性肺癌腫 c D N A ライブラリーをスクリーニングすることによって得た。c D N A ライブラリーから約 5 0 0, 0 0 0 個のクローンをスクリーニングし、そして 1. 3 5 k b の c D N A 挿入物を含有している 2 個の異なるクローンを単離した。全長の c D N A 配列を配列番号 1 8 7 3 に提供し、そしてコードされたアミノ酸配列を、配列番号 1 8 7 4 に提供する。驚くべきことに、L 9 8 5 P の単離された全長の c D N A 配列 (配列番号 1 8 7 3) の、C S I M M - 2 についての G e n e s e q データベース配列とのアラインメントは、L 9 8 5 P が、1 つのヌクレオチドによって G e n e S e q データベース配列とは異なることを示した。このヌクレオチドの差異は、配列番号 1 8 7 4 の 1 1 9 位での 1 つのアミノ酸残基の変化 (G から E) を生じる。

【0357】

L985Pの推定されるタンパク質構造および配列は、これが最近記載されたMS4A（膜を貫通する4個のドメイン、サブファミリーA）遺伝子ファミリー（LiangおよびTedder、2001、Genomics 72:119~127）のメンバーであることを示す。MS4A遺伝子ファミリーは、現在、少なくとも21個の異なるヒトおよびマウスのタンパク質から構成され、そのうちの9個のメンバーはヒト由来である。これらには、CD20、FcεRIβ、HTm4、MS4A4A、MS4A5、MS4A6A、MS4A7、MS4A8B（L985Pと同じ）、およびMS4A12が含まれる。MS4Aファミリーのメンバーは、4個の膜貫通ドメインを含有している、細胞表面で発現されるタンパク質であり、N末端領域およびC末端領域は細胞の細胞質側に面している。ヒトMS4Aファミリーのメンバーは、タンパク質レベルで20~40%の全体的な相同性を示し、これらは、たいてい膜貫通ドメインに限定される。L985Pの膜貫通ドメインは、CD20に対して最も高い相同性を共有し、この領域の2つのタンパク質配列間で、約40%の同一性および60%の類似性を有する。MS4Aファミリーのメンバーは、広範囲の組織分布を示し、発現は、造血組織または非造血組織中の多様な細胞型において観察される。しかし、いくつかのMS4Aファミリーのメンバー（例えば、CD20）の発現は、特定の細胞型に高く制限される。CD20は、B細胞上でのみ発現される。実施例1Cにもまた記載したように、L985P（MS4A8B）は、定量的リアルタイムPCRによって決定されるように、小細胞性肺癌腫（SCLC）において過剰発現される。低いレベルの発現が、肺、脳下垂体、胃、結腸、および気管を含むいくつかの正常組織中で見られ、一方、調べた他の正常組織中での発現は、無視できるかまたは検出不可能であった（実施例1Cを参照のこと）。

10

20

【0358】

MS4Aファミリーのメンバーのほとんどについての生理学的機能は、解明されるべきままである。以前の研究によって、CD20が細胞増殖および分化の調節、ならびにB細胞中でのシグナル伝達に機能的に重要であることが示された。CD20が、ホモ-またはヘテロ-四量体複合体を形成することによってカルシウムチャンネルとして機能し得るという証拠もまた、存在する。FcεRIβは、四量体レセプター複合体の一部である。これは、肥満細胞の顆粒消失のような細胞性応答を導く、IgE結合抗原との相互作用を媒介する。MS4Aファミリーのメンバーの間での配列および構造の相同性に起因して、これらは、重複している機能的特性を共有している可能性が高い。

30

【0359】

MS4Aファミリーのメンバーのいくつかは、癌に関係していることが見出されている。CD20は、B細胞非ホジキンリンパ腫の90%以上において発現され、このことによって、CD20はB細胞悪性腫瘍の処置のための免疫治療アプローチのための理想的な標的とされた。抗CD20モノクローナル抗体は、未処理の（naive）の形態および放射標識した形態の両方で、非ホジキンリンパ腫の処置において高い成功を伴って使用されている。抗CD20 MA bは、補体媒介性細胞崩壊、抗体媒介性細胞性細胞傷害性、ならびに抗体媒介性の細胞周期の停止およびアポトーシスを含む、いくつかの経路を通じてそれらの抗腫瘍効果を発揮することが示されている。ヒトESTデータベースの分析は、MS4A遺伝子ファミリーの他のメンバー（MS4A4A、A6A、A7、およびA8Bを含む）もまた、肺、乳房、膵臓、結腸、卵巣、腎臓、および脳を含む種々の癌において発現されることを示す。CD20との比較によって、MS4Aファミリーの遺伝子の1つ以上が、造血性および非造血性の悪性疾患の処置のための免疫治療アプローチについての標的として使用される。

40

【0360】

MS4Aタンパク質ファミリーのメンバーは、4個の膜貫通ドメインを有する他のメンバーのタンパク質ファミリーと構造的に類似している。これらには、テトラスパニン（Tetraspanin）タンパク質ファミリー（TM4SF）およびGABA-Aレセプタータンパク質ファミリーが含まれる。テトラスパニンは癌に関係し、そして腫瘍の進行を

50

制御することにおいて直接的な役割を果たし得る。CD9の発現は、B細胞の移動にポジティブな影響を与えるが、CD9の過剰発現は癌腫細胞においては運動性および転移を抑制し、そして癌腫中では転移と逆の相関関係が存在する。しかし、CD9はまた、非T細胞急性リンパ芽球性白血病細胞の90%、および慢性リンパ性白血病の50%について発現される。RT-PCR分析を使用する、バーキットリンパ腫細胞株中でのテトラスパニン発現の最近の研究は、これらの株のうちの90%が、CD53、CD81、CD63、CD82、およびSASを高いレベルで発現することを見出した。CD151/PETA3は、転移および細胞の移動のエフェクターであり、そしてこの活性をブロックするMAbが開発されている。同様に、テトラスパニンCO-029/D6.1の過剰発現は、細胞株の転移能力を増大させる。テトラスパニンは、MAbによって調節され得る多様なセットの生物学的機能を制御する。テトラスパニンの機能は、一般的には、細胞の活性化および増殖、ならびに接着および運動性に影響を与える作用に分類され得る。これらの機能は、インテグリンとのそれらの会合によって実行される傾向にある。テトラスパニンの機能的活性は、例えば、細胞増殖を制御するように、MAbを用いて調節され得る。例えば、CD81/TAPA-1は、B細胞の活性化および増大した増殖(MAbでブロックされ得る活性)に関係する。抗増殖活性を有するMAbは、テトラスパニンファミリーのメンバーであるCO-029/D6.1に対して作製されている。

【0361】

上記のように、L985Pは、小細胞性肺癌腫において特異的に過剰発現される。この事実、ならびにCD20、テトラスパニン、およびHer2(癌におけるそれらの過剰発現が癌治療の標的としてそれらを有効にする)との比較は、L985Pが小細胞性肺癌腫の処置のための免疫治療アプローチの良好な標的であり得ることを示す。

【0362】

L985Pに対するMAbの作成、精製、および評価を容易にするために、L985Pタンパク質の推定されるアミノ酸配列全体、L985P由来のペプチド、または化学的に産生された(合成の)L985Pペプチドに対するMAbが使用される。また、Ra12タンパク質(長い形態(Ra12-これは、Ra12の最初の128アミノ酸である)および/もしくは短い形態(Ra12S)のいずれか)に融合したか、またはポリヒスチジンペプチドに対して融合したL985Pタンパク質分子のキメラ形態に対して、あるいは、これらの分子の任意の組合せに対して惹起されたMAbを使用することもできる。Hisタグ化L985P-Ra12融合分子についての推定されるcDNA配列およびアミノ酸配列を提供する: Ra12-L985P__cDNA(配列番号1875)、Ra12-L985P__Protein(配列番号1876)、Ra12S-L985P__cDNA(配列番号1877)、およびRa12S-L985P__Protein(配列番号1878); ならびに、L985P由来ペプチド: his-tagged Ra12S-L985PEx__cDNA(配列番号1879)、his-tagged Ra12S-L985PEx__Protein(配列番号1880)、L985P__Extracellular__Loop-2__cDNA(配列番号1881)、およびL985P__Extracellular__Loop-2__Peptide(配列番号1882)。

【0363】

(実施例11)

(Hisタグ化Ra12-L985P融合タンパク質のE. coliにおける発現)

本実施例は、Ra12とL985Pとの間での融合の特定の実施形態、およびE. coliにおけるその発現を示す。長い形態のRa12と、L985Pの最初の3個のアミノ酸残基を除く全てとのHisタグ化融合タンパク質を、E. coli中で単一の組換えタンパク質として発現させた。長い形態のRa12を、MTA32Aのアミノ酸残基192~323のもともとの配列から改変した。ここでは、推定トロニン切断部位をHindIII制限部位で置換した。L985Pを、pCRX1ベクター中のRa12配列の下流に融合した。Ra12配列を、RBSの下流にクローニングした。結果として、Hisタグ化融合タンパク質は、組換えベクターがE. coli中で発現される場合に産生される。

長い形態の R a 1 2 と L 9 8 5 P との融合体の配列を、D N A 配列決定によって確認した。決定した c D N A 配列を、配列番号 1 8 7 5 に提供する。この配列は、実施例 1 0 で推定したものと同一である。

【0364】

(実施例 1 2)

(全長の L 1 4 2 8 P をコードする c D N A のクローニング)

実施例 5 に先に開示したように、種々の腫瘍および正常組織中でのそれらの発現プロフィールをさらに評価するために、リアルタイム P C R を、本明細書中に開示した肺腫瘍配列のサブセットに対して行った。結果を、上記の表 1 2 に提供する。分析した配列のうちの 1 つは、クローン # 5 9 3 1 6 由来であり (配列番号 1 1 8 0、L 1 4 2 8 P)、そして肺腺癌のサブセットにおいて発現されることが示された。 10

【0365】

さらなる研究によって、# 5 9 3 1 6 (配列番号 1 1 8 0、L 1 4 2 8 P) のクローニングした配列についての全長の c D N A を単離した。この遺伝子の転写物の大きさを決定するために、複数の組織のノーザンブロットを、放射標識した元々クローニングされた配列 (配列番号 1 1 8 0) を用いてプローブした。ノーザンブロットは、肺腺癌および正常組織のサンプルに由来する約 2 0 μ g の全 R N A を含んだ。露光したフィルムの視覚的分析によって、約 6 . 5 ~ 7 . 0 k b の単一の転写物が明らかとなった。このクローンの全長の配列を、肺腺癌初代腫瘍 c D N A ライブラリーを、元々クローニングされた配列 (配列番号 1 1 8 0) の放射標識プローブでスクリーニングすることによって得た。c D N A ライブラリー由来の約 5 0 0 , 0 0 0 個のクローンをスクリーニングし、そして 6 . 8 k b の c D N A 挿入物を含有している 5 個の独立のクローンを単離した。この挿入物の大きさは、ノーザンブロット分析によって概算した大きさと類似である。全長の配列を、配列番号 1 8 8 3 に提供する。異なる O R F を同定することはできなかったが、7 個の可能性のある O R F を推定することができ、そしてこれらの可能性のある O R F のアミノ酸配列 (L 1 4 2 8 P _ O R F 1 ~ O R F 7 と命名した) を、それぞれ、配列番号 1 8 8 4 ~ 1 8 9 0 に提供する。全長の L 1 4 2 8 P の発現を、実施例 5 に示すように、肺の腺癌および扁平上皮癌の両方の延長した c D N A のパネルについて、リアルタイム P C R によって再度分析した。肺腺癌の延長したパネルのリアルタイムの結果によって、腺癌中での L 1 4 2 8 P の発現を再び確認する。腺癌サンプルの約 4 0 ~ 5 0 % が、様々なレベルの発現を示す。扁平上皮細胞癌の延長したパネルからのリアルタイムの結果は、L 1 4 2 8 P がまた、肺扁平上皮癌においても発現されることを示す。しかし、発現は、ごく少ない肺扁平上皮癌サンプルにおいてであり、そして低いレベルであった。 20 30 40

【0366】

(実施例 1 3)

(マイクロアレイによって示されるような、肺腫瘍中で過剰発現される c D N A 配列のリアルタイム P C R 分析)

表 1 5 に列挙する以下のクローンが肺腫瘍中で過剰発現されることが、マイクロアレイ分析 (実施例 2 C、および実施例 6、表 1 1 を参照のこと) によって先に示された。このマイクロアレイ分析の結果を、以下の表 1 5 にまとめる。 40

【0367】

【表 2 1】

表 1 5

配列番号	クローン番号 #	比	平均シグナル 1	平均シグナル 2
1383	50096 (コンティグ156)	16.04	1.895	0.118
1560	54454 (コンティグ234)	2.86	0.2	0.07
1561	54463 (コンティグ235)	3.31	0.248	0.075
1707	59376	3.23	0.734	0.227
1733	61090	2.68	0.364	0.136
1735	61093	5.7	0.687	0.121
1758	61137	8.63	1.463	0.17
1761	61144	7.02	0.783	0.112
1766	61159	4.57	0.831	0.182
1771	61173	5.57	1.075	0.193
1775	61185	5.29	0.962	0.182
1786	61225	3.41	0.379	0.111

10

20

リアルタイムPCR分析を、実施例7に記載するように、小細胞性肺癌腫のパネル上でこれらの配列について行った。リアルタイムPCR分析から得た結果を、表16にまとめる。

【0368】

【表22】

表 16

配列番号	リアルタイムPCRの結果
1383	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、および6/6のSCLC細胞株において過剰発現される。いくらかの発現が、正常な下垂体、脊髄、および胸腺においてもまた見られる。
1560	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、および3/6のSCLC細胞株において過剰発現される。発現はまた、正常な脳、および下垂体においても見られる。
1561	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、6/6のSCLC細胞株、1/1の異型類癌腫転移、腺癌、および扁平上皮癌のプールにおいて過剰発現される。発現はまた、正常な骨髄、リンパ節、胸腺においても見られ、そして他の正常組織においては低いレベルで発現される。
1707	SCLCパネル上では、この遺伝子は、1/6のSCLC細胞株および0/2の初代小細胞性癌腫において過剰発現される。正常な胃でもまた発現され、そして正常な脳、唾液腺、および気管においては低いレベルで発現される。
1733	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、6/6のSCLC細胞株、1/1の異型類癌腫転移、腺癌、および扁平上皮癌のプールにおいて過剰発現される。いくらかの発現がまた、正常な骨髄、リンパ節、胸腺においても見られ、そして他の正常組織においては低いレベルで発現される。
1735	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、5/6のSCLC細胞株、0/1の異型類癌腫転移、および腺癌のプールにおいて過剰発現される。低いレベルの発現がまた、正常な骨髄および骨格筋においても見られる。
1758	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、6/6のSCLC細胞株、0/1の異型類癌腫転移、腺癌、および扁平上皮癌のプールにおいて過剰発現される。いくらかの発現がまた、正常な骨髄、甲状腺、および気管においても見られる。
1761	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、6/6のSCLC細胞株、0/1の異型類癌腫転移、腺癌、および扁平上皮癌のプールにおいて過剰発現される。発現はまた、正常な骨髄においても見られる。

10

20

30

(表 16 続き)

配列番号	リアルタイム PCR の結果
1766	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、6/6のSCLC細胞株、および0/1の異型類癌腫転移において過剰発現される。発現はまた、正常な下垂体、副腎、骨髄、胸腺、唾液腺においても見られ、そして種々の他の正常組織においては低いレベルで発現される。
1771	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、6/6のSCLC細胞株、0/1の異型類癌腫転移、および扁平上皮細胞癌のプールにおいて過剰発現される。いくつかの発現がまた、正常な下垂体においても見られる。
1775	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、腺癌、および扁平上皮細胞癌のプールにおいて過剰発現される。SCLC細胞株、および異型類癌腫転移においては発現は観察されない。いくつかの正常組織での発現が、肝臓、胃、甲状腺、リンパ節、および胸腺で見られる。
1776	SCLCパネル上では、この遺伝子は、2/2の初代小細胞性癌腫、6/6個のSCLC細胞株、腺癌、および扁平上皮細胞癌のプールにおいて過剰発現される。いくつかの発現がまた、正常な骨髄、下垂体、胃、気管、および胸腺においても見られる。

10

20

次いで、これらの配列を利用可能なデータベース（Genbank、GeneSeq、huESTなど）中の既知の配列と比較した。これらの配列のうちの9個が、利用可能なデータベース中の既知の配列に対していくつかの程度の類似性を示した。これらの9個のヒットの結果を、表17にまとめ、これらの一致のDNA配列およびそれぞれのコードされたアミノ酸配列（入手可能である場合）についての配列表識別子を一緒に提供する。配列番号1561および1786は、いずれの既知の配列に対しても有意な類似性を示さなかった。

【0369】

【表23】

30

表 17

配列 番号	GenBank データベースヒット	GeneSeq. ヒット	ヒットの cDNA 配列	ヒットの アミノ酸 配列 (既知の場合)
1383	Pr22タンパク質/Stathmin/ 腫瘍タンパク質 18	A16376, A08801, A01633	1891	1901
1560	CDNA DKFZp564O163	--	1892	--
1561	新規	ヒト分泌タンパク質 5' EST. (C30107)	--	--
1707	SOX21 (AF107044)	ヒト分泌タンパク質 遺伝子 7 クローン HE8CV18. (X27317)	1893	1902
1733	KIAA0166 遺伝子 (D79988)	ヒト遺伝子痕跡 HUMGS08725. (T26483)	1894	1903
1735	ユビキチン結合体化 酵素 E2 (AF160215)	ヒトユビキチン様結合体化 タンパク質をコードするDNA (UBCLE). (X81676)	1895	1904
1758	下垂体腫瘍トランスフォー ミング遺伝子タンパク質 1 (AF095287)	Z97293, X89295, V36964, V63198, Z97292, C00858, V88346, Z80287	1896	1905
1761	新規	腎臓損傷関連分子 HW051 cDNA クローン (V80605)	1897	1906
1766	サイクリン依存性キナーゼ インヒビター p18 (CDKN2C) (AF041248)	サイクリン依存性キナーゼ (CDK6) 阻害タンパク質 (T10925; T31456)	1898	1907
1771	CDK4-インヒビター (p16- INK4) (L27211)	複数の一致	1899	1908
1775	γインターフェロンによって 誘導されるモノカイン (MIG) (NM_002416)	γインターフェロンによって 誘導されるモノカイン (X14998; Z26088)	1900	1909
1776	新規	--	--	--

さらなる研究によって、クローン# 61093 (配列番号1735、L1437P) のクローニングされた配列の全長の cDNA 配列の単離を生じた。遺伝子の転写物の大きさを決定するために、複数の組織のノーザンブロットを、放射標識した元々クローニングされた配列 (配列番号1735) でプローブした。ノーザンブロットは、小細胞性肺癌腫および正常組織のサンプルに由来する 20 μg の全 RNA を含んだ。露光させたフィルムの視覚的分析によって、約 1.2 kb の単一の転写物が明らかとなった。全長の配列を、小細胞性肺癌腫腫瘍 cDNA ライブラリーを、元々クローニングされた配列 (配列番号1735) の放射標識プローブでスクリーニングすることによって得た。cDNA ライブラリー由来の約 120,000 個のクローンをスクリーニングし、そして 931 塩基の cDNA 挿入物を含有している 2 個の異なるクローンを単離した。この挿入物は、ノーザンブロット分析によって概算した大きさと類似である。全長の cDNA 配列を、配列番号1910に提供する。全長の cDNA と以前に公開された配列との間に 1 つのヌクレオチドの差異

10

20

30

40

50

が存在することを発見した。しかし、このヌクレオチドの変化は、推定アミノ酸配列には変化を生じない。全長のcDNAによってコードされる推定アミノ酸配列は、配列番号1904にすでに提供したものと同一であり、そしてこのcDNAが、既知のタンパク質である、ユビキチン結合酵素E2 (AF160215、配列番号1904)をコードするという以前の予測を確認する。配列番号1910 (L1437P)が肺小細胞性肺癌腫において過剰発現されることを、マイクロアレイ、リアルタイムPCR、およびノーザンブロット分析によって示した。

【0370】

(実施例14)

(L548S Hisタグ融合タンパク質のE. coli中での発現)

10

PCRを、以下のプライマーを用いてL548Sコード領域について行った：

順方向プライマー PDM-433 5' g c t a a a g g t g a c c c c a a g a a a c c a a a g 3' (配列番号1911) Tm 60℃

逆方向プライマー PDM-438 5' c t a t t a a c t c g a g g g a g a c a g a t a a a c a g t t t c t t t a 3' (配列番号1912) Tm 61℃。

【0371】

PCR条件は以下であった：

10 µl 10×Pfu緩衝液

1. 0 µl 10mMのdNTP

2. 0 µl 10 µMのそれぞれのプライマー

20

83 µl 滅菌水

1. 5 µl Pfu DNAポリメラーゼ (Stratagene, La Jolla, CA)

50 ng DNA

96℃で2分間、96℃で20秒間、61℃で15秒間、72℃で1分30秒間を40サイクル、次いで72℃で4分間。

【0372】

PCR産物を、XhoI制限酵素で消化し、ゲル精製し、次いでpPDM His (Eco72IおよびXhoI制限酵素で消化した、インフレーションでHisタグを有する改変型pET28ベクター)にクローニングした。正確な構築物を、DNA配列分析によって確認し、次いで発現のためにBL21 CodonPlus (Stratagene, La Jolla, CA)細胞およびBL21 pLys S (Novagen, Madison, WI)細胞中に形質転換した。

30

【0373】

発現した組換えL548Sのアミノ酸配列を、配列番号1913に示し、そしてDNAコード領域の配列を、配列番号1914に示す。

【0374】

(実施例15)

(L551S Hisタグ融合タンパク質のE. coli中での発現)

40

PCRを、以下のプライマーを用いてL551Sコード領域について行った：

順方向プライマー PDM-498 5' g t g a c g a t g g a g g a g c t g c g g g a g a t g g 3' (配列番号1915) Tm 67℃

逆方向プライマー PDM-499 5' c g c c t a a c t c g a g t c a c t a a c a g c t g g g a g 3' (配列番号1916) Tm 66℃。

【0375】

PCR条件は以下であった：

10 µl 10×Pfu緩衝液

1. 0 µl 10mMのdNTP

2. 0 µl 10 µMのそれぞれのプライマー

83 µl 滅菌水

50

1. 5 μ l Pfu DNAポリメラーゼ (Stratagene, La Jolla, CA)
50 ng DNA
96 $^{\circ}$ Cで2分間、96 $^{\circ}$ Cで20秒間、66 $^{\circ}$ Cで15秒間、72 $^{\circ}$ Cで2分間20秒間を40サイクル、次いで72 $^{\circ}$ Cで4分間。

【0376】

PCR産物を、XhoI制限酵素で消化し、ゲル精製し、次いでpPDM His中 (Eco72IおよびXhoI制限酵素で消化した、インフレームでHisタグを有する改変型pET28ベクター) にクローン化した。正確な構築物を、DNA配列分析によって確認し、次いで発現のためにBLR (DE3) pLys SおよびBLR (DE3) CodonPlus RP細胞中に形質転換した。 10

【0377】

発現した組換えL551Sのアミノ酸配列を、配列番号1917に示し、そしてDNAコード領域の配列を、配列番号1918に示す。

【0378】

(実施例16)

(L552S Hisタグ融合タンパク質のE. coli中での発現)

PCRを、以下のプライマーを用いてL552Sコード領域について行った：

順方向プライマー PDM-479 5' c g g t g c c a c g c c c a t g g a c c t t c 3' (配列番号1919) Tm 64 $^{\circ}$ C 20

逆方向プライマー PDM-480 5' c t g a g a a t t c a t t a a a c t t g t g g t t g c t c t t c a c c 3' (配列番号1920) Tm 62 $^{\circ}$ C。

【0379】

PCR条件は以下であった：

10 μ l 10 $\times$ Pfu緩衝液

1. 0 μ l 10mMのdNTP

2. 0 μ l 10 μ Mのそれぞれのプライマー

83 μ l 滅菌水

1. 5 μ l Pfu DNAポリメラーゼ (Stratagene, La Jolla, CA) 30

50 ng DNA

96 $^{\circ}$ Cで2分間、96 $^{\circ}$ Cで20秒間、63 $^{\circ}$ Cで15秒間、72 $^{\circ}$ Cで1分間を40サイクル、次いで72 $^{\circ}$ Cで4分間。

【0380】

PCR産物を、EcoRI制限酵素で消化し、ゲル精製し、次いでpPDM His中 (Eco72IおよびEcoRI制限酵素で消化した、インフレームでHisタグを有する改変型pET28ベクター) にクローン化した。正確な構築物を、DNA配列分析によって確認し、次いで発現のためにBL21 CodonPlus (Stratagene, La Jolla, CA) 細胞中に形質転換した。 40

【0381】

発現した組換えL552Sのアミノ酸配列を、配列番号1921に示し、そしてDNAコード領域の配列を、配列番号1922に示す。

【0382】

(実施例17)

(全長クローン#19069およびクローン#18965または#19002をコードするcDNAのクローニング)

2つの肺抗原 (クローン#19069 (配列番号90) および#18965または#19002 (両方とも、配列番号15)) の部分的な配列は、先に提供した。これらの部分的な配列を、公に利用可能なデータベースを検索することによって、単離しクローン化し配列決定した全長のcDNA配列を推定するための問い合わせ配列として使用した。配列番 50

号90の単離しクローン化した配列の推定される全長のcDNA配列を、配列番号1923に提供する。配列番号15の単離したクローン化した配列についての推定される全長のcDNA配列を、配列番号1924に提供する。2つの抗原の推定されるアミノ酸配列を、それぞれ、配列番号1925および1926に提供する。次いで、これらの配列を、GeneSeqデータベース中の既知の配列と比較した。両方の配列が、GeneSeqデータベース中の既知の配列に対していくらかの程度の類似性を示した。配列番号1923は、リボホスファチジン酸アセチルトランスフェラーゼ (GeneSeq Z25000およびZ65038) に対して類似性を示し、そして配列番号1924は、亜鉛／鉄調節輸送体様タンパク質 (GeneSeq Z38333およびA14995) に対して類似性を示す。

10

【0383】

(実施例18)

(L552S特異的抗体のエピトープマッピング)

候補の抗原のペプチドを、前臨床研究および臨床研究の両方において、抗体応答の評価のために使用し得る。これらのデータは、特定の候補の抗原に対する抗体応答をさらに確認することを可能にする。競合ペプチドを用いるおよびそれを用いないタンパク質に基づくELISA、ならびにペプチドに基づくELISAを、これらの抗体応答を評価するために使用し得る。ペプチドELISAは、それがタンパク質に基づくELISAにおいて観察される抗体力価の擬陽性をさらに排除し得ること、ならびに候補の抗原に対する抗体応答を試験するための最も単純なアッセイシステムを提供するので、特に有用である。本実施例において、個々の癌患者が、L552Sの以下の3個のエピトープを主に認識するL552S特異的抗体を産生することを示すデータを、L552Sペプチドを使用して得た：

20

(1) aa21-35: GPRSGGAQAKLGCCW (配列番号1927)

(2) aa116-135: KVICKSCISQTPGINLDLGS (配列番号1928)

(3) aa141-160: IIPKEEHCKMPEAGEEQPV (配列番号1929)。

【0384】

さらなる研究においては、配列番号1929から作成した親和性精製した抗体がL552Sタンパク質を認識し得、そして患者の肺の複数の滲出液 (LPE) の全イムノグロブリンκの約0.6%を占めることを見出した。配列番号1929がL552Sタンパク質に特異的なウサギポリクローナル抗体の優性エピトープであることもまた、見出した。

30

【0385】

(実施例19)

(L985Pの発現)

哺乳動物細胞中での組換え発現のために、全長のL985P cDNAを、哺乳動物発現ベクターpCEP4 (Invitrogen) 中に、FLAGエピトープタグを付けてまたはそれを付けずにサブクローン化した。両方の構築物を、HEK293細胞 (ATCC) 中に、Lipofectamine 2000試薬 (Invitrogen) を使用してトランスフェクトした。次いで、ウェスタンブロット分析を、組換えL985Pが一時的に発現されるかどうかを決定するために、これらのトランスフェクトした細胞について行った。

40

【0386】

簡潔には、トランスフェクションを以下のように行った。HEK細胞を、10%のFBS (Hyclone) を含有しているDMEM (Gibco) 中に350,000細胞／ウェル (6ウェルプレート) の密度でプレートし、そして一晚増殖させた。翌日、2μlのLipofectamine 2000 (Invitrogen) を、FBSを含まない50μlのOptiMen 1 (Invitrogen) に添加し、そして室温で5分間インキュベートした。別のチューブのセットにおいては、50μlのOptiMen 1

50

を、 $0.8 \mu\text{g}$ のL985P (FLAGとともにおよびそれを伴わずに) / プラスミドDNAと混合し、そして混合物をLipofectamine 2000 / Optimem混合物に移した。組合せた混合物を、室温で20分間インキュベートし、そして 0.5 ml のDMEM 10% FBSを含有しているHEK293細胞に移した。次いで、Lipofectamine 2000 / DNA混合物をHEK293細胞に添加し、そして 37°C で16~24時間の間、7%の CO_2 でインキュベートした。細胞をPBSでリンスし、次いで回収し、そして遠心分離によってペレット化した。

【0387】

ウェスタンブロット分析のために、全細胞溶解物を、氷上で30分間、Triton-X 100を含有している溶解緩衝液中で細胞をインキュベートすることによって作成した。次いで、溶解物を15,000 rpmで5分間、 4°C での遠心分離によって明澄化した。サンプルを、 β -メルカプトエタノールを含有しているSDS-PAGE充填緩衝液で希釈し、次いで10分間沸騰させ、その後SDS-PAGEゲル上に充填した。このタンパク質をニトロセルロースに移し、そして1:1000希釈で精製した抗L985Pウサギポリクローナル血清を使用してプローブした。ブロットを、HRPに結合させたロバ抗ウサギIg (Jackson ImmunoResearch) を用い、続くECL基質中でのインキュベーションによって明らかにした。ブロットの結果は、組換えL985P (FLAGを伴うおよびそれを伴わない) がHEK293細胞中で発現されることを示す。

【0388】

(実施例20)

(肺腫瘍抗原に対するポリクローナル抗体の作成)

3個の肺抗原 (L548S (配列番号789)、L552S (配列番号809)、およびL985Pペプチド#3482 (配列番号1930)) を発現させ、そして抗体作成における使用のために精製した。

【0389】

L548SおよびL552Sを、E. coli組換え発現システムにおいて発現させ、そして適切な抗生物質を有するLB培地中で、振盪インキュベーターの中で 37°C で一晩増殖させた。翌朝、 10 ml の一晩培養物を、2Lのバッフル三角フラスコ中の適切な抗生物質を有する500 mlの $2 \times \text{YT}$ に添加した。培養物の光学密度が 560 nm で $0.4 \sim 0.6$ に達したときに、細胞をIPTG (1 mM) で誘導した。IPTGでの誘導の4時間後、細胞を遠心分離によって回収した。

【0390】

次いで、細胞をリン酸緩衝化生理食塩水で洗浄し、そして再び遠心分離した。上清を廃棄し、そして細胞を、将来の使用のために凍結または即時処理のいずれかを行った。 20 ml の溶解緩衝液を細胞ペレットに添加し、そしてボルテックスした。次いで、E. coli細胞を破裂させて開くために、この混合物を $16,000 \text{ psi}$ の圧力のフレンチプレス (french press) を通過させた。次いで細胞を再び遠心分離し、そして上清およびペレットを、組換えタンパク質の分配のためにSDS-PAGEによってチェックした。

【0391】

細胞ペレットに局在化されたタンパク質については、ペレットを 10 mM のTris pH 8.0、1%のCHAPS中に再懸濁し、そして封入体ペレットを洗浄し、そして再び遠心分離した。この手順をさらに2回繰り返した。洗浄した封入体ペレットを、 8 M の尿素、または 10 mM のTris pH 8.0および 10 mM のイミダゾールを含有している 6 M のグアニジンHClのいずれかで可溶化した。可溶化したタンパク質を、 5 ml のニッケルキレート樹脂 (Qiaigen) に添加し、そして持続的に攪拌しながら室温で45分から1時間の間インキュベートした。

【0392】

インキュベーション後、この樹脂およびタンパク質混合物を、使い捨てカラムを通じて注ぎ、そしてフロースルーを回収した。次いで、カラムを、10~20倍のカラム容量の可

溶化緩衝液で洗浄した。次いで、抗原を、8 Mの尿素、10 mMのTris (pH 8.0)、および300 mMのイミダゾールを使用してカラムから溶出させ、そして3 mlの画分中に回収した。SDS-PAGEゲルを、どの画分をさらなる精製のためにプールするかを決定するために行った。

【0393】

最後の精製工程として、強力な陰イオン交換樹脂（この場合は、Hi-Prep Q (Bio-rad)）を適切な緩衝液を用いて平衡化し、そして上記からプールした画分をカラム上に充填した。それぞれの抗原を、漸増塩勾配を用いてカラムから溶出させた。画分をカラムを流出させるとともに回収し、そして別のSDS-PAGEゲルを、カラムに由来するどの画分をプールするかを決定するために行った。

10

【0394】

プールした画分を、10 mMのTris (pH 8.0) に対して透析した。放出基準は、SDS-PAGEまたはHPLCによって決定する場合には純度、Lowryアッセイまたはアミノ酸分析によって決定する場合には濃度、アミノ末端のタンパク質配列によって決定する場合には同一性であり、そして内毒素レベルをLimulus (LAL) アッセイによって決定した。次いで、タンパク質を、0.22ミクロンのフィルターを通じる濾過後にバイアルに入れ、そして抗原を免疫化のために必要とされるまで凍結させた。

【0395】

L985Pペプチド#3482を合成し、そしてKLHに結合させ、そして免疫化のために必要とされるまで凍結させた。

20

【0396】

ポリクローナル抗血清を、100 μgのムラミルジペプチド(MDP)と混合した400 μgのそれぞれの肺抗原を使用して作成した。等量の不完全なフロイトのアジュバント(IFA)を添加し、次いで混合し、そしてウサギに皮下(S.C.)注射した。4週間後、ウサギを、等量のIFAと混合した200 μgの抗原を用いて、S.C.で追加免疫した。その後、ウサギを、100 μgの抗原を用いてI.V.で追加免疫した。それぞれの追加免疫の7日後に、動物を放血させた。次いで、血液を4℃で12～24時間の間インキュベートし、その後遠心分離して血清を作成した。

【0397】

ポリクローナル抗血清を、抗原でコートした96ウェルプレートを使用し、そして50 μl (典型的には、1 μg/μl)のポリクローナル抗血清とともに4℃で20時間の間インキュベートすることによって特徴付けた。

30

【0398】

250 μlのBSAブロッキング緩衝液をウェルに添加し、そして室温で2時間インキュベートした。プレートを、PBS/0.1%のTweenで6回洗浄した。ウサギ血清をPBS/0.1%のTween/0.1%のBSA中に希釈した。50 μlの希釈した血清をそれぞれのウェルに添加し、そして室温で30分間インキュベートした。プレートを上記のように洗浄し、次いで、1:10000希釈の50 μlのヤギ抗ウサギ西洋ワサビペルオキシダーゼ(HRP)を添加し、そして室温で30分間インキュベーションした。

【0399】

プレートを上記のように洗浄し、そして100 μlのTMBマイクロウェルペルオキシダーゼ基質(TMB Microwell Peroxidase Substrate)をそれぞれのウェルに添加した。室温で暗所での15分のインキュベーション後、比色定量反応を、100 μlの1 NのH₂SO₄を用いて停止させ、そして直ちに450 nmで読み取った。全てのポリクローナル抗体が、適切な抗原に対する免疫反応性を示した。

40

【0400】

表18～20は、3個の肺抗原(L548S、L552S、およびL958Pペプチド#3482)に対する、段階希釈物中のウサギ抗血清の抗体反応性を示す。第1列目は抗体希釈を示す。列「予備免疫血清」は、予備免疫血清を使用した2つの実験についてのELISAのデータを示す。これらの結果を、4列目で平均する。列「抗L548S、L55

50

2 S、または# 3 4 8 2」は、それぞれの抗原（表中で L 5 4 8 S、L 5 5 2 S、または # 3 4 8 2 のいずれかと呼ぶ）を使用して、上記に記載したように免疫したウサギ由来の血清を使用した 2 つの実験の E L I S A データを示す。

【 0 4 0 1 】

【表 2 4】

表 18

抗体の希釈	予備免疫血清(1)	予備免疫血清(2)	平均	抗 - L548S (1)	抗 - L548S (2)	平均
1:1000	0.17	0.10	0.14	0.51	0.51	0.51
1:2000	0.12	0.09	0.11	0.30	0.30	0.30
1:4000	0.09	0.08	0.08	0.17	0.20	0.19
1:8000	0.09	0.07	0.08	0.12	0.13	0.12
1:16000	0.09	0.08	0.08	0.09	0.12	0.11
1:32000	0.09	0.08	0.09	0.08	0.09	0.09
1:64000	0.11	0.09	0.10	0.11	0.12	0.12
1:128000	0.10	0.08	0.09	0.08	0.09	0.08
1:256000	0.09	0.08	0.08	0.11	0.09	0.10
1:512000	0.11	0.09	0.10	0.08	0.08	0.08
1:1024000	0.10	0.08	0.09	0.08	0.10	0.09
1:2048000	0.10	0.09	0.09	0.08	0.09	0.09

10

20

【 0 4 0 2 】

【表 2 5】

表 19

抗体の希釈	予備免疫(1)	予備免疫(2)	平均	抗 - L552S (1)	抗 - L552S (2)	平均
1:1000	0.08	0.19	0.13	2.14	2.03	2.08
1:2000	0.08	0.06	0.07	1.93	1.97	1.95
1:4000	0.07	0.06	0.06	1.81	1.82	1.82
1:8000	0.08	0.06	0.07	1.63	1.64	1.64
1:16000	0.06	0.05	0.05	1.47	1.29	1.38
1:32000	0.06	0.05	0.06	1.03	1.10	1.06
1:64000	0.06	0.06	0.06	0.73	0.69	0.71
1:128000	0.06	0.05	0.06	0.44	0.48	0.46
1:256000	0.06	0.06	0.06	0.26	0.25	0.26
1:512000	0.07	0.06	0.06	0.16	0.15	0.16
1:1024000	0.06	0.07	0.06	0.12	0.10	0.11
0.00	0.06	0.06	0.06	0.06	0.06	0.06

30

40

【 0 4 0 3 】

【表 2 6】

表 20

抗体の希釈	予備免疫 (1)	予備免疫 (2)	平均	抗 - #3482 (1)	抗 - #3482 (2)	平均
1:1000	0.10	0.07	0.09	2.10	2.07	2.08
1:2000	0.07	0.07	0.07	1.80	1.84	1.82
1:4000	0.07	0.06	0.06	1.78	1.80	1.79
1:8000	0.06	0.06	0.06	1.94	1.72	1.83
1:16000	0.06	0.06	0.06	1.75	1.74	1.74
1:32000	0.06	0.06	0.06	1.42	1.47	1.44
1:64000	0.06	0.06	0.06	1.12	1.17	1.14
1:128000	0.06	0.06	0.06	0.79	0.87	0.83
1:256000	0.06	0.06	0.06	0.70	0.65	0.68
1:512000	0.06	0.06	0.06	0.41	0.41	0.41
1:1024000	0.06	0.06	0.06	0.25	0.25	0.25
0	0.06	0.06	0.06	0.06	0.06	0.06

表 2 1 は、肺抗原 L 5 4 8 S に対する抗体のプロテイン A 精製を示す。

【 0 4 0 4 】

【 表 2 7 】

表 21

精製した 抗体の濃度 ($\mu\text{g/ml}$)	Pro A 純度	Pro A 純度	平均
3.0	2.20	2.12	2.16
1.5	2.09	2.01	2.05
0.75	1.95	1.93	1.94
0.38	1.83	1.85	1.84
0.188	1.68	1.68	1.68
0.094	1.36	1.38	1.37
0.047	1.02	1.06	1.04
0.0234	0.68	0.73	0.71
0.0177	0.40	0.42	0.41
0.0059	0.24	0.25	0.24
0.0029	0.17	0.15	0.16
0.00	0.06	0.06	0.06

表 2 2 は、肺抗原 L 5 5 2 S に対する抗体の親和性精製を示す。

【 0 4 0 5 】

【 表 2 8 】

表 22

抗体の希釈	親和性 純度 (塩 ピーク)	親和性 純度 (塩 ピーク)	平均	抗体の希釈	親和性 純度 (酸 ピーク)	親和性 純度 (酸 ピーク)	平均
1:50	0.19	0.18	0.18	1:1000	2.22	2.20	2.21
1:100	0.10	0.09	0.10	1:2000	2.17	2.10	2.13
1:200	0.06	0.06	0.06	1:4000	2.05	2.06	2.06
1:400	0.06	0.06	0.06	1:8000	1.95	1.94	1.95
1:800	0.06	0.05	0.06	1:16000	1.86	1.82	1.84
1:1600	0.06	0.06	0.06	1:32000	1.64	1.57	1.61
1:3200	0.06	0.06	0.06	1:64000	1.28	1.26	1.27
1:6400	0.06	0.06	0.06	1:128000	0.88	0.87	0.88
1:12800	0.06	0.06	0.06	1:256000	0.56	0.56	0.56
1:25600	0.06	0.06	0.06	1:512000	0.35	0.34	0.34
1:51200	0.06	0.06	0.06	1:1024000	0.18	0.17	0.18
0.00	0.06	0.06	0.06	0.00	0.06	0.06	0.06

表 2 3 A および 2 3 B は、肺抗原 L 9 8 5 P ペプチド # 3 4 8 2 に対する抗体の親和性精製およびプロテイン A 精製を示す。

【 0 4 0 6 】

【 表 2 9 】

表 23A

精製した抗体の 濃度 ($\mu\text{g/ml}$)	親和性 純度	親和性 純度	平均
3.0	1.73	1.69	1.71
1.50	1.32	1.28	1.30
0.75	0.96	0.91	0.93
0.38	0.63	0.61	0.62
0.19	0.38	0.39	0.38
0.09	0.19	0.20	0.19
0.05	0.16	0.10	0.13
0.02	0.12	0.15	0.13

【 0 4 0 7 】

【 表 3 0 】

表 23B

精製した抗体の 濃度 ($\mu\text{g/ml}$)	Pro A 純度	Pro A 純度	平均
40.70	1.80	1.91	1.85
20.35	1.54	1.53	1.53
10.18	1.04	1.18	1.11
5.09	0.71	0.78	0.74
2.54	0.43	0.52	0.48
1.27	0.25	0.28	0.26
0.64	0.13	0.20	0.17
0.32	0.10		0.10

10

20

30

40

50

(実施例 21)

(L1439Pの延長したcDNA配列のクローニング)

クローン#61144の部分的なcDNA配列(配列番号1761、クローンL1439Dとも呼ぶ)を、肺チップ5についての小細胞性肺癌腫のPCRに基づくサブトラクションライブラリーから同定した。実施例4に先に開示したように、マイクロアレイ分析を、種々の腫瘍および正常組織中でのそれらの発現プロファイルをさらに評価するために、肺腫瘍配列のこのサブセットについて行い、その結果を表11に提供する。クローンL1439Pが、正常組織に対して、腫瘍プローブ群中での2倍以上の過剰発現を有し、そして正常組織中での平均の発現は0.2未満であることを見出した。さらなる研究は、この同じクローンが小細胞性肺癌腫のサブセットにおいて過剰発現されることを示した(表16を参照のこと)。

【0408】

さらなる研究によって、L1439Pのクローン化された配列の全長のcDNAの単離を得た。遺伝子の転写物の大きさを決定するために、複数の組織のノーザンブロットを、放射標識した最初にクローン化した配列(配列番号1180)でプローブした。ノーザンブロットには、小細胞性肺癌腫および正常組織のサンプルに由来する約10 μg の全RNAを含んだ。露光させたフィルムの視覚的な分析によって、約2.4kbの単一の転写物が明らかとなった。このクローンの延長した配列を、小細胞性肺癌腫腫瘍のcDNAライブラリーを、最初にクローン化した配列(配列番号1761)の放射標識プローブを用いてスクリーニングすることによって得た。cDNAライブラリー由来の約120,000個のクローンをスクリーニングし、そして1.5kbのcDNA挿入物を含有している1個の固有のクローンを単離した。この延長したcDNA配列を、配列番号1931に提供し、そしてこの延長したcDNA配列によってコードされる推定されるアミノ酸配列を、配列番号1932に提供する。この延長したcDNA配列を、Genbankデータベースに対して分析し、そしてNUF2R遺伝子(AF326731)に対して類似性を示すことを見出した。NUF2Rの全長のcDNA配列を、配列番号1933に提供する。そしてNUF2Rによってコードされる推定されるアミノ酸配列を、配列番号1934に提供

する。

【0409】

(実施例22)

(ヒスチジンタグを含まないL552S融合タンパク質のMegaterium中での発現)

PCRを、以下のプライマーを用いてL552Sコード領域について行った：

順方向プライマー PDM-737 5' ctatgttgcatgcgggtgccacgccc 3' (配列番号1935) Tm66℃

逆方向プライマー PDM-738 5' caccgacctaaagatcttcattaaaccttggtgtg 3' (配列番号1936) Tm60℃。

10

【0410】

PCR条件は以下であった：

10μl 10×Pfu緩衝液

1.0μl 10mMのdNTP

2.0μl 10μMのそれぞれのプライマー

83μl 滅菌水

1.5μl Pfu DNAポリメラーゼ (Stratagene, La Jolla, CA)

50ng DNA

96℃で2分間、96℃で20秒間、63℃で15秒間、72℃で1分間を40サイクル、次いで72℃で4分間。

20

【0411】

PCR産物を、SphIおよびBglII制限酵素で消化し、ゲル精製し、次いでpMEG-3中 (SphIおよびBglII制限酵素で消化した) にクローン化した。正確な構築物を、DNA配列分析によって確認し、次いで発現のためにMegaterium細胞中に形質転換した。

【0412】

発現した組換えL552Sのアミノ酸配列を、配列番号1937に示し、そしてDNAコード領域の配列を、配列番号1938に示す。

【0413】

30

(実施例23)

(ヒスチジンタグを含まないL552S融合タンパク質のE.coli中での発現)

PCRを、以下のプライマーを用いてL552Sコード領域について行った：

順方向プライマー PDM-736 5' ctatgttgcatatatgcgggtgccaccgccc 3' (配列番号1939) Tm64℃

逆方向プライマー PDM-480 5' ctgagaaattcattaaaccttggtgtgtgctcttcaccc 3' (配列番号1920) Tm62℃。

【0414】

PCR条件は以下であった：

10μl 10×Pfu緩衝液

1.0μl 10mMのdNTP

2.0μl 10μMのそれぞれのプライマー

83μl 滅菌水

1.5μl Pfu DNAポリメラーゼ (Stratagene, La Jolla, CA)

50ng DNA

96℃で2分間、96℃で20秒間、63℃で15秒間、72℃で1分間を40サイクル、次いで72℃で4分間。

40

【0415】

PCR産物を、NdeIおよびEcoRI制限酵素で消化し、ゲル精製し、次いでpPD 50

M中（Nde IおよびEcoR I制限酵素で消化した改変型pET28ベクター）にクローン化した。正確な構築物を、DNA配列分析によって確認し、次いで発現のためにBLR pLys SおよびHMS 174 pLys S細胞中に形質転換した。

【0416】

発現した組換えL552Sのアミノ酸配列を、配列番号1940に示し、そしてDNAコード領域の配列を、配列番号1941に示す。

【0417】

（実施例24）

（インビトロの全遺伝子感作を使用するL552S特異的CTL株の作成）

腫瘍抗原ワクシニア感染DCを用いるインビトロでの全遺伝子感作（Yeeら、The Journal of Immunology, 157（9）：4079-86、1996）を使用して、インターフェロン γ ELISPOT分析によって決定したL552S腫瘍抗原で形質導入した自己由来繊維芽細胞を特異的に認識する、ヒトCTL株を導いた。詳細には、樹状細胞（DC）を、10%のヒト血清、50ng/mlのヒトGM-CSF、および30ng/mlのヒトIL-4を含有しているRPMI培地中で5日間増殖させることによって正常なヒトドナーのPBMCから誘導した、Percoll精製した単球から分化させた。培養後、DCを、5の感染多重度（M.O.I）でL552Sを発現する組換えアデノウイルスとともに一晚感染させ、そして2 μ g/mlのCD40リガンドの添加によって一晚成熟させた。次いで、ウイルスを、UV照射によって不活化させた。CD8+細胞を、CD4+およびCD14+細胞の除去によって富化させた。CD8+T細胞を、磁気ビーズシステムを使用して単離し、そして培養物の感作を、サイトカインIL-6およびIL-12を用いて6個の96ウェルプレートの個々のウェル中で開始した。培養物を、L552SおよびIL-2の存在下での同時刺激分子CD80でレトロウイルスによって形質導入した自己由来の初代繊維芽細胞を使用して、7～10日毎に再刺激した。3回の刺激サイクルの後、2個のCD8+T細胞株（2-12Gおよび4-7A）を、インターフェロン γ ELISPOT分析（これにより、L552S腫瘍抗原で形質導入した自己由来繊維芽細胞で刺激した場合に、特異的にインターフェロン γ を生じるが、対照抗原を用いた場合には生じない）を使用して同定した。両方の株を再刺激し、そして特異的HLAに対する拘束を決定するために、抗体ブロッキングアッセイにおいて再度試験した。株2-12Gは、HLA-B/C拘束のようであるが、一方、株4-7AはHLA-A拘束のようである。株2-12Gを、抗CD3およびフィーダー細胞（feeder cell）を使用してクローン化し、14個の特異的クローンを回収した。これらのクローンは、抗体ブロッキングアッセイにおいて親のL2-12G CTL株と同じパターンの反応性を有する。さらに、L552Sを発現するベクターで形質導入したHLA不適合B-LCL株のパネルの使用、およびELISPOTアッセイにおいてCTL株によるインターフェロン γ の産生の測定によると、これらのCTLはHLA-B*4402によって拘束されるようである。

【0418】

（実施例25）

（L552S-およびXAGE-1-特異的抗体のエピトープマッピング）

L552SがXAGE-1の別のスプライシングアイソフォームであることを先に見出した。本実施例においては、L552SおよびXAGE-1についてそれぞれ特異的な抗体についての肺癌患者の血清をスクリーニングするために、L552SまたはXAGE-1のいずれかについて特異的な20マーのペプチドを使用して、データを得た。個々の癌患者が、L552SならびにXAGE-1に特異的な両方の抗体を産生することを見出した。これらの特異的抗体は、L552Sの以下のさらなるエピトープおよびXAGE-1の2個のエピトープを主に認識する：

L552S特異的：

aa31-50：LGCCWGYPSPRSTWND RPF（配列番号1942）

XAGE-1特異的：

10

20

30

40

50

a a 1 1 - 3 0 : C S L G V F P S A P S P V W G T R R S C (配列番号 1 9 4 3)

a a 4 1 - 5 0 : I L S P L L R H G G H T Q T Q N H T A S (配列番号 1 9 4 4)。

【0 4 1 9】

(実施例 2 6)

(L 5 5 2 S の免疫組織化学的分析)

種々の正常組織および肺癌組織中での L 5 5 2 S タンパク質の発現を決定するために、免疫組織化学的 (I H C) 分析を、親和性精製した L 5 5 2 S ポリクローナル抗体を使用して行った。詳細には、組織サンプルを。1 2 ~ 2 4 時間の間ホルマリン溶液中に固定し、そしてパラフィン中に包埋し、その後 8 ミクロンの切片に薄く切断した。0 . 1 M のクエン酸ナトリウム緩衝液 (p H 6 . 0) 中の蒸気熱で誘導したエピトープ読み出し (s t e a m h e a t i n d u c e d e p i t o p e r e t r i e v a l (S H I E R) 10) を、最適な染色条件に使用した。切片を 1 0 % の血清 / P B S とともに 5 分間インキュベートした。一次抗体を、示した濃度で 2 5 分間それぞれの切片に対して添加し、その後、抗ウサギまたは抗マウスビオチニル化抗体のいずれかとともに 2 5 分間インキュベートした。内因性のペルオキシダーゼ活性を、過酸化水素との 3 回の 1 . 5 分間のインキュベーションによってブロックした。アビジンビオチン複合体 / 西洋ワサビペルオキシダーゼ (A B C / H R P) システムを、L 5 5 2 S の発現を視覚化するために、D A B 色素原とともに使用した。スライドを、細胞核を視覚化するためにヘマトキシリンで対比染色した。

【0 4 2 0】

L 5 5 2 S の I H C 分析を表 2 4 に開示する：

【0 4 2 1】

【表 3 1】

20

表 24

組織型	染色 (陽性 ／全部)	注釈
アデノ肺ガン	10/12	強い染色(核「点様」)
鱗状肺ガン	4/12	明るい染色(細胞質「点様」)
副腎	0/1	
動脈内皮	0/1	明るい細胞質染色
血液(骨髓)	0/1	
脳(小脳)	0/1	
脳(皮質)	0/1	
胸部	0/1	
気管支	0/5	
結腸	1/4	非常に明るい染色(細胞質「点様」)
食道	0/1	
ファロピーオ氏管	0/1	
胆嚢	0/1	
心臓	0/1	
腎臓	2/4	非常に明るい染色(細胞質「点様」)
肝臓	1/4	非常に明るい染色(細胞質「点様」)
肺	0/13	
脾臓	0/1	
下垂体	1/1	非常に明るい染色(細胞質「点様」)
胎盤	1/1	非常に明るい染色(細胞質「点様」)
前立腺	0/1	
骨格筋	0/1	
皮膚	0/1	
小腸	0/1	
脊髄	0/1	
脾臓	0/1	
胃	0/1	
精巣	1/2	非常に少ない選択された細胞が陽性
胸腺	0/1	
甲状腺	0/1	
気管	0/1	
膀胱	0/1	
尿管	0/1	
子宮	0/1	

10

20

30

40

50

(実施例 27)

(L552S由来のCD4免疫原性T細胞エピトープの同定)

抗原L552S(配列番号809)に特異的なCD4 T細胞株を以下のように作成した。全長のL552S(配列番号809)のアミノ酸残基に対応している15個のアミノ酸によって重複している、全部で30個の20マーのペプチドを合成した。30個の20マーのペプチドのそれぞれのアミノ酸配列、およびこれらのペプチドのそれぞれをコードするそれぞれのDNA配列を、表25に提供する。樹状細胞(DC)を、可塑性接着剤によって正常な雄性ヒトドナーのPBMCから誘導したPerccoll精製した単球から分化させ、そして10%のヒト血清、50ngのヒトGM-CSF、および30ng/mlのヒトIL-4を含有しているRPMI培地中で5日間増殖させた。精製したCD4 T細胞を、MACビーズおよびPBMCのネガティブ選択を使用してDCとして同じドナー

から作成した。DCを、 $0.5 \mu\text{g}/\text{mL}$ の個々の濃度で各ペプチドを用いて、20マールのペプチドのプールを用いて一晩パルスした。パルスしたDCを洗浄し、そして96ウェル丸底プレートの1つのウェルあたり10,000個の細胞でプレートし、そして精製したCD4⁺T細胞を、1つのウェルあたり100,000個の細胞で添加した。培養物に、 $10 \text{ ng}/\text{mL}$ のIL-6および $5 \text{ ng}/\text{mL}$ のIL-12を補充し、そして37℃でインキュベートした。

【0422】

培養物を、 $10 \text{ U}/\text{mL}$ のIL-2および $5 \text{ ng}/\text{mL}$ のIL-7を補充した、APCについての上記と同様に作成しそしてパルスしたDCを使用して1週間の基準で上記のように再刺激した。3回のインビトロでの刺激サイクル（最初の感作+2回の再刺激）の後、株（それぞれの株が1つのウェルに対応する）を、分類した無関係な抗原に由来するペプチドの無関係なペプチドのプールに対して、刺激プールに応答する特異的増殖およびサイトカインの産生について試験した。個々のCD4⁺T細胞株の数（IFN γ によって49/576、増殖によって63/576）が、L552Sペプチドのプールに応答して有意なサイトカインの放出（IFN γ ）および増殖を示したが、対照ペプチドのプールについてはそうではなかった。特異的活性を示したT細胞株のうちの25個を、L552Sペプチドの適切なプールについて再刺激し、そしてE. coli中で作成した個々のペプチドまたは組換えタンパク質でパルスした自己由来DCについて再度アッセイした。約13個の免疫原性ペプチドが、試験したペプチド抗原の全てのセットに由来するT細胞によって認識された。これらの13個のペプチド（*）を表25に示す。

10

20

【0423】

いくつかの場合においては、T細胞株のペプチド反応性は、単一のペプチドに対してマップされ得るが、いくつかはそれぞれのプール中の1つ以上のペプチドに対してマップされ得る。この結果は、13個のペプチドの全てが、L552Sタンパク質の本来保有しているエピトープであり得ることを示す。

【0424】

【表32】

表 25

L552S CD4 重複しているペプチド			
ペプチド	配列	ペプチド 配列番号	DNA 配列番号
1	MRCHAHGPSCLVTAITREEG	1945	1974
2	HGPSCLVTAITREEGGPRSG	1946	1975
3	LVTAITREEGGPRSGGAQAK	1947	1976
4*	TREEGGPRSGGAQAKLGCCW	1948	1977
5*	GPRSGGAQAKLGCCWGYPSP	1949	1978
6*	GAQAKLGCCWGYPSPRSTWN	1950	1979
7*	LGCCWGYPSPRSTWNPDRRF	1951	1980
8	GYPSPRSTWNPDRRFWTPQT	1952	1981
9	RSTWNPDRRFWTPQTGPGE	1953	1982
10	PDRRFWTPQTGPGEGRHERH	1954	1983
11	WTPQTGPGEGRHERHTQTQN	1955	1984
12	GPGEGRHERHTQTQNHASP	1956	1985
13	RHERHTQTQNHASPRSPVM	1957	1986
14	TQTQNHASPRSPVMESPKK	1958	1987
15*	HTASPRSPVMESPKKKNQQL	1959	1988
16	RSPVMESPKKKNQQLKVGIL	1960	1989
17	ESPKKKNQQLKVGILHLGSR	1961	1990
18*	KNQQLKVGILHLGSRQKKIR	1962	1991

10

20

30

19*	KVGILHLGSRQKKIRIQLRS	1963	1992
20*	HLGSRQKKIRIQLRSQCATW	1964	1993
21*	RQKKIRIQLRSQCATWKVICK	1965	1994
22*	IQLRSQCATWKVICKSCISQ	1966	1995
23*	SQCATWKVICKSCISQTPGIN	1967	1996
24*	KVICKSCISQTPGINLDLGS	1968	1997
25	SCISQTPGINLDLGSGVKVK	1969	1998
26	TPGINLDLGSGVKVKIIPKE	1970	1999
27*	LDLGSGVKVKIIPKEEHCKM	1971	2000
28	GVKVKIIPKEEHCKMPEAGE	1972	2001
29	IIPKEEHCKMPEAGEEQPV	1973	2002

10

(実施例 28)

(L552SのHLA拘束)

HLA拘束を決定するために、1個または2個のHLA対立遺伝子と適合する繊維芽細胞のパネルを、pBIB L552で形質導入したか、または50:1のMOIでアデノウイルスL552で感染させた。これらの標的を、1ウェルあたり10000個の繊維芽細胞、10000個のT細胞、および5U/mLのIL-2で、ELISPOTアッセイにおいてD77 L552 CD8クローンに対して試験した。結果は、拘束が、B*4402またはCw*0501のいずれかであることを示す。

20

【0425】

2つの間を決定するために、COS-7細胞を、pcDNA3 L552Sと、pcDNA3 B*4402またはpBIB Cw*0501との組合せでトランスフェクトした。これらの標的を、再び、1ウェルあたり10000個のCOS-7、10000個のクローン14、および5U/mLのIL-2で、ELISPOTアッセイにおいて、上記のように同じD77 L552 CD8クローンに対して試験した。クローンは、L552およびB*4402でトランスフェクトしたCOS-7細胞を認識した。

30

【0426】

さらなる研究においては、3個の異なる腫瘍(659-22、3-90T、およびHTB183)(これらは、L552メッセージレベルについてリアルタイムRT-PCRによって先に分析した)を選択し、そしてB*4402、または対照としてのCw*0501で形質導入した。659-22および3-90Tは、L552メッセージを含むが、HTB183はネガティブコントロールであった。2回の選択の後、腫瘍を、それらのB44の発現レベルについてFACSによって分析した。659-22が、B44を内因的に高いレベルで発現することを見出したが、B44、B*4402、B*4404などのうちのどれが発現されるかは決定できなかった。約20%の3-90TがB44を発現した。HTB183はB44を極めて十分に発現する。上記のクローンを、これらの腫瘍に対して試験した場合には、B*4402で形質導入した659-22が認識された。従って、結果は、L552がHLA-B*4402によって拘束されることを示す。

40

【0427】

(実施例 29)

(L552SおよびXAGE-1の発現)

本実施例においては、肺癌のXAGE-1によるL552Sの発現を正確に説明するために、リアルタイムRT-PCR分析を行った。リアルタイムRT-PCR分析を、L552Sの5'固有の領域、XAGE-1の5'固有の領域、および3'共通の領域に配置さ

50

れる特異的プライマーを使用して行った。L 5 5 2 SおよびX A G E - 1 についての特異的なメッセージを、2つの肺腫瘍サンプル中で検出したが、正常な肺サンプル中では検出できなかった。

【0428】

同一の発現プロファイルを、L 5 5 2 Sの5'固有領域と、共通の3'配列との間で観察した。X A G E - 1の5'固有プライマーを使用して肺腫瘍中で検出したX A G E - 1のメッセージレベルは、L 5 5 2 Sと比較してはるかに低かった。しかし、X A G E - 1によって保有される極端な二次構造は、X A G E - 1に固有の5'配列のcDNA合成を阻止し得る。従って、X A G E - 1がより豊富なアイソフォームであり得ることが、ノーザン分析によって明らかである。

10

【0429】

(実施例30)

(L 5 5 2 SおよびX A G E - 1のEST発現プロファイル)

L 5 5 2 SおよびX A G E - 1の発現プロファイルをさらに評価するために、電気的発現プロファイリングを、それぞれの抗原について行った。これを、公開されているESTデータベースに対して、実施例29に開示するものと同じ特異的プライマーを用いる検索によって行った。このプロファイリングの結果によって、遺伝子の2つのアイソフォームが存在することを確認する。L 5 5 2 SとX A G E - 1との間での発現の比は、約1:5である。さらに、L 5 5 2 SおよびX A G E - 1は、肝細胞癌、CML、生殖細胞腫瘍、ユーイング肉腫、およびA l v e o l a r R h a b d o m y o s a r c o m a中で発現されるようである。

20

【0430】

(実施例31)

(L 9 8 5 P肺腫瘍抗原の表面発現)

小細胞性肺癌腫(SCLC)細胞株、NCI-H69、NCI-H128、HTB-171、HTB-173、HTB-175、およびDMS 79を、10%のFBS(Hyclone)を含有しているDMEM(Gibco)中で37℃にて、7%のCO₂を用いて増殖させた。次いで、これらの増殖しているSCLC細胞株を、L 9 8 5 PがこれらのSCLC細胞株の表面上で発現されるかどうかを決定するために、L 9 8 5 Pの予想される細胞外領域に対して惹起させたL 9 8 5 Pペプチドポリクローナル血清を使用してFACS分析に供した。

30

【0431】

FACS分析のために、細胞を回収し、そして氷冷した染色緩衝液(PBS+1%のBSA+Azide+10μg/mlのヒトIgG)で洗浄した。次に、細胞を、一次抗体なしで、無関係なウサギIgGとともに、最終濃度20μg/mlで全分子を用いて、またはL 9 8 5 Pペプチド#3482(配列番号1930)に対して惹起させた親和性精製したウサギポリクローナル血清を最終濃度20μg/mlで用いてのいずれかで、氷上で1時間インキュベートした。L 9 8 5 Pペプチド#3482の配列(配列番号1930)は、L 9 8 5 Pタンパク質の推定される細胞外領域を示す。細胞を、染色緩衝液で2回洗浄し、次いで1:100希釈のヤギ抗ウサギIg(H+L)-FITC試薬(Southern Biotechnology)とともに氷上で30分間インキュベートした。2回の洗浄後、細胞を、ヨウ化プロピジウム(PI)を含有している染色緩衝液中に再懸濁した、浸透性の細胞の同定を可能にするバイアル染色を行い、そしてFACSによって分析した。

40

【0432】

さらに、リアルタイムPCRを、L 9 8 5 P mRNAがこれらのSCLC細胞株中で発現されるかどうかを決定するために行った。FACS分析およびmRNAの発現のリアルタイムPCRの結果を、表26に示す。

【0433】

【表33】

50

表 26

SCLC 細胞株	リアルタイムPCRによる mRNAの発現	FACSによる表面発現
NCI-H69	+	+
HTB-173	+	+
HTB-175	-	-
HTB-171	-	-
NCI-H128	-	-
DMS 79	-	-

10

本発明の特定の実施形態が、説明の目的のために本明細書中に記載されているが、種々の
改変が、本発明の精神および範囲を逸脱することなく行われ得ることが、上記から明らか
である。従って、本発明は、添付される特許請求の範囲による場合を除いて、限定されな
い。

20

【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/04514 A2(51) International Patent Classification: **C07K 14/47**Street, Apt. 1101, Seattle, WA 98133 (US); MCNARR,
Andria [A]; 3602 N.E. 11th C.

(21) International Application Number: PCT/US91/22058

(22) International Filing Date: 10 July 2001 (10.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/614,121	11 July 2000 (11.07.2000)	US
09/651,563	29 August 2000 (29.08.2000)	US
09/658,824	8 September 2000 (08.09.2000)	US
09/671,323	26 September 2000 (26.09.2000)	US
09/677,119	6 October 2000 (06.10.2000)	US
09/702,705	30 October 2000 (30.10.2000)	US
09/736,457	13 December 2000 (13.12.2000)	US
09/849,626	3 May 2001 (03.05.2001)	US

(71) Applicant (for all designated States except US): **CORIXA CORPORATION** [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **WANG, Tongtong** [US/US]; 8049 N.E. 28th Street, Medina, WA 98059 (US); **WATANABE, Yoshitaka** [JP/US]; 2266 78th Avenue S.E., Mercer Island, WA 98040 (US); **HENDERSON, Robert, A.** [US/US]; 8904 192nd Street S.W., Edmonds, WA 98026 (US); **JOHNSON, Jeffrey, C.** [US/US]; 25828 15th Place Street, Des Moines, WA 98198 (US); **REITER, Marc, W.** [US/US]; 31402 N.E. 43rd Place, Carnation, WA 98011 (US); **MAHNERAKIS, Margarita** [US/US]; 3444 36th Avenue W., Seattle, WA 98199 (US); **CARTER, Darleek** [US/US]; 321 Summit Avenue E., Seattle, WA 98102 (US); **FANGER, Gary, Richard** [US/US]; 15906 - 29th Drive S.E., Mill Creek, WA 98012 (US); **VEDVICK, Thomas, S.** [US/US]; 134 S. 300th Place, Federal Way, WA 98003 (US); **BANGTOR, Chaitanya, S.** [IN/US]; 2102 N. 105th

(72) Inventors; and

(75) Inventors/Applicants (for US only): **WANG, Aijun** [CN/US]; 3106 213th Place S.E., Issaquah, WA 98029 (US); **FANGER, Neil** [US/US]; 3618 Wolman Avenue N., Allyn, Seattle, WA 98103 (US); **SWITZER, Ann** [US/US]; 3838 N.E. 85th Street, Seattle, WA 98115 (US); **MCNEILL, Patricia, D.** [US/US]; 1333 Smith 290th Place, Federal Way, WA 98003 (US); **CLAPPER, Jonathan, D.** [US/US]; 2119 Dexter Avenue N., #1, Seattle, WA 98109 (US).

(74) Agents: **VERNA, James, M.**, Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7012 et al. (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GL, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AL, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, NG, SN, TD, TG).

Published:
without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette

WO 02/04514 A2

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly lung cancer, are disclosed. Illustrative compositions comprise one or more lung tumor polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly lung cancer.

WO 02/04514

PCT/US01/22058

COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF LUNG CANCER

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of cancer, such as lung cancer. The invention is more specifically related to polypeptides, comprising at least a portion of a lung tumor protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides are useful in pharmaceutical compositions, e.g., vaccines, and other compositions for the diagnosis and treatment of lung cancer.

10

INCORPORATION OF SEQUENCE LISTING ON CD-ROM BY REFERENCE

The Sequence Listing associated with this application is provided on CD-ROM in lieu of a paper copy under AI § 801(a), and is hereby incorporated by reference into the specification. Four CD-ROMs are provided containing identical copies of the sequence listing: CD-ROM No. 1 is labeled "COPY 1 - SEQUENCE LISTING PART," contains the file 47802pc.app.txt which is 1.4 MB and created on July 10, 2001; CD-ROM No. 2 is labeled "COPY 2 - SEQUENCE LISTING," contains the file 47802pc.app.txt which is 1.4 MB and created on July 10, 2001; CD-ROM No. 3 is labeled "COPY 3 - SEQUENCE LISTING PART," contains the file 47802pc.app.txt which is 1.4 MB and created on July 10, 2001; CD-ROM No. 4 is labeled "CRF," contains the file 47802pc.app.txt which is 1.4 Mb and created on July 10, 2001.

15

20

BACKGROUND OF THE INVENTION

Cancer is a significant health problem throughout the world. Although advances have been made in detection and therapy of cancer, no vaccine or other universally successful method for prevention or treatment is currently available. Current therapies, which are generally based on a combination of chemotherapy or surgery and radiation, continue to prove inadequate in many patients.

25

Lung cancer is the primary cause of cancer death among both men and women in the U.S., with an estimated 172,000 new cases being reported in 1994. The

WO 02/04514

PCT/US01/22058

2

five-year survival rate among all lung cancer patients, regardless of the stage of disease at diagnosis, is only 13%. This contrasts with a five-year survival rate of 46% among cases detected while the disease is still localized. However, only 16% of lung cancers are discovered before the disease has spread.

5 Early detection is difficult since clinical symptoms are often not seen until the disease has reached an advanced stage. Currently, diagnosis is aided by the use of chest x-rays, analysis of the type of cells contained in sputum and fiberoptic examination of the bronchial passages. Treatment regimens are determined by the type and stage of the cancer, and include surgery, radiation therapy and/or chemotherapy.

10 In spite of considerable research into therapies for this and other cancers, lung cancer remains difficult to diagnose and treat effectively. Accordingly, there is a need in the art for improved methods for detecting and treating such cancers. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

15 In one aspect, the present invention provides polynucleotide compositions comprising a sequence selected from the group consisting of:

(a) sequences provided in SEQ ID NOs: 1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002;

20 (b) complements of the sequences provided in SEQ ID NOs: 1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002;

25 (c) sequences consisting of at least 20, 25, 30, 35, 40, 45, 50, 75 and 100 contiguous residues of a sequence provided in SEQ ID NOs: 1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881,

WO 02/04514

PCT/US01/22058

3

1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002;

(d) sequences that hybridize to a sequence provided in SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002, under moderate or highly stringent conditions;

(e) sequences having at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identity to a sequence of SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002;

(f) degenerate variants of a sequence provided in SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002.

In one preferred embodiment, the polynucleotide compositions of the invention are expressed in at least about 20%, more preferably in at least about 30%, and most preferably in at least about 50% of lung tumors samples tested, at a level that is at least about 2-fold, preferably at least about 5-fold, and most preferably at least about 10-fold higher than that for normal tissues.

The present invention, in another aspect, provides polypeptide compositions comprising an amino acid sequence that is encoded by a polynucleotide sequence described above.

The present invention further provides polypeptide compositions comprising an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NOs:324-340, 783, 786, 787, 789, 791, 793, 795, 797-799, 805, 806, 809, 827, 1667, 1670-1675, 1677-1679, 1806-1822, 1825, 1830-1833, 1834-1856,

WO 02/04514

PCT/US01/22058

4

1863, 1864, 1869-1872, 1874, 1876, 1878, 1880, 1882, 1884-1890, 1901-1909, 1913, 1917, 1921, 1925-1930, 1932, 1934, 1937, 1940 and 1942-1973.

In certain preferred embodiments, the polypeptides and/or polynucleotides of the present invention are immunogenic, *i.e.*, they are capable of eliciting an immune response, particularly a humoral and/or cellular immune response, as further described herein.

The present invention further provides fragments, variants and/or derivatives of the disclosed polypeptide and/or polynucleotide sequences, wherein the fragments, variants and/or derivatives preferably have a level of immunogenic activity of at least about 50%, preferably at least about 70% and more preferably at least about 90% of the level of immunogenic activity of a polypeptide sequence set forth in SEQ ID NOs:324-340, 783, 786, 787, 789, 791, 793, 795, 797-799, 805, 806, 809, 827, 1667, 1670-1675, 1677-1679, 1806-1822, 1825, 1830-1833, 1834-1856, 1863, 1864, 1869-1872, 1874, 1876, 1878, 1880, 1882, 1884-1890, 1901-1909, 1913, 1917, 1921, 1925-1930, 1932, 1934, 1937, 1940 and 1942-1973 or a polypeptide sequence encoded by a polynucleotide sequence set forth in SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002.

The present invention further provides polynucleotides that encode a polypeptide described above, expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, the pharmaceutical compositions, *e.g.*, vaccine compositions, are provided for prophylactic or therapeutic applications. Such compositions generally comprise an immunogenic polypeptide or polynucleotide of the invention and an immunostimulant, such as an adjuvant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to

WO 02/04514

PCT/US01/22058

5

a polypeptide of the present invention, or a fragment thereof; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as
5 described above and (b) a pharmaceutically acceptable carrier or excipient. Illustrative antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, pharmaceutical compositions are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above
10 and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins, typically in the form of pharmaceutical compositions, e.g., vaccine compositions, comprising a physiologically acceptable carrier and/or an
15 immunostimulant. The fusions proteins may comprise multiple immunogenic polypeptides or portions/variants thereof, as described herein, and may further comprise one or more polypeptide segments for facilitating the expression, purification and/or immunogenicity of the polypeptide(s).

Within further aspects, the present invention provides methods for
20 stimulating an immune response in a patient, preferably a T cell response in a human patient, comprising administering a pharmaceutical composition described herein. The patient may be afflicted with lung cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

Within further aspects, the present invention provides methods for
25 inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition as recited above. The patient may be afflicted with lung cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

30 The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological

WO 02/04514

PCT/US01/22058

6

sample with T cells that specifically react with a polypeptide of the present invention, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a polypeptide of the present invention, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of polypeptide disclosed herein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for determining the presence or absence of a cancer, preferably a lung cancer, in a patient comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence

WO 02/04514

PCT/US01/22058

7

or absence of a cancer in the patient. Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of: (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a polypeptide of the present invention; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

WO 02/04514

PCT/US01/22058

8

(d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as
 5 monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description. All references disclosed herein are
 10 hereby incorporated by reference in their entirety as if each was incorporated individually.

SEQUENCE IDENTIFIERS

SEQ ID NO:1 is the determined cDNA sequence for clone #19038, also referred to as L845P.

15 SEQ ID NO:2 is the determined cDNA sequence for clone #19036.

SEQ ID NO:3 is the determined cDNA sequence for clone #19034.

SEQ ID NO:4 is the determined cDNA sequence for clone #19033.

SEQ ID NO:5 is the determined cDNA sequence for clone #19032.

20 SEQ ID NO:6 is the determined cDNA sequence for clone #19030, also referred to as L559S.

SEQ ID NO:7 is the determined cDNA sequence for clone #19029.

SEQ ID NO:8 is the determined cDNA sequence for clone #19025.

SEQ ID NO:9 is the determined cDNA sequence for clone #19023.

SEQ ID NO:10 is the determined cDNA sequence for clone #18929.

25 SEQ ID NO:11 is the determined cDNA sequence for clone #19010.

SEQ ID NO:12 is the determined cDNA sequence for clone #19009.

SEQ ID NO:13 is the determined cDNA sequence for clones #19005, 19007, 19016 and 19017.

SEQ ID NO:14 is the determined cDNA sequence for clone #19004.

30 SEQ ID NO:15 is the determined cDNA sequence for clones #19002 and 18965.

WO 02/04514

PCT/US01/22058

9

- SEQ ID NO:16 is the determined cDNA sequence for clone #18998.
SEQ ID NO:17 is the determined cDNA sequence for clone #18997.
SEQ ID NO:18 is the determined cDNA sequence for clone #18996.
SEQ ID NO:19 is the determined cDNA sequence for clone #18995.
- 5 SEQ ID NO:20 is the determined cDNA sequence for clone #18994, also known as L846P.
SEQ ID NO:21 is the determined cDNA sequence for clone #18992.
SEQ ID NO:22 is the determined cDNA sequence for clone #18991.
SEQ ID NO:23 is the determined cDNA sequence for clone #18990, also referred to as
- 10 clone #20111.
SEQ ID NO:24 is the determined cDNA sequence for clone #18987.
SEQ ID NO:25 is the determined cDNA sequence for clone #18985, also referred to as L839P.
SEQ ID NO:26 is the determined cDNA sequence for clone #18984, also referred to as
- 15 L847P.
SEQ ID NO:27 is the determined cDNA sequence for clone #18983.
SEQ ID NO:28 is the determined cDNA sequence for clones #18976 and 18980.
SEQ ID NO:29 is the determined cDNA sequence for clone #18975.
SEQ ID NO:30 is the determined cDNA sequence for clone #18974.
- 20 SEQ ID NO:31 is the determined cDNA sequence for clone #18973.
SEQ ID NO:32 is the determined cDNA sequence for clone #18972.
SEQ ID NO:33 is the determined cDNA sequence for clone #18971, also referred to as L801P.
SEQ ID NO:34 is the determined cDNA sequence for clone #18970.
- 25 SEQ ID NO:35 is the determined cDNA sequence for clone #18966.
SEQ ID NO:36 is the determined cDNA sequence for clones #18964, 18968 and 19039.
SEQ ID NO:37 is the determined cDNA sequence for clone #18960.
SEQ ID NO:38 is the determined cDNA sequence for clone #18959.
SEQ ID NO:39 is the determined cDNA sequence for clones #18958 and 18982.
- 30 SEQ ID NO:40 is the determined cDNA sequence for clones #18956 and 19015.

WO 02/04514

PCT/US01/22058

10

- SEQ ID NO:41 is the determined cDNA sequence for clone #18954, also referred to L848P.
- SEQ ID NO:42 is the determined cDNA sequence for clone #18951.
- SEQ ID NO:43 is the determined cDNA sequence for clone #18950.
- 5 SEQ ID NO:44 is the determined cDNA sequence for clones #18949 and 19024, also referred to as L844P.
- SEQ ID NO:45 is the determined cDNA sequence for clone #18948.
- SEQ ID NO:46 is the determined cDNA sequence for clone #18947, also referred to as L840P.
- 10 SEQ ID NO:47 is the determined cDNA sequence for clones #18946, 18953, 18969 and 19027.
- SEQ ID NO:48 is the determined cDNA sequence for clone #18942.
- SEQ ID NO:49 is the determined cDNA sequence for clone #18940, 18962, 18963, 19006, 19008, 19000, and 19031.
- 15 SEQ ID NO:50 is the determined cDNA sequence for clone #18939.
- SEQ ID NO:51 is the determined cDNA sequence for clones #18938 and 18952.
- SEQ ID NO:52 is the determined cDNA sequence for clone #18938.
- SEQ ID NO:53 is the determined cDNA sequence for clone #18937.
- SEQ ID NO:54 is the determined cDNA sequence for clones #18934, 18935, 18993 and
- 20 19022, also referred to as L548S.
- SEQ ID NO:55 is the determined cDNA sequence for clone #18932.
- SEQ ID NO:56 is the determined cDNA sequence for clones #18931 and 18936.
- SEQ ID NO:57 is the determined cDNA sequence for clone #18930.
- SEQ ID NO:58 is the determined cDNA sequence for clone #19014 (this sequence has
- 25 homology to clone L773P, which is also described in co-pending U.S. application 09/285,479, filed April 2, 1999).
- SEQ ID NO:59 is the determined cDNA sequence for clone #19127.
- SEQ ID NO:60 is the determined cDNA sequence for clones #19057 and 19064.
- SEQ ID NO:61 is the determined cDNA sequence for clone #19122.
- 30 SEQ ID NO:62 is the determined cDNA sequence for clones #19120 and 18121.
- SEQ ID NO:63 is the determined cDNA sequence for clone #19118.

WO 02/04514

PCT/US01/22058

11

- SEQ ID NO:64 is the determined cDNA sequence for clone #19117.
SEQ ID NO:65 is the determined cDNA sequence for clone #19116.
SEQ ID NO:66 is the determined cDNA sequence for clone #19114.
SEQ ID NO:67 is the determined cDNA sequence for clone #19112, also known as
5 L561S.
SEQ ID NO:68 is the determined cDNA sequence for clone #19110.
SEQ ID NO:69 is the determined cDNA sequence for clone #19107, also referred to as
L552S.
SEQ ID NO:70 is the determined cDNA sequence for clone #19106, also referred to as
10 L547S.
SEQ ID NO:71 is the determined cDNA sequence for clones #19105 and 19111.
SEQ ID NO:72 is the determined cDNA sequence for clone #19099.
SEQ ID NO:73 is the determined cDNA sequence for clones #19095, 19104 and 19125,
also referred to as L549S.
15 SEQ ID NO:74 is the determined cDNA sequence for clone #19094.
SEQ ID NO:75 is the determined cDNA sequence for clones #19089 and 19101.
SEQ ID NO:76 is the determined cDNA sequence for clone #19088.
SEQ ID NO:77 is the determined cDNA sequence for clones #19087, 19092, 19096,
19100 and 19119.
20 SEQ ID NO:78 is the determined cDNA sequence for clone #19086.
SEQ ID NO:79 is the determined cDNA sequence for clone #19085, also referred to as
L550S.
SEQ ID NO:80 is the determined cDNA sequence for clone #19084, also referred to as
clone #19079.
25 SEQ ID NO:81 is the determined cDNA sequence for clone #19082.
SEQ ID NO:82 is the determined cDNA sequence for clone #19080.
SEQ ID NO:83 is the determined cDNA sequence for clone #19077.
SEQ ID NO:84 is the determined cDNA sequence for clone #19076, also referred to as
L551S.
30 SEQ ID NO:85 is the determined cDNA sequence for clone #19074, also referred to as
clone #20102.

WO 02/04514

PCT/US01/22058

12

- SEQ ID NO:86 is the determined cDNA sequence for clone #19073, also referred to as L560S.
- SEQ ID NO:87 is the determined cDNA sequence for clones #19072 and 19115.
- SEQ ID NO:88 is the determined cDNA sequence for clone #19071.
- 5 SEQ ID NO:89 is the determined cDNA sequence for clone #19070.
- SEQ ID NO:90 is the determined cDNA sequence for clone #19069.
- SEQ ID NO:91 is the determined cDNA sequence for clone #19068, also referred to L563S.
- SEQ ID NO:92 is the determined cDNA sequence for clone #19066.
- 10 SEQ ID NO:93 is the determined cDNA sequence for clone #19065.
- SEQ ID NO:94 is the determined cDNA sequence for clone #19063.
- SEQ ID NO:95 is the determined cDNA sequence for clones #19061, 19081, 19108 and 19109.
- SEQ ID NO:96 is the determined cDNA sequence for clones #19060, 19067 and 19083,
- 15 also referred to as L548S.
- SEQ ID NO:97 is the determined cDNA sequence for clones #19059 and 19062.
- SEQ ID NO:98 is the determined cDNA sequence for clone #19058.
- SEQ ID NO:99 is the determined cDNA sequence for clone #19124.
- SEQ ID NO:100 is the determined cDNA sequence for clone #18929.
- 20 SEQ ID NO:101 is the determined cDNA sequence for clone #18422.
- SEQ ID NO:102 is the determined cDNA sequence for clone #18425.
- SEQ ID NO:103 is the determined cDNA sequence for clone #18431.
- SEQ ID NO:104 is the determined cDNA sequence for clone #18433.
- SEQ ID NO:105 is the determined cDNA sequence for clone #18444.
- 25 SEQ ID NO:106 is the determined cDNA sequence for clone #18449.
- SEQ ID NO:107 is the determined cDNA sequence for clone #18451.
- SEQ ID NO:108 is the determined cDNA sequence for clone #18452.
- SEQ ID NO:109 is the determined cDNA sequence for clone #18455.
- SEQ ID NO:110 is the determined cDNA sequence for clone #18457.
- 30 SEQ ID NO:111 is the determined cDNA sequence for clone #18466.
- SEQ ID NO:112 is the determined cDNA sequence for clone #18468.

WO 02/04514

PCT/US01/22058

13

- SEQ ID NO:113 is the determined cDNA sequence for clone #18471.
SEQ ID NO:114 is the determined cDNA sequence for clone #18475.
SEQ ID NO:115 is the determined cDNA sequence for clone #18476.
SEQ ID NO:116 is the determined cDNA sequence for clone #18477.
5 SEQ ID NO:117 is the determined cDNA sequence for clone #20631.
SEQ ID NO:118 is the determined cDNA sequence for clone #20634.
SEQ ID NO:119 is the determined cDNA sequence for clone #20635.
SEQ ID NO:120 is the determined cDNA sequence for clone #20637.
SEQ ID NO:121 is the determined cDNA sequence for clone #20638.
10 SEQ ID NO:122 is the determined cDNA sequence for clone #20643.
SEQ ID NO:123 is the determined cDNA sequence for clone #20652.
SEQ ID NO:124 is the determined cDNA sequence for clone #20653.
SEQ ID NO:125 is the determined cDNA sequence for clone #20657.
SEQ ID NO:126 is the determined cDNA sequence for clone #20658.
15 SEQ ID NO:127 is the determined cDNA sequence for clone #20660.
SEQ ID NO:128 is the determined cDNA sequence for clone #20661.
SEQ ID NO:129 is the determined cDNA sequence for clone #20663.
SEQ ID NO:130 is the determined cDNA sequence for clone #20665.
SEQ ID NO:131 is the determined cDNA sequence for clone #20670.
20 SEQ ID NO:132 is the determined cDNA sequence for clone #20671.
SEQ ID NO:133 is the determined cDNA sequence for clone #20672.
SEQ ID NO:134 is the determined cDNA sequence for clone #20675.
SEQ ID NO:135 is the determined cDNA sequence for clone #20679.
SEQ ID NO:136 is the determined cDNA sequence for clone #20681.
25 SEQ ID NO:137 is the determined cDNA sequence for clone #20682.
SEQ ID NO:138 is the determined cDNA sequence for clone #20684.
SEQ ID NO:139 is the determined cDNA sequence for clone #20685.
SEQ ID NO:140 is the determined cDNA sequence for clone #20689.
SEQ ID NO:141 is the determined cDNA sequence for clone #20699.
30 SEQ ID NO:142 is the determined cDNA sequence for clone #20701.
SEQ ID NO:143 is the determined cDNA sequence for clone #20702.

WO 02/04514

PCT/US01/22058

14

- SEQ ID NO:144 is the determined cDNA sequence for clone #20708.
SEQ ID NO:145 is the determined cDNA sequence for clone #20715.
SEQ ID NO:146 is the determined cDNA sequence for clone #20716.
SEQ ID NO:147 is the determined cDNA sequence for clone #20719.
5 SEQ ID NO:148 is the determined cDNA sequence for clone #19129.
SEQ ID NO:149 is the determined cDNA sequence for clone #19131.1.
SEQ ID NO:150 is the determined cDNA sequence for clone #19132.2.
SEQ ID NO:151 is the determined cDNA sequence for clone #19133.
SEQ ID NO:152 is the determined cDNA sequence for clone #19134.2.
10 SEQ ID NO:153 is the determined cDNA sequence for clone #19135.2.
SEQ ID NO:154 is the determined cDNA sequence for clone #19137.
SEQ ID NO:155 is a first determined cDNA sequence for clone #19138.1.
SEQ ID NO:156 is a second determined cDNA sequence for clone #19138.2.
SEQ ID NO:157 is the determined cDNA sequence for clone #19139.
15 SEQ ID NO:158 is a first determined cDNA sequence for clone #19140.1.
SEQ ID NO:159 is a second determined cDNA sequence for clone #19140.2.
SEQ ID NO:160 is the determined cDNA sequence for clone #19141.
SEQ ID NO:161 is the determined cDNA sequence for clone #19143.
SEQ ID NO:162 is the determined cDNA sequence for clone #19144.
20 SEQ ID NO:163 is a first determined cDNA sequence for clone #19145.1.
SEQ ID NO:164 is a second determined cDNA sequence for clone #19145.2.
SEQ ID NO:165 is the determined cDNA sequence for clone #19146.
SEQ ID NO:166 is the determined cDNA sequence for clone #19149.1.
SEQ ID NO:167 is the determined cDNA sequence for clone #19152.
25 SEQ ID NO:168 is a first determined cDNA sequence for clone #19153.1.
SEQ ID NO:169 is a second determined cDNA sequence for clone #19153.2.
SEQ ID NO:170 is the determined cDNA sequence for clone #19155.
SEQ ID NO:171 is the determined cDNA sequence for clone #19157.
SEQ ID NO:172 is the determined cDNA sequence for clone #19159.
30 SEQ ID NO:173 is the determined cDNA sequence for clone #19160.
SEQ ID NO:174 is a first determined cDNA sequence for clone #19161.1.

WO 02/04514

PCT/US01/22058

15

- SEQ ID NO:175 is a second determined cDNA sequence for clone #19161.2.
SEQ ID NO:176 is the determined cDNA sequence for clone #19162.1.
SEQ ID NO:177 is the determined cDNA sequence for clone #19166.
SEQ ID NO:178 is the determined cDNA sequence for clone #19169.
5 SEQ ID NO:179 is the determined cDNA sequence for clone #19171.
SEQ ID NO:180 is a first determined cDNA sequence for clone #19173.1.
SEQ ID NO:181 is a second determined cDNA sequence for clone #19173.2.
SEQ ID NO:182 is the determined cDNA sequence for clone #19174.1.
SEQ ID NO:183 is the determined cDNA sequence for clone #19175.
10 SEQ ID NO:184 is the determined cDNA sequence for clone #19177.
SEQ ID NO:185 is the determined cDNA sequence for clone #19178.
SEQ ID NO:186 is the determined cDNA sequence for clone #19179.1.
SEQ ID NO:187 is the determined cDNA sequence for clone #19179.2.
SEQ ID NO:188 is the determined cDNA sequence for clone #19180.
15 SEQ ID NO:189 is a first determined cDNA sequence for clone #19182.1.
SEQ ID NO:190 is a second determined cDNA sequence for clone #19182.2.
SEQ ID NO:191 is the determined cDNA sequence for clone #19183.1.
SEQ ID NO:192 is the determined cDNA sequence for clone #19185.1.
SEQ ID NO:193 is the determined cDNA sequence for clone #19187.
20 SEQ ID NO:194 is the determined cDNA sequence for clone #19188.
SEQ ID NO:195 is the determined cDNA sequence for clone #19190.
SEQ ID NO:196 is the determined cDNA sequence for clone #19191.
SEQ ID NO:197 is the determined cDNA sequence for clone #19192.
SEQ ID NO:198 is the determined cDNA sequence for clone #19193.
25 SEQ ID NO:199 is a first determined cDNA sequence for clone #19194.1.
SEQ ID NO:200 is a second determined cDNA sequence for clone #19194.2.
SEQ ID NO:201 is the determined cDNA sequence for clone #19197.
SEQ ID NO:202 is a first determined cDNA sequence for clone #19200.1.
SEQ ID NO:203 is a second determined cDNA sequence for clone #19200.2.
30 SEQ ID NO:204 is the determined cDNA sequence for clone #19202.
SEQ ID NO:205 is a first determined cDNA sequence for clone #19204.1.

WO 02/04514

PCT/US01/22058

16

- SEQ ID NO:206 is a second determined cDNA sequence for clone #19204.2.
SEQ ID NO:207 is the determined cDNA sequence for clone #19205.
SEQ ID NO:208 is a first determined cDNA sequence for clone #19206.1.
SEQ ID NO:209 is a second determined cDNA sequence for clone #19206.2.
5 SEQ ID NO:210 is the determined cDNA sequence for clone #19207.
SEQ ID NO:211 is the determined cDNA sequence for clone #19208.
SEQ ID NO:212 is a first determined cDNA sequence for clone #19211.1.
SEQ ID NO:213 is a second determined cDNA sequence for clone #19211.2.
SEQ ID NO:214 is a first determined cDNA sequence for clone #19214.1.
10 SEQ ID NO:215 is a second determined cDNA sequence for clone #19214.2.
SEQ ID NO:216 is the determined cDNA sequence for clone #19215.
SEQ ID NO:217 is a first determined cDNA sequence for clone #19217. 2.
SEQ ID NO:218 is a second determined cDNA sequence for clone #19217.2.
SEQ ID NO:219 is a first determined cDNA sequence for clone #19218.1.
15 SEQ ID NO:220 is a second determined cDNA sequence for clone #19218.2.
SEQ ID NO:221 is a first determined cDNA sequence for clone #19220.1.
SEQ ID NO:222 is a second determined cDNA sequence for clone #19220.2.
SEQ ID NO:223 is the determined cDNA sequence for clone #22015.
SEQ ID NO:224 is the determined cDNA sequence for clone #22017.
20 SEQ ID NO:225 is the determined cDNA sequence for clone #22019.
SEQ ID NO:226 is the determined cDNA sequence for clone #22020.
SEQ ID NO:227 is the determined cDNA sequence for clone #22023.
SEQ ID NO:228 is the determined cDNA sequence for clone #22026.
SEQ ID NO:229 is the determined cDNA sequence for clone #22027.
25 SEQ ID NO:230 is the determined cDNA sequence for clone #22028.
SEQ ID NO:231 is the determined cDNA sequence for clone #22032.
SEQ ID NO:232 is the determined cDNA sequence for clone #22037.
SEQ ID NO:233 is the determined cDNA sequence for clone #22045.
SEQ ID NO:234 is the determined cDNA sequence for clone #22048.
30 SEQ ID NO:235 is the determined cDNA sequence for clone #22050.
SEQ ID NO:236 is the determined cDNA sequence for clone #22052.

WO 02/04514

PCT/US01/22058

17

- SEQ ID NO:237 is the determined cDNA sequence for clone #22053.
SEQ ID NO:238 is the determined cDNA sequence for clone #22057.
SEQ ID NO:239 is the determined cDNA sequence for clone #22066.
SEQ ID NO:240 is the determined cDNA sequence for clone #22077.
5 SEQ ID NO:241 is the determined cDNA sequence for clone #22085.
SEQ ID NO:242 is the determined cDNA sequence for clone #22105.
SEQ ID NO:243 is the determined cDNA sequence for clone #22108.
SEQ ID NO:244 is the determined cDNA sequence for clone #22109.
SEQ ID NO:245 is the determined cDNA sequence for clone #24842.
10 SEQ ID NO:246 is the determined cDNA sequence for clone #24843.
SEQ ID NO:247 is the determined cDNA sequence for clone #24845.
SEQ ID NO:248 is the determined cDNA sequence for clone #24851.
SEQ ID NO:249 is the determined cDNA sequence for clone #24852.
SEQ ID NO:250 is the determined cDNA sequence for clone #24853.
15 SEQ ID NO:251 is the determined cDNA sequence for clone #24854.
SEQ ID NO:252 is the determined cDNA sequence for clone #24855.
SEQ ID NO:253 is the determined cDNA sequence for clone #24860.
SEQ ID NO:254 is the determined cDNA sequence for clone #24864.
SEQ ID NO:255 is the determined cDNA sequence for clone #24866.
20 SEQ ID NO:256 is the determined cDNA sequence for clone #24867.
SEQ ID NO:257 is the determined cDNA sequence for clone #24868.
SEQ ID NO:258 is the determined cDNA sequence for clone #24869.
SEQ ID NO:259 is the determined cDNA sequence for clone #24870.
SEQ ID NO:260 is the determined cDNA sequence for clone #24872.
25 SEQ ID NO:261 is the determined cDNA sequence for clone #24873.
SEQ ID NO:262 is the determined cDNA sequence for clone #24875.
SEQ ID NO:263 is the determined cDNA sequence for clone #24882.
SEQ ID NO:264 is the determined cDNA sequence for clone #24885.
SEQ ID NO:265 is the determined cDNA sequence for clone #24886.
30 SEQ ID NO:266 is the determined cDNA sequence for clone #24887.
SEQ ID NO:267 is the determined cDNA sequence for clone #24888.

WO 02/04514

PCT/US01/22058

18

- SEQ ID NO:268 is the determined cDNA sequence for clone #24890.
SEQ ID NO:269 is the determined cDNA sequence for clone #24896.
SEQ ID NO:270 is the determined cDNA sequence for clone #24897.
SEQ ID NO:271 is the determined cDNA sequence for clone #24899.
5 SEQ ID NO:272 is the determined cDNA sequence for clone #24901.
SEQ ID NO:273 is the determined cDNA sequence for clone #24902.
SEQ ID NO:274 is the determined cDNA sequence for clone #24906.
SEQ ID NO:275 is the determined cDNA sequence for clone #24912.
SEQ ID NO:276 is the determined cDNA sequence for clone #24913.
10 SEQ ID NO:277 is the determined cDNA sequence for clone #24920.
SEQ ID NO:278 is the determined cDNA sequence for clone #24927.
SEQ ID NO:279 is the determined cDNA sequence for clone #24930.
SEQ ID NO:280 is the determined cDNA sequence for clone #26938.
SEQ ID NO:281 is the determined cDNA sequence for clone #26939.
15 SEQ ID NO:282 is the determined cDNA sequence for clone #26943.
SEQ ID NO:283 is the determined cDNA sequence for clone #26948.
SEQ ID NO:284 is the determined cDNA sequence for clone #26951.
SEQ ID NO:285 is the determined cDNA sequence for clone #26955.
SEQ ID NO:286 is the determined cDNA sequence for clone #26956.
20 SEQ ID NO:287 is the determined cDNA sequence for clone #26959.
SEQ ID NO:288 is the determined cDNA sequence for clone #26961.
SEQ ID NO:289 is the determined cDNA sequence for clone #26962.
SEQ ID NO:290 is the determined cDNA sequence for clone #26964.
SEQ ID NO:291 is the determined cDNA sequence for clone #26966.
25 SEQ ID NO:292 is the determined cDNA sequence for clone #26968.
SEQ ID NO:293 is the determined cDNA sequence for clone #26972.
SEQ ID NO:294 is the determined cDNA sequence for clone #26973.
SEQ ID NO:295 is the determined cDNA sequence for clone #26974.
SEQ ID NO:296 is the determined cDNA sequence for clone #26976.
30 SEQ ID NO:297 is the determined cDNA sequence for clone #26977.
SEQ ID NO:298 is the determined cDNA sequence for clone #26979.

WO 02/04514

PCT/US01/22058

19

- SEQ ID NO:299 is the determined cDNA sequence for clone #26980.
SEQ ID NO:300 is the determined cDNA sequence for clone #26981.
SEQ ID NO:301 is the determined cDNA sequence for clone #26984.
SEQ ID NO:302 is the determined cDNA sequence for clone #26985.
5 SEQ ID NO:303 is the determined cDNA sequence for clone #26986.
SEQ ID NO:304 is the determined cDNA sequence for clone #26993.
SEQ ID NO:305 is the determined cDNA sequence for clone #26994.
SEQ ID NO:306 is the determined cDNA sequence for clone #26995.
SEQ ID NO:307 is the determined cDNA sequence for clone #27003.
10 SEQ ID NO:308 is the determined cDNA sequence for clone #27005.
SEQ ID NO:309 is the determined cDNA sequence for clone #27010.
SEQ ID NO:310 is the determined cDNA sequence for clone #27011.
SEQ ID NO:311 is the determined cDNA sequence for clone #27013.
SEQ ID NO:312 is the determined cDNA sequence for clone #27016.
15 SEQ ID NO:313 is the determined cDNA sequence for clone #27017.
SEQ ID NO:314 is the determined cDNA sequence for clone #27019.
SEQ ID NO:315 is the determined cDNA sequence for clone #27028.
SEQ ID NO:316 is the full-length cDNA sequence for clone #19060.
SEQ ID NO:317 is the full-length cDNA sequence for clone #18964.
20 SEQ ID NO:318 is the full-length cDNA sequence for clone #18929.
SEQ ID NO:319 is the full-length cDNA sequence for clone #18991.
SEQ ID NO:320 is the full-length cDNA sequence for clone #18996.
SEQ ID NO:321 is the full-length cDNA sequence for clone #18966.
SEQ ID NO:322 is the full-length cDNA sequence for clone #18951.
25 SEQ ID NO:323 is the full-length cDNA sequence for clone #18973 (also known as L516S).
SEQ ID NO:324 is the amino acid sequence for clone #19060.
SEQ ID NO:325 is the amino acid sequence for clone #19063.
SEQ ID NO:326 is the amino acid sequence for clone #19077.
30 SEQ ID NO:327 is the amino acid sequence for clone #19110.
SEQ ID NO:328 is the amino acid sequence for clone #19122.

WO 02/04514

PCT/US01/22058

20

- SEQ ID NO:329 is the amino acid sequence for clone #19118.
SEQ ID NO:330 is the amino acid sequence for clone #19080.
SEQ ID NO:331 is the amino acid sequence for clone #19127.
SEQ ID NO:332 is the amino acid sequence for clone #19117.
- 5 SEQ ID NO:333 is the amino acid sequence for clone #19095, also referred to L549S.
SEQ ID NO:334 is the amino acid sequence for clone #18964.
SEQ ID NO:335 is the amino acid sequence for clone #18929.
SEQ ID NO:336 is the amino acid sequence for clone #18991.
SEQ ID NO:337 is the amino acid sequence for clone #18996.
- 10 SEQ ID NO:338 is the amino acid sequence for clone #18966.
SEQ ID NO:339 is the amino acid sequence for clone #18951.
SEQ ID NO:340 is the amino acid sequence for clone #18973.
SEQ ID NO:341 is the determined cDNA sequence for clone 26461.
SEQ ID NO:342 is the determined cDNA sequence for clone 26462.
- 15 SEQ ID NO:343 is the determined cDNA sequence for clone 26463.
SEQ ID NO:344 is the determined cDNA sequence for clone 26464.
SEQ ID NO:345 is the determined cDNA sequence for clone 26465.
SEQ ID NO:346 is the determined cDNA sequence for clone 26466.
SEQ ID NO:347 is the determined cDNA sequence for clone 26467.
- 20 SEQ ID NO:348 is the determined cDNA sequence for clone 26468.
SEQ ID NO:349 is the determined cDNA sequence for clone 26469.
SEQ ID NO:350 is the determined cDNA sequence for clone 26470.
SEQ ID NO:351 is the determined cDNA sequence for clone 26471.
SEQ ID NO:352 is the determined cDNA sequence for clone 26472.
- 25 SEQ ID NO:353 is the determined cDNA sequence for clone 26474.
SEQ ID NO:354 is the determined cDNA sequence for clone 26475.
SEQ ID NO:355 is the determined cDNA sequence for clone 26476.
SEQ ID NO:356 is the determined cDNA sequence for clone 26477.
SEQ ID NO:357 is the determined cDNA sequence for clone 26478.
- 30 SEQ ID NO:358 is the determined cDNA sequence for clone 26479.
SEQ ID NO:359 is the determined cDNA sequence for clone 26480.

WO 02/04514

PCT/US01/22058

21

- SEQ ID NO:360 is the determined cDNA sequence for clone 26481.
SEQ ID NO:361 is the determined cDNA sequence for clone 26482.
SEQ ID NO:362 is the determined cDNA sequence for clone 26483.
SEQ ID NO:363 is the determined cDNA sequence for clone 26484.
5 SEQ ID NO:364 is the determined cDNA sequence for clone 26485.
SEQ ID NO:365 is the determined cDNA sequence for clone 26486.
SEQ ID NO:366 is the determined cDNA sequence for clone 26487.
SEQ ID NO:367 is the determined cDNA sequence for clone 26488.
SEQ ID NO:368 is the determined cDNA sequence for clone 26489.
10 SEQ ID NO:369 is the determined cDNA sequence for clone 26490.
SEQ ID NO:370 is the determined cDNA sequence for clone 26491.
SEQ ID NO:371 is the determined cDNA sequence for clone 26492.
SEQ ID NO:372 is the determined cDNA sequence for clone 26493.
SEQ ID NO:373 is the determined cDNA sequence for clone 26494.
15 SEQ ID NO:374 is the determined cDNA sequence for clone 26495.
SEQ ID NO:375 is the determined cDNA sequence for clone 26496.
SEQ ID NO:376 is the determined cDNA sequence for clone 26497.
SEQ ID NO:377 is the determined cDNA sequence for clone 26498.
SEQ ID NO:378 is the determined cDNA sequence for clone 26499.
20 SEQ ID NO:379 is the determined cDNA sequence for clone 26500.
SEQ ID NO:380 is the determined cDNA sequence for clone 26501.
SEQ ID NO:381 is the determined cDNA sequence for clone 26502.
SEQ ID NO:382 is the determined cDNA sequence for clone 26503.
SEQ ID NO:383 is the determined cDNA sequence for clone 26504.
25 SEQ ID NO:384 is the determined cDNA sequence for clone 26505.
SEQ ID NO:385 is the determined cDNA sequence for clone 26506.
SEQ ID NO:386 is the determined cDNA sequence for clone 26507.
SEQ ID NO:387 is the determined cDNA sequence for clone 26508.
SEQ ID NO:388 is the determined cDNA sequence for clone 26509.
30 SEQ ID NO:389 is the determined cDNA sequence for clone 26511.
SEQ ID NO:390 is the determined cDNA sequence for clone 26513.

WO 02/04514

PCT/US01/22058

22

- SEQ ID NO:391 is the determined cDNA sequence for clone 26514.
SEQ ID NO:392 is the determined cDNA sequence for clone 26515.
SEQ ID NO:393 is the determined cDNA sequence for clone 26516.
SEQ ID NO:394 is the determined cDNA sequence for clone 26517.
5 SEQ ID NO:395 is the determined cDNA sequence for clone 26518.
SEQ ID NO:396 is the determined cDNA sequence for clone 26519.
SEQ ID NO:397 is the determined cDNA sequence for clone 26520.
SEQ ID NO:398 is the determined cDNA sequence for clone 26521.
SEQ ID NO:399 is the determined cDNA sequence for clone 26522.
10 SEQ ID NO:400 is the determined cDNA sequence for clone 26523.
SEQ ID NO:401 is the determined cDNA sequence for clone 26524.
SEQ ID NO:402 is the determined cDNA sequence for clone 26526.
SEQ ID NO:403 is the determined cDNA sequence for clone 26527.
SEQ ID NO:404 is the determined cDNA sequence for clone 26528.
15 SEQ ID NO:405 is the determined cDNA sequence for clone 26529.
SEQ ID NO:406 is the determined cDNA sequence for clone 26530.
SEQ ID NO:407 is the determined cDNA sequence for clone 26532.
SEQ ID NO:408 is the determined cDNA sequence for clone 26533.
SEQ ID NO:409 is the determined cDNA sequence for clone 26534.
20 SEQ ID NO:410 is the determined cDNA sequence for clone 26535.
SEQ ID NO:411 is the determined cDNA sequence for clone 26536.
SEQ ID NO:412 is the determined cDNA sequence for clone 26537.
SEQ ID NO:413 is the determined cDNA sequence for clone 26538.
SEQ ID NO:414 is the determined cDNA sequence for clone 26540.
25 SEQ ID NO:415 is the determined cDNA sequence for clone 26541.
SEQ ID NO:416 is the determined cDNA sequence for clone 26542.
SEQ ID NO:417 is the determined cDNA sequence for clone 26543.
SEQ ID NO:418 is the determined cDNA sequence for clone 26544.
SEQ ID NO:419 is the determined cDNA sequence for clone 26546.
30 SEQ ID NO:420 is the determined cDNA sequence for clone 26547.
SEQ ID NO:421 is the determined cDNA sequence for clone 26548.

WO 02/04514

PCT/US01/22058

23

- SEQ ID NO:422 is the determined cDNA sequence for clone 26549.
SEQ ID NO:423 is the determined cDNA sequence for clone 26550.
SEQ ID NO:424 is the determined cDNA sequence for clone 26551.
SEQ ID NO:425 is the determined cDNA sequence for clone 26552.
5 SEQ ID NO:426 is the determined cDNA sequence for clone 26553.
SEQ ID NO:427 is the determined cDNA sequence for clone 26554.
SEQ ID NO:428 is the determined cDNA sequence for clone 26556.
SEQ ID NO:429 is the determined cDNA sequence for clone 26557.
SEQ ID NO:430 is the determined cDNA sequence for clone 27631.
10 SEQ ID NO:431 is the determined cDNA sequence for clone 27632.
SEQ ID NO:432 is the determined cDNA sequence for clone 27633.
SEQ ID NO:433 is the determined cDNA sequence for clone 27635.
SEQ ID NO:434 is the determined cDNA sequence for clone 27636.
SEQ ID NO:435 is the determined cDNA sequence for clone 27637.
15 SEQ ID NO:436 is the determined cDNA sequence for clone 27638.
SEQ ID NO:437 is the determined cDNA sequence for clone 27639.
SEQ ID NO:438 is the determined cDNA sequence for clone 27640.
SEQ ID NO:439 is the determined cDNA sequence for clone 27641.
SEQ ID NO:440 is the determined cDNA sequence for clone 27642.
20 SEQ ID NO:441 is the determined cDNA sequence for clone 27644.
SEQ ID NO:442 is the determined cDNA sequence for clone 27646.
SEQ ID NO:443 is the determined cDNA sequence for clone 27647.
SEQ ID NO:444 is the determined cDNA sequence for clone 27649.
SEQ ID NO:445 is the determined cDNA sequence for clone 27650.
25 SEQ ID NO:446 is the determined cDNA sequence for clone 27651.
SEQ ID NO:447 is the determined cDNA sequence for clone 27652.
SEQ ID NO:448 is the determined cDNA sequence for clone 27654.
SEQ ID NO:449 is the determined cDNA sequence for clone 27655.
SEQ ID NO:450 is the determined cDNA sequence for clone 27657.
30 SEQ ID NO:451 is the determined cDNA sequence for clone 27659.
SEQ ID NO:452 is the determined cDNA sequence for clone 27665.

WO 02/04514

PCT/US01/22058

24

- SEQ ID NO:453 is the determined cDNA sequence for clone 27666.
SEQ ID NO:454 is the determined cDNA sequence for clone 27668.
SEQ ID NO:455 is the determined cDNA sequence for clone 27670.
SEQ ID NO:456 is the determined cDNA sequence for clone 27671.
5 SEQ ID NO:457 is the determined cDNA sequence for clone 27672.
SEQ ID NO:458 is the determined cDNA sequence for clone 27674.
SEQ ID NO:459 is the determined cDNA sequence for clone 27677.
SEQ ID NO:460 is the determined cDNA sequence for clone 27681.
SEQ ID NO:461 is the determined cDNA sequence for clone 27682.
10 SEQ ID NO:462 is the determined cDNA sequence for clone 27683.
SEQ ID NO:463 is the determined cDNA sequence for clone 27686.
SEQ ID NO:464 is the determined cDNA sequence for clone 27688.
SEQ ID NO:465 is the determined cDNA sequence for clone 27689.
SEQ ID NO:466 is the determined cDNA sequence for clone 27690.
15 SEQ ID NO:467 is the determined cDNA sequence for clone 27693.
SEQ ID NO:468 is the determined cDNA sequence for clone 27699.
SEQ ID NO:469 is the determined cDNA sequence for clone 27700.
SEQ ID NO:470 is the determined cDNA sequence for clone 27702.
SEQ ID NO:471 is the determined cDNA sequence for clone 27705.
20 SEQ ID NO:472 is the determined cDNA sequence for clone 27706.
SEQ ID NO:473 is the determined cDNA sequence for clone 27707.
SEQ ID NO:474 is the determined cDNA sequence for clone 27708.
SEQ ID NO:475 is the determined cDNA sequence for clone 27709.
SEQ ID NO:476 is the determined cDNA sequence for clone 27710.
25 SEQ ID NO:477 is the determined cDNA sequence for clone 27711.
SEQ ID NO:478 is the determined cDNA sequence for clone 27712.
SEQ ID NO:479 is the determined cDNA sequence for clone 27713.
SEQ ID NO:480 is the determined cDNA sequence for clone 27714.
SEQ ID NO:481 is the determined cDNA sequence for clone 27715.
30 SEQ ID NO:482 is the determined cDNA sequence for clone 27716.
SEQ ID NO:483 is the determined cDNA sequence for clone 27717.

WO 02/04514

PCT/US01/22058

25

- SEQ ID NO:484 is the determined cDNA sequence for clone 27718.
SEQ ID NO:485 is the determined cDNA sequence for clone 27719.
SEQ ID NO:486 is the determined cDNA sequence for clone 27720.
SEQ ID NO:487 is the determined cDNA sequence for clone 27722.
5 SEQ ID NO:488 is the determined cDNA sequence for clone 27723.
SEQ ID NO:489 is the determined cDNA sequence for clone 27724.
SEQ ID NO:490 is the determined cDNA sequence for clone 27726.
SEQ ID NO:491 is the determined cDNA sequence for clone 25015.
SEQ ID NO:492 is the determined cDNA sequence for clone 25016.
10 SEQ ID NO:493 is the determined cDNA sequence for clone 25017.
SEQ ID NO:494 is the determined cDNA sequence for clone 25018.
SEQ ID NO:495 is the determined cDNA sequence for clone 25030.
SEQ ID NO:496 is the determined cDNA sequence for clone 25033.
SEQ ID NO:497 is the determined cDNA sequence for clone 25034.
15 SEQ ID NO:498 is the determined cDNA sequence for clone 25035.
SEQ ID NO:499 is the determined cDNA sequence for clone 25036.
SEQ ID NO:500 is the determined cDNA sequence for clone 25037.
SEQ ID NO:501 is the determined cDNA sequence for clone 25038.
SEQ ID NO:502 is the determined cDNA sequence for clone 25039.
20 SEQ ID NO:503 is the determined cDNA sequence for clone 25040.
SEQ ID NO:504 is the determined cDNA sequence for clone 25042.
SEQ ID NO:505 is the determined cDNA sequence for clone 25043.
SEQ ID NO:506 is the determined cDNA sequence for clone 25044.
SEQ ID NO:507 is the determined cDNA sequence for clone 25045.
25 SEQ ID NO:508 is the determined cDNA sequence for clone 25047.
SEQ ID NO:509 is the determined cDNA sequence for clone 25048.
SEQ ID NO:510 is the determined cDNA sequence for clone 25049.
SEQ ID NO:511 is the determined cDNA sequence for clone 25185.
SEQ ID NO:512 is the determined cDNA sequence for clone 25186.
30 SEQ ID NO:513 is the determined cDNA sequence for clone 25187.
SEQ ID NO:514 is the determined cDNA sequence for clone 25188.

WO 02/04514

PCT/US01/22058

26

- SEQ ID NO:515 is the determined cDNA sequence for clone 25189.
SEQ ID NO:516 is the determined cDNA sequence for clone 25190.
SEQ ID NO:517 is the determined cDNA sequence for clone 25193.
SEQ ID NO:518 is the determined cDNA sequence for clone 25194.
5 SEQ ID NO:519 is the determined cDNA sequence for clone 25196.
SEQ ID NO:520 is the determined cDNA sequence for clone 25198.
SEQ ID NO:521 is the determined cDNA sequence for clone 25199.
SEQ ID NO:522 is the determined cDNA sequence for clone 25200.
SEQ ID NO:523 is the determined cDNA sequence for clone 25202.
10 SEQ ID NO:524 is the determined cDNA sequence for clone 25364.
SEQ ID NO:525 is the determined cDNA sequence for clone 25366.
SEQ ID NO:526 is the determined cDNA sequence for clone 25367.
SEQ ID NO:527 is the determined cDNA sequence for clone 25368.
SEQ ID NO:528 is the determined cDNA sequence for clone 25369.
15 SEQ ID NO:529 is the determined cDNA sequence for clone 25370.
SEQ ID NO:530 is the determined cDNA sequence for clone 25371.
SEQ ID NO:531 is the determined cDNA sequence for clone 25372.
SEQ ID NO:532 is the determined cDNA sequence for clone 25373.
SEQ ID NO:533 is the determined cDNA sequence for clone 25374.
20 SEQ ID NO:534 is the determined cDNA sequence for clone 25376.
SEQ ID NO:535 is the determined cDNA sequence for clone 25377.
SEQ ID NO:536 is the determined cDNA sequence for clone 25378.
SEQ ID NO:537 is the determined cDNA sequence for clone 25379.
SEQ ID NO:538 is the determined cDNA sequence for clone 25380.
25 SEQ ID NO:539 is the determined cDNA sequence for clone 25381.
SEQ ID NO:540 is the determined cDNA sequence for clone 25382.
SEQ ID NO:541 is the determined cDNA sequence for clone 25383.
SEQ ID NO:542 is the determined cDNA sequence for clone 25385.
SEQ ID NO:543 is the determined cDNA sequence for clone 25386.
30 SEQ ID NO:544 is the determined cDNA sequence for clone 25387.
SEQ ID NO:545 is the determined cDNA sequence for clone 26013.

WO 02/04514

PCT/US01/22058

27

- SEQ ID NO:546 is the determined cDNA sequence for clone 26014.
SEQ ID NO:547 is the determined cDNA sequence for clone 26016.
SEQ ID NO:548 is the determined cDNA sequence for clone 26017.
SEQ ID NO:549 is the determined cDNA sequence for clone 26018.
5 SEQ ID NO:550 is the determined cDNA sequence for clone 26019.
SEQ ID NO:551 is the determined cDNA sequence for clone 26020.
SEQ ID NO:552 is the determined cDNA sequence for clone 26021.
SEQ ID NO:553 is the determined cDNA sequence for clone 26022.
SEQ ID NO:554 is the determined cDNA sequence for clone 26027.
10 SEQ ID NO:555 is the determined cDNA sequence for clone 26197.
SEQ ID NO:556 is the determined cDNA sequence for clone 26199.
SEQ ID NO:557 is the determined cDNA sequence for clone 26201.
SEQ ID NO:558 is the determined cDNA sequence for clone 26202.
SEQ ID NO:559 is the determined cDNA sequence for clone 26203.
15 SEQ ID NO:560 is the determined cDNA sequence for clone 26204.
SEQ ID NO:561 is the determined cDNA sequence for clone 26205.
SEQ ID NO:562 is the determined cDNA sequence for clone 26206.
SEQ ID NO:563 is the determined cDNA sequence for clone 26208.
SEQ ID NO:564 is the determined cDNA sequence for clone 26211.
20 SEQ ID NO:565 is the determined cDNA sequence for clone 26212.
SEQ ID NO:566 is the determined cDNA sequence for clone 26213.
SEQ ID NO:567 is the determined cDNA sequence for clone 26214.
SEQ ID NO:568 is the determined cDNA sequence for clone 26215.
SEQ ID NO:569 is the determined cDNA sequence for clone 26216.
25 SEQ ID NO:570 is the determined cDNA sequence for clone 26217.
SEQ ID NO:571 is the determined cDNA sequence for clone 26218.
SEQ ID NO:572 is the determined cDNA sequence for clone 26219.
SEQ ID NO:573 is the determined cDNA sequence for clone 26220.
SEQ ID NO:574 is the determined cDNA sequence for clone 26221.
30 SEQ ID NO:575 is the determined cDNA sequence for clone 26224.
SEQ ID NO:576 is the determined cDNA sequence for clone 26225.

WO 02/04514

PCT/US01/22058

28

- SEQ ID NO:577 is the determined cDNA sequence for clone 26226.
SEQ ID NO:578 is the determined cDNA sequence for clone 26227.
SEQ ID NO:579 is the determined cDNA sequence for clone 26228.
SEQ ID NO:580 is the determined cDNA sequence for clone 26230.
5 SEQ ID NO:581 is the determined cDNA sequence for clone 26231.
SEQ ID NO:582 is the determined cDNA sequence for clone 26234.
SEQ ID NO:583 is the determined cDNA sequence for clone 26236.
SEQ ID NO:584 is the determined cDNA sequence for clone 26237.
SEQ ID NO:585 is the determined cDNA sequence for clone 26239.
10 SEQ ID NO:586 is the determined cDNA sequence for clone 26240.
SEQ ID NO:587 is the determined cDNA sequence for clone 26241.
SEQ ID NO:588 is the determined cDNA sequence for clone 26242.
SEQ ID NO:589 is the determined cDNA sequence for clone 26246.
SEQ ID NO:590 is the determined cDNA sequence for clone 26247.
15 SEQ ID NO:591 is the determined cDNA sequence for clone 26248.
SEQ ID NO:592 is the determined cDNA sequence for clone 26249.
SEQ ID NO:593 is the determined cDNA sequence for clone 26250.
SEQ ID NO:594 is the determined cDNA sequence for clone 26251.
SEQ ID NO:595 is the determined cDNA sequence for clone 26252.
20 SEQ ID NO:596 is the determined cDNA sequence for clone 26253.
SEQ ID NO:597 is the determined cDNA sequence for clone 26254.
SEQ ID NO:598 is the determined cDNA sequence for clone 26255.
SEQ ID NO:599 is the determined cDNA sequence for clone 26256.
SEQ ID NO:600 is the determined cDNA sequence for clone 26257.
25 SEQ ID NO:601 is the determined cDNA sequence for clone 26259.
SEQ ID NO:602 is the determined cDNA sequence for clone 26260.
SEQ ID NO:603 is the determined cDNA sequence for clone 26261.
SEQ ID NO:604 is the determined cDNA sequence for clone 26262.
SEQ ID NO:605 is the determined cDNA sequence for clone 26263.
30 SEQ ID NO:606 is the determined cDNA sequence for clone 26264.
SEQ ID NO:607 is the determined cDNA sequence for clone 26265.

WO 02/04514

PCT/US01/22058

29

- SEQ ID NO:608 is the determined cDNA sequence for clone 26266.
SEQ ID NO:609 is the determined cDNA sequence for clone 26268.
SEQ ID NO:610 is the determined cDNA sequence for clone 26269.
SEQ ID NO:611 is the determined cDNA sequence for clone 26271.
5 SEQ ID NO:612 is the determined cDNA sequence for clone 26273.
SEQ ID NO:613 is the determined cDNA sequence for clone 26810.
SEQ ID NO:614 is the determined cDNA sequence for clone 26811.
SEQ ID NO:615 is the determined cDNA sequence for clone 26812.1.
SEQ ID NO:616 is the determined cDNA sequence for clone 26812.2.
10 SEQ ID NO:617 is the determined cDNA sequence for clone 26813.
SEQ ID NO:618 is the determined cDNA sequence for clone 26814.
SEQ ID NO:619 is the determined cDNA sequence for clone 26815.
SEQ ID NO:620 is the determined cDNA sequence for clone 26816.
SEQ ID NO:621 is the determined cDNA sequence for clone 26818.
15 SEQ ID NO:622 is the determined cDNA sequence for clone 26819.
SEQ ID NO:623 is the determined cDNA sequence for clone 26820.
SEQ ID NO:624 is the determined cDNA sequence for clone 26821.
SEQ ID NO:625 is the determined cDNA sequence for clone 26822.
SEQ ID NO:626 is the determined cDNA sequence for clone 26824.
20 SEQ ID NO:627 is the determined cDNA sequence for clone 26825.
SEQ ID NO:628 is the determined cDNA sequence for clone 26826.
SEQ ID NO:629 is the determined cDNA sequence for clone 26827.
SEQ ID NO:630 is the determined cDNA sequence for clone 26829.
SEQ ID NO:631 is the determined cDNA sequence for clone 26830.
25 SEQ ID NO:632 is the determined cDNA sequence for clone 26831.
SEQ ID NO:633 is the determined cDNA sequence for clone 26832.
SEQ ID NO:634 is the determined cDNA sequence for clone 26835.
SEQ ID NO:635 is the determined cDNA sequence for clone 26836.
SEQ ID NO:636 is the determined cDNA sequence for clone 26837.
30 SEQ ID NO:637 is the determined cDNA sequence for clone 26839.
SEQ ID NO:638 is the determined cDNA sequence for clone 26841.

WO 02/04514

PCT/US01/22058

30

- SEQ ID NO:639 is the determined cDNA sequence for clone 26843.
SEQ ID NO:640 is the determined cDNA sequence for clone 26844.
SEQ ID NO:641 is the determined cDNA sequence for clone 26845.
SEQ ID NO:642 is the determined cDNA sequence for clone 26846.
5 SEQ ID NO:643 is the determined cDNA sequence for clone 26847.
SEQ ID NO:644 is the determined cDNA sequence for clone 26848.
SEQ ID NO:645 is the determined cDNA sequence for clone 26849.
SEQ ID NO:646 is the determined cDNA sequence for clone 26850.
SEQ ID NO:647 is the determined cDNA sequence for clone 26851.
10 SEQ ID NO:648 is the determined cDNA sequence for clone 26852.
SEQ ID NO:649 is the determined cDNA sequence for clone 26853.
SEQ ID NO:650 is the determined cDNA sequence for clone 26854.
SEQ ID NO:651 is the determined cDNA sequence for clone 26856.
SEQ ID NO:652 is the determined cDNA sequence for clone 26857.
15 SEQ ID NO:653 is the determined cDNA sequence for clone 26858.
SEQ ID NO:654 is the determined cDNA sequence for clone 26859.
SEQ ID NO:655 is the determined cDNA sequence for clone 26860.
SEQ ID NO:656 is the determined cDNA sequence for clone 26862.
SEQ ID NO:657 is the determined cDNA sequence for clone 26863.
20 SEQ ID NO:658 is the determined cDNA sequence for clone 26864.
SEQ ID NO:659 is the determined cDNA sequence for clone 26865.
SEQ ID NO:660 is the determined cDNA sequence for clone 26867.
SEQ ID NO:661 is the determined cDNA sequence for clone 26868.
SEQ ID NO:662 is the determined cDNA sequence for clone 26871.
25 SEQ ID NO:663 is the determined cDNA sequence for clone 26873.
SEQ ID NO:664 is the determined cDNA sequence for clone 26875.
SEQ ID NO:665 is the determined cDNA sequence for clone 26876.
SEQ ID NO:666 is the determined cDNA sequence for clone 26877.
SEQ ID NO:667 is the determined cDNA sequence for clone 26878.
30 SEQ ID NO:668 is the determined cDNA sequence for clone 26880.
SEQ ID NO:669 is the determined cDNA sequence for clone 26882.

WO 02/04514

PCT/US01/22058

31

- SEQ ID NO:670 is the determined cDNA sequence for clone 26883.
SEQ ID NO:671 is the determined cDNA sequence for clone 26884.
SEQ ID NO:672 is the determined cDNA sequence for clone 26885.
SEQ ID NO:673 is the determined cDNA sequence for clone 26886.
5 SEQ ID NO:674 is the determined cDNA sequence for clone 26887.
SEQ ID NO:675 is the determined cDNA sequence for clone 26888.
SEQ ID NO:676 is the determined cDNA sequence for clone 26889.
SEQ ID NO:677 is the determined cDNA sequence for clone 26890.
SEQ ID NO:678 is the determined cDNA sequence for clone 26892.
10 SEQ ID NO:679 is the determined cDNA sequence for clone 26894.
SEQ ID NO:680 is the determined cDNA sequence for clone 26895.
SEQ ID NO:681 is the determined cDNA sequence for clone 26897.
SEQ ID NO:682 is the determined cDNA sequence for clone 26898.
SEQ ID NO:683 is the determined cDNA sequence for clone 26899.
15 SEQ ID NO:684 is the determined cDNA sequence for clone 26900.
SEQ ID NO:685 is the determined cDNA sequence for clone 26901.
SEQ ID NO:686 is the determined cDNA sequence for clone 26903.
SEQ ID NO:687 is the determined cDNA sequence for clone 26905.
SEQ ID NO:688 is the determined cDNA sequence for clone 26906.
20 SEQ ID NO:689 is the determined cDNA sequence for clone 26708.
SEQ ID NO:690 is the determined cDNA sequence for clone 26709.
SEQ ID NO:691 is the determined cDNA sequence for clone 26710.
SEQ ID NO:692 is the determined cDNA sequence for clone 26711.
SEQ ID NO:693 is the determined cDNA sequence for clone 26712.
25 SEQ ID NO:694 is the determined cDNA sequence for clone 26713.
SEQ ID NO:695 is the determined cDNA sequence for clone 26714.
SEQ ID NO:696 is the determined cDNA sequence for clone 26715.
SEQ ID NO:697 is the determined cDNA sequence for clone 26716.
SEQ ID NO:698 is the determined cDNA sequence for clone 26717.
30 SEQ ID NO:699 is the determined cDNA sequence for clone 26718.
SEQ ID NO:700 is the determined cDNA sequence for clone 26719.

WO 02/04514

PCT/US01/22058

32

- SEQ ID NO:701 is the determined cDNA sequence for clone 26720.
SEQ ID NO:702 is the determined cDNA sequence for clone 26721.
SEQ ID NO:703 is the determined cDNA sequence for clone 26722.
SEQ ID NO:704 is the determined cDNA sequence for clone 26723.
5 SEQ ID NO:705 is the determined cDNA sequence for clone 26724.
SEQ ID NO:706 is the determined cDNA sequence for clone 26725.
SEQ ID NO:707 is the determined cDNA sequence for clone 26726.
SEQ ID NO:708 is the determined cDNA sequence for clone 26727.
SEQ ID NO:709 is the determined cDNA sequence for clone 26728.
10 SEQ ID NO:710 is the determined cDNA sequence for clone 26729.
SEQ ID NO:711 is the determined cDNA sequence for clone 26730.
SEQ ID NO:712 is the determined cDNA sequence for clone 26731.
SEQ ID NO:713 is the determined cDNA sequence for clone 26732.
SEQ ID NO:714 is the determined cDNA sequence for clone 26733.1.
15 SEQ ID NO:715 is the determined cDNA sequence for clone 26733.2.
SEQ ID NO:716 is the determined cDNA sequence for clone 26734.
SEQ ID NO:717 is the determined cDNA sequence for clone 26735.
SEQ ID NO:718 is the determined cDNA sequence for clone 26736.
SEQ ID NO:719 is the determined cDNA sequence for clone 26737.
20 SEQ ID NO:720 is the determined cDNA sequence for clone 26738.
SEQ ID NO:721 is the determined cDNA sequence for clone 26739.
SEQ ID NO:722 is the determined cDNA sequence for clone 26741.
SEQ ID NO:723 is the determined cDNA sequence for clone 26742.
SEQ ID NO:724 is the determined cDNA sequence for clone 26743.
25 SEQ ID NO:725 is the determined cDNA sequence for clone 26744.
SEQ ID NO:726 is the determined cDNA sequence for clone 26745.
SEQ ID NO:727 is the determined cDNA sequence for clone 26746.
SEQ ID NO:728 is the determined cDNA sequence for clone 26747.
SEQ ID NO:729 is the determined cDNA sequence for clone 26748.
30 SEQ ID NO:730 is the determined cDNA sequence for clone 26749.
SEQ ID NO:731 is the determined cDNA sequence for clone 26750.

WO 02/04514

PCT/US01/22058

33

- SEQ ID NO:732 is the determined cDNA sequence for clone 26751.
SEQ ID NO:733 is the determined cDNA sequence for clone 26752.
SEQ ID NO:734 is the determined cDNA sequence for clone 26753.
SEQ ID NO:735 is the determined cDNA sequence for clone 26754.
5 SEQ ID NO:736 is the determined cDNA sequence for clone 26755.
SEQ ID NO:737 is the determined cDNA sequence for clone 26756.
SEQ ID NO:738 is the determined cDNA sequence for clone 26757.
SEQ ID NO:739 is the determined cDNA sequence for clone 26758.
SEQ ID NO:740 is the determined cDNA sequence for clone 26759.
10 SEQ ID NO:741 is the determined cDNA sequence for clone 26760.
SEQ ID NO:742 is the determined cDNA sequence for clone 26761.
SEQ ID NO:743 is the determined cDNA sequence for clone 26762.
SEQ ID NO:744 is the determined cDNA sequence for clone 26763.
SEQ ID NO:745 is the determined cDNA sequence for clone 26764.
15 SEQ ID NO:746 is the determined cDNA sequence for clone 26765.
SEQ ID NO:747 is the determined cDNA sequence for clone 26766.
SEQ ID NO:748 is the determined cDNA sequence for clone 26767.
SEQ ID NO:749 is the determined cDNA sequence for clone 26768.
SEQ ID NO:750 is the determined cDNA sequence for clone 26769.
20 SEQ ID NO:751 is the determined cDNA sequence for clone 26770.
SEQ ID NO:752 is the determined cDNA sequence for clone 26771.
SEQ ID NO:753 is the determined cDNA sequence for clone 26772.
SEQ ID NO:754 is the determined cDNA sequence for clone 26773.
SEQ ID NO:755 is the determined cDNA sequence for clone 26774.
25 SEQ ID NO:756 is the determined cDNA sequence for clone 26775.
SEQ ID NO:757 is the determined cDNA sequence for clone 26776.
SEQ ID NO:758 is the determined cDNA sequence for clone 26777.
SEQ ID NO:759 is the determined cDNA sequence for clone 26778.
SEQ ID NO:760 is the determined cDNA sequence for clone 26779.
30 SEQ ID NO:761 is the determined cDNA sequence for clone 26781.
SEQ ID NO:762 is the determined cDNA sequence for clone 26782.

WO 02/04514

PCT/US01/22058

34

- SEQ ID NO:763 is the determined cDNA sequence for clone 26783.
SEQ ID NO:764 is the determined cDNA sequence for clone 26784.
SEQ ID NO:765 is the determined cDNA sequence for clone 26785.
SEQ ID NO:766 is the determined cDNA sequence for clone 26786.
5 SEQ ID NO:767 is the determined cDNA sequence for clone 26787.
SEQ ID NO:768 is the determined cDNA sequence for clone 26788.
SEQ ID NO:769 is the determined cDNA sequence for clone 26790.
SEQ ID NO:770 is the determined cDNA sequence for clone 26791.
SEQ ID NO:771 is the determined cDNA sequence for clone 26792.
10 SEQ ID NO:772 is the determined cDNA sequence for clone 26793.
SEQ ID NO:773 is the determined cDNA sequence for clone 26794.
SEQ ID NO:774 is the determined cDNA sequence for clone 26795.
SEQ ID NO:775 is the determined cDNA sequence for clone 26796.
SEQ ID NO:776 is the determined cDNA sequence for clone 26797.
15 SEQ ID NO:777 is the determined cDNA sequence for clone 26798.
SEQ ID NO:778 is the determined cDNA sequence for clone 26800.
SEQ ID NO:779 is the determined cDNA sequence for clone 26801.
SEQ ID NO:780 is the determined cDNA sequence for clone 26802.
SEQ ID NO:781 is the determined cDNA sequence for clone 26803.
20 SEQ ID NO:782 is the determined cDNA sequence for clone 26804.
SEQ ID NO:783 is the amino acid sequence for L773P.
SEQ ID NO:784 is the determined DNA sequence of the L773P expression construct.
SEQ ID NO:785 is the determined DNA sequence of the L773PA expression construct.
SEQ ID NO:786 is a predicted amino acid sequence for L552S.
25 SEQ ID NO:787 is a predicted amino acid sequence for L840P.
SEQ ID NO:788 is the full-length cDNA sequence for L548S.
SEQ ID NO:789 is the amino acid sequence encoded by SEQ ID NO:788.
SEQ ID NO:790 is an extended cDNA sequence for L552S.
SEQ ID NO:791 is the predicted amino acid sequence encoded by the cDNA sequence
30 of SEQ ID NO:790.
SEQ ID NO:792 is the determined cDNA sequence for an isoform of L552S.

WO 02/04514

PCT/US01/22058

35

- SEQ ID NO:793 is the predicted amino acid sequence encoded by SEQ ID NO:792.
SEQ ID NO:794 is an extended cDNA sequence for L840P.
SEQ ID NO:795 is the predicted amino acid sequence encoded by SEQ ID NO:794.
SEQ ID NO:796 is an extended cDNA sequence for L801P.
- 5 SEQ ID NO:797 is a first predicted amino acid sequence encoded by SEQ ID NO:796.
SEQ ID NO:798 is a second predicted amino acid sequence encoded by SEQ ID NO:796.
SEQ ID NO:799 is a third predicted amino acid sequence encoded by SEQ ID NO:796.
SEQ ID NO:800 is the determined full-length sequence for L844P.
- 10 SEQ ID NO:801 is the 5' consensus cDNA sequence for L551S.
SEQ ID NO:802 is the 3' consensus cDNA sequence for L551S.
SEQ ID NO:803 is the cDNA sequence for STY8.
SEQ ID NO:804 is an extended cDNA sequence for L551S.
SEQ ID NO:805 is the amino acid sequence for STY8.
- 15 SEQ ID NO:806 is the extended amino acid sequence for L551S.
SEQ ID NO:807 is the determined full length cDNA sequence for L773P.
SEQ ID NO:808 is the full-length cDNA sequence of L552S.
SEQ ID NO:809 is the full-length amino acid sequence of L552S.
SEQ ID NO:810 is the determined cDNA sequence of clone 50989.
- 20 SEQ ID NO:811 is the determined cDNA sequence of clone 50990.
SEQ ID NO:812 is the determined cDNA sequence of clone 50992.
SEQ ID NO:813-824 are the determined cDNA sequences for clones isolated from lung tumor tissue.
SEQ ID NO:825 is the determined cDNA sequence for the full-length L551S clone 54305.
- 25 SEQ ID NO:826 is the determined cDNA sequence for the full-length L551S clone 54298.
SEQ ID NO:827 is the full-length amino acid sequence for L551S.

WO 02/04514

PCT/US01/22058

36

Tables 1-6 contain the sequence identifiers for SEQ ID NO:828-1664.

Table 1A

5

SEQ ID NO:	CLONE IDENTIFIER	SEQ ID NO:	CLONE IDENTIFIER
828	R0126:A02	869	R0126:D12
829	R0126:A03	870	R0126:E01
830	R0126:A05	871	R0126:E02
831	R0126:A06	872	R0126:E03
832	R0126:A08	873	R0126:E04
833	R0126:A09	874	R0126:E05
834	R0126:A10	875	R0126:E06
835	R0126:A11	876	R0126:E07
836	R0126:A12	877	R0126:E08
837	R0126:B01	878	R0126:E09
838	R0126:B03	879	R0126:E10
839	R0126:B04	880	R0126:E11
840	R0126:B05	881	R0126:E12
841	R0126:B06	882	R0126:F01
842	R0126:B07	883	R0126:F02
843	R0126:B08	884	R0126:F03
844	R0126:B09	885	R0126:F04
845	R0126:B11	886	R0126:F05
846	R0126:B12	887	R0126:F06
847	R0126:C01	888	R0126:F07
848	R0126:C02	889	R0126:F08
849	R0126:C03	890	R0126:F10
850	R0126:C05	891	R0126:F11
851	R0126:C06	892	R0126:F12
852	R0126:C07	893	R0126:G01
853	R0126:C08	894	R0126:G02
854	R0126:C09	895	R0126:G03
855	R0126:C10	896	R0126:G04
856	R0126:C11	897	R0126:G05
857	R0126:C12	898	R0126:G06
858	R0126:D01	899	R0126:G07
859	R0126:D02	900	R0126:G09
860	R0126:D03	901	R0126:G10
861	R0126:D04	902	R0126:G11
862	R0126:D05	903	R0126:G12
863	R0126:D06	904	R0126:H01
864	R0126:D07	905	R0126:H02

WO 02/04514

PCT/US01/22058

37

SEQ ID NO:	CLONE IDENTIFIER	SEQ ID NO:	CLONE IDENTIFIER
865	R0126:D08	906	R0126:H03
866	R0126:D09	907	R0126:H04
867	R0126:D10	908	R0126:H05
868	R0126:D11	909	R0126:H06

Table 1B

SEQ ID NO	CLONE IDENTIFIER	SEQ ID NO	CLONE IDENTIFIER
910	R0126:H07	951	R0127:D10
911	R0126:H09	952	R0127:D11
912	R0126:H10	953	R0127:D12
913	R0126:H11	954	R0127:E02
914	R0127:A02	955	R0127:E03
915	R0127:A05	956	R0127:E04
916	R0127:A06	957	R0127:E05
917	R0127:A07	958	R0127:E06
918	R0127:A08	959	R0127:E07
919	R0127:A09	960	R0127:E08
920	R0127:A10	961	R0127:E09
921	R0127:A11	962	R0127:E10
922	R0127:A12	963	R0127:E11
923	R0127:B01	964	R0127:F01
924	R0127:B03	965	R0127:F02
925	R0127:B04	966	R0127:F03
926	R0127:B05	967	R0127:F04
927	R0127:B06	968	R0127:F05
928	R0127:B07	969	R0127:F06
929	R0127:B08	970	R0127:F07
930	R0127:B09	971	R0127:F08
931	R0127:B10	972	R0127:F10
932	R0127:B11	973	R0127:F11
933	R0127:B12	974	R0127:F12
934	R0127:C01	975	R0127:G01
935	R0127:C03	976	R0127:G02
936	R0127:C04	977	R0127:G03
937	R0127:C05	978	R0127:G04
938	R0127:C07	979	R0127:G05
939	R0127:C08	980	R0127:G06
940	R0127:C09	981	R0127:G07
941	R0127:C10	982	R0127:G08
942	R0127:C11	983	R0127:G09
943	R0127:D01	984	R0127:G10
944	R0127:D02	985	R0127:G11
945	R0127:D03	986	R0127:G12
946	R0127:D04	987	R0127:H01
947	R0127:D05	988	R0127:H02
948	R0127:D06	989	R0127:H03
949	R0127:D07	990	R0127:H04
950	R0127:D01	991	R0127:H05

Table 1C

SEQ ID NO	CLONE IDENTIFIER	SEQ ID NO	CLONE IDENTIFIER
992	R0127:H06	1034	R0128:D11
993	R0127:H07	1035	R0128:D12
994	R0127:H08	1036	R0128:E01
995	R1027:H09	1037	R0128:E02
996	R1027:H10	1038	R0128:E03
997	R1027:H11	1039	R0128:E04
998	R1028:A02	1040	R0128:E05
999	R1028:A05	1041	R0128:E06
1000	R1028:A06	1042	R0128:E07
1001	R1028:A07	1043	R0128:E08
1002	R1028:A08	1044	R0128:E09
1003	R1028:A09	1045	R0128:E10
1004	R1028:A10	1046	R0128:F12
1005	R1028:B01	1047	R0128:F01
1006	R1028:B02	1048	R0128:F02
1007	R1028:B03	1049	R0128:F03
1008	R1028:B04	1050	R0128:F04
1009	R1028:B05	1051	R0128:F06
1010	R1028:B08	1052	R0128:F07
1011	R1028:B09	1053	R0128:F08
1012	R1028:B10	1054	R0128:F09
1013	R1028:B11	1055	R0128:F10
1014	R1028:B12	1056	R0128:F12
1015	R1028:C01	1057	R0128:G01
1016	R1028:C03	1058	R0128:G02
1017	R1028:C04	1059	R0128:G03
1018	R1028:C05	1060	R0128:G04
1019	R1028:C06	1061	R0128:G05
1020	R1028:C07	1062	R0128:G06
1021	R1028:C08	1063	R0128:G07
1022	R1028:C10	1064	R0128:G09
1023	R1028:C11	1065	R0128:G10
1024	R1028:C12	1066	R0128:G11
1025	R1028:D01	1067	R0128:G12
1026	R1028:D02	1068	R0128:H01
1027	R1028:D04	1069	R0128:H02
1028	R1028:D05	1070	R0128:H03
1029	R1028:D06	1071	R0128:H04
1030	R1028:D07	1072	R0128:H05
1031	R1028:D08	1073	R0128:H06
1032	R1028:D09	1074	R0128:H07
1033	R0128:D10	1075	R0128:H08

WO 02/04514

PCT/US01/22058

40

Table 1D

SEQ ID NO	CLONE IDENTIFIER	SEQ ID NO	CLONE IDENTIFIER
1076	R0128:H09	1118	R0130:D12
1077	R0128:H10	1119	R0130:E01
1078	R0128:H11	1120	R0130:E02
1079	R0130:A02	1121	R0130:E03
1080	R0130:A05	1122	R0130:E04
1081	R0130:A06	1123	R0130:E05
1082	R0130:A08	1124	R0130:E06
1083	R0130:A09	1125	R0130:E07
1084	R0130:A10	1126	R0130:E08
1085	R0130:A11	1127	R0130:E09
1086	R0130:A12	1128	R0130:E10
1087	R0130:B01	1129	R0130:E11
1088	R0130:B02	1130	R0130:E12
1089	R0130:B03	1131	R0130:F02
1090	R0130:B04	1132	R0130:F03
1091	R0130:B05	1133	R0130:F05
1092	R0130:B06	1134	R0130:F06
1093	R0130:B08	1135	R0130:F07
1094	R0130:B09	1136	R0130:F08
1095	R0130:B10	1137	R0130:F09
1096	R0130:B11	1138	R0130:F10
1097	R0130:B12	1139	R0130:F11
1098	R0130:C02	1140	R0130:F12
1099	R0130:C03	1141	R0130:G01
1100	R0130:C04	1142	R0130:G02
1101	R0130:C05	1143	R0130:G03
1102	R0130:C06	1144	R0130:G04
1103	R0130:C07	1145	R0130:G05
1104	R0130:C08	1146	R0130:G06
1105	R0130:C09	1147	R0130:G07
1106	R0130:C10	1148	R0130:G08
1107	R0130:C11	1149	R0130:G09
1108	R0130:C12	1150	R0130:G10
1109	R0130:D02	1151	R0130:G11
1110	R0130:D03	1152	R0130:G12
1111	R0130:D04	1153	R0130:H01
1112	R0130:D05	1154	R0130:H02
1113	R0130:D06	1155	R0130:H04
1114	R0130:D07	1156	R0130:H05
1115	R0130:D09	1157	R0130:H06
1116	R0130:D10	1158	R0130:H07
1117	R0130:D11	1159	R0130:H08

WO 02/04514

PCT/US01/22058

41

Table 1E

SEQ ID NO	CLONE IDENTIFIER	SEQ ID NO	CLONE IDENTIFIER
1160	R0130:H09	1200	R0131:E01
1161	R0130:H10	1201	R0131:E02
1162	R0130:H11	1202	R0131:E03
1163	R0131:A02	1203	R0131:E04
1164	R0131:A05	1204	R0131:E06
1165	R0131:A06	1205	R0131:E07
1166	R0131:A07	1206	R0131:E08
1167	R0131:A08	1207	R0131:E10
1168	R0131:A09	1208	R0131:E11
1169	R0131:A11	1209	R0131:E12
1170	R0131:A12	1210	R0131:F02
1171	R0131:B02	1211	R0131:F04
1172	R0131:B03	1212	R0131:F05
1173	R0131:B04	1213	R0131:F06
1174	R0131:B05	1214	R0131:F07
1175	R0131:B07	1215	R0131:F08
1176	R0131:B08	1216	R0131:F09
1177	R0131:B09	1217	R0131:F10
1178	R0131:B10	1218	R0131:F11
1179	R0131:B11	1219	R0131:F12
1180	R0131:C01	1220	R0131:G01
1181	R0131:C02	1221	R0131:G02
1182	R0131:C03	1222	R0131:G03
1183	R0131:C04	1223	R0131:G04
1184	R0131:C06	1224	R0131:G05
1185	R0131:C07	1225	R0131:G06
1186	R0131:C08	1226	R0131:G07
1187	R0131:C10	1227	R0131:G08
1188	R0131:C11	1228	R0131:G09
1189	R0131:C12	1229	R0131:G10
1190	R0131:D02	1230	R0131:G11
1191	R0131:D03	1231	R0131:G12
1192	R0131:D04	1232	R0131:H01
1193	R0131:D05	1233	R0131:H02
1194	R0131:D06	1234	R0131:H05
1195	R0131:D07	1235	R0131:H06
1196	R0131:D09	1236	R0131:H07
1197	R0131:D10	1237	R0131:H08
1198	R0131:D11	1238	R0131:H09
1199	R0131:D12	1239	R0131:H11

Table 2:
Clone names for NSCLC-SQL1 and corresponding SEQ ID NOs

SEQ ID NO	CLONE IDENTIFIER
1240	Contig 54
1241	Contig 55
1242	Contig 57
1243	Contig 58
1244	Contig 60
1245	Contig 62
1246	Contig 63
1247	Contig 64
1248	Contig 65
1249	Contig 66
1250	Contig 67
1251	Contig 68
1252	Contig 69
1253	Contig 70
1254	Contig 71
1255	Contig 72
1256	Contig 73
1257	Contig 74
1258	Contig 75
1259	Contig 77
1260	Contig 78
1261	Contig 79
1262	Contig 80
1263	Contig 81
1264	Contig 83
1265	Contig 84
1266	Contig 86
1267	Contig 87
1268	Contig 88
1269	Contig 89
1270	Contig 90
1271	Contig 91
1272	Contig 92
1273	Contig 94
1274	Contig 95
1275	Contig 96
1276	Contig 97
1277	Contig 98
1278	Contig 99
1279	Contig 100

Table 3:
Clone names for NSCLC-SCLI and corresponding SEQ ID NOs

SEQ ID NO	CLONE IDENTIFIER
1280	Contig 38
1281	Contig 39
1282	Contig 40
1283	Contig 41
1284	Contig 42
1285	Contig 43
1286	Contig 44
1287	Contig 45
1288	Contig 46
1289	Contig 47
1290	Contig 48
1291	Contig 49
1292	Contig 51
1293	Contig 52
1294	Contig 53
1295	Contig 54
1296	Contig 55
1297	Contig 56
1298	Contig 57
1299	Contig 58
1300	Contig 59
1301	Contig 60
1302	Contig 62
1303	Contig 63
1304	Contig 64
1305	Contig 65
1306	Contig 66
1307	Contig 67
1308	Contig 68
1309	Contig 69
1310	Contig 70
1311	Contig 72
1312	Contig 73
1313	Contig 75
1314	Contig 76
1315	Contig 77
1316	Contig 78
1317	Contig 79
1318	Contig 80
1319	Contig 81
1320	Contig 82

Table 4A:
Clone names for NSCLC-SCL3-SCL4 and corresponding SEQ ID NOs

SEQ ID NO	CLONE IDENTIFIER	SEQ ID NO	CLONE IDENTIFIER
1321	Contig 94	1363	Contig 136
1322	Contig 95	1364	Contig 137
1323	Contig 96	1365	Contig 138
1324	Contig 97	1366	Contig 139
1325	Contig 98	1367	Contig 140
1326	Contig 99	1368	Contig 141
1327	Contig 100	1369	Contig 142
1328	Contig 101	1370	Contig 143
1329	Contig 102	1371	Contig 144
1330	Contig 103	1372	Contig 145
1331	Contig 104	1373	Contig 146
1332	Contig 105	1374	Contig 147
1333	Contig 106	1375	Contig 148
1334	Contig 107	1376	Contig 149
1335	Contig 108	1377	Contig 150
1336	Contig 109	1378	Contig 151
1337	Contig 110	1379	Contig 152
1338	Contig 111	1380	Contig 153
1339	Contig 112	1381	Contig 154
1340	Contig 113	1382	Contig 155
1341	Contig 114	1383	Contig 156
1342	Contig 115	1384	Contig 157
1343	Contig 116	1385	Contig 158
1344	Contig 117	1386	Contig 159
1345	Contig 118	1387	Contig 160
1346	Contig 119	1388	Contig 161
1347	Contig 120	1389	Contig 162
1348	Contig 121	1390	Contig 163
1349	Contig 122	1391	Contig 164
1350	Contig 123	1392	Contig 165
1351	Contig 124	1393	Contig 166
1352	Contig 125	1394	Contig 167
1353	Contig 126	1395	Contig 168
1354	Contig 127	1396	Contig 169
1355	Contig 128	1397	Contig 170
1356	Contig 129	1398	Contig 171
1357	Contig 130	1399	Contig 172
1358	Contig 131	1400	Contig 173
1359	Contig 132	1401	Contig 174
1360	Contig 133	1402	Contig 175
1361	Contig 134	1403	Contig 176

WO 02/04514

PCT/US01/22058

45

SEQ ID NO	CLONE IDENTIFIER	SEQ ID NO	CLONE IDENTIFIER
1362	Contig 135		

Table 4B:

5 Clone names for NSCLC-SCL3-SCL4 and corresponding SEQ ID NOs

SEQ ID NO	CLONE IDENTIFIER
1404	Contig 177
1405	Contig 178
1406	Contig 179
1407	Contig 180
1408	Contig 181
1409	Contig 182
1410	Contig 183
1411	Contig 184
1412	Contig 185
1413	Contig 186
1414	Contig 187

Table 5:
Clone names for SCLC-SQL1 and corresponding SEQ ID NOs

SEQ ID NO	CLONE IDENTIFIER
1415	Contig 17
1416	Contig 18
1417	Contig 20
1418	Contig 23
1419	Contig 24
1420	Contig 25
1421	Contig 26
1422	Contig 27
1423	Contig 28
1424	Contig 29
1425	Contig 30
1426	Contig 31
1427	Contig 20
1428	Contig 21
1429	Contig 22
1430	Contig 23
1431	Contig 24
1432	Contig 25
1433	Contig 26
1434	Contig 27
1435	Contig 28
1436	Contig 29
1437	Contig 30
1438	Contig 31
1439	Contig 32
1440	Contig 33
1441	Contig 34
1442	Contig 35
1443	Contig 36
1444	Contig 37
1445	Contig 38

WO 02/04514

PCT/US01/22058

47

Table 6A:
Clone names for SCLC-SCL3-SCL4 and corresponding SEQ ID NOs

SEQ ID NO	CLONE IDENTIFIER	SEQ ID NO	CLONE IDENTIFIER
1446	Contig 116	1488	Contig 160
1447	Contig 117	1489	Contig 161
1448	Contig 118	1490	Contig 162
1449	Contig 119	1491	Contig 163
1450	Contig 120	1492	Contig 164
1451	Contig 122	1493	Contig 165
1452	Contig 123	1494	Contig 166
1453	Contig 124	1495	Contig 167
1454	Contig 125	1496	Contig 168
1455	Contig 126	1497	Contig 169
1456	Contig 127	1498	Contig 170
1457	Contig 128	1499	Contig 171
1458	Contig 129	1500	Contig 172
1459	Contig 130	1501	Contig 173
1460	Contig 131	1502	Contig 174
1461	Contig 132	1503	Contig 175
1462	Contig 133	1504	Contig 176
1463	Contig 135	1505	Contig 177
1464	Contig 136	1506	Contig 178
1465	Contig 137	1507	Contig 179
1466	Contig 138	1508	Contig 181
1467	Contig 139 (L985P)	1509	Contig 182
1468	Contig 140	1510	Contig 183
1469	Contig 141	1511	Contig 184
1470	Contig 142	1512	Contig 185
1471	Contig 143	1513	Contig 186
1472	Contig 144	1514	Contig 187
1473	Contig 145	1515	Contig 189
1474	Contig 146	1516	Contig 190
1475	Contig 147	1517	Contig 191
1476	Contig 148	1518	Contig 192
1477	Contig 149	1519	Contig 193
1478	Contig 150	1520	Contig 194
1479	Contig 151	1521	Contig 195
1480	Contig 152	1522	Contig 196
1481	Contig 153	1523	Contig 197
1482	Contig 154	1524	Contig 198
1483	Contig 155	1525	Contig 199
1484	Contig 156	1526	Contig 200
1485	Contig 157	1527	Contig 201
1486	Contig 158	1528	Contig 202

WO 02/04514

PCT/US01/22058

48

SEQ ID NO	CLONE IDENTIFIER	SEQ ID NO	CLONE IDENTIFIER
1487	Contig 159		

Table 6B:
Clone names for SCLC-SCL3-SCL4 and corresponding SEQ ID NOs

SEQ ID NO	CLONE IDENTIFIER
1529	Contig 203
1530	Contig 204
1531	Contig 205
1532	Contig 206
1533	Contig 207
1534	Contig 208
1535	Contig 209
1536	Contig 210
1537	Contig 211
1538	Contig 212
1539	Contig 213
1540	Contig 214
1541	Contig 215
1542	Contig 216
1543	Contig 217
1544	Contig 218
1545	Contig 219
1546	Contig 220
1547	Contig 221
1548	Contig 222
1549	Contig 223
1550	Contig 224
1551	Contig 225
1552	Contig 226
1553	Contig 227
1554	Contig 228
1555	Contig 229
1556	Contig 230
1557	Contig 231
1558	Contig 232
1559	Contig 233
1560	Contig 234
1561	Contig 235
1562	Contig 236
1563	Contig 237

Table 7.

SEQ ID NO:	CLONE IDENTIFIER	SEQ ID NO:	CLONE IDENTIFIER
1564	R0124:E05	1609	R0129:D09
1565	R0124:E06	1610	R0129:D10
1566	R0124:E08	1611	R0129:D11
1567	R0124:F07	1612	R0129:E02
1568	R0124:F08	1613	R0129:E03
1569	R0124:F09	1614	R0129:E04
1570	R0124:G04	1615	R0129:E05
1571	R0129:A02	1616	R0129:E06
1572	R0129:A03	1617	R0129:E07
1573	R0129:A06	1618	R0129:E08
1574	R0129:A07	1619	R0129:E09
1575	R0129:A08	1620	R0129:E11
1576	R0129:A09	1621	R0129:F12
1577	R0129:A10	1622	R0129:F01
1578	R0129:A11	1623	R0129:F02
1579	R0129:A12	1624	R0129:F03
1580	R0129:B02	1625	R0129:F04
1581	R0129:B03	1626	R0129:F06
1582	R0129:B04	1627	R0129:F07
1583	R0129:B05	1628	R0129:F08
1584	R0129:B06	1629	R0129:F09
1585	R0129:B07	1630	R0129:F10
1586	R0129:B08	1631	R0129:F11
1587	R0129:B09	1632	R0129:F12
1588	R0129:B10	1633	R0129:G01
1589	R0129:B11	1634	R0129:G02
1590	R0129:B12	1635	R0129:G03
1591	R0129:C01	1636	R0129:G04
1592	R0129:C02	1637	R0129:G05
1593	R0129:C03	1638	R0129:G06
1594	R0129:C04	1639	R0129:G07
1595	R0129:C06	1640	R0129:G08
1596	R0129:C07	1641	R0129:G09
1597	R0129:C08	1642	R0129:G10
1598	R0129:C09	1643	R0129:G11
1599	R0129:C10	1644	R0129:G12
1600	R0129:C11	1645	R0129:H01
1601	R0129:C12	1646	R0129:H02
1602	R0129:D01	1647	R0129:H03
1603	R0129:D03	1648	R0129:H04
1604	R0129:D04	1649	R0129:H05
1605	R0129:D05	1650	R0129:H08

WO 02/04514

PCT/US01/22058

50

1606	R0129:D06	1651	R0129:I109
1607	R0129:D07	1652	R0129:I110
1608	R0129:D08	1653	R0129:I111

Table 8

5

SEQ ID NO:	CLONE IDENTIFIER
1654	26484
1655	26496
1656	26517
1657	26531
1658	26022
1659	26026
1660	26810
1661	26815
1662	26869
1663	26883
1664	26902

SEQ ID NO:1665 and 1666 are primers used in the amplification of the coding region of L548S

SEQ ID NO:1667 is the protein sequence of expressed recombinant L7548S.

10 SEQ ID NO:1668 is the cDNA sequence of expressed recombinant L7548S.

SEQ ID NO:1669 is the extended cDNA sequence of clone #18971 (L801P).

SEQ ID NO:1670 is the amino acid sequence of open reading frame ORF4 encoded by SEQ ID NO:1669.

SEQ ID NO:1671 is the amino acid sequence of open reading frame ORF5 encoded by

15 SEQ ID NO:1669.

SEQ ID NO:1672 is the amino acid sequence of open reading frame ORF6 encoded by SEQ ID NO:1669.

SEQ ID NO:1673 is the amino acid sequence of open reading frame ORF7 encoded by SEQ ID NO:1669.

20 SEQ ID NO:1674 is the amino acid sequence of open reading frame ORF8 encoded by SEQ ID NO:1669.

SEQ ID NO:1675 is the amino acid sequence of open reading frame ORF9 encoded by SEQ ID NO:1669.

WO 02/04514

PCT/US01/22058

51

SEQ ID NO:1676 is the extended cDNA for contig 139 (SEQ ID NO:1467), also known as L985P.

SEQ ID NO:1677 is the L985P amino acid sequence encoded by SEQ ID NO:1676.

SEQ ID NO:1678 is the amino acid sequence of open reading frame ORF5X of SEQ ID

5 NO:1669.

SEQ ID NO:1679 is the amino acid sequence of an open reading frame for contig 139 (SEQ ID NO:1467).

SEQ ID NO:1680-1788, set forth in the Table 9, represent cDNA clones identified by microarray analysis of the SQU1, SCL1, SCL3 and SCL4 libraries on lung chip 5.

10

Table 9

SEQ ID NO:	CLONE IDENTIFIER
1680	58456
1681	58458
1682	58462
1683	58469
1684	58470
1685	58482
1686	58485
1687	58501
1688	58502
1689	58505
1690	58507
1691	58509
1692	58512
1693	58527
1694	58529
1695	58531
1696	58537
1697	58539
1698	58545
1699	59319
1700	59322
1701	59348
1702	59350
1703	59363
1704	59365
1705	59370
1706	59373
1707	59376
1708	61050
1709	61051
1710	61052
1711	61054
1712	61056
1713	61057
1714	61060
1715	61062
1716	61063
1717	61064
1718	61065
1719	61066
1720	61069

WO 02/04514

PCT/US01/22058

53

SEQ ID NO:	CLONE IDENTIFIER
1721	61070
1722	61071
1723	61074
1724	61075
1725	61077
1726	61079
1727	61080
1728	61081
1729	61083
1730	61085
1731	61086
1732	61088
1733	61090
1734	61091
1735	61093
1736	61094
1737	61096
1738	61097
1739	61099
1740	61100
1741	61103
1742	61105
1743	61106
1744	61110
1745	61113
1746	61115
1747	61117
1748	61118
1749	61119
1750	61120
1751	61122
1752	61125
1753	61126
1754	61130
1755	61133
1756	61134
1757	61135
1758	61137
1759	61139
1760	61143
1761	61144
1762	61148
1763	61151

WO 02/04514

PCT/US01/22058

54

SEQ ID NO:	CLONE IDENTIFIER
1764	61155
1765	61156
1766	61159
1767	61160
1768	61163
1769	61167
1770	61172
1771	61173
1772	61176
1773	61177
1774	61183
1775	61185
1776	61188
1777	61192
1778	61198
1779	61201
1780	61202
1781	61204
1782	61206
1783	61210
1784	61212
1785	61216
1786	61225
1787	61226
1788	61227

SEQ ID NO:1789 is the cDNA sequence of clone #47988 (L972P).

SEQ ID NO:1790 is the cDNA sequence of clone #48005 (L979P).

SEQ ID NO:1791 is an extended cDNA sequence for clone #48005 (L979P).

- 5 SEQ ID NO:1792 is an extended cDNA sequence for clone #49826 (SEQ ID NO:1279; L980P).

SEQ ID NO:1793 is an extended cDNA sequence for clone #20631 (SEQ ID NO:117; L973P).

SEQ ID NO:1794 is an extended cDNA sequence for clone #20661 (SEQ ID NO:128; L974P).

- 10 SEQ ID NO:1795 is an extended cDNA sequence for clone #50430 (SEQ ID NO:1442; L996P).

WO 02/04514

PCT/US01/22058

55

- SEQ ID NO:1796 is an extended cDNA sequence for clone #26961 (SEQ ID NO:288; L977P).
- SEQ ID NO:1797 is an extended cDNA sequence for clone #24928 (SEQ ID NO:1339; L978P).
- 5 SEQ ID NO:1798 is an extended cDNA sequence for clone #50507 (SEQ ID NO:1446; L984P).
- SEQ ID NO:1799 is an extended cDNA sequence for clone #50645 (SEQ ID NO:1531; L988P).
- SEQ ID NO:1800 is an extended cDNA sequence for clone #50628 (SEQ ID NO:1533; L1423P).
- 10 L1423P).
- SEQ ID NO:1801 is an extended cDNA sequence for clone #50560 (SEQ ID NO:1527; L987P).
- SEQ ID NO:1802 is an extended cDNA sequence for clone #27699 (SEQ ID NO:468; L998P).
- 15 SEQ ID NO:1803 is an extended cDNA sequence for clone #59303 (SEQ ID NO:949; L1425P).
- SEQ ID NO:1804 is an extended cDNA sequence for clone #59314 (SEQ ID NO:1156; L1426P).
- SEQ ID NO:1805 is an extended cDNA sequence for clone #59298 (SEQ ID NO:921; L1427P).
- 20 L1427P).
- SEQ ID NO:1806 is an amino acid sequence encoded by SEQ ID NO:1791.
- SEQ ID NO:1807 is an amino acid sequence encoded by SEQ ID NO:1792.
- SEQ ID NO:1808 is an amino acid sequence encoded by SEQ ID NO:1793.
- SEQ ID NO:1809 is an amino acid sequence encoded by SEQ ID NO:1794.
- 25 SEQ ID NO:1810 is an amino acid sequence encoded by SEQ ID NO:1795.
- SEQ ID NO:1811 is an amino acid sequence encoded by SEQ ID NO:1796.
- SEQ ID NO:1812 is an amino acid sequence encoded by SEQ ID NO:1797.
- SEQ ID NO:1813 is an amino acid sequence encoded by SEQ ID NO:1798.
- SEQ ID NO:1814 is an amino acid sequence encoded by SEQ ID NO:1799.
- 30 SEQ ID NO:1815 is an amino acid sequence encoded by SEQ ID NO:1800.
- SEQ ID NO:1816 is an amino acid sequence encoded by SEQ ID NO:1527 (L987P).
- SEQ ID NO:1817 is an amino acid sequence encoded by SEQ ID NO:1823.

WO 02/04514

PCT/US01/22058

56

- SEQ ID NO:1818 is an amino acid sequence encoded by SEQ ID NO:1801.
SEQ ID NO:1819 is an amino acid sequence encoded by SEQ ID NO:1802.
SEQ ID NO:1820 is an amino acid sequence encoded by SEQ ID NO:1803.
SEQ ID NO:1821 is an amino acid sequence encoded by SEQ ID NO:1804.
5 SEQ ID NO:1822 is an amino acid sequence encoded by SEQ ID NO:1805.
SEQ ID NO:1823 is an extended cDNA sequence for clone #50560 (SEQ ID NO:1527; L987P).
SEQ ID NO:1824 is a full length cDNA sequence for clone L872P (SEQ ID NO:34).
SEQ ID NO:1825 is the amino acid sequence encoded by SEQ ID NO:1824.
10 SEQ ID NO:1826 is the cDNA sequence encoding the N-terminal portion of L552S.
SEQ ID NO:1827-1829 are cDNA sequences of portions of L552S.
SEQ ID NO:1830 is the N-terminal portion of L552S.
SEQ ID NO:1831-1833 are the amino acid sequences encoded by SEQ ID NO:1827-1829, respectively.
15 SEQ ID NO:1834-1856 are the amino acid sequences of peptides of L548S.
SEQ ID NO:1857-1860 are PCR primers.
SEQ ID NO:1861 is the determined DNA sequence for a fusion of Ra12 and ORF4 of P801P.
SEQ ID NO:1862 is the determined DNA sequence for a fusion of Ra12 and ORF5 of
20 P801P.
SEQ ID NO:1863 is the amino acid sequence of the fusion of Ra12 and ORF4 of P801P.
SEQ ID NO:1864 is the amino acid sequence of the fusion of Ra12 and ORF5 of P801P.
25 SEQ ID NO:1865 is the determined cDNA sequence for clone L984P (573A).
SEQ ID NO:1866 is the determined cDNA sequence for clone L984P (512A).
SEQ ID NO:1867 is the determined cDNA sequence for clone L984P (NCI-H128).
SEQ ID NO:1868 is the determined cDNA sequence for clone L984P (DMS-79).
SEQ ID NO:1869 is the amino acid sequence encoded by SEQ ID NO:1865.
30 SEQ ID NO:1870 is the amino acid sequence encoded by SEQ ID NO:1866.
SEQ ID NO:1871 is the amino acid sequence encoded by SEQ ID NO:1867.
SEQ ID NO:1872 is the amino acid sequence encoded by SEQ ID NO:1868.

WO 02/04514

PCT/US01/22058

57

- SEQ ID NO:1873 is a full length cDNA sequence for clone L985P (partial sequence given in SEQ ID NO:1467).
- SEQ ID NO:1874 is the amino acid sequence for L985P encoded by SEQ ID NO:1873.
- SEQ ID NO:1875 is the predicted and determined cDNA sequence for a fusion of Ra12 and L985P.
- SEQ ID NO:1876 is the predicted amino acid sequence of a fusion of Ra12 and L985P encoded by SEQ ID NO:1875.
- SEQ ID NO:1877 is the predicted cDNA sequence for a fusion of Ra12S and L985P.
- SEQ ID NO:1878 is the predicted amino acid sequence of a fusion of Ra12S and L985P encoded by SEQ ID NO:1877.
- SEQ ID NO:1879 is the predicted cDNA sequence for a fusion of Ra12S and L985PEx.
- SEQ ID NO:1880 is the predicted amino acid sequence of a fusion of Ra12S and L985PEx encoded by SEQ ID NO:1879.
- SEQ ID NO:1881 is the predicted cDNA sequence the extracellular loop 2 peptide of L985P.
- SEQ ID NO:1882 is the predicted amino acid sequence for the extracellular loop 2 peptide of L985P encoded by SEQ ID NO:1875.
- SEQ ID NO:1883 is an extended cDNA sequence for clone #59316 (SEQ ID NO:1180; L1428P).
- SEQ ID NO:1884 is a first predicted amino acid sequence encoded by SEQ ID NO:1883 and designated L1428P-ORF1.
- SEQ ID NO:1885 is a second predicted amino acid sequence encoded by SEQ ID NO:1883 and designated L1428P-ORF2.
- SEQ ID NO:1886 is a third predicted amino acid sequence encoded by SEQ ID NO:1883 and designated L1428P-ORF3.
- SEQ ID NO:1887 is a fourth predicted amino acid sequence encoded by SEQ ID NO:1883 and designated L1428P-ORF4.
- SEQ ID NO:1888 is a fifth predicted amino acid sequence encoded by SEQ ID NO:1883 and designated L1428P-ORF5.
- SEQ ID NO:1889 is a sixth predicted amino acid sequence encoded by SEQ ID NO:1883 and designated L1428P-ORF6.

WO 02/04514

PCT/US01/22058

58

- SEQ ID NO:1890 is a seventh predicted amino acid sequence encoded by SEQ ID NO:1883 and designated L1428P-ORF7.
- SEQ ID NO:1891-1900 are the nucleotide sequences for the database hits described in Table 17.
- 5 SEQ ID NO:1901-1909 are the deduced amino acid sequences encoded by the nucleotide sequences described in Table 17.
- SEQ ID NO:1910 is the full-length cDNA for clone L1437P (partial sequence given in SEQ ID NO:1896).
- SEQ ID NO:1911 is the forward primer PDM-433 for the coding region of clone
- 10 L548S.
- SEQ ID NO:1912 is the reverse primer PDM-438 for the coding region of clone L548S.
- SEQ ID NO:1913 is the amino acid sequence for the expressed recombinant L548S.
- SEQ ID NO:1914 is the DNA coding sequence for the recombinant L548S.
- SEQ ID NO:1915 is the forward primer PDM-498 for the coding region of clone L551S
- 15 SEQ ID NO:1916 is the reverse primer PDM-499 for the coding region of clone L551S
- SEQ ID NO:1917 is the amino acid sequence for the expressed recombinant L551S.
- SEQ ID NO:1918 is the DNA coding sequence for the recombinant L551S.
- SEQ ID NO:1919 is the forward primer PDM-479 for the coding region of clone L552S
- SEQ ID NO:1920 is the reverse primer PDM-480 for the coding region of clone L552S
- 20 SEQ ID NO:1921 is the amino acid sequence for the expressed recombinant L552S.
- SEQ ID NO:1922 is the DNA coding sequence for the recombinant L552S.
- SEQ ID NO:1923 is the predicted full-length cDNA sequence for clone #19069 (partial sequence given in SEQ ID NO:90).
- SEQ ID NO:1924 is the predicted full-length cDNA sequence for clone #18965 or
- 25 #19002 (partial sequence given in SEQ ID NO:15).
- SEQ ID NO:1925 is the deduced amino acid sequence encoded by SEQ ID NO:1923.
- SEQ ID NO:1926 is the deduced amino acid sequence encoded by SEQ ID NO:1924.
- SEQ ID NO:1927 is the determined amino acid sequence of a first L552S epitope.
- SEQ ID NO:1928 is the determined amino acid sequence of a second L552S epitope.
- 30 SEQ ID NO:1929 is the determined amino acid sequence of a third L552S epitope.
- SEQ ID NO:1930 is the amino acid sequence for L985P peptide #3482.

WO 02/04514

PCT/US01/22058

59

- SEQ ID NO:1931 is an extended cDNA sequence for clone #61144 (SEQ ID NO:1761, T1439P).
- SEQ ID NO:1932 is the deduced amino acid sequence encoded by SEQ ID NO:1931.
- SEQ ID NO:1933 is the full-length cDNA of the NUF2R gene to which SEQ ID
- 5 NO:1931 shows some sequence similarity.
- SEQ ID NO:1934 is the deduced amino acid sequence encoded by SEQ ID NO:1933.
- SEQ ID NO:1935 is a forward primer PDM-737 for the coding region of clone L552S.
- SEQ ID NO:1936 is a reverse primer PDM-738 for the coding region of clone L552S.
- SEQ ID NO:1937 is the amino acid sequence for the expressed recombinant L552S.
- 10 SEQ ID NO:1938 is the DNA coding sequence for the recombinant L552S.
- SEQ ID NO:1939 is another forward primer PDM-736 for the coding region of clone L552S.
- SEQ ID NO:1940 is the amino acid sequence for a second expressed recombinant L552S.
- 15 SEQ ID NO:1941 is the DNA coding sequence for a second recombinant L552S.
- SEQ ID NO:1942 is the determined amino acid sequence of a fourth L552S epitope.
- SEQ ID NO:1943 is the determined amino acid sequence of a first XAGE-1 epitope.
- SEQ ID NO:1944 is the determined amino acid sequence of a second XAGE-1 epitope.
- SEQ ID NO:1945 is the determined amino acid sequence of a first 20-mer peptide
- 20 corresponding to amino acid residues 1-20 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1946 is the determined amino acid sequence of a second 20-mer peptide corresponding to amino acid residues 6-25 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1947 is the determined amino acid sequence of a third 20-mer peptide corresponding to amino acid residues 11-30 of full-length L552S (SEQ ID NO:809).
- 25 SEQ ID NO:1948 is the determined amino acid sequence of a fourth 20-mer peptide corresponding to amino acid residues 16-35 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1949 is the determined amino acid sequence of a fifth 20-mer peptide corresponding to amino acid residues 21-40 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1950 is the determined amino acid sequence of a sixth 20-mer peptide
- 30 corresponding to amino acid residues 26-45 of full-length L552S (SEQ ID NO:809).

- SEQ ID NO:1951 is the determined amino acid sequence of a seventh 20-mer peptide corresponding to amino acid residues 31-50 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1952 is the determined amino acid sequence of a eighth 20-mer peptide corresponding to amino acid residues 36-55 of full-length L552S (SEQ ID NO:809).
- 5 SEQ ID NO:1953 is the determined amino acid sequence of a ninth 20-mer peptide corresponding to amino acid residues 41-60 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1954 is the determined amino acid sequence of a tenth 20-mer peptide corresponding to amino acid residues 46-65 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1955 is the determined amino acid sequence of a eleventh 20-mer peptide
- 10 corresponding to amino acid residues 51-70 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1955 is the determined amino acid sequence of a twelfth 20-mer peptide corresponding to amino acid residues 56-75 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1956 is the determined amino acid sequence of a thirth 20-mer peptide corresponding to amino acid residues 61-80 of full-length L552S (SEQ ID NO:809).
- 15 SEQ ID NO:1957 is the determined amino acid sequence of a fourteenth 20-mer peptide corresponding to amino acid residues 66-85 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1958 is the determined amino acid sequence of a fifteenth 20-mer peptide corresponding to amino acid residues 71-90 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1959 is the determined amino acid sequence of a sixteenth 20-mer peptide
- 20 corresponding to amino acid residues 76-95 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1961 is the determined amino acid sequence of a seventeen 20-mer peptide corresponding to amino acid residues 81-100 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1962 is the determined amino acid sequence of a eightth 20-mer peptide corresponding to amino acid residues 86-105 of full-length L552S (SEQ ID NO:809).
- 25 SEQ ID NO:1963 is the determined amino acid sequence of a nineteenth 20-mer peptide corresponding to amino acid residues 91-110 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1964 is the determined amino acid sequence of a twentieth 20-mer peptide corresponding to amino acid residues 96-115 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1965 is the determined amino acid sequence of a twenty-first 20-mer
- 30 peptide corresponding to amino acid residues 101-120 of full-length L552S (SEQ ID NO:809).

WO 02/04514

PCT/US01/22058

61

- SEQ ID NO:1966 is the determined amino acid sequence of a twenty-second 20-mer peptide corresponding to amino acid residues 106-125 of full-length L552S (SEQ ID NO:809).
- 5 SEQ ID NO:1967 is the determined amino acid sequence of a twenty-third 20-mer peptide corresponding to amino acid residues 111-130 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1968 is the determined amino acid sequence of a twenty-fourth 20-mer peptide corresponding to amino acid residues 116-135 of full-length L552S (SEQ ID NO:809).
- 10 SEQ ID NO:1969 is the determined amino acid sequence of a twenty-fifth 20-mer peptide corresponding to amino acid residues 121-140 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1970 is the determined amino acid sequence of a twenty-sixth 20-mer peptide corresponding to amino acid residues 126-145 of full-length L552S (SEQ ID
- 15 NO:809).
- SEQ ID NO:1971 is the determined amino acid sequence of a twentieth-seventh 20-mer peptide corresponding to amino acid residues 131-150 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1972 is the determined amino acid sequence of a twenty-eighth 20-mer
- 20 peptide corresponding to amino acid residues 136-155 of full-length L552S (SEQ ID NO:809).
- SEQ ID NO:1973 is the determined amino acid sequence of a twenty-ninth 20-mer peptide corresponding to amino acid residues 141-160 of full-length L552S (SEQ ID NO:809).
- 25 SEQ ID NO:1974 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1945.
- SEQ ID NO:1975 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1946.
- SEQ ID NO:1976 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1947.
- SEQ ID NO:1977 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1948.
- SEQ ID NO:1978 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1949.
- 30 SEQ ID NO:1979 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1950.
- SEQ ID NO:1980 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1951.

WO 02/04514

PCT/US01/22058

62

- SEQ ID NO:1981 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1952.
SEQ ID NO:1982 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1953.
SEQ ID NO:1983 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1954.
SEQ ID NO:1984 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1955.
5 SEQ ID NO:1985 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1956.
SEQ ID NO:1986 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1957.
SEQ ID NO:1987 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1958.
SEQ ID NO:1988 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1959.
SEQ ID NO:1989 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1960.
10 SEQ ID NO:1990 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1961.
SEQ ID NO:1991 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1962.
SEQ ID NO:1992 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1963.
SEQ ID NO:1993 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1964.
SEQ ID NO:1994 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1965.
15 SEQ ID NO:1995 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1966.
SEQ ID NO:1996 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1967.
SEQ ID NO:1997 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1968.
SEQ ID NO:1998 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1969.
SEQ ID NO:1999 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1970.
20 SEQ ID NO:2000 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1971.
SEQ ID NO:2001 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1972.
SEQ ID NO:2002 is the DNA sequence which encodes the 20-mer of SEQ ID NO:1973.

25 DETAILED DESCRIPTION OF THE INVENTION

- The present invention is directed generally to compositions and their use in the therapy and diagnosis of cancer, particularly lung cancer. As described further below, illustrative compositions of the present invention include, but are not restricted to, polypeptides, particularly immunogenic polypeptides, polynucleotides encoding such
30 polypeptides, antibodies and other binding agents, antigen presenting cells (APCs) and immune system cells (e.g., T cells).

WO 02/04514

PCT/US01/22058

63

The practice of the present invention will employ, unless indicated specifically to the contrary, conventional methods of virology, immunology, microbiology, molecular biology and recombinant DNA techniques within the skill of the art, many of which are described below for the purpose of illustration. Such techniques are explained fully in the literature. See, e.g., Sambrook, et al. Molecular Cloning: A Laboratory Manual (2nd Edition, 1989); Maniatis et al. Molecular Cloning: A Laboratory Manual (1982); DNA Cloning: A Practical Approach, vol. I & II (D. Glover, ed.); Oligonucleotide Synthesis (N. Gait, ed., 1984); Nucleic Acid Hybridization (B. Hames & S. Higgins, eds., 1985); Transcription and Translation (B. Hames & S. Higgins, eds., 1984); Animal Cell Culture (R. Freshney, ed., 1986); Perbal, A Practical Guide to Molecular Cloning (1984).

All publications, patents and patent applications cited herein, whether supra or infra, are hereby incorporated by reference in their entirety.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise.

Polypeptide Compositions

As used herein, the term "polypeptide" is used in its conventional meaning, i.e., as a sequence of amino acids. The polypeptides are not limited to a specific length of the product; thus, peptides, oligopeptides, and proteins are included within the definition of polypeptide, and such terms may be used interchangeably herein unless specifically indicated otherwise. This term also does not refer to or exclude post-expression modifications of the polypeptide, for example, glycosylations, acetylations, phosphorylations and the like, as well as other modifications known in the art, both naturally occurring and non-naturally occurring. A polypeptide may be an entire protein, or a subsequence thereof. Particular polypeptides of interest in the context of this invention are amino acid subsequences comprising epitopes, i.e., antigenic determinants substantially responsible for the immunogenic properties of a polypeptide and being capable of evoking an immune response.

Particularly illustrative polypeptides of the present invention comprise those encoded by a polynucleotide sequence set forth in any one of SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 5 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002, or a sequence that hybridizes under moderately stringent conditions, or, alternatively, under highly stringent conditions, to a polynucleotide sequence set forth in any one of SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 10 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002. Certain other illustrative polypeptides of the invention comprise amino acid sequences as set forth in any one of SEQ ID NOs:324-340, 786, 787, 789, 791, 793, 795, 797-799, 805, 806, 809, 827, 1667, 1670-1675, 1677-1679, 1806-1822, 1825, 1830-1833, 1834-1856, 1863, 1864, 1869-1872, 15 1874, 1876, 1878, 1880, 1882, 1884-1890, 1901-1909, 1913, 1917, 1921, 1925-1930, 1932, 1934, 1937, 1940 and 1942-1973.

The polypeptides of the present invention are sometimes herein referred to as lung tumor proteins or lung tumor polypeptides, as an indication that their identification has been based at least in part upon their increased levels of expression in 20 lung tumor samples. Thus, a "lung tumor polypeptide" or "lung tumor protein," refers generally to a polypeptide sequence of the present invention, or a polynucleotide sequence encoding such a polypeptide, that is expressed in a substantial proportion of lung tumor samples, for example preferably greater than about 20%, more preferably greater than about 30%, and most preferably greater than about 50% or more of lung 25 tumor samples tested, at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in normal tissues, as determined using a representative assay provided herein. A lung tumor polypeptide sequence of the invention, based upon its increased level of expression in tumor cells, has particular utility both as a diagnostic marker as well as a therapeutic target, as further described 30 below.

WO 02/04514

PCT/US01/22058

65

In certain preferred embodiments, the polypeptides of the invention are immunogenic, *i.e.*, they react detectably within an immunoassay (such as an ELISA or T-cell stimulation assay) with antisera and/or T-cells from a patient with lung cancer. Screening for immunogenic activity can be performed using techniques well known to the skilled artisan. For example, such screens can be performed using methods such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In one illustrative example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ¹²⁵I-labeled Protein A.

As would be recognized by the skilled artisan, immunogenic portions of the polypeptides disclosed herein are also encompassed by the present invention. An "immunogenic portion," as used herein, is a fragment of an immunogenic polypeptide of the invention that itself is immunologically reactive (*i.e.*, specifically binds) with the B-cells and/or T-cell surface antigen receptors that recognize the polypeptide. Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well-known techniques.

In one preferred embodiment, an immunogenic portion of a polypeptide of the present invention is a portion that reacts with antisera and/or T-cells at a level that is not substantially less than the reactivity of the full-length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Preferably, the level of immunogenic activity of the immunogenic portion is at least about 50%, preferably at least about 70% and most preferably greater than about 90% of the immunogenicity for the full-length polypeptide. In some instances, preferred immunogenic portions will be identified that have a level of immunogenic activity greater than that of the corresponding full-length

WO 02/04514

PCT/US01/22058

66

polypeptide, e.g., having greater than about 100% or 150% or more immunogenic activity.

In certain other embodiments, illustrative immunogenic portions may include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other illustrative immunogenic portions will contain a small N- and/or C-terminal deletion (e.g., 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

In another embodiment, a polypeptide composition of the invention may also comprise one or more polypeptides that are immunologically reactive with T cells and/or antibodies generated against a polypeptide of the invention, particularly a polypeptide having an amino acid sequence disclosed herein, or to an immunogenic fragment or variant thereof.

In another embodiment of the invention, polypeptides are provided that comprise one or more polypeptides that are capable of eliciting T cells and/or antibodies that are immunologically reactive with one or more polypeptides described herein, or one or more polypeptides encoded by contiguous nucleic acid sequences contained in the polynucleotide sequences disclosed herein, or immunogenic fragments or variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency.

The present invention, in another aspect, provides polypeptide fragments comprising at least about 5, 10, 15, 20, 25, 50, or 100 contiguous amino acids, or more, including all intermediate lengths, of a polypeptide compositions set forth herein, such as those set forth in SEQ ID NOs:324-340, 786, 787, 789, 791, 793, 795, 797-799, 805, 806, 809, 827, 1667, 1670-1675, 1677-1679, 1806-1822, 1825, 1830-1833, 1834-1856, 1863, 1864, 1869-1872, 1874, 1876, 1878, 1880, 1882, 1884-1890, 1901-1909, 1913, 1917, 1921, 1925-1930, 1932, 1934, 1937, 1940 and 1942-1973, or those encoded by a polynucleotide sequence set forth in a sequence of SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002.

In another aspect, the present invention provides variants of the polypeptide compositions described herein. Polypeptide variants generally encompassed by the present invention will typically exhibit at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described below), along its length, to a polypeptide sequences set forth herein.

In one preferred embodiment, the polypeptide fragments and variants provide by the present invention are immunologically reactive with an antibody and/or T-cell that reacts with a full-length polypeptide specifically set forth herein.

10 In another preferred embodiment, the polypeptide fragments and variants provided by the present invention exhibit a level of immunogenic activity of at least about 50%, preferably at least about 70%, and most preferably at least about 90% or more of that exhibited by a full-length polypeptide sequence specifically set forth herein.

15 A polypeptide "variant," as the term is used herein, is a polypeptide that typically differs from a polypeptide specifically disclosed herein in one or more substitutions, deletions, additions and/or insertions. Such variants may be naturally occurring or may be synthetically generated, for example, by modifying one or more of the above polypeptide sequences of the invention and evaluating their immunogenic activity as described herein and/or using any of a number of techniques well known in
20 the art.

For example, certain illustrative variants of the polypeptides of the invention include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other illustrative variants
25 include variants in which a small portion (e.g., 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

In many instances, a variant will contain conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydrophobic nature of the
30 polypeptide to be substantially unchanged. As described above, modifications may be

WO 02/04514

PCT/US01/22058

68

made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a variant or derivative polypeptide with desirable characteristics, *e.g.*, with immunogenic characteristics. When it is desired to alter the amino acid sequence of a polypeptide to create an equivalent, or
5 even an improved, immunogenic variant or portion of a polypeptide of the invention, one skilled in the art will typically change one or more of the codons of the encoding DNA sequence according to Table 10.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with
10 structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus
15 contemplated that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

TABLE 10

Amino Acids			Codons						
Alanine	Ala	A	GCA	GCC	GCG	GCU			
Cysteine	Cys	C	UGC	UGU					
Aspartic acid	Asp	D	GAC	GAU					
Glutamic acid	Glu	E	GAA	GAG					
Phenylalanine	Phe	F	UUC	UUU					
Glycine	Gly	G	GGA	GGC	GGG	GGU			
Histidine	His	H	CAC	CAU					
Isoleucine	Ile	I	AUA	AUC	AUU				
Lysine	Lys	K	AAA	AAG					
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
Methionine	Met	M	AUG						
Asparagine	Asn	N	AAC	AAU					
Proline	Pro	P	CCA	CCC	CCG	CCU			
Glutamine	Gln	Q	CAA	CAG					
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU	
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU	
Threonine	Thr	T	ACA	ACC	ACG	ACU			
Valine	Val	V	GUA	GUC	GUG	GUU			
Tryptophan	Trp	W	UGG						
Tyrosine	Tyr	Y	UAC	UAU					

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are:

isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

5 It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydrophobic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydrophobic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

15 As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 $\pm$ 1); glutamate (+3.0 $\pm$ 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 $\pm$ 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

25 As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

WO 02/04514

PCT/US01/22058

71

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of
 5 nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

Amino acid substitutions may further be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic
 10 nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may
 15 represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or
 20 alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydrophobic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally
 25 directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

When comparing polypeptide sequences, two sequences are said to be
 30 "identical" if the sequence of amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two

WO 02/04514

PCT/US01/22058

72

sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad. Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Adv. Appl. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402

WO 02/04514

PCT/US01/22058

73

and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for
5 Biotechnology Information. For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is
10 reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment.

In one preferred approach, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polypeptide sequence in
15 the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of
20 matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Within other illustrative embodiments, a polypeptide may be a fusion polypeptide that comprises multiple polypeptides as described herein, or that comprises
25 at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological
30 and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the polypeptide or to enable the polypeptide to be targeted to

desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the polypeptide.

Fusion polypeptides may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion polypeptide is expressed as a recombinant polypeptide, allowing the production of increased levels, relative to a non-fused polypeptide, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion polypeptide that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements

responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

5 The fusion polypeptide can comprise a polypeptide as described herein together with an unrelated immunogenic protein, such as an immunogenic protein capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (see, for example, Stoute et al. *New Engl. J. Med.*, 336:86-91, 1997).

10 In one preferred embodiment, the immunological fusion partner is derived from a *Mycobacterium* sp., such as a *Mycobacterium tuberculosis*-derived Ra12 fragment. Ra12 compositions and methods for their use in enhancing the expression and/or immunogenicity of heterologous polynucleotide/polypeptide sequences is described in U.S. Patent Application 60/158,585, the disclosure of which is
15 incorporated herein by reference in its entirety. Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid. MTB32A is a serine protease of 32 KD molecular weight encoded by a gene in virulent and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been described (for example, U.S. Patent Application
20 60/158,585; see also, Skeiky et al., *Infection and Immunity*, (1999) 67:3998-4007, incorporated herein by reference). Surprisingly, it was discovered that a 14KD C-terminal fragment of the MTB32A coding sequence expresses at high levels on its own and remains as a soluble polypeptide throughout the purification process. Moreover, this fragment may enhance the immunogenicity of heterologous antigenic polypeptides
25 with which it is fused. This 14 KD C-terminal fragment is referred to herein as Ra12 and represents a fragment comprising some or all of amino acid residues 192 to 323 of MTB32A.

Other preferred Ra12 polynucleotides generally comprise at least about
15 consecutive nucleotides, at least about 30 nucleotides, at least about 60 nucleotides,
30 at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides that encode a portion of a Ra12 polypeptide.

Ra12 polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a Ra12 polypeptide or a portion thereof) or may comprise a variant of such a sequence. Ra12 polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded fusion polypeptide is not substantially diminished, relative to a fusion polypeptide comprising a native Ra12 polypeptide. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a polynucleotide sequence that encodes a native Ra12 polypeptide or a portion thereof.

Within other preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see Biotechnology* 10:795-798,

WO 02/04514

PCT/US01/22058

77

1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion polypeptide. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

Yet another illustrative embodiment involves fusion polypeptides, and the polynucleotides encoding them, wherein the fusion partner comprises a targeting signal capable of directing a polypeptide to the endosomal/lysosomal compartment, as described in U.S. Patent No. 5,633,234. An immunogenic polypeptide of the invention, when fused with this targeting signal, will associate more efficiently with MHC class II molecules and thereby provide enhanced in vivo stimulation of CD4⁺ T-cells specific for the polypeptide.

Polypeptides of the invention are prepared using any of a variety of well known synthetic and/or recombinant techniques, the latter of which are further described below. Polypeptides, portions and other variants generally less than about 150 amino acids can be generated by synthetic means, using techniques well known to those of ordinary skill in the art. In one illustrative example, such polypeptides are synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. See Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

In general, polypeptide compositions (including fusion polypeptides) of the invention are isolated. An "isolated" polypeptide is one that is removed from its original environment. For example, a naturally-occurring protein or polypeptide is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are also purified, e.g., are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure.

Polynucleotide Compositions

The present invention, in other aspects, provides polynucleotide compositions. The terms "DNA" and "polynucleotide" are used essentially interchangeably herein to refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. "Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA molecule does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA molecule as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be understood by those skilled in the art, the polynucleotide compositions of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

As will be also recognized by the skilled artisan, polynucleotides of the invention may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules may include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a polypeptide/protein of the invention or a portion thereof) or may comprise a sequence that encodes a variant or derivative, preferably and immunogenic variant or derivative, of such a sequence.

Therefore, according to another aspect of the present invention, polynucleotide compositions are provided that comprise some or all of a polynucleotide sequence set forth in any one of SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-

WO 02/04514

PCT/US01/22058

79

1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002, complements of a polynucleotide sequence set forth in any one of SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 5 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002, and degenerate variants of a polynucleotide sequence set forth in any one of SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 10 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002. In certain preferred embodiments, the polynucleotide sequences set forth herein encode immunogenic polypeptides, as described above.

In other related embodiments, the present invention provides polynucleotide variants having substantial identity to the sequences disclosed herein in 15 SEQ ID NOs:1-57, 59-323, 341-782, 784-785, 788, 790, 792, 794, 796, 800-804, 807, 808, 810-826, 828-1664, 1668, 1669, 1676, 1680-1805, 1824, 1826-1829, 1865-1868, 1873, 1875, 1877, 1879, 1881, 1883, 1891-1900, 1910, 1914, 1918, 1922-1924, 1931, 1933, 1938, 1941 and 1974-2002, for example those comprising at least 70% sequence identity, preferably at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or 20 higher, sequence identity compared to a polynucleotide sequence of this invention using the methods described herein, (e.g., BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, 25 reading frame positioning and the like.

Typically, polynucleotide variants will contain one or more substitutions, additions, deletions and/or insertions, preferably such that the immunogenicity of the polypeptide encoded by the variant polynucleotide is not substantially diminished relative to a polypeptide encoded by a polynucleotide sequence specifically set forth 30 herein). The term "variants" should also be understood to encompass homologous genes of xenogenic origin.

WO 02/04514

PCT/US01/22058

80

In additional embodiments, the present invention provides polynucleotide fragments comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 10, 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

In another embodiment of the invention, polynucleotide compositions are provided that are capable of hybridizing under moderate to high stringency conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-60°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS. One skilled in the art will understand that the stringency of hybridization can be readily manipulated, such as by altering the salt content of the hybridization solution and/or the temperature at which the hybridization is performed. For example, in another embodiment, suitable highly stringent hybridization conditions include those described above, with the exception that the temperature of hybridization is increased, *e.g.*, to 60-65°C or 65-70°C.

In certain preferred embodiments, the polynucleotides described above, *e.g.*, polynucleotide variants, fragments and hybridizing sequences, encode polypeptides that are immunologically cross-reactive with a polypeptide sequence specifically set forth herein. In other preferred embodiments, such polynucleotides encode polypeptides that have a level of immunogenic activity of at least about 50%, preferably

at least about 70%, and more preferably at least about 90% of that for a polypeptide sequence specifically set forth herein.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative polynucleotide segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

When comparing polynucleotide sequences, two sequences are said to be "identical" if the sequence of nucleotides in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins - Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989)

WO 02/04514

PCT/US01/22058

82

- CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy - the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and
- 5 Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988)

10 *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining

15 percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al. (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul et al. (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides of the invention. Software for performing

20 BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). Extension of the word hits in each direction are halted when: the cumulative alignment score falls off

25 by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of

30 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl.*

Acad. Sci. USA 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

It will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

Therefore, in another embodiment of the invention, a mutagenesis approach, such as site-specific mutagenesis, is employed for the preparation of immunogenic variants and/or derivatives of the polypeptides described herein. By this approach, specific modifications in a polypeptide sequence can be made through mutagenesis of the underlying polynucleotides that encode them. These techniques provides a straightforward approach to prepare and test sequence variants, for example,

incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the polynucleotide.

Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the immunogenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded

vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is
5 then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in
10 which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kubly, 1994; and
15 Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable
20 signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known
25 rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

30 In another approach for the production of polypeptide variants of the present invention, recursive sequence recombination, as described in U.S. Patent No.

WO 02/04514

PCT/US01/22058

86

5,837,458, may be employed. In this approach, iterative cycles of recombination and screening or selection are performed to "evolve" individual polynucleotide variants of the invention having, for example, enhanced immunogenic activity.

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequences set forth herein, or to any continuous portion of the sequences, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, e.g., one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at

WO 02/04514

PCT/US01/22058

88

temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to
5 prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species
10 can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a
15 method of choice depending on the desired results.

According to another embodiment of the present invention, polynucleotide compositions comprising antisense oligonucleotides are provided. Antisense oligonucleotides have been demonstrated to be effective and targeted inhibitors of protein synthesis, and, consequently, provide a therapeutic approach by
20 which a disease can be treated by inhibiting the synthesis of proteins that contribute to the disease. The efficacy of antisense oligonucleotides for inhibiting protein synthesis is well established. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829).
25 Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA_A receptor and human EGF (Jaskulski *et al.*, Science, 1988 Jun 10;240(4858):1544-6; Vasanthakumar and Ahmed, Cancer Commun. 1989;1(4):225-32; Peris *et al.*, Brain Res Mol Brain Res. 1998 Jun 15;57(2):310-20; U. S. Patent
30 5,801,154; U.S. Patent 5,789,573; U. S. Patent 5,718,709 and U.S. Patent 5,610,288). Antisense constructs have also been described that inhibit and can be used to treat a

WO 02/04514

PCT/US01/22058

89

variety of abnormal cellular proliferations, e.g. cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683).

Therefore, in certain embodiments, the present invention provides oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein. Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence and determination of secondary structure, T_m , binding energy, and relative stability. Antisense compositions may be selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell. Highly preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which are substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations can be performed, for example, using v.4 of the OLIGO primer analysis software and/or the BLASTN 2.0.5 algorithm software (Altschul *et al.*, Nucleic Acids Res. 1997, 25(17):3389-402).

The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, Nucleic Acids Res. 1997 Jul 15;25(14):2730-6). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%).

Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane.

According to another embodiment of the invention, the polynucleotide compositions described herein are used in the design and preparation of ribozyme molecules for inhibiting expression of the tumor polypeptides and proteins of the present invention in tumor cells. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, *Proc Natl Acad Sci U S A*. 1987 Dec;84(24):8788-92; Forster and Symons, *Cell*. 1987 Apr 24;49(2):211-20). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, *Cell*. 1981 Dec;27(3 Pt 2):487-96; Michel and Westhof, *J Mol Biol*. 1990 Dec 5;216(3):585-610; Reinhold-Hurek and Shub, *Nature*. 1992 May 14;357(6374):173-6). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme

WO 02/04514

PCT/US01/22058

91

necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, Proc Natl Acad Sci U S A. 1992 Aug 15;89(16):7305-9). Thus, the specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* Nucleic Acids Res. 1992 Sep 11;20(17):4559-65. Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz, Biochemistry 1989 Jun 13;28(12):4929-33; Hampel *et al.*, Nucleic Acids Res. 1990 Jan 25;18(2):299-304 and U. S. Patent 5,631,359. An example of the hepatitis δ virus motif is described by Perrotta and Becn, Biochemistry. 1992 Dec 1;31(47):11843-52; an example of the RNaseP motif is described by Guerrier-Takada *et al.*, Cell. 1983 Dec;35(3 Pt 2):849-57; Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, Cell. 1990 May 18;61(4):685-96; Saville and Collins, Proc Natl Acad Sci U S A. 1991 Oct 1;88(19):8826-30; Collins and Olive, Biochemistry. 1993 Mar 23;32(11):2795-9); and an example of the Group I intron is described in (U. S. Patent 4,987,071). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

WO 02/04514

PCT/US01/22058

92

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific
5 examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see e.g., Int. Pat. Appl. Publ. No. WO 92/07065; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO
10 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis
15 times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by
20 incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other
25 routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated
30 herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, etc.) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells. Ribozymes expressed from such promoters have been shown to function in mammalian cells. Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

In another embodiment of the invention, peptide nucleic acids (PNAs) compositions are provided. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, *Antisense Nucleic Acid Drug Dev.* 1997 7(4) 431-37). PNA is able to be utilized in a number of methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (*Trends Biotechnol* 1997 Jun;15(6):224-9). As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, *Science* 1991 Dec 6;254(5037):1497-500; Hanvey *et al.*, *Science*. 1992 Nov 27;258(5087):1481-5; Hyrup and Nielsen, *Bioorg Med Chem.* 1996 Jan;4(1):5-23). This chemistry has three important

consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc or Fmoc protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used.

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, Bioorg Med Chem. 1995 Apr;3(4):437-45). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography, providing yields and purity of product similar to those observed during the synthesis of peptides.

Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (for example, Norton *et al.*, Bioorg Med Chem. 1995 Apr;3(4):437-45; Petersen *et al.*, J Pept Sci. 1995 May-Jun;1(3):175-83; Orum *et al.*, Biotechniques. 1995 Sep;19(3):472-80; Footer *et al.*, Biochemistry. 1996 Aug 20;35(33):10673-9; Griffith *et al.*, Nucleic Acids Res. 1995 Aug 11;23(15):3003-8;

WO 02/04514

PCT/US01/22058

95

- Pardridge *et al.*, *Proc Natl Acad Sci U S A*. 1995 Jun 6;92(12):5592-6; Boffa *et al.*, *Proc Natl Acad Sci U S A*. 1995 Mar 14;92(6):1901-5; Gambacorti-Passerini *et al.*, *Blood*. 1996 Aug 15;88(4):1411-7; Armitage *et al.*, *Proc Natl Acad Sci U S A*. 1997 Nov 11;94(23):12320-5; Seeger *et al.*, *Biotechniques*. 1997 Sep;23(3):512-7. U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (*Anal Chem*. 1993 Dec 15;65(24):3545-9) and Jensen *et al.* (10 *Biochemistry*. 1997 Apr 22;36(16):5072-7). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIAcore™ technology.

Other applications of PNAs that have been described and will be 15 apparent to the skilled artisan include use in DNA strand invasion, antisense inhibition, mutational analysis, enhancers of transcription, nucleic acid purification, isolation of transcriptionally active genes, blocking of transcription factor binding, genome cleavage, biosensors, *in situ* hybridization, and the like.

Polynucleotide Identification, Characterization and Expression

- 20 Polynucleotides compositions of the present invention may be identified, prepared and/or manipulated using any of a variety of well established techniques (see generally, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989, and other like references). For example, a polynucleotide may be identified, as described in more detail below, by 25 screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using the microarray technology of Affymetrix, Inc. (Santa Clara, CA) according to the manufacturer's instructions (and essentially as described by Schena *et al.*, *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller *et al.*, *Proc. Natl. Acad. Sci. USA* 30

WO 02/04514

PCT/US01/22058

96

94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as tumor cells.

Many template dependent processes are available to amplify a target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCRTM) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCRTM, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCRTM amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Any of a number of other template dependent processes, many of which are variations of the PCRTM amplification technique, are readily known and available in the art. Illustratively, some such methods include the ligase chain reaction (referred to as LCR), described, for example, in Eur. Pat. Appl. Publ. No. 320,308 and U.S. Patent No. 4,883,750; Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880; Strand Displacement Amplification (SDA) and Repair Chain Reaction (RCR). Still other amplification methods are described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025. Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (PCT Intl. Pat. Appl. Publ. No. WO 88/10315), including nucleic acid sequence based amplification (NASBA) and 3SR. Eur. Pat. Appl. Publ. No. 329,822 describes a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA). PCT Intl. Pat. Appl.

Publ. No. WO 89/06700 describes a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. Other amplification methods such as "RACE" (Trohman, 1990), and "one-sided PCR" (Ohara, 5 1989) are also well-known to those of skill in the art.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (e.g., a tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. 10 Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (e.g., by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or 15 bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (see Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may 20 be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can 25 then be assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, amplification techniques, such as those described above, can be useful for obtaining a full length coding sequence from a partial cDNA sequence. One such amplification technique is inverse PCR (see Triglia et al., *Nucl. Acids Res.* 30 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and

used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of
5 amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer,
10 which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res. 19*:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

15 In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (e.g., NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences
20 may also be obtained by analysis of genomic fragments.

In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of
25 the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing
30 non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression

WO 02/04514

PCT/US01/22058

99

or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. et al. (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) *Proteins, Structures*

WO 02/04514

PCT/US01/22058

100

and Molecular Principles, WII Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described, for example, in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription

WO 02/04514

PCT/US01/22058

101

and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In
5 mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, any of a number of expression vectors may be
10 selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence
15 encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced; pIN vectors (Van Hecke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion
20 proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at
25 will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.* 153:516-544.

30 In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For

WO 02/04514

PCT/US01/22058

102

example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. et al. (1994) *Proc. Natl. Acad. Sci.* 91:3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the

WO 02/04514

PCT/US01/22058

103

ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, COS, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, aminometabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chloresulfuron and phosphinotricin acetyltransferase, respectively (Murty, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells that contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include,

for example, membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med.* 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other

WO 02/04514

PCT/US01/22058

106

recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow
5 purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAG extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to
10 facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase
15 cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; *DNA Cell Biol.* 12:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using
20 solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length
25 molecule.

Antibody Compositions, Fragments Thereof and Other Binding Agents

According to another aspect, the present invention further provides binding agents, such as antibodies and antigen-binding fragments thereof, that exhibit immunological binding to a tumor polypeptide disclosed herein, or to a portion, variant
30 or derivative thereof. An antibody, or antigen-binding fragment thereof, is said to

WO 02/04514

PCT/US01/22058

107

"specifically bind," "immunologically bind," and/or is "immunologically reactive" to a polypeptide of the invention if it reacts at a detectable level (within, for example, an ELISA assay) with the polypeptide, and does not react detectably with unrelated polypeptides under similar conditions.

5 Immunological binding, as used in this context, generally refers to the non-covalent interactions of the type which occur between an immunoglobulin molecule and an antigen for which the immunoglobulin is specific. The strength, or affinity of immunological binding interactions can be expressed in terms of the dissociation constant (K_d) of the interaction, wherein a smaller K_d represents a greater
10 affinity. Immunological binding properties of selected polypeptides can be quantified using methods well known in the art. One such method entails measuring the rates of antigen-binding site/antigen complex formation and dissociation, wherein those rates depend on the concentrations of the complex partners, the affinity of the interaction, and on geometric parameters that equally influence the rate in both directions. Thus, both
15 the "on rate constant" (K_{on}) and the "off rate constant" (K_{off}) can be determined by calculation of the concentrations and the actual rates of association and dissociation. The ratio of K_{off}/K_{on} enables cancellation of all parameters not related to affinity, and is thus equal to the dissociation constant K_d . See, generally, Davies et al. (1990) Annual Rev. Biochem. 59:439-473.

20 An "antigen-binding site," or "binding portion" of an antibody refers to the part of the immunoglobulin molecule that participates in antigen binding. The antigen binding site is formed by amino acid residues of the N-terminal variable ("V") regions of the heavy ("H") and light ("L") chains. Three highly divergent stretches within the V regions of the heavy and light chains are referred to as "hypervariable
25 regions" which are interposed between more conserved flanking stretches known as "framework regions," or "FRs". Thus the term "FR" refers to amino acid sequences which are naturally found between and adjacent to hypervariable regions in immunoglobulins. In an antibody molecule, the three hypervariable regions of a light chain and the three hypervariable regions of a heavy chain are disposed relative to each
30 other in three dimensional space to form an antigen-binding surface. The antigen-binding surface is complementary to the three-dimensional surface of a bound antigen,

WO 02/04514

PCT/US01/22058

108

and the three hypervariable regions of each of the heavy and light chains are referred to as "complementarity-determining regions," or "CDRs."

Binding agents may be further capable of differentiating between patients with and without a cancer, such as lung cancer, using the representative assays provided herein. For example, antibodies or other binding agents that bind to a tumor protein will preferably generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, more preferably at least about 30% of patients. Alternatively, or in addition, the antibody will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (e.g., blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. Preferably, a statistically significant number of samples with and without the disease will be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (e.g., mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier

WO 02/04514

PCT/US01/22058

109

protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

WO 02/04514

PCT/US01/22058

110

A number of therapeutically useful molecules are known in the art which comprise antigen-binding sites that are capable of exhibiting immunological binding properties of an antibody molecule. The proteolytic enzyme papain preferentially cleaves IgG molecules to yield several fragments, two of which (the "F(ab)" fragments) each comprise a covalent heterodimer that includes an intact antigen-binding site. The enzyme pepsin is able to cleave IgG molecules to provide several fragments, including the "F(ab')₂" fragment which comprises both antigen-binding sites. An "Fv" fragment can be produced by preferential proteolytic cleavage of an IgM, and on rare occasions IgG or IgA immunoglobulin molecule. Fv fragments are, however, more commonly derived using recombinant techniques known in the art. The Fv fragment includes a non-covalent V_H::V_L heterodimer including an antigen-binding site which retains much of the antigen recognition and binding capabilities of the native antibody molecule. Inbar et al. (1972) Proc. Nat. Acad. Sci. USA 69:2659-2662; Hochman et al. (1976) Biochem 15:2706-2710; and Ehrlich et al. (1980) Biochem 19:4091-4096.

A single chain Fv ("scFv") polypeptide is a covalently linked V_H::V_L heterodimer which is expressed from a gene fusion including V_H- and V_L-encoding genes linked by a peptide-encoding linker. Huston et al. (1988) Proc. Nat. Acad. Sci. USA 85(16):5879-5883. A number of methods have been described to discern chemical structures for converting the naturally aggregated--but chemically separated--light and heavy polypeptide chains from an antibody V region into an scFv molecule which will fold into a three dimensional structure substantially similar to the structure of an antigen-binding site. See, e.g., U.S. Pat. Nos. 5,091,513 and 5,132,405, to Huston et al.; and U.S. Pat. No. 4,946,778, to Ladner et al.

Each of the above-described molecules includes a heavy chain and a light chain CDR set, respectively interposed between a heavy chain and a light chain FR set which provide support to the CDRs and define the spatial relationship of the CDRs relative to each other. As used herein, the term "CDR set" refers to the three hypervariable regions of a heavy or light chain V region. Proceeding from the N-terminus of a heavy or light chain, these regions are denoted as "CDR1," "CDR2," and "CDR3" respectively. An antigen-binding site, therefore, includes six CDRs, comprising the CDR set from each of a heavy and a light chain V region. A polypeptide

WO 02/04514

PCT/US01/22058

111

comprising a single CDR, (e.g., a CDR1, CDR2 or CDR3) is referred to herein as a "molecular recognition unit." Crystallographic analysis of a number of antigen-antibody complexes has demonstrated that the amino acid residues of CDRs form extensive contact with bound antigen, wherein the most extensive antigen contact is with the heavy chain CDR3. Thus, the molecular recognition units are primarily responsible for the specificity of an antigen-binding site.

As used herein, the term "FR set" refers to the four flanking amino acid sequences which frame the CDRs of a CDR set of a heavy or light chain V region. Some FR residues may contact bound antigen; however, FRs are primarily responsible for folding the V region into the antigen-binding site, particularly the FR residues directly adjacent to the CDRs. Within FRs, certain amino residues and certain structural features are very highly conserved. In this regard, all V region sequences contain an internal disulfide loop of around 90 amino acid residues. When the V regions fold into a binding-site, the CDRs are displayed as projecting loop motifs which form an antigen-binding surface. It is generally recognized that there are conserved structural regions of FRs which influence the folded shape of the CDR loops into certain "canonical" structures--regardless of the precise CDR amino acid sequence. Further, certain FR residues are known to participate in non-covalent interdomain contacts which stabilize the interaction of the antibody heavy and light chains.

A number of "humanized" antibody molecules comprising an antigen-binding site derived from a non-human immunoglobulin have been described, including chimeric antibodies having rodent V regions and their associated CDRs fused to human constant domains (Winter et al. (1991) *Nature* 349:293-299; Lobuglio et al. (1989) *Proc. Nat. Acad. Sci. USA* 86:4220-4224; Shaw et al. (1987) *J Immunol.* 138:4534-4538; and Brown et al. (1987) *Cancer Res.* 47:3577-3583), rodent CDRs grafted into a human supporting FR prior to fusion with an appropriate human antibody constant domain (Riechmann et al. (1988) *Nature* 332:323-327; Verhoeven et al. (1988) *Science* 239:1534-1536; and Jones et al. (1986) *Nature* 321:522-525), and rodent CDRs supported by recombinantly veneered rodent FRs (European Patent Publication No. 519,596, published Dec. 23, 1992). These "humanized" molecules are designed to minimize unwanted immunological response toward rodent antihuman antibody

WO 02/04514

PCT/US01/22058

112

molecules which limits the duration and effectiveness of therapeutic applications of those moieties in human recipients.

As used herein, the terms "veneered FRs" and "recombinantly veneered FRs" refer to the selective replacement of FR residues from, e.g., a rodent heavy or light chain V region, with human FR residues in order to provide a xenogeneic molecule comprising an antigen-binding site which retains substantially all of the native FR polypeptide folding structure. Veneering techniques are based on the understanding that the ligand binding characteristics of an antigen-binding site are determined primarily by the structure and relative disposition of the heavy and light chain CDR sets within the antigen-binding surface. Davies et al. (1990) *Ann. Rev. Biochem.* 59:439-473. Thus, antigen binding specificity can be preserved in a humanized antibody only wherein the CDR structures, their interaction with each other, and their interaction with the rest of the V region domains are carefully maintained. By using veneering techniques, exterior (e.g., solvent-accessible) FR residues which are readily encountered by the immune system are selectively replaced with human residues to provide a hybrid molecule that comprises either a weakly immunogenic, or substantially non-immunogenic veneered surface.

The process of veneering makes use of the available sequence data for human antibody variable domains compiled by Kabat et al., in *Sequences of Proteins of Immunological Interest*, 4th ed., (U.S. Dept. of Health and Human Services, U.S. Government Printing Office, 1987), updates to the Kabat database, and other accessible U.S. and foreign databases (both nucleic acid and protein). Solvent accessibilities of V region amino acids can be deduced from the known three-dimensional structure for human and murine antibody fragments. There are two general steps in veneering a murine antigen-binding site. Initially, the FRs of the variable domains of an antibody molecule of interest are compared with corresponding FR sequences of human variable domains obtained from the above-identified sources. The most homologous human V regions are then compared residue by residue to corresponding murine amino acids. The residues in the murine FR which differ from the human counterpart are replaced by the residues present in the human moiety using recombinant techniques well known in the art. Residue switching is only carried out with moieties which are at least partially

WO 02/04514

PCT/US01/22058

113

exposed (solvent accessible), and care is exercised in the replacement of amino acid residues which may have a significant effect on the tertiary structure of V region domains, such as proline, glycine and charged amino acids.

In this manner, the resultant "vencered" murine antigen-binding sites are thus designed to retain the murine CDR residues, the residues substantially adjacent to the CDRs, the residues identified as buried or mostly buried (solvent inaccessible), the residues believed to participate in non-covalent (e.g., electrostatic and hydrophobic) contacts between heavy and light chain domains, and the residues from conserved structural regions of the FRs which are believed to influence the "canonical" tertiary structures of the CDR loops. These design criteria are then used to prepare recombinant nucleotide sequences which combine the CDRs of both the heavy and light chain of a murine antigen-binding site into human-appearing FRs that can be used to transfect mammalian cells for the expression of recombinant human antibodies which exhibit the antigen specificity of the murine antibody molecule.

In another embodiment of the invention, monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y , ^{125}I , ^{124}I , ^{131}I , ^{186}Re , ^{188}Re , ^{211}At , and ^{212}Bi . Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (e.g., covalently bonded) to a suitable monoclonal antibody either directly or indirectly (e.g., via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (e.g., a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an

WO 02/04514

PCT/US01/22058

114

antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which
5 otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups,
10 sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of
15 different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by
20 serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody.
25 Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent
30 bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides

WO 02/04514

PCT/US01/22058

115

such as aminodextran (e.g., U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (e.g., U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

T Cell Compositions

The present invention, in another aspect, provides T cells specific for a tumor polypeptide disclosed herein, or for a variant or derivative thereof. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a polypeptide, polynucleotide encoding a polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide of interest. Preferably, a tumor polypeptide or polynucleotide of the invention is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a polypeptide of the present invention if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For

example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen et al., *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (e.g., by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a tumor polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days will typically result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (e.g., TNF or IFN-γ) is indicative of T cell activation (see Coligan et al., *Current Protocols in Immunology*, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a tumor polypeptide, polynucleotide or polypeptide-expressing APC may be CD4⁺ and/or CD8⁺. Tumor polypeptide-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, CD4⁺ or CD8⁺ T cells that proliferate in response to a tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a tumor polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of the tumor polypeptide can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

T Cell Receptor Compositions

The T cell receptor (TCR) consists of 2 different, highly variable polypeptide chains, termed the T-cell receptor α and β chains, that are linked by a disulfide bond (Janeway, Travers, Walport. *Immunobiology*. Fourth Ed., 148-159. Elsevier Science Ltd/Garland Publishing, 1999). The α/β heterodimer complexes with the invariant CD3 chains at the cell membrane. This complex recognizes specific antigenic peptides bound to MHC molecules. The enormous diversity of TCR specificities is generated much like immunoglobulin diversity, through somatic gene rearrangement. The β chain genes contain over 50 variable (V), 2 diversity (D), over 10 joining (J) segments, and 2 constant region segments (C). The α chain genes contain over 70 V segments, and over 60 J segments but no D segments, as well as one C segment. During T cell development in the thymus, the D to J gene rearrangement of the β chain occurs, followed by the V gene segment rearrangement to the DJ. This functional VDJ _{β} exon is transcribed and spliced to join to a C _{β} . For the α chain, a V _{α} gene segment rearranges to a J _{α} gene segment to create the functional exon that is then transcribed and spliced to the C _{α} . Diversity is further increased during the recombination process by the random addition of P and N-nucleotides between the V, D, and J segments of the β chain and between the V and J segments in the α chain (Janeway, Travers, Walport. *Immunobiology*. Fourth Ed., 98 and 150. Elsevier Science Ltd/Garland Publishing, 1999).

The present invention, in another aspect, provides TCRs specific for a polypeptide disclosed herein, or for a variant or derivative thereof. In accordance with the present invention, polynucleotide and amino acid sequences are provided for the V-J or V-D-J junctional regions or parts thereof for the alpha and beta chains of the T-cell receptor which recognize tumor polypeptides described herein. In general, this aspect of the invention relates to T-cell receptors which recognize or bind tumor polypeptides presented in the context of MHC. In a preferred embodiment the tumor antigens recognized by the T-cell receptors comprise a polypeptide of the present invention. For example, cDNA encoding a TCR specific for a tumor peptide can be isolated from T cells specific for a tumor polypeptide using standard molecular biological and recombinant DNA techniques.

WO 02/04514

PCT/US01/22058

118

This invention further includes the T-cell receptors or analogs thereof having substantially the same function or activity as the T-cell receptors of this invention which recognize or bind tumor polypeptides. Such receptors include, but are not limited to, a fragment of the receptor, or a substitution, addition or deletion mutant of a T-cell receptor provided herein. This invention also encompasses polypeptides or peptides that are substantially homologous to the T-cell receptors provided herein or that retain substantially the same activity. The term "analog" includes any protein or polypeptide having an amino acid residue sequence substantially identical to the T-cell receptors provided herein in which one or more residues, preferably no more than 5 residues, more preferably no more than 25 residues have been conservatively substituted with a functionally similar residue and which displays the functional aspects of the T-cell receptor as described herein.

The present invention further provides for suitable mammalian host cells, for example, non-specific T cells, that are transfected with a polynucleotide encoding TCRs specific for a polypeptide described herein, thereby rendering the host cell specific for the polypeptide. The α and β chains of the TCR may be contained on separate expression vectors or alternatively, on a single expression vector that also contains an internal ribosome entry site (IRES) for cap-independent translation of the gene downstream of the IRES. Said host cells expressing TCRs specific for the polypeptide may be used, for example, for adoptive immunotherapy of lung cancer as discussed further below.

In further aspects of the present invention, cloned TCRs specific for a polypeptide recited herein may be used in a kit for the diagnosis of lung cancer. For example, the nucleic acid sequence or portions thereof, of tumor-specific TCRs can be used as probes or primers for the detection of expression of the rearranged genes encoding the specific TCR in a biological sample. Therefore, the present invention further provides for an assay for detecting messenger RNA or DNA encoding the TCR specific for a polypeptide. Pharmaceutical Compositions

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions

WO 02/04514

PCT/US01/22058

119

disclosed herein in pharmaceutically-acceptable carriers for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will be understood that, if desired, a composition as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other
5 proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from
10 host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Therefore, in another aspect of the present invention, pharmaceutical compositions are provided comprising one or more of the polynucleotide, polypeptide,
15 antibody, and/or T-cell compositions described herein in combination with a physiologically acceptable carrier. In certain preferred embodiments, the pharmaceutical compositions of the invention comprise immunogenic polynucleotide and/or polypeptide compositions of the invention for use in prophylactic and therapeutic vaccine applications. Vaccine preparation is generally described in, for example, M.F.
20 Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Generally, such compositions will comprise one or more polynucleotide and/or polypeptide compositions of the present invention in combination with one or more immunostimulants.

It will be apparent that any of the pharmaceutical compositions described
25 herein can contain pharmaceutically acceptable salts of the polynucleotides and polypeptides of the invention. Such salts can be prepared, for example, from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts).

30 In another embodiment, illustrative immunogenic compositions, *e.g.*, vaccine compositions, of the present invention comprise DNA encoding one or more of

WO 02/04514

PCT/US01/22058

120

the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the polynucleotide may be administered within any of a variety of delivery systems known to those of ordinary skill in the art. Indeed, numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate polynucleotide expression systems will, of course, contain the necessary regulatory DNA regulatory sequences for expression in a patient (such as a suitable promoter and terminating signal). Alternatively, bacterial delivery systems may involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope.

Therefore, in certain embodiments, polynucleotides encoding immunogenic polypeptides described herein are introduced into suitable mammalian host cells for expression using any of a number of known viral-based systems. In one illustrative embodiment, retroviruses provide a convenient and effective platform for gene delivery systems. A selected nucleotide sequence encoding a polypeptide of the present invention can be inserted into a vector and packaged in retroviral particles using techniques known in the art. The recombinant virus can then be isolated and delivered to a subject. A number of illustrative retroviral systems have been described (e.g., U.S. Pat. No. 5,219,740; Miller and Rosman (1989) *BioTechniques* 7:980-990; Miller, A. D. (1990) *Human Gene Therapy* 1:5-14; Scarpa et al. (1991) *Virology* 180:849-852; Burns et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:8033-8037; and Boris-Lawrie and Temin (1993) *Cur. Opin. Genet. Develop.* 3:102-109.

In addition, a number of illustrative adenovirus-based systems have also been described. Unlike retroviruses which integrate into the host genome, adenoviruses persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham (1986) *J. Virol.* 57:267-274; Bett et al. (1993) *J. Virol.* 67:5911-5921; Mittereder et al. (1994) *Human Gene Therapy* 5:717-729; Seth et al. (1994) *J. Virol.* 68:933-940; Barr et al. (1994) *Gene Therapy* 1:51-58; Berkner, K. L. (1988) *BioTechniques* 6:616-629; and Rich et al. (1993) *Human Gene Therapy* 4:461-476).

WO 02/04514

PCT/US01/22058

121

Various adeno-associated virus (AAV) vector systems have also been developed for polynucleotide delivery. AAV vectors can be readily constructed using techniques well known in the art. See, e.g., U.S. Pat. Nos. 5,173,414 and 5,139,941; International Publication Nos. WO 92/01070 and WO 93/03769; Lebikowski et al. (1988) *Molec. Cell. Biol.* 8:3988-3996; Vincent et al. (1990) *Vaccines 90* (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) *Current Opinion in Biotechnology* 3:533-539; Muzyczka, N. (1992) *Current Topics in Microbiol. and Immunol.* 158:97-129; Kotin, R. M. (1994) *Human Gene Therapy* 5:793-801; Shelling and Smith (1994) *Gene Therapy* 1:165-169; and Zhou et al. (1994) *J. Exp. Med.* 179:1867-1875.

Additional viral vectors useful for delivering the polynucleotides encoding polypeptides of the present invention by gene transfer include those derived from the pox family of viruses, such as vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the novel molecules can be constructed as follows. The DNA encoding a polypeptide is first inserted into an appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells which are simultaneously infected with vaccinia. Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the polypeptide of interest into the viral genome. The resulting TK^{sup}(-) recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

A vaccinia-based infection/transfection system can be conveniently used to provide for inducible, transient expression or coexpression of one or more polypeptides described herein in host cells of an organism. In this particular system, cells are first infected in vitro with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the polynucleotide or polynucleotides of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into polypeptide by the host translational machinery. The method provides for high level,

WO 02/04514

PCT/US01/22058

122

transient, cytoplasmic production of large quantities of RNA and its translation products. See, e.g., Elroy-Stein and Moss, *Proc. Natl. Acad. Sci. USA* (1990) 87:6743-6747; Fuerst et al. *Proc. Natl. Acad. Sci. USA* (1986) 83:8122-8126.

Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the coding sequences of interest. Recombinant avipox viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an Avipox vector is particularly desirable in human and other mammalian species since members of the Avipox genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant Avipoxviruses are known in the art and employ genetic recombination, as described above with respect to the production of vaccinia viruses. See, e.g., WO 91/12882; WO 89/03429; and WO 92/03545.

Any of a number of alphavirus vectors can also be used for delivery of polynucleotide compositions of the present invention, such as those vectors described in U.S. Patent Nos. 5,843,723; 6,015,686; 6,008,035 and 6,015,694. Certain vectors based on Venezuelan Equine Encephalitis (VEE) can also be used, illustrative examples of which can be found in U.S. Patent Nos. 5,505,947 and 5,643,576.

Moreover, molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al. *J. Biol. Chem.* (1993) 268:6866-6869 and Wagner et al. *Proc. Natl. Acad. Sci. USA* (1992) 89:6099-6103, can also be used for gene delivery under the invention.

Additional illustrative information on these and other known viral-based delivery systems can be found, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993.

WO 02/04514

PCT/US01/22058

123

In certain embodiments, a polynucleotide may be integrated into the genome of a target cell. This integration may be in the specific location and orientation via homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the polynucleotide may be stably maintained in the cell as a separate, episomal segment of DNA. Such polynucleotide segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. The manner in which the expression construct is delivered to a cell and where in the cell the polynucleotide remains is dependent on the type of expression construct employed.

In another embodiment of the invention, a polynucleotide is administered/delivered as "naked" DNA, for example as described in Ulmer et al., *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

In still another embodiment, a composition of the present invention can be delivered via a particle bombardment approach, many of which have been described. In one illustrative example, gas-driven particle acceleration can be achieved with devices such as those manufactured by Powderject Pharmaceuticals PLC (Oxford, UK) and Powderject Vaccines Inc. (Madison, WI), some examples of which are described in U.S. Patent Nos. 5,846,796; 6,010,478; 5,865,796; 5,584,807; and EP Patent No. 0500 799. This approach offers a needle-free delivery approach wherein a dry powder formulation of microscopic particles, such as polynucleotide or polypeptide particles, are accelerated to high speed within a helium gas jet generated by a hand held device, propelling the particles into a target tissue of interest.

In a related embodiment, other devices and methods that may be useful for gas-driven needle-less injection of compositions of the present invention include those provided by Bioject, Inc. (Portland, OR), some examples of which are described in U.S. Patent Nos. 4,790,824; 5,064,413; 5,312,335; 5,383,851; 5,399,163; 5,520,639 and 5,993,412.

According to another embodiment, the pharmaceutical compositions described herein will comprise one or more immunostimulants in addition to the immunogenic polynucleotide, polypeptide, antibody, T-cell and/or APC compositions of this invention. An immunostimulant refers to essentially any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. One preferred type of immunostimulant comprises an adjuvant. Many adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Certain adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF, interleukin-2, -7, -12, and other like growth factors, may also be used as adjuvants.

Within certain embodiments of the invention, the adjuvant composition is preferably one that induces an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (e.g., IFN- γ , TNF α , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (e.g., IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Certain preferred adjuvants for eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A, together with an aluminum salt. MPL[®] adjuvants are available from Corixa Corporation (Seattle, WA; see, for example, US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato et al., *Science* 273:352, 1996. Another preferred adjuvant comprises a saponin, such as Quil A, or derivatives thereof, including QS21 and QS7 (Aquila Biopharmaceuticals Inc., Framingham, MA); Escin; Digitonin; or *Gypsophila* or *Chenopodium quinoa* saponins. Other preferred formulations include more than one saponin in the adjuvant combinations of the present invention, for example combinations of at least two of the following group comprising QS21, QS7, Quil A, β -escin, or digitonin.

Alternatively the saponin formulations may be combined with vaccine vehicles composed of chitosan or other polycationic polymers, polylactide and polylactide-co-glycolide particles, poly-N-acetyl glucosamine-based polymer matrix, particles composed of polysaccharides or chemically modified polysaccharides, liposomes and lipid-based particles, particles composed of glycerol monoesters, etc. The saponins may also be formulated in the presence of cholesterol to form particulate structures such as liposomes or ISCOMs. Furthermore, the saponins may be formulated together with a polyoxyethylene ether or ester, in either a non-particulate solution or suspension, or in a particulate structure such as a paucilamellar liposome or ISCOM. The saponins may also be formulated with excipients such as Carbopol[®] to increase viscosity, or may be formulated in a dry powder form with a powder excipient such as lactose.

In one preferred embodiment, the adjuvant system includes the combination of a monophosphoryl lipid A and a saponin derivative, such as the combination of QS21 and 3D-MPL[®] adjuvant, as described in WO 94/00153, or a less

WO 02/04514

PCT/US01/22058

126

reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. Another particularly preferred adjuvant formulation employing QS21, 3D-MPL[®] adjuvant and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Another enhanced adjuvant system involves the combination of a CpG-containing oligonucleotide and a saponin derivative particularly the combination of CpG and QS21 is disclosed in WO 00/09159. Preferably the formulation additionally comprises an oil in water emulsion and tocopherol.

Additional illustrative adjuvants for use in the pharmaceutical compositions of the invention include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (e.g., SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Enhancyn[®]) (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties, and polyoxyethylene ether adjuvants such as those described in WO 99/52549A1.

Other preferred adjuvants include adjuvant molecules of the general formula



wherein, n is 1-50, A is a bond or $-\text{C}(\text{O})-$, R is C_{1-50} alkyl or Phenyl C_{1-50} alkyl.

One embodiment of the present invention consists of a vaccine formulation comprising a polyoxyethylene ether of general formula (I), wherein n is between 1 and 50, preferably 4-24, most preferably 9; the R component is C_{1-50} , preferably $\text{C}_4\text{-C}_{20}$ alkyl and most preferably C_{12} alkyl, and A is a bond. The concentration of the polyoxyethylene ethers should be in the range 0.1-20%, preferably from 0.1-10%, and most preferably in the range 0.1-1%. Preferred polyoxyethylene ethers are selected from the following group: polyoxyethylene-9-lauryl ether, polyoxyethylene-9-stearyl ether, polyoxyethylene-8-stearyl ether, polyoxyethylene-4-lauryl ether, polyoxyethylene-35-lauryl ether, and polyoxyethylene-23-lauryl ether.

WO 02/04514

PCT/US01/22058

127

Polyoxyethylene ethers such as polyoxyethylene lauryl ether are described in the Merck index (12th edition: entry 7717). These adjuvant molecules are described in WO 99/52549.

5 The polyoxyethylene ether according to the general formula (I) above may, if desired, be combined with another adjuvant. For example, a preferred adjuvant combination is preferably with CpG as described in the pending UK patent application GB 9820956.2.

According to another embodiment of this invention, an immunogenic composition described herein is delivered to a host via antigen presenting cells (APCs),
10 such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs
15 may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent
20 APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up,
25 process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called
30 exosomes) may be used within a vaccine (*see* Zitvogel et al., *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of
5 cytokines such as GM-CSF, IL-4, IL-13 and/or TNF α to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF α , CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation,
10 maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are
15 characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fc γ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (e.g., CD54 and CD11) and costimulatory molecules
20 (e.g., CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide of the invention (or portion or other variant thereof) such that the encoded polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a pharmaceutical composition comprising such transfected cells
25 may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997.
30 Antigen loading of dendritic cells may be achieved by incubating dendritic cells or

WO 02/04514

PCT/US01/22058

129

progenitor cells with the tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (e.g., vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (e.g., a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will typically vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, mucosal, intravenous, intracranial, intraperitoneal, subcutaneous and intramuscular administration.

Carriers for use within such pharmaceutical compositions are biocompatible, and may also be biodegradable. In certain embodiments, the formulation preferably provides a relatively constant level of active component release. In other embodiments, however, a more rapid rate of release immediately upon administration may be desired. The formulation of such compositions is well within the level of ordinary skill in the art using known techniques. Illustrative carriers useful in this regard include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other illustrative delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (e.g., a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (see e.g., U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

In another illustrative embodiment, biodegradable microspheres (e.g., polylactate polyglycolate) are employed as carriers for the compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763;

WO 02/04514

PCT/US01/22058

130

5,814,344, 5,407,609 and 5,942,252. Modified hepatitis B core protein carrier systems such as described in WO/99 40934, and references cited therein, will also be useful for many applications. Another illustrative carrier/delivery system employs a carrier comprising particulate-protein complexes, such as those described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

The pharmaceutical compositions of the invention will often further comprise one or more buffers (e.g., neutral buffered saline or phosphate buffered saline), carbohydrates (e.g., glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (e.g., aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate.

The pharmaceutical compositions described herein may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are typically sealed in such a way to preserve the sterility and stability of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

The development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including e.g., oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation, is well known in the art, some of which are briefly discussed below for general purposes of illustration.

In certain applications, the pharmaceutical compositions disclosed herein may be delivered via oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (see, for example, Mathiowitz *et al.*, Nature 1997 Mar 27;386(6623):410-4; Hwang *et al.*, Crit Rev Ther Drug Carrier Syst 5 1998;15(3):243-84; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451). Tablets, troches, pills, capsules and the like may also contain any of a variety of additional components, for example, a binder, such as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may 10 be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, 15 tablets, pills, or capsules may be coated with shellac, sugar, or both. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations will contain at least about 0.1% of the 20 active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared is such a way that a suitable dosage will be obtained in any given unit dose of the compound. 25 Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may 30 alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation.

WO 02/04514

PCT/US01/22058

132

Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants.

5 Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein *parenterally*, intravenously, intramuscularly, or even intraperitoneally. Such approaches are well known to the skilled artisan, some of which are further described, for example, in U. S. Patent 5,543,158; U. S. Patent 5,641,515
10 and U. S. Patent 5,399,363. In certain embodiments, solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils.
15 Under ordinary conditions of storage and use, these preparations generally will contain a preservative to prevent the growth of microorganisms.

Illustrative pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (for example, see U. S. Patent
20 5,466,468). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (e.g., glycerol, propylene glycol, and
25 liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and/or by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be
30 preferable to include isotonic agents, for example, sugars or sodium chloride.

WO 02/04514

PCT/US01/22058

133

Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

In one embodiment, for parenteral administration in an aqueous solution, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. Moreover, for human administration, preparations will of course preferably meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

In another embodiment of the invention, the compositions disclosed herein may be formulated in a neutral or salt form. Illustrative pharmaceutically-acceptable salts include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective.

The carriers can further comprise any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art.

Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions. The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human.

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described, *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent 5,804,212. Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, *J Controlled Release* 1998 Mar 2;52(1-2):81-7) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871) are also well-known in the pharmaceutical arts. Likewise, illustrative transmucosal drug delivery in the form of a polytetrafluoroethylene support matrix is described in U. S. Patent 5,780,045.

In certain embodiments, liposomes, nanocapsules, microparticles, lipid particles, vesicles, and the like, are used for the introduction of the compositions of the present invention into suitable host cells/organisms. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like. Alternatively, compositions of the present invention can be bound, either covalently or non-covalently, to the surface of such carrier vehicles.

The formation and use of liposome and liposome-like preparations as potential drug carriers is generally known to those of skill in the art (see for example, Lasic, *Trends Biotechnol* 1998 Jul;16(7):307-21; Takakura, *Nippon Rinsho* 1998 Mar;56(3):691-5; Chandran *et al.*, *Indian J Exp Biol* 1997 Aug;35(8):801-9; Margalit, *Crit Rev Ther Drug Carrier Syst* 1995;12(2-3):233-61; U.S. Patent 5,567,434; U.S. Patent 5,552,157; U.S. Patent 5,565,213; U.S. Patent 5,738,868 and U.S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally difficult to transfect by other procedures, including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, *J Biol Chem* 1990 Sep

25;265(27):16337-42; Muller *et al.*, DNA Cell Biol. 1990 Apr;9(3):221-9). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, various drugs, radiotherapeutic agents, enzymes, viruses, transcription factors, allosteric effectors and
5 the like, into a variety of cultured cell lines and animals. Furthermore, the use of liposomes does not appear to be associated with autoimmune responses or unacceptable toxicity after systemic delivery.

In certain embodiments, liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric
10 bilayer vesicles (also termed multilamellar vesicles (MLVs)).

Alternatively, in other embodiments, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (see, for example, Quintanar-Guerrero *et al.*, Drug Dev Ind Pharm.
15 1998 Dec;24(12):1113-28). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) may be designed using polymers able to be degraded *in vivo*. Such particles can be made as described, for example, by Couvreur *et al.*, Crit Rev Ther Drug Carrier Syst. 1988;5(1):1-20; zur Muhlen *et al.*, Eur J Pharm Biopharm. 1998 Mar;45(2):149-55; Zambaux *et al.* J
20 Controlled Release. 1998 Jan 2;50(1-3):31-40; and U. S. Patent 5,145,684.

Cancer Therapeutic Methods

Immunologic approaches to cancer therapy are based on the recognition that cancer cells can often evade the body's defenses against aberrant or foreign cells and molecules, and that these defenses might be therapeutically stimulated to regain the
25 lost ground, e.g. pgs. 623-648 in Klein, Immunology (Wiley-Interscience, New York, 1982). Numerous recent observations that various immune effectors can directly or indirectly inhibit growth of tumors has led to renewed interest in this approach to cancer therapy, e.g. Jager, et al., Oncology 2001;60(1):1-7; Renner, et al., Ann Hematol 2000 Dec;79(12):651-9.

WO 02/04514

PCT/US01/22058

136

Four basic cell types whose function has been associated with antitumor cell immunity and the elimination of tumor cells from the body are: i) B-lymphocytes which secrete immunoglobulins into the blood plasma for identifying and labeling the nonself invader cells; ii) monocytes which secrete the complement proteins that are responsible for lysing and processing the immunoglobulin-coated target invader cells; iii) natural killer lymphocytes having two mechanisms for the destruction of tumor cells, antibody-dependent cellular cytotoxicity and natural killing; and iv) T-lymphocytes possessing antigen-specific receptors and having the capacity to recognize a tumor cell carrying complementary marker molecules (Schreiber, H., 1989, in Fundamental Immunology (ed). W. E. Paul, pp. 923-955).

Cancer immunotherapy generally focuses on inducing humoral immune responses, cellular immune responses, or both. Moreover, it is well established that induction of CD4⁺ T helper cells is necessary in order to secondarily induce either antibodies or cytotoxic CD8⁺ T cells. Polypeptide antigens that are selective or ideally specific for cancer cells, particularly lung cancer cells, offer a powerful approach for inducing immune responses against lung cancer, and are an important aspect of the present invention.

Therefore, in further aspects of the present invention, the pharmaceutical compositions described herein may be used for the treatment of cancer, particularly for the immunotherapy of lung cancer. Within such methods, the pharmaceutical compositions described herein are administered to a patient, typically a warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. As discussed above, administration of the pharmaceutical compositions may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

WO 02/04514

PCT/US01/22058

137

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided
5 herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host
10 immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8⁺ cytotoxic T lymphocytes and CD4⁺ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody
15 receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

20 Monoclonal antibodies may be labeled with any of a variety of labels for desired selective usages in detection, diagnostic assays or therapeutic applications (as described in U.S. Patent Nos. 6,090,365; 6,015,542; 5,843,398; 5,595,721; and 4,708,930, hereby incorporated by reference in their entirety as if each was incorporated individually). In each case, the binding of the labelled monoclonal antibody to the
25 determinant site of the antigen will signal detection or delivery of a particular therapeutic agent to the antigenic determinant on the non-normal cell. A further object of this invention is to provide the specific monoclonal antibody suitably labelled for achieving such desired selective usages thereof.

Effector cells may generally be obtained in sufficient quantities for
30 adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with

retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand
5 antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a
10 polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented
15 with IL-2 (*see, for example, Cheever et al., Immunological Reviews 157:177, 1997*).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by
20 intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous,
25 intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when
30 administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response

can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (e.g., more frequent remissions, complete or
5 partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 µg to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

10 In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (e.g., more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in
15 preexisting immune responses to a tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

Cancer Detection and Diagnostic Compositions, Methods and Kits

20 In general, a cancer may be detected in a patient based on the presence of one or more lung tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as lung cancer. In addition, such
25 proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample.

Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of
30 a cancer. In general, a tumor sequence should be present at a level that is at least two-

fold, preferably three-fold, and more preferably five-fold or higher in tumor tissue than in normal tissue of the same type from which the tumor arose. Expression levels of a particular tumor sequence in tissue types different from that in which the tumor arose are irrelevant in certain diagnostic embodiments since the presence of tumor cells can
5 be confirmed by observation of predetermined differential expression levels, e.g., 2-fold, 5-fold, etc., in tumor tissue to expression levels in normal tissue of the same type.

Other differential expression patterns can be utilized advantageously for diagnostic purposes. For example, in one aspect of the invention, overexpression of a tumor sequence in tumor tissue and normal tissue of the same type, but not in other
10 normal tissue types, e.g. PBMCs, can be exploited diagnostically. In this case, the presence of metastatic tumor cells, for example in a sample taken from the circulation or some other tissue site different from that in which the tumor arose, can be identified and/or confirmed by detecting expression of the tumor sequence in the sample, for example using RT-PCR analysis. In many instances, it will be desired to enrich for
15 tumor cells in the sample of interest, e.g., PBMCs, using cell capture or other like techniques.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory,
20 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent
25 immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent
30 that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a

WO 02/04514

PCT/US01/22058

141

polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length lung tumor proteins and polypeptide portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 µg, and preferably about 100 ng to about 1 µg, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports

WO 02/04514

PCT/US01/22058

142

having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (see, e.g., Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

5 In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody
10 complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as
15 described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as
20 phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (i.e., incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with lung cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of
25 ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support
30 with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second

antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as lung cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered

positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

5 In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution
10 containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent.
15 Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the
20 biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 µg, and more preferably from about 50 ng to about
25 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to
30 those of ordinary skill in the art that the above protocols may be readily modified to use tumor polypeptides to detect antibodies that bind to such polypeptides in a biological

sample. The detection of such tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a tumor protein in a biological sample. Within
5 certain methods, a biological sample comprising CD4⁺ and/or CD8⁺ T cells isolated from a patient is incubated with a tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For
10 example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (e.g., 5 - 25 µg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of tumor polypeptide to serve as a control. For CD4⁺ T cells, activation is
15 preferably detected by evaluating proliferation of the T cells. For CD8⁺ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on
20 the level of mRNA encoding a tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the tumor protein. The amplified cDNA is
25 then separated and detected using techniques well known in the art, such as gel electrophoresis.

Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

30 To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%,

preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a tumor protein of the invention that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence as disclosed herein. Techniques for both PCR based assays and hybridization assays are well known in the art (see, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another aspect of the present invention, cell capture technologies may be used in conjunction, with, for example, real-time PCR to provide a more sensitive tool for detection of metastatic cells expressing lung tumor antigens. Detection of lung cancer cells in biological samples, e.g., bone marrow samples, peripheral blood, and small needle aspiration samples is desirable for diagnosis and prognosis in lung cancer patients.

Immunomagnetic beads coated with specific monoclonal antibodies to surface cell markers, or tetrameric antibody complexes, may be used to first enrich or positively select cancer cells in a sample. Various commercially available kits may be used, including Dynabeads® Epithelial Enrich (DynaL Biotech, Oslo, Norway),
 5 StemSep™ (StemCell Technologies, Inc., Vancouver, BC), and RosetteSep (StemCell Technologies). A skilled artisan will recognize that other methodologies and kits may also be used to enrich or positively select desired cell populations. Dynabeads® Epithelial Enrich contains magnetic beads coated with MAbs specific for two glycoprotein membrane antigens expressed on normal and neoplastic epithelial tissues.
 10 The coated beads may be added to a sample and the sample then applied to a magnet, thereby capturing the cells bound to the beads. The unwanted cells are washed away and the magnetically isolated cells eluted from the beads and used in further analyses.

RosetteSep can be used to enrich cells directly from a blood sample and consists of a cocktail of tetrameric antibodies that targets a variety of unwanted cells
 15 and crosslinks them to glycophorin A on red blood cells (RBC) present in the sample, forming rosettes. When centrifuged over Ficoll, targeted cells pellet along with the free RBC. The combination of antibodies in the depletion cocktail determines which cells will be removed and consequently which cells will be recovered. Antibodies that are available include, but are not limited to: CD2, CD3, CD4, CD5, CD8, CD10, CD11b,
 20 CD14, CD15, CD16, CD19, CD20, CD24, CD25, CD29, CD33, CD34, CD36, CD38, CD41, CD45, CD45RA, CD45RO, CD56, CD66B, CD66c, HLA-DR, IgE, and TCRαβ.

Additionally, it is contemplated in the present invention that MAbs specific for lung tumor antigens can be generated and used in a similar manner. For example, MAbs that bind to tumor-specific cell surface antigens may be conjugated to
 25 magnetic beads, or formulated in a tetrameric antibody complex, and used to enrich or positively select metastatic lung tumor cells from a sample. Once a sample is enriched or positively selected, cells may be lysed and RNA isolated. RNA may then be subjected to RT-PCR analysis using lung tumor-specific primers in a real-time PCR assay as described herein. One skilled in the art will recognize that enriched or selected
 30 populations of cells may be analyzed by other methods (*e.g. in situ* hybridization or flow cytometry).

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

10 Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

15 As noted above, to improve sensitivity, multiple tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided
20 herein may be combined with assays for other known tumor antigens.

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a
25 monoclonal antibody or fragment thereof that specifically binds to a tumor protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain
30 a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a tumor protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

10

EXAMPLE I

IDENTIFICATION AND CHARACTERIZATION OF LUNG TUMOR cDNAs

This Example illustrates the identification of cDNA molecules encoding lung tumor proteins.

15

A. Isolation of cDNA Sequences from Lung Adenocarcinoma Libraries using Conventional cDNA Library Subtraction

A human lung adenocarcinoma cDNA expression library was constructed from poly A⁺ RNA from patient tissues (# 40031486) using a Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning kit (BRL Life Technologies, Gaithersburg, MD) following the manufacturer's protocol. Specifically, lung carcinoma tissues were homogenized with polytron (Kinematica, Switzerland) and total RNA was extracted using Trizol reagent (BRL Life Technologies) as directed by the manufacturer. The poly A⁺ RNA was then purified using an oligo dT cellulose column as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989. First-strand cDNA was synthesized using the NotI/Oligo-dT18 primer. Double-stranded cDNA was synthesized, ligated with BstXI/EcoRI adaptors (Invitrogen, San Diego, CA) and digested with NotI. Following size fractionation with cDNA size fractionation columns (BRL Life Technologies), the cDNA was ligated into the BstXI/NotI site of pcDNA3.1 (Invitrogen) and transformed

20
25
30

WO 02/04514

PCT/US01/22058

150

into ElectroMax *E. coli* DH10B cells (BRL Life Technologies) by electroporation. A total of 3×10^6 independent colonies were generated.

Using the same procedure, a normal human cDNA expression library was prepared from a panel of normal tissue specimens, including lung, liver, pancreas, skin, kidney, brain and resting PBMC.

cDNA library subtraction was performed using the above lung adenocarcinoma and normal tissue cDNA libraries, as described by Hara *et al.* (*Blood*, 84:189-199, 1994) with some modifications. Specifically, a lung adenocarcinoma-specific subtracted cDNA library was generated as follows. The normal tissue cDNA library (80 µg) was digested with BamHI and XhoI, followed by a filling-in reaction with DNA polymerase Klenow fragment. After phenol-chloroform extraction and ethanol precipitation, the DNA was dissolved in 133 µl of H₂O, heat-denatured and mixed with 133 µl (133 µg) of Photoprobe biotin (Vector Laboratories, Burlingame, CA). As recommended by the manufacturer, the resulting mixture was irradiated with a 270 W sunlamp on ice for 20 minutes. Additional Photoprobe biotin (67 µl) was added and the biotinylation reaction was repeated. After extraction with butanol five times, the DNA was ethanol-precipitated and dissolved in 23 µl H₂O. The resulting DNA, plus other highly redundant cDNA clones that were frequently recovered in previous lung subtractions formed the driver DNA.

To form the tracer DNA, 10 µg lung adenocarcinoma cDNA library was digested with NotI and SpeI, phenol chloroform extracted and passed through Chroma spin-400 columns (Clontech, Palo Alto, CA). Typically, 5 µg of cDNA was recovered after the sizing column. Following ethanol precipitation, the tracer DNA was dissolved in 5 µl H₂O. Tracer DNA was mixed with 15 µl driver DNA and 20 µl of 2 x hybridization buffer (1.5 M NaCl/10 mM EDTA/50 mM HEPES pH 7.5/0.2% sodium dodecyl sulfate), overlaid with mineral oil, and heat-denatured completely. The sample was immediately transferred into a 68 °C water bath and incubated for 20 hours (long hybridization [LH]). The reaction mixture was then subjected to a streptavidin treatment followed by phenol/chloroform extraction. This process was repeated three more times. Subtracted DNA was precipitated, dissolved in 12 µl H₂O, mixed with 8 µl driver DNA and 20 µl of 2 x hybridization buffer, and subjected to a hybridization at

68°C for 2 hours (short hybridization [SH]). After removal of biotinylated double-stranded DNA, subtracted cDNA was ligated into NotI/SpeI site of chloramphenicol resistant pBCSK⁺ (Stratagene, La Jolla, CA) and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate a lung adenocarcinoma specific subtracted cDNA library, referred to as LAT-S1. Similarly, LAT-S2 was generated by including 23 genes that were over-expressed in the tracer as additional drivers.

A second human lung adenocarcinoma cDNA expression library was constructed using adenocarcinoma tissue from a second patient (# 86-66) and used to prepare a second lung adenocarcinoma-specific subtracted cDNA library (referred to as LAT2-S2), as described above, using the same panel of normal tissues and the additional genes over-expressed in LAT-S1.

A third human metastatic lung adenocarcinoma library was constructed from a pool of two lung pleural effusions with lung and gastric adenocarcinoma origins. The subtracted cDNA library, referred to as Mets-sub2, was generated as described above using the same panel of normal tissues. The Mets-sub3 subtracted library was constructed by including 51 additional genes as drivers. These 51 genes were recovered in Mets-sub2, representing over-expressed housekeeping genes in the testers.

A total of 16 cDNA fragments isolated from LAT-S1, 585 cDNA fragments isolated from LAT-S2, 568 cDNA clones from LAT2-S2, 15 cDNA clones from Mets-sub2 and 343 cDNA clones from Mets-sub3, described above, were colony PCR amplified and their mRNA expression levels in lung tumor, normal lung, and various other normal and tumor tissues were determined using microarray technology (Incyte, Palo Alto, CA). Briefly, the PCR amplification products were dotted onto slides in an array format, with each product occupying a unique location in the array. mRNA was extracted from the tissue sample to be tested, reverse transcribed, and fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured. This intensity correlates with the hybridization intensity. Seventy-three non-redundant cDNA clones, of which 42 were found to be unique, showed over-expression in lung tumors, with expression in normal tissues tested (lung, skin, lymph node, colon, liver, pancreas, breast, heart, bone marrow, large intestine, kidney, stomach, brain, small

WO 02/04514

PCT/US01/22058

152

intestine, bladder and salivary gland) being either undetectable, or at significantly lower levels compared to lung adenocarcinoma tumors. These clones were further characterized by DNA sequencing with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A and/or Model 377 (Foster City, CA).

5 The sequences were compared to known sequences in the gene bank using the EMBL GenBank databases (release 96). No significant homologies were found to the sequence provided in SEQ ID NO:67, with no apparent homology to previously identified expressed sequence tags (ESTs). The sequences of SEQ ID NO:60, 62, 65, 66, 69-71, 74, 76, 79, 80, 84, 86, 89-92, 95, 97 and 98 were found to
10 show some homology to previously identified expressed sequence tags (ESTs). The cDNA sequences of SEQ ID NO:59, 61, 63, 64, 67, 68, 72, 73, 75, 77, 78, 81-83, 85, 87, 88, 93, 94, 96, 99 and 100 showed homology to previously identified genes. The full-length cDNA sequences for the clones of SEQ ID NO:96 and 100 are provided in SEQ ID NO:316 and 318, respectively. The amino acid sequences for the clones of
15 SEQ ID NO: 59, 61, 63, 64, 68, 73, 82, 83, 94, 96 and 100 are provided in SEQ ID NO:331, 328, 329, 332, 327, 333, 330, 326, 325, 324 and 335, respectively. The amino acid sequence encoded by the sequence of SEQ ID NO:69 (referred to as L552S) is provided in SEQ ID NO:786.

Further studies led to the isolation of an extended cDNA sequence and
20 the open reading frame for L552S (SEQ ID NO:790). The amino acid sequence encoded by the cDNA sequence of SEQ ID NO:790 is provided in SEQ ID NO:791. Subsequent studies led to the isolation of the full-length cDNA sequence of L552S (SEQ ID NO:808). The full-length cDNA of L552S has an open-reading frame of 480 base pairs (SEQ ID NO:790) and encodes a putative polypeptide of 160 amino acids
25 (SEQ ID NO:809).

Initial database searches failed to detect any sequence homology with proteins in the database, suggesting that L552S encodes a novel protein of unknown function. Recently, a cancer-testis antigen, XAGE-1, was found to be over-expressed in Ewing's Sarcoma (Liu *et al.*, 2000 Cancer Res. 60:4752-4755). The determined cDNA
30 sequence of XAGE-1 is provided in SEQ ID NO:792 with the corresponding amino acid sequence being provided in SEQ ID NO:793. A sequence comparison of L552S and

XAGE-1 reveals striking identities as well as differences. The majority of the C-terminal sequences are identical to each other. The polypeptides predicted from L552S and XAGE-1 have diverged N-terminal sequences. Hydrophilicity analysis of the L552S amino acids suggested a very hydrophilic protein with no transmembrane domains predicted. PSORT analysis of L552S revealed the same result. Since L552S is also localized in chromosome X, it is likely to be a new isoform of cancer testis antigen, XAGE-1. The genomic sequence analysis revealed that both genes localized in the same region of the X chromosome and both have four exons. The last three exons are identical for L552S and XAGE-1. However, the first exon for XAGE-1 is upstream of the first exon for L552S and this results in distinct 5' nucleotide and amino acid sequences for L552S and XAGE-1. Therefore, L552S and XAGE-1 are alternatively spliced isoforms.

The amino acid sequence of this N-terminal portion of L552S is provided in SEQ ID NO:1830 with the corresponding cDNA sequence being provided in SEQ ID NO:1826. The cDNA sequences provided in SEQ ID NO:1827-1829 represent bp 394-681 of SEQ ID NO:808, bp394-534 of SEQ ID NO:808 and bp 214-394 of SEQ ID NO:792, respectively, with the corresponding amino acid sequences being provided in SEQ ID NO:1831-1833, respectively.

Full-length cloning efforts on L552S also led to the isolation of three additional cDNA sequences (SEQ ID NO:810-812; referred to as clones 50989, 50990 and 50992, respectively) from a metastatic lung adenocarcinoma library. The sequence of SEQ ID NO:810 was found to show some homology to previously identified human DNA sequences. The sequence of SEQ ID NO:811 was found to show some homology to a previously identified DNA sequence. The sequence of SEQ ID NO:812 was found to show some homology to previously identified ESTs.

The gene of SEQ ID NO:84 (referred to as L551S) was determined by real-time RT-PCR analysis to be over-expressed in 2/9 primary adenocarcinomas and to be expressed at lower levels in 2/2 metastatic adenocarcinomas and 1/2 squamous cell carcinomas. No expression was observed in normal tissues, with the exception of very low expression in normal stomach. Further studies on L551S led to the isolation of the 5' and 3' cDNA consensus sequences provided in SEQ ID NO:801 and 802,

WO 02/04514

PCT/US01/22058

154

respectively. The L551S 5' sequence was found to show some homology to the previously identified gene STY8 (cDNA sequence provided in SEQ ID NO:803; corresponding amino acid sequence provided in SEQ ID NO:805), which is a mitogen activated protein kinase phosphatase. However, no significant homologies were found to the 3' sequence of L551S. Subsequently, an extended cDNA sequence for L551S was isolated (SEQ ID NO:804). The corresponding amino acid sequence is provided in SEQ ID NO:806. Further studies led to the isolation of two independent full-length clones for L551S (referred to as 54298 and 54305). These two clones have five nucleotide differences compared to the STY8 DNA sequence. Two of these differences are single nucleotide polymorphisms which do not effect the encoded amino acid sequences. The other three nucleotide differences are consistent between the two L551S clones but lead to encoded amino acid sequences that are different from the STY8 protein sequence. The determined cDNA sequences for the L551S full-length clones 54305 and 54298 are provided in SEQ ID NO:825 and 826, respectively, with the amino acid sequence for L551S being provided in SEQ ID NO:827.

B. Isolation of cDNA Sequences from Lung Adenocarcinoma Libraries using PCR-Based cDNA Library Subtraction

cDNA clones from a subtracted library, containing cDNA from a pool of two human lung primary adenocarcinomas subtracted against a pool of nine normal human tissue cDNAs including skin, colon, lung, esophagus, brain, kidney, spleen, pancreas and liver, (Clontech, Palo Alto, CA) were submitted to a first round of PCR amplification. This library (referred to as ALT-1) was subjected to a second round of PCR amplification, following the manufacturer's protocol. The expression levels of 760 cDNA clones in lung tumor, normal lung, and various other normal and tumor tissues, were examined using microarray technology as described above. A total of 118 clones, of which 55 were unique, were found to be over-expressed in lung tumor tissue, with expression in normal tissues tested (lung, skin, lymph node, colon, liver, pancreas, breast, heart, bone marrow, large intestine, kidney, stomach, brain, small intestine, bladder and salivary gland) being either undetectable, or at significantly lower levels. The sequences were compared to known sequences in the gene bank using the EMBL

WO 02/04514

PCT/US01/22058

155

and GenBank databases (release 96). No significant homologies (including ESTs) were found to the sequence provided in SEQ ID NO:44. The sequences of SEQ ID NO:1, 11, 13, 15, 20, 23-27, 29, 30, 33, 34, 39, 41, 43, 45, 46, 51 and 57 were found to show some homology to previously identified expressed sequence tags (ESTs). The cDNA sequences of SEQ ID NO:2-10, 12, 14, 16-19, 21, 22, 28, 31, 32, 35-38, 40, 42, 44, 47-50, 52-56 and 58 showed homology to previously identified genes. The full-length cDNA sequences for the clones of SEQ ID NO:18, 22, 31, 35, 36 and 42 are provided in SEQ ID NO:320, 319, 323, 321, 317, 321 and 322, respectively, with the corresponding amino acid sequences being provided in SEQ ID NO: 337, 336, 340, 338, 334, and 339, respectively.

Further studies led to the isolation of an extended cDNA sequence for the clone of SEQ ID NO:33 (referred to as L801P). This extended cDNA sequence (provided in SEQ ID NO:796), was found to contain three potential open reading frames (ORFs). The predicted amino acid sequences encoded by these three ORFs are provided in SEQ ID NO:797-799, respectively. Additional full-length cloning efforts led to still further extended cDNA sequence for L801P, set forth in SEQ ID NO:1669, in addition to five potential open reading frames (referred to as ORFs 4-9; SEQ ID NO:1670-1675, respectively) encoded by the extended cDNA sequence. L801P was mapped to chromosomal region 20p13 and a 137 amino acid ORF from this genomic region was identified that corresponds to ORF4 (SEQ ID NO:1670), suggesting that this is likely an authentic ORF for L801P.

By microarray analysis, L801P was found to be overexpressed by 2-fold or greater in lung tumor tissue compared to normal tissue. By real-time PCR analysis, greater than 50% of lung adenocarcinoma and greater than 30% of lung squamous cell carcinoma tumor samples tested had elevated L801P expression relative to normal lung tissue. Of those that displayed elevated L801P, the level of expression was greater than 10-fold higher than in normal lung tissue samples. Moreover, low or no expression of L801P was detected in an extensive panel of normal tissue RNAs.

L801P expression was also detected in a number of other tumor types, including breast, prostate, ovarian and colon tumors, and thus may have diagnostic and/or therapeutic utility in these cancer types.

WO 02/04514

PCT/US01/22058

156

In subsequent studies, a full-length cDNA sequence for the clone of SEQ ID NO:44 (referred to as L844P) was isolated (provided in SEQ ID NO:800). Comparison of this sequence with those in the public databases revealed that the 470 bases at the 5' end of the sequence show homology to the known gene dihydrodiol dehydrogenase, thus indicating that L844P is a novel transcript of the dihydrodiol dehydrogenase family having 2007 base pairs of previously unidentified 3' untranslated region.

The predicted amino acid sequence encoded by the sequence of SEQ ID NO:46 (referred to as L840P) is provided in SEQ ID NO:787. An extended cDNA sequence for L840P, which was determined to include an open reading frame, is provided in SEQ ID NO:794. The predicted amino acid sequence encoded by the cDNA sequence of SEQ ID NO:794 is provided in SEQ ID NO:795. The full-length cDNA sequence for the clone of SEQ ID NO:54 (referred to as L548S) is provided in SEQ ID NO:788, with the corresponding amino acid sequence being provided in SEQ ID NO:789.

Northern blot analyses of the genes of SEQ ID NO:25 and 46 (referred to as L839P and L840P, respectively) were remarkably similar. Both genes were expressed in 1/2 lung adenocarcinomas as two bands of 3.6 kb and 1.6 kb. No expression of L839P was observed in normal lung or trachea. No expression of L840P was observed in normal bone marrow, resting or activated PBMC, esophagus, or normal lung. Given the similar expression patterns, L839P and L840P may be derived from the same gene.

Additional lung adenocarcinoma cDNA clones were isolated as follows. A cDNA library was prepared from a pool of two lung adenocarcinomas and subtracted against cDNA from a panel of normal tissues including lung, brain, liver, kidney, pancreas, skin, heart and spleen. The subtraction was performed using a PCR-based protocol (Clontech), which was modified to generate larger fragments. Within this protocol, tester and driver double stranded cDNA were separately digested with five restriction enzymes that recognize six-nucleotide restriction sites (MluI, MscI, PvuII, SalI and StuI). This digestion resulted in an average cDNA size of 600 bp, rather than the average size of 300 bp that results from digestion with RsaI according to the

Clontech protocol. The ends of the restriction digested tester cDNA were filled in to generate blunt ends for adapter ligation. This modification did not affect the subtraction efficiency. Two tester populations were then created with different adapters, and the driver library remained without adapters. The tester and driver libraries were then
5 hybridized using excess driver cDNA. In the first hybridization step, driver was separately hybridized with each of the two tester cDNA populations. This resulted in populations of (a) unhybridized tester cDNAs, (b) tester cDNAs hybridized to other tester cDNAs, (c) tester cDNAs hybridized to driver cDNAs and (d) unhybridized driver cDNAs. The two separate hybridization reactions were then combined, and
10 rehybridized in the presence of additional denatured driver cDNA. Following this second hybridization, in addition to populations (a) through (d), a fifth population (e) was generated in which tester cDNA with one adapter hybridized to tester cDNA with the second adapter. Accordingly, the second hybridization step resulted in enrichment of differentially expressed sequences which could be used as templates for PCR
15 amplification with adaptor-specific primers.

The ends were then filled in, and PCR amplification was performed using adaptor-specific primers. Only population (e), which contained tester cDNA that did not hybridize to driver cDNA, was amplified exponentially. A second PCR amplification step was then performed, to reduce background and further enrich
20 differentially expressed sequences.

Fifty-seven cDNA clones were isolated from the subtracted library (referred to as LAP1) and sequenced. The determined cDNA sequences for 16 of these clones are provided in SEQ ID NO:101-116. The sequences of SEQ ID NO:101 and 114 showed no significant homologies to previously identified sequences. The
25 sequences of SEQ ID NO:102-109 and 112 showed some similarity to previously identified sequences, while the sequences of SEQ ID NO:113, 115 and 116 showed some similarity to previously isolated ESTs.

An additional 502 clones analyzed from the LAP1 library were sequenced and the determined cDNA sequences are shown in SEQ ID NO:828-1239
30 and 1564-1653.

C. Isolation of cDNA Sequences from Small Cell Lung Carcinoma

Libraries using PCR-Based cDNA Library Subtraction

A subtracted cDNA library for small cell lung carcinoma (referred to as SCL1) was prepared essentially using the modified PCR-based subtraction process described above. cDNA from small cell lung carcinoma was subtracted against cDNA from a panel of normal tissues, including normal lung, brain, kidney, liver, pancreas, skin, heart, lymph node and spleen. Both tester and driver poly A⁺ RNA were initially amplified using SMART PCR cDNA synthesis kit (Clontech, Palo Alto, CA). The tester and driver double stranded cDNA were separately digested with five restriction enzymes (DraI, MscI, PvuII, SmaI, and StuI). These restriction enzymes generated blunt end cuts and the digestion resulted in an average insert size of 600 bp. Digestion with this set of restriction enzymes eliminates the step required to generate blunt ends by filling in of the cDNA ends. These modifications did not affect subtraction efficiency.

Eighty-five clones were isolated and sequenced. The determined cDNA sequences for 31 of these clones are provided in SEQ ID NO:117-147. The sequences of SEQ ID NO:122, 124, 126, 127, 130, 131, 133, 136, 139 and 147 showed no significant homologies to previously identified sequences. The sequences of SEQ ID NO:120, 129, 135, 137, 140, 142, 144 and 145 showed some similarity to previously identified gene sequences, while the sequences of SEQ ID NO:114, 118, 119, 121, 123, 125, 128, 132, 134, 138, 141, 143 and 147 showed some similarity to previously isolated ESTs.

In further studies, three additional cDNA libraries were generated from poly A⁺ RNA from a single small cell lung carcinoma sample subtracted against a pool of poly A⁺ RNA from nine normal tissues (lung, brain, kidney, liver, pancreas, skin, heart pituitary gland and spleen). For the first library (referred to as SCL2), the subtraction was carried out essentially as described above for the LAP1 library, with the exception that the tester and driver were digested with PvuII, StuI, MscI and DraI. The ratio of tester and driver cDNA used was as recommended by Clontech. For the second library (referred to as SCL3), subtraction was performed essentially as for SCL2 except that cDNA for highly redundant clones identified from the SCL2 library was included in the driver cDNA. Construction of the SCL4 library was performed essentially as

WO 02/04514

PCT/US01/22058

159

described for the SCL3 library except that a higher ratio of driver to tester was employed.

Each library was characterized by DNA sequencing and database analyses. The determined cDNA sequence for 35 clones isolated from the SCL2 library are provided in SEQ ID NO:245-279, with the determined cDNA sequences for 21 clones isolated from the SCL3 library and for 15 clones isolated from the SCL4 library being provided in SEQ ID NO:280-300 and 301-315, respectively. The sequences of SEQ ID NO:246, 254, 261, 262, 304, 309 and 311 showed no significant homologies to previously identified sequences. The sequence of SEQ ID NO:245, 248, 255, 266, 270, 275, 280, 282, 283, 288-290, 292, 295, 301 and 303 showed some homology to previously isolated ESTs, while the sequences of SEQ ID NO:247, 249-253, 256-260, 263-265, 267-269, 271-274, 276-279, 281, 284-287, 291, 293, 294, 296-300, 302, 305-308, 310 and 312-315 showed some homology to previously identified gene sequences.

Sequences disclosed herein were further evaluated for overexpression in specific tumor tissues by microarray analysis. Using this approach, cDNA sequences were PCR amplified and their mRNA expression profiles in tumor and normal tissues were examined using cDNA microarray technology essentially as described (Shena, M. *et al.*, 1995 Science 270:467-70). In brief, the clones were arrayed onto glass slides as multiple replicas, with each location corresponding to a unique cDNA clone (as many as 5500 clones can be arrayed on a single slide or chip). Each chip was hybridized with a pair of cDNA probes that are fluorescence-labeled with Cy3 and Cy5, respectively. Typically, 1µg of polyA⁺ RNA was used to generate each cDNA probe. After hybridization, the chips were scanned and the fluorescence intensity recorded for both Cy3 and Cy5 channels. There were multiple built-in quality control steps. First, the probe quality was monitored using a panel of ubiquitously expressed genes. Secondly, the control plate also can include yeast DNA fragments of which complementary RNA may be spiked into the probe synthesis for measuring the quality of the probe and the sensitivity of the analysis. Currently, the technology offers a sensitivity of 1 in 100,000 copies of mRNA. Finally, the reproducibility of this technology can be ensured by including duplicated control cDNA elements at different locations.

WO 02/04514

PCT/US01/22058

160

3264 cDNA clones from three PCR-based subtracted cDNA libraries were analyzed by the above cDNA microarray technology. The cDNA clones were arrayed on Lung Chip 5. Of these cDNA clones, 960 clones came from SQL1 library, 768 clones came from SCL1 library, and 1536 clones came from SCL3 and SCL4 libraries. Thirty-five pairs of fluorescent labeled cDNA probes were used for the microarray analysis. Each probe pair included a lung tumor probe paired with a normal tissue probe. The expression data was analyzed. 498 cDNA clones were found to be overexpressed by 2-fold or greater in the small cell and/or non-small cell lung tumor probe groups compared to the normal tissue probe group. Also, the mean expression values for these clones in normal tissues were below 0.1 (range of expression is from 0.001 to 10). The cDNA sequences disclosed in SEQ ID NO:1240-1563 represent 324 non-redundant clones.

The following sequences were novel based on database analysis including GenBank and GeneSeq: SEQ ID NOs:1240, 1243, 1247, 1269, 1272, 1280, 1283, 1285, 1286, 1289, 1300, 1309, 1318, 1319, 1327, 1335, 1339, 1346, 1359, 1369, 1370, 1371, 1393, 1398, 1405, 1408, 1413, 1414, 1417, 1422, 1429, 1432, 1435, 1436, 1438-1442, 1447, 1450, 1453, 1463, 1467, 1470, 1473, 1475, 1482, 1486, 1491-1494, 1501, 1505, 1506, 1514-1517, 1520, 1522, 1524, 1535, 1538, 1542, 1543, 1547, 1554, 1557, 1559, 1561, and 1563.

The extended cDNA sequence of the partial sequence of contig 139 (SEQ ID NO:1467), also known as L985P, was predicted by searching public databases using SEQ ID NO:1467 as a query. By this approach, it was found that SEQ ID NO:1467 had homology to a cDNA sequence (SEQ ID NO:1676) which encodes the cell surface immunomodulator-2 (CSIMM-2). The cDNA sequence of SEQ ID NO:1676 encodes a protein having the sequence set forth in SEQ ID NO:1677.

By microarray analysis, L985P was overexpressed by 2-fold or greater in the lung tumor probe groups compared to the normal tissue probe group. Moreover, the mean expression values for L985P in normal tissues was below 0.2 (range of expression was from 0.01 to 10). By real-time PCR analysis, greater than 40% of small cell lung carcinoma lung tumor samples tested had elevated L985P expression relative to normal lung tissue. Of those that displayed elevated L985P, the level of expression was greater

than 3-fold higher than in normal lung tissue samples. Low or no expression of L985P was detected in an extensive panel of normal tissue RNAs. These findings for L985P support its use both as a diagnostic marker for detecting the presence of lung cancer in a patient and/or as an immunotherapeutic target for the treatment of lung cancer.

5 D. Isolation of cDNA Sequences from a Neuroendocrine Library using
 PCR-Based cDNA Library Subtraction

 Using the modified PCR-based subtraction process, essentially as described above for the LAP1 subtracted library, a subtracted cDNA library (referred to as MLN1) was derived from a lung neuroendocrine carcinoma that had metastasized to the subcarinal lymph node, by subtraction with a panel of nine normal tissues, including
10 normal lung, brain, kidney, liver, pancreas, skin, heart, lymph node and spleen.

 Ninety-one individual clones were isolated and sequenced. The determined cDNA sequences for 58 of these clones are provided in SEQ ID NO:147-222. The sequences of SEQ ID NO:150, 151, 154, 157, 158, 159, 160, 163, 174, 175,
15 178, 186-190, 192, 193, 195-200, 208-210, 212-215 and 220 showed no significant homologies to previously identified sequences. The sequences of SEQ ID NO:152, 155, 156, 161, 165, 166, 176, 179, 182, 184, 185, 191, 194, 221 and 222 showed some similarity to previously identified gene sequences, while the sequences of SEQ ID NO:148, 149, 153, 164, 167-173, 177, 180, 181, 183, 201-207, 211 and 216-219
20 showed some similarity to previously isolated ESTs.

 The determined cDNA sequences of an additional 442 clones isolated from the MLN1 library are provided in SEQ ID NO:341-782, with the determined cDNA sequences of an additional 11 clones isolated from the MLN1 library are provided in SEQ ID NO:1654-1664.

25 E. Isolation of cDNA Sequences from a Squamous Cell Lung Carcinoma
 Library using PCR-Based cDNA Library Subtraction

 A subtracted cDNA library for squamous cell lung carcinoma (referred to as SQL1) was prepared essentially using the modified PCR-based subtraction process described above, except the tester and driver double stranded cDNA were separately

WO 02/04514

PCT/US01/22058

162

digested with four restriction enzymes (DraI, MseI, PvuII and StuI). cDNA from a pool of two squamous cell lung carcinomas was subtracted against cDNA from a pool of 10 normal tissues, including normal lung, brain, kidney, liver, pancreas, skin, heart, spleen, esophagus and trachea.

5 Seventy-four clones were isolated and sequenced. The determined cDNA sequences for 22 of these clones are provided in SEQ ID NO:223-244. The sequence of SEQ ID NO:241 showed no significant homologies to previously identified sequences. The sequences of SEQ ID NO:223, 225, 232, 233, 235, 238, 239, 242 and 243 showed some similarity to previously identified gene sequences, while the
10 sequences of SEQ ID NO:224, 226-231, 234, 236, 237, 240, 241 and 244 showed some similarity to previously isolated ESTs.

 The sequences of an additional 12 clones isolated during characterization of cDNA libraries prepared from lung tumor tissue are provided in SEQ ID NO:813-824. Comparison of these sequences with those in the GenBank database and the
15 GeneSeq DNA database revealed no significant homologies to previously identified sequences.

EXAMPLE 2

SYNTHESIS OF POLYPEPTIDES

20

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems Division 430A peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a
25 method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water :phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing
30 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water

WO 02/04514

PCT/US01/22058

163

(containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

5

EXAMPLE 3

EXPRESSION IN *E. COLI* OF L548S HIS TAG FUSION PROTEIN

The L548S coding region was PCR amplified with the following
10 primers:

Forward primer starting at amino acid 2:

PDM-433: 5' gctaaaggtagaccocnagaaaccaang 3' Tm 60°C (SEQ ID NO:1665)

Reverse primer creating a XhoI site after the stop codon:

15 PDM-438: 5' ctattaactcgaggagacagataaacagttcttta 3' Tm 61°C (SEQ ID NO:1666)

The PCR product was then digested with XhoI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and XhoI restriction enzymes. The correct construct was
20 confirmed by DNA sequence analysis and then transformed into BL21 (DE3) pLys S and BL21 (DE3) CodonPlus RIL expression hosts. The protein sequence of expressed recombinant L548S is shown in SEQ ID NO:1667, and the DNA sequence of expressed recombinant L7548S is shown in SEQ ID NO:1668.

25

EXAMPLE 4

ADDITIONAL ANALYSES OF LUNG CHIP 5 SQL1, SCL1, SCL3 AND SCL4 LIBRARIES

Additional analyses were performed on lung chip 5 using a criteria of
30 greater than or equal to 2-fold over-expression in tumor probe groups versus normal tissues and an average expression in normal tissues of less than or equal to 0.2. This resulted in the identification of 109 non-redundant clones that are over-expressed in

WO 02/04514

PCT/US01/22058

164

lung carcinomas. As summarized in the Table 11 below, 19 cDNA clones were recovered from the lung squamous cell carcinoma subtracted library SQL1, 9 cDNA clones were recovered from the small cell lung carcinoma library SCL1, and 81 cDNA clones were recovered from the small cell lung carcinoma libraries SCL3 and SCL4.

5

Table 11

SEQ ID NO:	Seq. Ref.	Element (384)	Element (96)	Ratio	Mean Signal 1	Mean Signal 2	Library
1680	58456	p0003r03c13	R0001 E7	3.09	0.424	0.137	SQL1
1681	58458	p0003r03c10	R0001 F5	2.31	0.408	0.176	SQL1
1682	58462	p0003r04c16	R0001 H8	2.22	0.257	0.116	SQL1
1683	58469	p0003r07c12	R0002 F6	2.1	0.289	0.133	SQL1
1684	58470	p0003r08c21	R0003 A11	2.55	0.493	0.194	SQL1
1685	58482	p0003r12c19	R0003 G10	2.16	0.36	0.167	SQL1
1686	58485	p0003r12c10	R0003 H5	2.48	0.273	0.11	SQL1
1687	58501	p0004r04c23	R0005 G12	2.04	0.26	0.128	SQL1
1688	58502	p0004r04c03	R0005 G2	2.17	0.289	0.133	SQL1
1689	58505	p0004r05c23	R0005 A12	3.08	0.454	0.148	SQL1
1690	58507	p0004r05c11	R0005 C6	3.22	0.49	0.152	SQL1
1691	58509	p0004r07c15	R0006 E8	3.26	0.421	0.129	SQL1
1692	58512	p0004r09c03	R0007 A2	3.16	0.559	0.177	SQL1
1693	58527	p0004r12c22	R0007 H11	2.03	0.278	0.137	SQL1
1694	58529	p0004r14c09	R0008 C5	2.26	0.45	0.199	SQL1
1695	58531	p0004r16c01	R0008 G1	2.94	0.387	0.136	SQL1
1696	58537	p0005r02c08	R0009 D4	2.03	0.355	0.175	SQL1
1697	58539	p0005r03c08	R0009 F4	2.34	0.42	0.18	SQL1
1698	58545	p0005r07c21	R0010 E11	2.96	0.361	0.122	SQL1
1699	59319	p0005r10c04	R0011 D2	3.1	0.478	0.154	SCL1
1700	59322	p0005r12c01	R0011 G1	2.16	0.255	0.118	SCL1
1701	59348	p0006r11c12	R0015 F6	2.33	0.269	0.116	SCL1
1702	59350	p0006r14c13	R0016 C7	2.41	0.447	0.185	SCL1
1703	59363	p0007r02c16	R0017 D8	2.12	0.421	0.199	SCL1
1704	59365	p0007r03c20	R0017 F10	3.07	0.584	0.19	SCL1
1705	59370	p0007r04c10	R0017 H5	2.06	0.284	0.138	SCL1
1706	59373	p0007r05c23	R0018 A12	2.95	0.472	0.16	SCL1
1707	59376	p0007r08c02	R0018 D1	2.13	0.246	0.116	SCL1
1708	61050	p0011r02c10	R0033 D5	2.23	0.306	0.137	SCL3/4
1709	61051	p0011r03c23	R0033 E12	2.9	0.298	0.103	SCL3/4
1710	61052	p0011r03c08	R0033 F4	2.18	0.265	0.122	SCL3/4
1711	61054	p0011r03c16	R0033 F8	2.11	0.415	0.197	SCL3/4
1712	61056	p0011r04c13	R0033 G7	2.73	0.314	0.115	SCL3/4
1713	61057	p0011r04c10	R0033 H5	2.45	0.463	0.189	SCL3/4
1714	61060	p0011r05c11	R0034 A6	3.28	0.536	0.164	SCL3/4
1715	61062	p0011r06c21	R0034 C11	2.73	0.526	0.192	SCL3/4
1716	61063	p0011r06c05	R0034 C3	3.61	0.513	0.142	SCL3/4

WO 02/04514

PCT/US01/22058

165

SEQ ID NO:	Seq. Ref.	Element (384)	Element (96)	Ratio	Mean Signal 1	Mean Signal 2	Library
1717	61064	p0011r06c04	R0034 D2	2.58	0.477	0.185	SCL3/4
1718	61065	p0011r06c14	R0034 D7	4.91	0.55	0.112	SCL3/4
1719	61066	p0011r06c18	R0034 D9	2.38	0.285	0.12	SCL3/4
1720	61069	p0011r07c16	R0034 F8	2.25	0.426	0.189	SCL3/4
1721	61070	p0011r08c21	R0034 G11	2	0.234	0.117	SCL3/4
1722	61071	p0011r08c03	R0034 G2	2.78	0.321	0.116	SCL3/4
1723	61074	p0011r08c16	R0034 H8	3.02	0.399	0.132	SCL3/4
1724	61075	p0011r09c05	R0035 A3	3.83	0.498	0.13	SCL3/4
1725	61077	p0011r10c21	R0035 C11	2.12	0.306	0.144	SCL3/4
1726	61079	p0011r11c23	R0035 E12	2.04	0.22	0.108	SCL3/4
1727	61080	p0011r11c15	R0035 E8	2.76	0.299	0.108	SCL3/4
1728	61081	p0011r11c14	R0035 F7	2.37	0.303	0.128	SCL3/4
1729	61083	p0011r12c15	R0035 G8	2.29	0.351	0.153	SCL3/4
1730	61085	p0011r13c05	R0036 A3	2.62	0.43	0.154	SCL3/4
1731	61086	p0011r13c09	R0036 A5	2.53	0.398	0.157	SCL3/4
1732	61088	p0011r14c05	R0036 C3	4.26	0.702	0.185	SCL3/4
1733	61090	p0011r15c07	R0036 E4	3.16	0.429	0.136	SCL3/4
1734	61091	p0011r15c16	R0036 H8	3.54	0.634	0.178	SCL3/4
1735	61093	p0012r02c03	R0037 C2	2.2	0.265	0.121	SCL3/4
1736	61094	p0012r02c11	R0037 C6	15.17	1.79	0.116	SCL3/4
1737	61096	p0012r02c08	R0037 D4	2.44	0.27	0.111	SCL3/4
1738	61097	p0012r02c10	R0037 D5	4.52	0.81	0.179	SCL3/4
1739	61099	p0012r03c02	R0037 F1	3.34	0.39	0.117	SCL3/4
1740	61100	p0012r03c06	R0037 F3	2.03	0.233	0.114	SCL3/4
1741	61103	p0012r04c17	R0037 G9	2.48	0.413	0.167	SCL3/4
1742	61105	p0012r05c11	R0038 A6	3.26	0.501	0.154	SCL3/4
1743	61106	p0012r05c06	R0038 B4	2.46	0.354	0.144	SCL3/4
1744	61110	p0012r06c15	R0038 C8	2.18	0.41	0.188	SCL3/4
1745	61113	p0012r07c09	R0038 E5	2.47	0.376	0.162	SCL3/4
1746	61115	p0012r07c13	R0038 E7	2.57	0.463	0.188	SCL3/4
1747	61117	p0012r07c24	R0038 F12	2.18	0.235	0.109	SCL3/4
1748	61118	p0012r07c18	R0038 F9	4.44	0.605	0.136	SCL3/4
1749	61119	p0012r08c03	R0038 G2	2.97	0.35	0.118	SCL3/4
1750	61120	p0012r08c07	R0038 G4	2.23	0.323	0.144	SCL3/4
1751	61122	p0012r08c18	R0038 H9	2.23	0.373	0.168	SCL3/4
1752	61125	p0012r10c17	R0039 C9	2.1	0.22	0.105	SCL3/4
1753	61126	p0012r10c16	R0039 D8	2.47	0.345	0.14	SCL3/4
1754	61130	p0012r12c12	R0039 H6	2.66	0.282	0.106	SCL3/4
1755	61133	p0012r13c24	R0040 B12	2.25	0.27	0.12	SCL3/4
1756	61134	p0012r14c23	R0040 C12	2.23	0.228	0.102	SCL3/4
1757	61135	p0012r14c03	R0040 C2	2.05	0.298	0.146	SCL3/4
1758	61137	p0012r14c02	R0040 D1	8.63	1.463	0.17	SCL3/4
1759	61139	p0012r14c14	R0040 D7	2.89	0.3	0.111	SCL3/4
1760	61143	p0012r16c02	R0040 H1	2.55	0.318	0.125	SCL3/4
1761	61144	p0012r16c18	R0040 H9	2.85	0.318	0.112	SCL3/4
1762	61148	p0013r02c19	R0041 C10	2.33	0.463	0.189	SCL3/4
1763	61151	p0013r02c03	R0041 C2	2.25	0.336	0.149	SCL3/4
1764	61155	p0013r04c07	R0041 G4	2.13	0.366	0.171	SCL3/4

WO 02/04514

PCT/US01/22058

166

SEQ ID NO:	Seq. Ref.	Element (354)	Element (96)	Ratio	Mean Signal 1	Mean Signal 2	Library
1765	61156	p0013r05c05	R0042 A3	2.73	0.38	0.139	SCL3/4
1766	61159	p0013r06c24	R0042 D12	4.57	0.831	0.182	SCL3/4
1767	61160	p0013r07c19	R0042 E10	8.6	1.191	0.138	SCL3/4
1768	61163	p0013r07c18	R0042 F9	2.18	0.278	0.128	SCL3/4
1769	61167	p0013r10c12	R0043 D6	3.13	0.38	0.124	SCL3/4
1770	61172	p0013r12c03	R0043 G2	2	0.398	0.198	SCL3/4
1771	61173	p0013r12c07	R0043 G4	3.73	0.72	0.193	SCL3/4
1772	61176	p0013r13c04	R0044 B2	2.34	0.446	0.19	SCL3/4
1773	61177	p0013r14c01	R0044 C1	3.9	0.539	0.138	SCL3/4
1774	61183	p0013r15c14	R0044 F7	5.49	0.959	0.175	SCL3/4
1775	61185	p0013r16c24	R0044 H12	2.25	0.409	0.182	SCL3/4
1776	61188	p0014r01c07	R0045 A4	2.14	0.271	0.127	SCL3/4
1777	61192	p0014r02c19	R0045 C10	2.33	0.321	0.138	SCL3/4
1778	61198	p0014r04c24	R0045 H12	2.3	0.321	0.14	SCL3/4
1779	61201	p0014r06c22	R0046 D11	2.43	0.269	0.111	SCL3/4
1780	61202	p0014r06c09	R0046 D4	2.57	0.346	0.135	SCL3/4
1781	61204	p0014r07c07	R0046 E4	4.27	0.516	0.121	SCL3/4
1782	61208	p0014r07c12	R0046 F6	2.18	0.364	0.167	SCL3/4
1783	61210	p0015r09c02	R0051 B1	2.43	0.483	0.19	SCL3/4
1784	61212	p0015r10c15	R0051 C8	2.64	0.406	0.164	SCL3/4
1785	61216	p0015r11c16	R0051 F8	2.28	0.278	0.122	SCL3/4
1786	61225	p0015r14c12	R0052 D6	2.25	0.25	0.111	SCL3/4
1787	61226	p0015r14c14	R0052 D7	2.54	0.3	0.118	SCL3/4
1788	61227	p0015r16c18	R0062 H9	2.06	0.312	0.151	SCL3/4

The ratio of signal 1 to signal 2 in the table above provides a measure of the level of expression of the identified sequences in tumor versus normal tissues. For example, for SEQ ID NO:1669, the tumor-specific signal was 3.09 times that of the signal for the normal tissues tested; for SEQ ID NO:1670, the tumor-specific signal was 2.31 times that of the signal for normal tissues, etc.

EXAMPLE 5

REAL-TIME PCR ANALYSES OF LUNG TUMOR SEQUENCES

10

Real-time PCR was performed on a subset of the lung tumor sequences disclosed herein in order to further evaluate their expression profiles in various tumor and normal tissues. Briefly, quantitation of PCR product relies on the few cycles where the amount of DNA amplifies logarithmically from barely above the background to the plateau. Using continuous fluorescence monitoring, the threshold cycle number where

15

WO 02/04514

PCT/US01/22058

167

DNA amplification is easily determined in each PCR reaction. There are two fluorescence detecting systems. One is based upon a double-strand DNA specific binding dye SYBR Green I dye. The other uses TaqMan probe containing a Reporter dye at the 5' end (FAM) and a Quencher dye at the 3' end (TAMRA) (Perkin Elmer/Applied Biosystems Division, Foster City, CA). Target-specific PCR amplification results in cleavage and release of the Reporter dye from the Quencher-containing probe by the nuclease activity of AmpliTaq Gold™ (Perkin Elmer/Applied Biosystems Division, Foster City, CA). Thus, fluorescence signal generated from released reporter dye is proportional to the amount of PCR product. Both detection methods have been found to generate comparable results. To compare the relative level of gene expression in multiple tissue samples, a panel of cDNAs is constructed using RNA from tissues and/or cell lines, and real-time PCR is performed using gene specific primers to quantify the copy number in each cDNA sample. Each cDNA sample is generally performed in duplicate and each reaction repeated in duplicated plates. The final Real-time PCR result is typically reported as an average of copy number of a gene of interest normalized against internal actin number in each cDNA sample. Real-time PCR reactions may be performed on a GeneAmp 5700 Detector using SYBR Green I dye or an ABI PRISM 7700 Detector using the TaqMan probe (Perkin Elmer/Applied Biosystems Division, Foster City, CA).

Results obtained from real-time PCR analysis of a number of lung tumor-specific sequences disclosed herein are summarized in the table below. In addition, extended cDNA sequences for many of these clones were obtained by searching public sequence databases. The extended sequences, and the proteins encoded by those sequences, are identified by SEQ ID NO: in Table 12 below.

Table 12

Library	Clone Name	Clone No.	SEQ ID NO:	Real-Time PCR Results	Extended cDNA Sequence	Encoded Polypeptide Sequence
SQL1	L972P	47988	1789	Overexpressed in 1/7 lung squamous tumor, 1/8 HN squamous tumor. Low or no expression in normal tissues.		

WO 02/04514

PCT/US01/22058

168

Library	Clone Name	Clone No.	SEQ ID NO.	Real-Time PCR Results	Extended cDNA Sequence	Encoded Polypeptide Sequence
SQL1	L979P	49005	1790	Over-expressed in 2/7 squamous lung tumors, 1/3 HN squamous tumors, 1/2 adeno lung tumors. Low or no expression in normal tissues.	1791	1806
SQL1	L970P	49853	1269	Highly overexpressed in 1/7 lung squamous tumors and 1/3 HN squamous tumor. Low or no expression in normal tissues.		
SQL1	L981P	49965	1272	Over-expressed in 1/8 squamous lung and 1/3 HN squamous tumors. Low or no expression in normal tissues.		
SQL1	L980P	49826	1279	Over-expressed in 3/7 squamous lung tumors, 1/3 HN squamous tumors, 1/2 adeno lung tumors. Low or no expression in normal tissues.	1792	1807
SCL1	L973P	20631	117	Over-expressed in atypical carcinoid METs and adenocarcinoma. Expression in several normal tissues.	1793	1808
SCL1	L974P	20661	128	Over-expressed in primary small cell, squamous and adenocarcinomas. Expression observed in several normal tissues.	1794	1809
SCL1	L986P	50430	1442	Over-expressed in 2/2 Primary Small Cell, 5/8 Small Cell Cell Lines, 1/1 Atypical Carc. METs, 1/1 Adeno, 1/1 Squamous. Very low or no expression in normal tissues.	1795	1810
SCL3	L977P	26961	288	Over-expressed in 1/2 Primary Small Cell, 1/8 Small Cell-Cell Line, and 1/1 Carcinoid Meta. Very low or no expression in normal tissues.	1796	1811
SCL2	L978P	24928	1339	Over-expressed in 2/2 primary small cell, 3/8 small cell-cell lines, 1/1 carcinoid mets, adeno and squamous tumor pools. Low or no expression in normal tissues.	1797	1812
SCL3/4	L984P	50507	1446	Highly expressed in 1/2 primary small cell tumors and 4/8 small cell tumor cell lines. Low or no expression in normal tissues.	1798	1813
SCL3/4	L580S	50536	1449	Over-expressed in select small cell and squamous tumors. Some expression observed normal brain, bronchiot, soft palate and trachea.		
SCL3/4	L988P	50646	1531	Over-expressed in 1/2 Primary Small Cell, 1/2 Primary Small Cell, 6/8 Small Cell-Cell Lines, 1/1 Carcinoid Mets., Adeno & Squamous Tumor pool. Expressed in some normal tissues (brain, adrenal gland, salivary gland, trachea, thymus).	1799	1814
SCL3/4	L1423P	50626	1533	Over-expressed in 1/2 primary small cell, 5/8 small cell-cell lines, 1/1 carcinoid mets. Also expressed in normal brain and pituitary gland.	1800	1815

WO 02/04514

PCT/US01/22058

169

Library	Clone Name	Clone No.	SEQ ID NO:	Real-Time PCR Results	Extended cDNA Sequence	Encoded Polypeptide Sequence
SCL3/4	L988P	50493	1490	Over-expressed in 1/2 primary small cell, 5/6 small cell-cell lines, 1/1 carcinoid mets., adeno and squamous tumor pool. Expressed in normal brain, pituitary gland and spinal cord.		
SCL3/4	L987P	50560	1527	Over-expression in 1/2 Primary Small Cell, 8/6 Small Cell-Cell Lines, 1/1 Carcinoid Mets. Expression in normal pituitary and adrenal glands.	1801	1816-1818
SCL3/4	L1424P	50839	1547	Over-expression in 1/2 Primary Small Cell and 1/1 Carcinoid Mets. Low or no expression in normal tissues.		
MLN1	L997P	26749	730	Over-expression in 1/1 atypical carcinoid METs. No expression in normal tissues.		
MLN1	L990P	26762	733	Over-expressed in 2/2 Primary Small Cell, 6/6 Small Cell Cell Lines, 1/1 Atypical Carc. METs. Expression in several normal tissues.		
MLN1	L1400P	26520	405	Over-expressed in 2/6 Small Cell Cell Lines and 2/2 Primary Small Cell. Moderate to low expression in several normal tissues.		
MLN1	L998P	27699	466	Over-expression in 1/1 Atypical Carcinoid METs. Low expression in normal tissues.	1802	1819
LAP1	L1425P	59303	949	Over-expressed in 4/7 squamous tumors, 1/2 adenocarcinoma tumors and in a pool of six small cell lung carcinomas. Moderate to high expression observed in normal brain, kidney and skeletal muscle.	1803	1820
LAP1	L1426P	59314	1156	Highly overexpressed in one lung squamous tumor and one HN squamous tumor. Very low or no expression observed in normal tissues.	1804	1821
LAP1	L1427P	59298	921	Highly over-expressed in 3/12 adenocarcinoma tumors. Very low or no expression in normal tissues.	1805	1822
LAP1	L1428P	59316	1180	Over-expressed in 4/12 adenocarcinoma tumors and lower level expression in several other adenocarcinoma tumors. Very low or no expression in normal tissues.		

EXAMPLE 6IDENTIFICATION OF CD4 IMMUNOGENIC T CELL EPITOPES
DERIVED FROM LUNG TUMOR ANTIGENS

5

CD4 T cell lines specific for the antigen L548S (SEQ ID NO:789) were generated as follows.

A series of overlapping 20-mer peptides were synthesized that spanned the entire L548S sequence (SEQ ID NO:1834-1856, respectively). For priming, peptides were combined into pools of 4-5 peptides and cultured at 2 micrograms/ml with dendritic cells and purified CD4+ T cells in 96 well U-bottomed plates. One hundred cultures were generated for each peptide pool. Cultures were restimulated weekly with fresh dendritic cells loaded with peptide pools. Following a total of 3 stimulation cycles, cells were rested for an additional week and tested for specificity to antigen presenting cells (APC) pulsed with peptide pools using interferon-gamma ELISA and proliferation assays. For these assays, adherent monocytes loaded with either the relevant peptide pool, recombinant L548S or an irrelevant peptide were used as antigen presenting cells (APC). As shown in Table 13, below, a number of CD4 T cell lines demonstrated reactivity with the priming peptides as well as recombinant L548S protein. These lines were further expanded to be tested for recognition of individual peptides from the pools, as well as for recognition of recombinant L548S.

The dominant reactivity of these lines appeared to be with peptide 21 (SEQ ID NO:1854), which corresponds to amino acids 161-180 of L548S.

WO 02/04514

PCT/US01/22058

171

Table 13

Stimulation		Proliferation (CPM)			Stimulation Index	
Peptides	Positive cell lines	No antigen	Peptides	L548S	Peptides	L548S
p1-5	B1	700	14891	8791	21	13
	B2	1135	48724	53944	42	47
	G2	1227	7609	3193	6	3
p6-10						
p11-15	F2	8315	33723	13391	4	2
	E4	22097	100040	44171	5	2
p16-19						
p20-23	E3	3937	45367	15524	11	4
	F4	2648	130947	12927	49	5

5

EXAMPLE 7DETECTION OF ANTIBODIES AGAINST LUNG TUMOR ANTIGENS
IN PATIENT SERA

10 Antibodies specific for the lung tumor antigens L548S (SEQ ID NO:789), and L552S (SEQ ID NO:809) were shown to be present in effusion fluid or sera of lung cancer patients but not in normal donors. More specifically, the presence of antibodies against L548S, L551S (SEQ ID NO:827) and L552S in effusion fluid obtained from lung cancer patients and in sera from normal donors was examined by

15 ELISA using recombinant proteins and HRP-conjugated anti-human Ig. Briefly, each protein (100 ng) was coated in a 96-well plate at pH 9.5. In parallel, BSA (bovine serum albumin) was also coated as a control protein. The signals ([S], absorbance measured at 405 nm) against BSA ([N]) were determined. The results of these studies are shown in Table 14, wherein - represents [S]/[N] < 2; +/- represents [S]/[N] > 2; ++

20 represents [S]/[N] > 3; and +++ represents [S]/[N] > 5.

Table 14

Detection of Antibodies against Lung Tumor Antigens

	L548S	L551S	L552S
Effusion fluid			
#1	+/-	-	-
#2	-	-	-
#3	-	-	+++
#4	-	-	-
#5	+/-	-	-
#7	-	-	-
#8	-	-	+/-
#10	-	-	+/-
#11	-	-	-
#12	=	-	+/-
#13	-	-	-
#14	+/-	-	+/-
#15	-	-	-
#17	-	-	++
#18	-	-	-
#19	-	-	++
#20	-	-	-
Normal sera			
#21	-	-	-
#22	-	-	-
#23	-	-	-
#24	-	-	-
#25	-	-	-

- 5 Using Western blot analyses, antibodies against L552S were found to present in 1 out of 4 samples of effusion fluid from lung cancer patients, with no L552S antibodies being detected in the three samples of normal sera tested.

EXAMPLE 8

10 FUSION PROTEINS OF LUNG TUMOR ANTIGENS

Fusion proteins of full-length Ra12 with either L801P ORF4 (SEQ ID NO:1670) or L801P ORF5 (SEQ ID NO:1671) were prepared and expressed as single recombinant proteins in *E. coli* as follows.

WO 02/04514

PCT/US01/22058

173

The cDNA for ORF4 of L801P was obtained by PCR with a cDNA for the full length L801P and the primers of SEQ ID NO:1857 and 1858. The cDNA for ORF5 of L801P was obtained by PCR with a cDNA for the full length L801P and the primers of SEQ ID NO:1859 and 1860. The PCR products with expected size were recovered from agarose gel, digested with restriction enzymes EcoRI and XhoI, and cloned into the corresponding sites in the expression vector pCRX1 for subsequent expression in *E. coli*. For the fusion of Ra12 with ORF4, the best expression was obtained in HMS174(DE3)pLysS in 2xYS media, with recombinant protein being induced using IPTG at 37 °C for approximately 3 hours. For the fusion of Ra12 with ORF5, the best expression was obtained in HMS174(DE3)pLysS in 2xYS media, again with recombinant protein being induced using IPTG at 37 °C for approximately 3 hours. The plasmids used for the fusion protein production were confirmed by DNA sequencing. The determined cDNA sequences for the ORF4 and ORF5 fusions are provided in SEQ ID NO:1861 and 1862, respectively, with the corresponding amino acid sequences being provided in SEQ ID NO:1863 and 1864, respectively.

EXAMPLE 9

CLONING OF CDNA ENCODING FULL-LENGTH L984P

The example illustrates the isolation of cDNA sequences encoding L984P by PCR amplification from four separate cDNA sources. Briefly, an earlier isolated cDNA sequence of clone L984P was identified as having homology to a DNA sequence that encodes human achaete-scute homolog 1 (ASH1). Gene specific primers were made using the sequence information present in the public domain for ASH1 (genbank acc. NM-004316). Using these gene specific primers in PCR amplification, L984P was cloned from four separate cDNA sources. The four cDNA sources were a small cell lung carcinoma primary tumor sample (RNA Id. 573A), a METs neuroendocrine atypical carcinoid sample (RNA Id. 512A), and two small cell lung carcinoma cell lines (cell-line Id. NCI H128 and DMS79). The determined cDNA sequences for these four clones are provided in SEQ ID NO:1865-1868, respectively. The coding region of the cDNA contains a repeat of the triplet CAG that exhibits

WO 02/04514

PCT/US01/22058

174

polymorphism in the human genomic DNA. This polymorphism can be observed in L984P cloned from the four different sources as well as from the cDNA of ASH1 and that derived from the human chromosome 4. The cDNA of ASH1 (genbank Acc. NM-006688) deposited in the genbank database contains 14 copies of the triple CAG, whereas the cDNA sequence derived from the human chromosome 4 sequence (genbank Acc. XM_006688) contains 12 copies of the triplet. The cDNA cloned from 573A and 512A both contain 12 copies of the CAG triplet. The cDNA cloned from the small cell lung carcinoma cell line DMS79 contains 13 copies of the triplet CAG, while the cDNA cloned from the small cell lung carcinoma cell line NCI H128 contains only 10 copies of the triplet CAG. As the polymorphism is present in the coding region, this results in polymorphisms in the protein sequences as well (see, SEQ ID NO:1869-1872).

15

EXAMPLE 10

CLONING OF CDNA ENCODING FULL-LENGTH L985P

As previously disclosed in Example 1C, a search of the public databases using the sequence for contig 139 (SEQ ID NO:1467, also known as L985P) as the query was conducted and showed that this sequence had homology to the cDNA (SEQ ID NO:1676) encoding the cell surface immunomodulator-2 (CSIMM-2, sequence obtained from the Geneseq database). The full-length sequence of the clone was obtained by screening a small cell lung carcinoma cDNA library with a radioactively labeled probe of the original cloned sequence (SEQ ID NO:1467). Approximately 500,000 clones from the cDNA library were screened and 2 independent clones containing cDNA insert of 1.35 kb were isolated. The full-length cDNA sequence is provided in SEQ ID NO:1873 and the encoded amino acid sequence is provided in SEQ ID NO:1874. Surprisingly, an alignment of the isolated full-length cDNA sequence of L985P (SEQ ID NO:1873) with the Geneseq database sequence for CSIMM-2 showed that L985P differs from the GeneSeq database sequence by one nucleotide. This nucleotide difference results in a change of one amino acid residue at position 119 (G to F) of SEQ ID NO:1874.

The predicted protein structure and sequence of L985P indicates that it is a member of the recently described MS4A (membrane-spanning 4-domain, subfamily A) gene family (Liang and Tedder, 2001 Genomics 72:119-127). The MS4A gene family currently consists of at least 21 distinct human and mouse proteins of which nine members are from humans. These include CD20, FcεRIβ, HTm4, MS4A4A, MS4A5, MS4A6A, MS4A7, MS4A8B (same as L985P) and MS4A12. The MS4A family members are cell surface expressed proteins containing four transmembrane spanning domains, with N- and C-terminal regions facing the cytoplasmic side of the cells. The human MS4A family members exhibit 20-40% overall homology at the protein level, which is confined mostly to the transmembrane domains. The transmembrane domains of L985P share the highest homology to CD20 with approximately 40% identity and 60% similarity between the two protein sequences in this region. The MS4A family members demonstrate a broad tissue distribution with expression observed in diverse cell types in hematopoietic and nonhematopoietic tissues. However, expression of some MS4A family members is highly restricted to a particular cell type, such as CD20, which is only expressed on B-cells. As also mentioned in Example 1C, L985P (MS4A8B) is over-expressed in small cell lung carcinoma (SCLC) as determined by quantitative real-time PCR. Low level expression is seen in some normal tissues including lung, pituitary gland, stomach, colon, and trachea, while expression in other normal tissues checked was negligible or undetectable (*see*, Example 1C).

The physiological function for most of the MS4A family members remains to be elucidated. Previous studies have shown that CD20 is functionally important for the regulation of cell growth and differentiation, and signal transduction in B-cells. There is also evidence that CD20 may serve as a calcium channel by forming a homo- or heterotetrameric complex. FcεRIβ is part of a tetrameric receptor complex, which mediates interaction with IgE-bound antigens that lead to cellular responses such as the degranulation of mast cells. Because of the sequence and structural homologies between the MS4A family members, it is highly likely that they will share overlapping functional properties.

Some of the MS4A family members have been found to be associated with cancer. CD20 is expressed in more than 90% of B-cell non-Hodgkin lymphomas,

which has made it an ideal target for immunotherapeutic approaches for the treatment of B-cell malignancies. Anti-CD20 monoclonal antibodies have been used with high success in the treatment of non-Hodgkin-lymphomas in both naïve and radiolabeled forms. The anti-CD20 MAbs have been shown to exert their antitumor effects through
5 several pathways, which include complement-mediated cytotoxicity, antibody-mediated cellular cytotoxicity and antibody-mediated cell cycle arrest and apoptosis. Analysis of the human EST databases indicates that other members of the MS4A gene family including MS4A4A, A6A, A7 and A8B are also expressed in various cancers including lung, breast, pancreas, colon, ovary, kidney and brain. By comparison to CD20, one or
10 more of the MS4A family genes would be used as targets for immunotherapeutic approaches for the treatment of hematopoietic and nonhematopoietic malignancies.

The MS4A protein family members are structurally similar to other membrane protein families with four transmembrane domains. These include the Tetraspanin protein family (TM4SF) and the GABA-A receptor protein family.
15 Tetraspanins are associated with cancer and may play a direct role in controlling tumor progression. Although CD9 expression will positively influence B cell migration, CD9 overexpression suppresses motility and metastasis in carcinoma cells and there is an inverse correlation with metastasis in melanoma. However, CD9 is also expressed on 90% of non-T cell acute lymphoblastic leukemia cells and 50% of chronic lymphocytic
20 leukemias. A recent study using RT-PCR analysis of tetraspanin expression in Burkitt lymphoma cell lines found that 90% of the lines express CD53, CD81, CD63, CD82 and SAS at high levels. CD151/TETRA3 is an effector of metastasis and cell migration and MAbs that block this activity have been developed. Similarly, overexpression of the tetraspanin CO-029/D6.1 will increase the metastatic potential of cell lines. The
25 tetraspanins control a diverse set of biological functions that can be regulated by MAbs. The functions of the tetraspanins, in general, can be grouped into actions that affect cell activation and proliferation, as well as adhesion and motility. These functions tend to be carried out by their association with integrins. The functional activity of tetraspanins can be modulated with MAbs in such a way as to control cell proliferation. For
30 example, CD81/TAPA-I is associated with B cell activation and increased proliferation,

WO 02/04514

PCT/US01/22058

177

an activity that can be blocked with MAbs. MAbs with anti-proliferative activity have been generated to the tetraspanin family member CO-029/D6.1..

As mentioned above, L985P is specifically over-expressed in small cell lung carcinomas. This fact and a comparison to CD20, tetraspanins and Her2 whose over-expression in cancers makes them effective cancer therapeutic targets, indicate that L985P may be a good target for immunotherapeutic approaches for the treatment of small cell lung carcinomas.

To facilitate the generation, purification, and evaluation of MAb against L985P, MAbs against the entire deduced amino acid sequence of the L985P protein, peptides derived from L985P or chemically produced (synthetic) L985P peptides will be used. Also, one can use MAbs raised against chimeric forms of L985P protein molecule fused to Ra12 protein, either the long form (Ra12- which is the first 128 amino acids of Ra12) and/or the short form (Ra12S), or fused to a polyhistidine peptide or any combination of these molecules. Provided are the predicted cDNA and amino acid sequences for the his-tagged L985P-Ra12 fusion molecules: Ra12-L985P_cDNA (SEQ ID NO:1875), Ra12-L985P_Protein (SEQ ID NO:1876), Ra12S-L985P_cDNA (SEQ ID NO:1877) and Ra12S-L985P_Protein (SEQ ID NO:1878); and the L985P derived peptides: his-tagged Ra12S-L985PEx_cDNA (SEQ ID NO:1879), his-tagged Ra12S-L985PEx_Protein (SEQ ID NO:1880), L985P_Extracellular_Loop-2_cDNA (SEQ ID NO:1881) and L985P_Extracellular_Loop-2_Peptide (SEQ ID NO:1882).

EXAMPLE 11

EXPRESSION IN *E. COLI* OF A HIS-TAGGED RA12-L985P FUSION PROTEIN

This example sets forth a specific embodiment of a fusion between Ra12 and L985P and its expression in *E. coli*. A his-tagged fusion protein of the long-form of Ra12 and all but the first three amino acid residues of L985P was expressed as a single recombinant protein in *E. coli*. The long-form of Ra12 was modified from the original sequence of amino acid residues 192-323 of MTA32A in that a putative thrombin cleavage site was replaced with a HindIII restriction site. The L985P was fused downstream of the Ra12 sequence in a pCRX1 vector. The Ra12 sequence was cloned

WO 02/04514

PCT/US01/22058

178

downstream of the RBS. As a result, a his-tagged fusion protein is produced when the recombinant vector is expressed in *E. coli*. The sequence for the fusion of the long-form of Ra12 and L985P was confirmed by DNA sequencing. The determined cDNA sequence is provided in SEQ ID NO:1875 as this sequence is the same as that predicted in Example 10.

EXAMPLE 12

CLONING OF CDNA ENCODING FULL-LENGTH L1428P

As previously disclosed in Example 5, real-time PCR was performed on a subset of the lung tumor sequences disclosed herein in order to further evaluate their expression profiles in various tumor and normal tissues. The results are provided above in Table 12. One of the sequences analyzed was from clone #59316 (SEQ ID NO:1180, L1428P) and was shown to be expressed in a subset of lung adenocarcinomas.

Further studies have isolated the full-length cDNA for the cloned sequence of #59316 (SEQ ID NO:1180, L1428P). In order to determine the transcript size of the gene, a multiple tissue Northern blot was probed with the radioactively labeled original cloned sequence (SEQ ID NO:1180). The Northern blot included about 20 ug of total RNA from lung adenocarcinoma and normal tissues samples. Visual analysis of the exposed film revealed a single transcript of approximately 6.5 - 7.0 kb. The full-length sequence of the clone was obtained by screening a lung adenocarcinoma primary tumor cDNA library with a radioactively labeled probe of the original cloned sequence (SEQ ID NO:1180). Approximately 500,000 clones from the cDNA library were screened and 5 independent clones containing a cDNA insert of 6.8 kb were isolated. This insert size is similar to the size estimated by Northern-blot analysis. The full-length sequence is provided in SEQ ID NO:1883. Although no distinct ORF could be identified, seven potential ORFs can be predicted and the amino acid sequences of these potential ORFs, designated L1428P_ORF1 to ORF7, are provided in SEQ ID NO:1884-1890, respectively. The expression of full-length L1428P was re-analyzed by real-time PCR as set forth in Example 5 on extended cDNA panels for both lung adenocarcinomas and squamous cell carcinomas. The lung adenocarcinoma extended

WO 02/04514

PCT/US01/22058

179

panel real-time results again confirm the expression of L1428P in adenocarcinoma with about 40-50% of the adenocarcinoma samples showing varying levels of expression. The real-time results from the squamous cell carcinoma extended panel shows that L1428P is also expressed in lung squamous cell carcinoma. However, the expression was in fewer lung squamous cell carcinoma samples and at a lower level.

EXAMPLE 13

REAL-TIME PCR ANALYSIS OF CDNA SEQUENCES WHICH ARE OVER-EXPRESSED IN LUNG TUMORS AS SHOWN BY MICROARRAY

10

The following clones listed in Table 15 were shown previously to be over-expressed in lung tumors by microarray analysis (see, Example 2C and Example 6, Table 11). The results of this microarray analysis are summarized below in Table 15.

Table 15

SEQ ID NO:	CLONE ID #	Ratio	Mean Signal 1	Mean Signal 2
1383	50096 (contig 156)	16.04	1.895	0.118
1560	54454 (contig 234)	2.86	0.2	0.07
1561	54463 (contig 235)	3.31	0.248	0.075
1707	59376	3.23	0.734	0.227
1733	61090	2.68	0.364	0.136
1735	61093	5.7	0.687	0.121
1758	61137	8.63	1.463	0.17
1761	61144	7.02	0.783	0.112
1766	61159	4.57	0.831	0.182
1771	61173	5.57	1.075	0.193
1775	61185	5.29	0.962	0.182
1786	61225	3.41	0.379	0.111

15

WO 02/04514

PCT/US01/22058

180

Real-time PCR analysis was performed on these sequences on a small cell lung carcinoma panel as described in Example 7. The results obtained from the real-time PCR analysis are summarized in Table 16.

Table 16

SEQ ID NO:	Real-Time PCR Results
1383	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinoma and 6/6 SCLC cell lines. Some expression is also seen in normal brain, pituitary gland, spinal cord, and thymus.
1560	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas and 3/6 SCLC cell lines. Expression is also seen in normal brain and pituitary gland.
1561	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas, 6/6 SCLC cell lines, 1/1 atypical carcinoid metastases, adenocarcinoma, and squamous cell carcinoma pools. Expression is also seen in normal bone marrow, lymph node, thymus, and at lower levels in other normal tissues.
1707	On the SCLC panel, this gene is overexpressed in 1/6 SCLC cell lines and 0/2 primary small cell carcinomas. It is also expressed in normal stomach and at lower levels in normal brain, salivary gland, and trachea.
1733	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas, 6/6 SCLC cell lines, 1/1 atypical carcinoid metastases, adenocarcinoma, and squamous cell carcinoma pools. Some expression is also seen in normal bone marrow, lymph node, thymus, and at lower levels in other normal tissues.
1735	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas, 5/6 SCLC cell lines, 0/1 atypical carcinoid metastases, and adenocarcinoma pool. Lower level expression is also seen in normal bone marrow and skeletal muscle.
1758	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas, 6/6 SCLC cell lines, 0/1 atypical carcinoid metastases, adenocarcinoma, and squamous cell carcinoma pools. Some expression is also seen in normal bone marrow, thyroid gland, and trachea.
1761	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas, 6/6 SCLC cell lines, 0/1 atypical carcinoid metastases, adenocarcinoma, and squamous cell carcinoma pools. Expression is also seen in normal bone marrow.

WO 02/04514

PCT/US01/22058

181

SEQ ID NO:	Real-Time PCR Results
1766	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas, 6/6 SCLC cell lines, and 0/1 atypical carcinoid metastases. Expression is also seen in normal pituitary gland, adrenal gland, bone marrow, thymus, salivary gland, and at lower levels in a variety of other normal tissues.
1771	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas, 6/6 SCLC cell lines, 0/1 atypical carcinoid metastases, and squamous cell carcinoma pools. Some expression is also seen in normal pituitary gland.
1775	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas, adenocarcinoma, and squamous cell carcinoma pools. No expression is observed in the SCLC cell lines and the atypical carcinoid metastases. Some normal tissue expression is seen in liver, stomach, thyroid gland, lymph node, and thymus.
1776	On the SCLC panel, this gene is overexpressed in 2/2 primary small cell carcinomas, 6/6 SCLC cell lines, adenocarcinoma, and squamous cell carcinoma pools. Some expression is also seen in normal bone marrow, pituitary gland, stomach, trachea, and thymus.

These sequences were then compared to known sequences in the available databases (Genbank, GeneSeq, huEST, etc.). Nine of these sequences showed some degree of similarity to known sequences in the available databases. The results of

5 these nine hits are summarized in Table 17 along with providing sequence listing identifiers for the DNA sequences and the respective encoded amino acid sequences (where available) of these matches. SEQ ID NO:1561 and 1786 showed no significant similarity to any known sequences.

WO 02/04514

PCT/US01/22058

182

Table 17

SEQ ID NO:	GenBank Database Hit	GeneSeq. Hit	cDNA Seq. of Hit	Amino Acid Seq. Of Hit (if Known)
1383	Pr22 Protein/Stathmin/Oncoprotein 18	A16376, A08801, A01633	1891	1901
1560	CDNA DKFZp564O163	--	1892	--
1561	Novel	Human secreted protein 5' EST. (C30107)	--	--
1707	SOX21 (AF107044)	Human secreted protein gene 7 clone HE8CV18. (X27317)	1893	1902
1733	KIAA0166 gene (D79988)	Human gene signature HUMGS08725. (T26483)	1894	1903
1735	Ubiquitin-conjugating enzyme E2 (AF160215)	DNA encoding human ubiquitin-like conjugating protein (UBCLL). (X81676)	1895	1904
1758	Pituitary tumor transforming gene protein 1 (AF095287)	Z97293, X89295, V36964, V63198, Z97292, C00858, V88346, Z80287	1896	1905
1761	Novel	Kidney injury associated molecule HW051 cDNA clone. (V80605)	1897	1906
1766	Cyclin-dependent kinase inhibitor p18 (CDKN2C) (AF041248)	Cyclin-dependent kinase (CDK6) inhibiting protein. (T10925; T31456)	1898	1907
1771	CDK4-inhibitor (p16-INK4) (L27211)	Multiple Matches	1899	1908
1775	Monokine induced by gamma interferon (MIG) (NM_002416)	Monokine induced by gamma-interferon. (X14998; Z26088)	1900	1909
1776	Novel	--	--	--

Further studies have resulted in isolation of the full-length cDNA sequence for the cloned sequence of clone #61093 (SEQ ID NO:1735, L1437P). In

WO 02/04514

PCT/US01/22058

183

order to determine the transcript size of the gene, a multiple tissue Northern blot was probed with the radioactively labeled original cloned sequence (SEQ ID NO:1735). The Northern blot included about 20 ug of total RNA from small cell lung carcinoma and normal tissues samples. Visual analysis of the exposed film revealed a single transcript of approximately 1.2 kb. The full-length sequence was obtained by screening a small cell lung carcinoma tumor cDNA library with the radioactively labeled probe of the original cloned sequence (SEQ ID NO:1735). Approximately 120,000 clones from the cDNA library were screened and 2 independent clones containing a cDNA insert of 931 bases were isolated. The inserts are similar in size to that estimated by Northern Blot analysis. The full-length cDNA sequence is provided in SEQ ID NO:1910. It was discovered that there was one nucleotide difference between the full-length cDNA and a previously published sequence. However, this nucleotide change does not result in a change in the deduced amino acid sequence. The deduced amino acid sequence encoded by the full-length cDNA is the same as already provided in SEQ ID NO:1904, and confirms earlier predictions that this cDNA encodes a known protein, ubiquitin-conjugated enzyme E2 (AF160215, SEQ ID NO:1904). SEQ ID NO:1910 (L1437P) was shown to be over-expressed in lung small cell lung carcinoma by microarray, real-time PCR and Northern Blot analysis.

20

EXAMPLE 14

EXPRESSION IN *E. COLI* OF A L548S HIS TAG FUSION PROTEIN

PCR was performed on the L548S coding region with the following primers:

Forward primer PDM-433 5' gctaaaggtgacccaagaaccnaag 3' (SEQ ID NO:1911) Tm 60°C.

Reverse primer PDM-438 5' ctattaactcgaggagacagaataacagttcttta 3' (SEQ ID NO:1912) Tm 61°C.

The PCR conditions were as follows:

10µl 10X Pfu buffer
1.0µl 10mM dNTPs

WO 02/04514

PCT/US01/22058

184

2.0µl 10µM each primer

83µl sterile water

1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)

50ng DNA

5 96°C for 2 minutes, 96°C for 20 seconds, 61°C for 15 seconds, 72°C for 1 minute 30 seconds with 40 cycles and then 72°C for 4 minutes.

The PCR product was digested with XhoI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and XhoI restriction enzymes. The correct
10 construct was confirmed by DNA sequence analysis and then transformed into BL21 CodonPlus (Stratagene, La Jolla, CA) and BL21 pLys S (Novagen, Madison, WI) cells for expression.

The amino acid sequence of expressed recombinant L548S is shown in SEQ ID NO:1913, and the DNA coding region sequence is shown in SEQ ID NO:1914.

15

EXAMPLE 15EXPRESSION IN *E. COLI* OF A L551S HIS TAG FUSION PROTEIN

PCR was performed on the L551S coding region with the following
20 primers:

Forward primer PDM-498 5' gtgacgatggagagctgcggagatgg 3' (SEQ ID NO:1915) Tm 67°C.

Reverse primer PDM-499 5' cgcctaaactcgagtcactaacagctggag 3' (SEQ ID NO:1916) Tm 66°C.

25 The PCR conditions were as follows:

10µl 10X Pfu buffer

1.0µl 10mM dNTPs

2.0µl 10µM each primer

83µl sterile water

30 1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)

50ng DNA

WO 02/04514

PCT/US01/22058

185

96°C for 2 minutes, 96°C for 20 seconds, 66°C for 15 seconds, 72°C for 2 minutes 20 seconds with 40 cycles and then 72°C for 4 minutes.

The PCR product was digested with XhoI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and XhoI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into BLR (DE3) pLys S and BLR (DE3) CodonPlus RP cells for expression.

The amino acid sequence of expressed recombinant I.551S is shown in SEQ ID NO:1917, and the DNA coding region sequence is shown in SEQ ID NO:1918.

EXAMPLE 16

EXPRESSION IN *E. COLI* OF A I.552S HIS TAG FUSION PROTEIN

PCR was performed on the I.552S coding region with the following primers:

Forward primer PDM-479 5' cgggccacgcccattggaccttc 3' (SEQ ID NO:1919) Tm 64°C.

Reverse primer PDM-480 5' ctgagaattcattaaactgtgtgtgctcttcacc 3' (SEQ ID NO:1920) Tm 62°C.

The PCR conditions were as follows:

10µl 10X Pfu buffer

1.0µl 10mM dNTPs

2.0µl 10µM each primer

83µl sterile water

1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)

50ng DNA

96°C for 2 minutes, 96°C for 20 seconds, 63°C for 15 seconds, 72°C for 1 minute with 40 cycles and then 72°C for 4 minutes.

The PCR product was digested with EcoRI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The

WO 02/04514

PCT/US01/22058

186

correct construct was confirmed by DNA sequence analysis and then transformed into BL21 CodonPlus (Stratagene, La Jolla, CA) cells for expression.

The amino acid sequence of expressed recombinant L552S is shown in SEQ ID NO:1921, and the DNA coding region sequence is shown in SEQ ID NO:1922.

5

EXAMPLE 17

CLONING OF CDNA ENCODING FULL-LENGTH CLONES #19069 AND CLONE
#18965 OR #19002

10 Partial sequences of two lung antigens, clones #19069 (SEQ ID NO:90)
and #18965 or #19002 (both SEQ ID NO:15), were previously provided. These partial
sequences were used as a query to predict the full-length cDNA sequences for the
isolated cloned sequenced by searching the publicly available databases. The predicted
full-length cDNA sequence for the isolated cloned sequence of SEQ ID NO:90 is
15 provided in SEQ ID NO:1923. The predicted full-length cDNA sequence for the
isolated cloned sequence of SEQ ID NO:15 is provided in SEQ ID NO:1924. The
deduced amino acid sequences of the two antigens are provided in SEQ ID NO:1925
and 1926, respectively. These sequences were then compared to known sequences in the
GeneSeq database. Both sequences showed some degree of similarity to known
20 sequences in the GeneSeq database. SEQ ID NO:1923 shows similarity to a
lipophosphatic acid acyltransferase (GeneSeq Z25000 and Z65038) and SEQ ID
NO:1924 shows similarity to a zinc/iron regulated transporter-like protein (Geneseq
Z38333 and A14995).

25

EXAMPLE 18

EPITOPE-MAPPING OF L552S-SPECIFIC ANTIBODIES

Peptides of candidate antigens can be used for the evaluation of antibody
responses in both preclinical and clinical studies. These data allow one to further
30 confirm the antibody response against a certain candidate antigen. Protein-based ELISA
with and without competitive peptides and peptide-based ELISA can be used to

WO 02/04514

PCT/US01/22058

187

evaluate these antibody responses. Peptide ELISA is especially useful since it can further exclude the false positive of the antibody titer observed in protein-based ELISA as well as to provide the simplest assay system to test antibody responses to candidate antigens. In this example, data was obtained using L552S-peptides that show that individual cancer patients produce L552S-specific antibodies recognizing primarily the following three epitopes of L552S:

- (1) aa21-35: GPRSGGAQAKLGCCW (SEQ ID NO:1927)
- (2) aa116-135: KVICKSCISQTPGINLDLGS (SEQ ID NO:1928)
- (3) aa141-160: IIPKEDHCKMPEAGEEQPQV (SEQ ID NO:1929)

10

In further studies, it was found that affinity-purified antibodies generated by SEQ ID NO:1929 can recognize the L552S protein, and occupy about 0.6% of the total immunoglobulin kappa of a patient's lung plural effusion (LPE) fluid. It was also found that SEQ ID NO:1929 is the dominant epitope of the rabbit polyclonal antibodies specific for L552S protein.

15

EXAMPLE 19

L985P EXPRESSION

20

For recombinant expression in mammalian cells, the full length L985P cDNA was subcloned into the mammalian expression vector pCEP4 (Invitrogen) with and without a FLAG epitope tag. Both constructs were transfected into HEK293 cells (ATCC) using Lipofectamine 2000 reagent (Invitrogen). Western blot analysis was then performed on these transfected cells to determine if recombinant L985P was being transiently expressed.

25

Briefly the transfection was carried out as follows. HEK cells were plated at a density of 350,000 cells/well (6 well plate) in DMEM (Gibco) containing 10% FBS (Hyclone) and grown overnight. The following day, 2µl of Lipofectamine 2000 (Invitrogen) was added to 50µl of Optimem 1 (Invitrogen) containing no FBS and incubated for 5 minutes at room temperature. In a different set of tubes 50µl of

30

WO 02/04514

PCT/US01/22058

188

Optimem 1 was mixed with 0.8µg of L985P (with and without FLAG)/plasmid DNA and the mixture was transferred to the Lipofectamine 2000/Optimem mix. The combined mixture was incubated for 20 minutes at room temperature and transferred to the HEK293 cells containing 0.5ml of DMEM 10% FBS. The Lipofectamine
5 2000/DNA mix was then added to the HEK293 cells and incubated for 16-24 hrs at 37°C with 7% CO₂. Cells were rinsed with PBS then collected and pelleted by centrifugation.

For Western blot analysis, whole cell lysates were generated by incubating the cells in Triton-X100 containing lysis buffer for 30 minutes on ice.
10 Lysates were then cleared by centrifugation at 15,000 rpm for 5 minutes at 4°C. Samples were diluted with SDS-PAGE loading buffer containing beta-mercaptoethanol, then boiled for 10 minutes prior to loading on the SDS-PAGE gel. The protein was transferred to nitrocellulose and probed using a purified anti-L985P rabbit polyclonal sera at a dilution of 1:1000. The blot was revealed with a donkey anti-rabbit Ig coupled
15 to HRP (Jackson ImmunoResearch) followed by incubation in ECL substrate. Results of the blot indicate that recombinant L985P (with and without the FLAG) was expressed in the HEK293 cells.

EXAMPLE 20

20 GENERATION OF POLYCLONAL ANTIBODIES TO LUNG TUMOR ANTIGENS

Three lung antigens, L548S (SEQ ID NO:789), L552S (SEQ ID NO:809) and L985P peptide #3482 (SEQ ID NO:1930), were expressed and purified for use in antibody generation.

25 L548S and L552S were expressed in an *E. coli* recombinant expression system and grown overnight in LB Broth with the appropriate antibiotics at 37°C in a shaking incubator. The next morning, 10 ml of the overnight culture was added to 500 ml of 2x YT with the appropriate antibiotics in a 2L-baffled Erlenmeyer flask. When the optical density of the culture reached 0.4-0.6 at 560 nanometers, the cells were
30 induced with IPTG (1 mM). Four hours after induction with IPTG, the cells were harvested by centrifugation.

WO 02/04514

PCT/US01/22058

189

The cells were then washed with phosphate buffered saline and centrifuged again. The supernatant was discarded and the cells were either frozen for future use or immediately processed. Twenty milliliters of lysis buffer was added to the cell pellets and vortexed. To break open the *E. coli* cells, this mixture was then run through a french press at a pressure of 16,000 psi. The cells were then centrifuged again and the supernatant and pellet were checked by SDS-PAGE for the partitioning of the recombinant protein.

For proteins that localized to the cell pellet, the pellet was resuspended in 10 mM Tris pH 8.0, 1% CHAPS and the inclusion body pellet was washed and centrifuged again. This procedure was repeated twice more. The washed inclusion body pellet was solubilized with either 8M urea or 6M guanidine HCl containing 10 mM Tris pH 8.0 plus 10 mM imidazole. The solubilized protein was added to 5 ml of nickel-chelate resin (Qiagen) and incubated for 45 minutes to 1 hour at room temperature with continuous agitation.

After incubation, the resin and protein mixture was poured through a disposable column and the flow through was collected. The column was then washed with 10-20 column volumes of the solubilization buffer. The antigen was then eluted from the column using 8M urea, 10 mM Tris pH 8.0 and 300 mM imidazole and collected in 3 ml fractions. A SDS-PAGE gel was run to determine which fractions to pool for further purification.

As a final purification step, a strong anion exchange resin, in this case Hi-Prep Q (Biorad), was equilibrated with the appropriate buffer and the pooled fractions from above were loaded onto the column. Each antigen was eluted off the column with an increasing salt gradient. Fractions were collected as the column was run and another SDS-PAGE gel was run to determine which fractions from the column to pool.

The pooled fractions were dialyzed against 10 mM Tris pH 8.0. The release criteria were purity as determined by SDS-PAGE or HPLC, concentration as determined by Lowry assay or Amino Acid Analysis, identity as determined by amino terminal protein sequence, and endotoxin level was determined by the Limulus (LAL)

WO 02/04514

PCT/US01/22058

190

assay. The proteins were then put in vials after filtration through a 0.22-micron filter and the antigens were frozen until needed for immunization.

The L985P peptide #3482 was synthesized and conjugated to KLH and frozen until needed for immunization.

5 The polyclonal antisera were generated using 400 micrograms of each lung antigen combined with 100 micrograms of muramyl dipeptide (MDP). An equal volume of Incomplete Freund's Adjuvant (IFA) was added and then mixed and injected subcutaneously (S.C.) into a rabbit. After four weeks, the rabbit was S.C. boosted with 200 micrograms of antigen mixed with an equal volume of IFA. Thereafter the rabbit
10 was I.V. boosted with 100 micrograms of antigen. The animal was bled seven days following each boost. The blood was then incubated at 4°C for 12-24 hours followed by centrifugation to generate the sera.

The polyclonal antisera were characterized using 96 well plates coated with antigen and incubating with 50 microliters (typically 1 microgram/microliter) of
15 the polyclonal antisera at 4°C for 20 hours.

250 microliters of BSA blocking buffer was added to the wells and incubated at room temperature for 2 hours. Plates were washed 6 times with PBS/0.1% Tween. The rabbit sera were diluted in PBS/0.1% Tween/0.1% BSA. 50 microliters of diluted sera was added to each well and incubated at room temperature for 30 minutes.
20 The plates were washed as described above, and then 50 microliters of goat anti-rabbit horseradish peroxidase (HRP) at a 1:10000 dilution was added and incubated at room temperature for 30 minutes.

The plates were washed as described above, and 100 microliters of TMB Microwell Peroxidase Substrate was added to each well. Following a 15-minute
25 incubation in the dark at room temperature, the colorimetric reaction was stopped with 100 microliters of 1N H₂SO₄ and read immediately at 450 nm. All the polyclonal antibodies showed immunoreactivity to the appropriate antigen.

Tables 18-20 show the antibody reactivity of rabbit antisera in serial dilution to the three lung antigens, L548S, L552S and L985P peptide #3482. The first
30 column shows the antibody dilutions. The columns "Pre-immune sera" indicate ELISA data for two experiments using pre-immune sera. These results are averaged in the

WO 02/04514

PCT/US01/22058

191

fourth column. The columns "anti-L548S, L552S or #3482" indicate ELISA data for two experiments using sera from rabbits immunized as described above, using the respective antigen, referred to as either L548S, L552S or #3482 in the tables.

WO 02/04514

PCT/US01/22058

192

Table 18

Antibody dilution	Pre-immune sera (1)	Pre-immune sera (2)	Average	Anti-L548S (1)	Anti-L548S (2)	Average
1:1000	0.17	0.10	0.14	0.51	0.51	0.51
1:2000	0.12	0.09	0.11	0.30	0.30	0.30
1:4000	0.09	0.08	0.08	0.17	0.20	0.19
1:8000	0.09	0.07	0.08	0.12	0.13	0.12
1:16000	0.09	0.08	0.08	0.09	0.12	0.11
1:32000	0.09	0.08	0.09	0.08	0.09	0.09
1:64000	0.11	0.09	0.10	0.11	0.12	0.12
1:128000	0.10	0.08	0.09	0.08	0.09	0.08
1:256000	0.09	0.08	0.08	0.11	0.09	0.10
1:512000	0.11	0.09	0.10	0.08	0.08	0.08
1:1024000	0.10	0.08	0.09	0.08	0.10	0.09
1:2048000	0.10	0.09	0.09	0.08	0.09	0.09

Table 19

5

Antibody dilution	Pre-immune sera (1)	Pre-immune sera (2)	Average	Anti-L552S (1)	Anti-L552S (2)	Average
1:1000	0.08	0.19	0.13	2.14	2.03	2.08
1:2000	0.08	0.06	0.07	1.93	1.97	1.95
1:4000	0.07	0.06	0.06	1.81	1.82	1.82
1:8000	0.08	0.06	0.07	1.63	1.64	1.64
1:16000	0.06	0.05	0.05	1.47	1.29	1.38
1:32000	0.06	0.05	0.06	1.03	1.10	1.06
1:64000	0.06	0.06	0.06	0.73	0.69	0.71
1:128000	0.06	0.05	0.06	0.44	0.48	0.46
1:256000	0.06	0.06	0.06	0.26	0.25	0.26
1:512000	0.07	0.06	0.06	0.16	0.15	0.16
1:1024000	0.06	0.07	0.06	0.12	0.10	0.11
0.00	0.06	0.06	0.06	0.06	0.06	0.06

WO 02/04514

PCT/US01/22058

193

Table 20

Antibody dilution	Pre-immune sera (1)	Pre-immune sera (2)	Average	Anti-#3482 (1)	Anti-#3482 (2)	Average
1:1000	0.10	0.07	0.09	2.10	2.07	2.08
1:2000	0.07	0.07	0.07	1.80	1.84	1.82
1:4000	0.07	0.06	0.06	1.78	1.80	1.79
1:8000	0.06	0.06	0.06	1.94	1.72	1.83
1:16000	0.06	0.06	0.06	1.75	1.74	1.74
1:32000	0.06	0.06	0.06	1.42	1.47	1.44
1:64000	0.06	0.06	0.06	1.12	1.17	1.14
1:128000	0.06	0.06	0.06	0.79	0.87	0.83
1:256000	0.06	0.06	0.06	0.70	0.65	0.68
1:512000	0.06	0.06	0.06	0.41	0.41	0.41
1:1024000	0.06	0.06	0.06	0.25	0.25	0.25
0	0.06	0.06	0.06	0.06	0.06	0.06

5 Table 21 shows the Protein A purification of the antibodies to the lung antigen, L548S.

Table 21

Purified Antibody Concentration (μg/ml)	Pro A pure	Pro A pure	Average
3.0	2.20	2.12	2.16
1.5	2.09	2.01	2.05
0.75	1.95	1.93	1.94
0.38	1.83	1.85	1.84
0.188	1.68	1.68	1.68
0.094	1.36	1.38	1.37
0.047	1.02	1.06	1.04
0.0234	0.68	0.73	0.71
0.0177	0.40	0.42	0.41
0.0059	0.24	0.25	0.24
0.0029	0.17	0.15	0.16
0.00	0.06	0.06	0.06

WO 02/04514

PCT/US01/22058

194

Table 22 shows the affinity purification of the antibodies to the lung antigen, L552S.

Table 22

Antibody dilution	Affinity pure (salt peak)	Affinity pure (salt peak)	Average	Antibody dilution	Affinity pure (acid peak)	Affinity pure (acid peak)	Average
1:50	0.19	0.18	0.18	1:1000	2.22	2.20	2.21
1:100	0.10	0.09	0.10	1:2000	2.17	2.10	2.13
1:200	0.06	0.06	0.06	1:4000	2.05	2.06	2.06
1:400	0.06	0.06	0.06	1:8000	1.95	1.94	1.95
1:800	0.06	0.05	0.06	1:16000	1.86	1.82	1.84
1:1600	0.06	0.06	0.06	1:32000	1.64	1.57	1.61
1:3200	0.06	0.06	0.06	1:64000	1.28	1.26	1.27
1:6400	0.06	0.06	0.06	1:128000	0.88	0.87	0.88
1:12800	0.06	0.06	0.06	1:256000	0.56	0.56	0.56
1:25600	0.06	0.06	0.06	1:512000	0.35	0.34	0.34
1:51200	0.06	0.06	0.06	1:1024000	0.18	0.17	0.18
0.00	0.06	0.06	0.06	0.00	0.06	0.06	0.06

5

Tables 23A and 23B show the affinity purification and the Protein A purification of the antibodies to the lung antigen L985P peptide #3482.

Table 23A

10

Purified Antibody Concentration (μg/ml)	Affinity pure	Affinity pure	Average
3.0	1.73	1.69	1.71
1.50	1.32	1.28	1.30
0.75	0.96	0.91	0.93
0.38	0.63	0.61	0.62
0.19	0.38	0.39	0.38
0.09	0.19	0.20	0.19
0.05	0.16	0.10	0.13
0.02	0.12	0.15	0.13

Table 23B

Purified Antibody Concentration (µg/ml)	Pro A pure	Pro A pure	Average
40.70	1.80	1.91	1.85
20.35	1.54	1.53	1.53
10.18	1.04	1.18	1.11
5.09	0.71	0.78	0.74
2.54	0.43	0.52	0.48
1.27	0.25	0.28	0.26
0.64	0.13	0.20	0.17
0.32	0.10		0.10

5

EXAMPLE 21

CLONING OF EXTENDED CDNA SEQUENCE OF L1439P

The partial cDNA sequence for clone #61144 (SEQ ID NO:1761, referred to as clone L1439P) was identified from a small cell lung carcinoma PCR based subtraction library on lung chip 5. As previously disclosed in Example 4, microarray analysis was performed on this subset of the lung tumor sequences in order to further evaluate their expression profiles in various tumor and normal tissues, the results of which are provided in Table 11. Clone L1439P was shown to have a greater than 2-fold over-expression in the tumor probe group versus normal tissues and an average expression in normal tissues of less than 0.2. Additional studies showed this same clone to be over-expressed in a subset of small cell lung carcinomas (see, Table 16).

Further studies have resulted in the isolation of the full-length cDNA for the cloned sequence of L1439P. In order to determine the transcript size of the gene, a multiple tissue Northern blot was probed with the radioactively labeled original cloned sequence (SEQ ID NO:1180). The Northern blot included about 10 µg of total RNA from small cell lung carcinoma and normal tissues samples. Visual analysis of the

20

WO 02/04514

PCT/US01/22058

196

exposed film revealed a single transcript of approximately 2.4 kb. The extended sequence of the clone was obtained by screening a small cell lung carcinoma tumor cDNA library with a radioactively labeled probe of the original cloned sequence (SEQ ID NO:1761). Approximately 120,000 clones from the cDNA library were screened
5 and one independent clone containing a cDNA insert of 1.5 kb was isolated. This extended cDNA sequence is provided in SEQ ID NO:1931, and the deduced amino acid sequence encoded by this extended cDNA sequence is provided in SEQ ID NO:1932. This extended cDNA sequence was analyzed against the Genbank database and was found to show similarity to the NUF2R gene (AF326731). The full-length cDNA
10 sequence of NUF2R is provided in SEQ ID NO:1933, and the deduced amino acid sequence encoded by NUF2R is provided in SEQ ID NO:1934.

EXAMPLE 2215 EXPRESSION IN MEGATERIUM OF A HISTIDINE TAG-FREE L552S FUSION PROTEIN

PCR was performed on the L552S coding region with the following primers:

20 Forward primer PDM-737 5' ctatgtggcatgcgggtccacgccc 3' (SEQ ID NO:1935) Tm 66°C.

Reverse primer PDM-738 5' cagcctaagatcttcattaaactgtggftg 3' (SEQ ID NO:1936) Tm 60°C.

The PCR conditions were as follows:

25 10µl 10X Pfu buffer
1.0µl 10mM dNTPs
2.0µl 10µM each primer
83µl sterile water
1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)
30 50ng DNA

WO 02/04514

PCT/US01/22058

197

96°C for 2 minutes, 96°C for 20 seconds, 63°C for 15 seconds, 72°C for 1 minute with 40 cycles and then 72°C for 4 minutes.

The PCR product was digested with SphI and BglII restriction enzymes, gel purified and then cloned into pMEG-3, which had been digested with SphI and BglII restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into Megaterium cells for expression.

The amino acid sequence of expressed recombinant L552S is shown in SEQ ID NO:1937, and the DNA coding region sequence is shown in SEQ ID NO:1938.

10

EXAMPLE 23

EXPRESSION IN *E. COLI* OF A HISTIDINE TAG-FREE L552S FUSION PROTEIN

PCR was performed on the L552S coding region with the following primers:

Forward primer PDM-736 5' ctatgttgcatatgatcggtgccacgc 3' (SEQ ID NO:1939) Tm 64°C.

Reverse primer PDM-480 5' ctgagaattcaataaacttggttgctcttccac 3' (SEQ ID NO:1920) Tm 62°C.

The PCR conditions were as follows:

10µl 10X Pfu buffer

1.0µl 10mM dNTPs

2.0µl 10µM each primer

83µl sterile water

1.5µl Pfu DNA polymerase (Stratagene, La Jolla, CA)

50ng DNA

96°C for 2 minutes, 96°C for 20 seconds, 63°C for 15 seconds, 72°C for 1 minute with 40 cycles and then 72°C for 4 minutes.

The PCR product was digested with NdeI and EcoRI restriction enzymes, gel purified and then cloned into pPDM, a modified pET28 vector, which had been digested with NdeI and EcoRI restriction enzymes. The correct construct was

WO 02/04514

PCT/US01/22058

198

confirmed by DNA sequence analysis and then transformed into BLR pLys S and HMS 174 pLysS cells for expression.

The amino acid sequence of expressed recombinant L552S is shown in SEQ ID NO:1940, and the DNA coding region sequence is shown in SEQ ID NO:1941.

5

EXAMPLE 24

GENERATION OF L552S-SPECIFIC CTL LINES USING IN VITRO WHOLE-GENE PRIMING

10 Using *in vitro* whole-gene priming with tumor antigen-vaccinia infected DC (Yee et al, *The Journal of Immunology*, 157(9):4079-86, 1996), human CTL lines were derived that specifically recognize autologous fibroblasts transduced with the L552S tumor antigen, as determined by interferon-gamma ELISPOT analysis. Specifically, dendritic cells (DC) were differentiated from Percoll-purified monocytes
15 derived from PBMC of normal human donors by growing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, DC were infected overnight with a recombinant adenovirus that expresses L552S at a multiplicity of infection (M.O.I) of five, and matured overnight by the addition of 2 µg/ml CD40 ligand. Virus was then inactivated by UV
20 irradiation. CD8+ cells were enriched for by the depletion of CD4+ and CD14+ cells. CD8+ T cells were isolated using a magnetic bead system, and priming cultures were initiated in individual wells of six 96-well plates with the cytokines IL-6 and IL-12. Cultures were restimulated every 7-10 days using autologous primary fibroblasts retrovirally transduced with L552S, and the costimulatory molecule CD80 in the
25 presence of IL-2. Following three stimulation cycles, two CD8+ T cell lines, 2-12G and 4-7A, were identified using interferon-gamma ELISPOT analysis that specifically produce interferon-gamma when stimulated with the L552S tumor antigen-transduced autologous fibroblasts, but not with a control antigen. Both lines were restimulated and tested again in an antibody blocking assay to determine restriction to specific HLAs.
30 Line 2-12G appears to be HLA-B/C restricted, while Line 4-7A appears to be HLA-A restricted. Line 2-12G was cloned using anti-CD3 and feeder cells, with fourteen

WO 02/04514

PCT/US01/22058

199

specific clones being recovered. These clones have the same pattern of reactivity in antibody blocking assays as the parental L2-12G CTL line. In addition, using a panel of HLA-mismatched B-LCL lines transduced with a vector expressing L552S, and measuring interferon-gamma production by the CTL lines in an ELISPOT assay, these

5 CTLs appear to be restricted by HLA-B*4402.

EXAMPLE 25

EPITOPE-MAPPING OF L552S- AND XAGE-1-SPECIFIC ANTIBODIES

10

It was previously found that L552S is an alternative splicing isoform of XAGE-1. In this example, data was obtained using 20mer peptides specific for either L552S or XAGE-1 to screen the sera of lung cancer patients for antibodies specific for L552S and XAGE-1, respectively. It was found that individual cancer patients produce

15 both antibodies specific for L552S as well as for XAGE-1. It was determined that these specific antibodies recognize primarily the following additional epitope of L552S and two epitopes of XAGE-1:

L552S-specific:

aa31-50: LGCCWGYSPRSTWDRPF (SEQ ID NO:1942)

20

XAGE-1-specific:

aa11-30: CSLGVFPSAPSPVWGTRRSC (SEQ ID NO:1943)

aa41-50: ILSPLLRIGGHTQTQNHITAS (SEQ ID NO:1944)

25

EXAMPLE 26

IMMUNOHISTOCHEMISTRY ANALYSIS OF L552S

In order to determine L552S protein expression in various normal and lung cancer tissues, immunohistochemistry (IHC) analysis was performed using an

30 affinity purified L552S polyclonal antibody. Specifically, tissue samples were fixed in a

WO 02/04514

PCT/US01/22058

200

formalin solution for 12-24 hrs and embedded in paraffin before being sliced into 8 micron sections. Steam heat induced epitope retrieval (SHIER) in 0.1 M sodium citrate buffer (pH 6.0) was used for optimal staining conditions. Sections were incubated with 10% serum/PBS for 5 minutes. Primary antibody was added to each section for 25
5 minutes at indicated concentrations followed by a 25 minute incubation with either anti-rabbit or anti-mouse biotinylated antibody. Endogenous peroxidase activity was blocked by three-1.5 minute incubations with hydrogen peroxidase. The avidin biotin complex/horse radish peroxidase (ABC/HRP) system was used along with DAB chromogen to visualize L552S expression. Slides were counterstained with
10 hematoxylin to visualize cell nuclei.

IHC analysis of L552S is disclosed in Table 24:

WO 02/04514

PCT/US01/22058

201

Table 24

Tissue Type	Staining (pos/total)	Comments
Adeno Lung Cancer	10/12	Strong staining (nuclear "dot-like")
Squamous Lung Cancer	4/12	Light staining (cytoplasmic "dot-like")
Adrenal	0/1	
Artery-endothelium	0/1	Light cytoplasmic staining
Blood (bone marrow)	0/1	
Brain (cerebellum)	0/1	
Brain (cortex)	0/1	
Breast	0/1	
Bronchus	0/5	
Colon	1/4	Very light staining (cytoplasmic "dot-like")
Esophagus	0/1	
Fallopian Tube	0/1	
Gall Bladder	0/1	
Heart	0/1	
Kidney	2/4	Very light staining (cytoplasmic "dot-like")
Liver	1/4	Very light staining (cytoplasmic "dot-like")
Lung	0/13	
Pancreas	0/1	
Pituitary	1/1	Very light staining (cytoplasmic "dot-like")
Placenta	1/1	Very light staining (cytoplasmic "dot-like")
Prostate	0/1	
Skeletal Muscle	0/1	
Skin	0/1	
Small Bowel	0/1	
Spinal Cord	0/1	
Spleen	0/1	
Stomach	0/1	
Testis	1/2	Very few selected cells positive
Thymus	0/1	
Thyroid	0/1	
Trachea	0/1	
Urinary Bladder	0/1	
Ureter	0/1	
Uterus	0/1	

WO 02/04514

PCT/US01/22058

202

EXAMPLE 27IDENTIFICATION OF CD4 IMMUNOGENIC T CELL EPTIOPES
DERIVED FROM L552S

5 CD4 T cell lines specific for the antigen L552S (SEQ ID NO:809) were generated as follows. A total of thirty 20-mer peptides overlapping by 15 amino acids corresponding to the amino acid residues of full-length L552S (SEQ ID NO:809) were synthesized. The amino acid sequence of each of the thirty 20-mer peptides and the
10 respective DNA sequence which encodes each of these peptides is provided in Table 25. Dendritic cells (DC) were differentiated from Percoll-purified monocytes derived from PBMC of normal male human donors by plastic adherence and growing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Purified CD4 T cells were generated from the same donor as the DCs by
15 using MACs beads and negative selection of PBMCs. The DCs were pulsed overnight with pools of the 20-mer peptides with each peptide at an individual concentration of 0.5 µg/ml. The pulsed DCs were washed and plated at 10,000 cells per well of 96-well round bottom plates, and purified CD4 T cells were added at 100,000 cells per well. Cultures were supplemented with 10 ng/ml IL-6 and 5 ng/ml IL-12 and incubated at
20 37°C

Cultures were restimulated as above on a weekly basis using DCs made and pulsed as above as the APC, supplemented with 10 U/ml IL-2 and 5 ng/ml IL-7. Following three in vitro stimulation cycles (the initial priming + two restimulations), lines (each line corresponds to one well) were tested for specific proliferation and
25 cytokine production in response to the stimulating pool versus an irrelevant peptide pool of peptides derived from assorted unrelated antigens. A number of individual CD4 T cell lines (49/576 by IFN-gamma and 63/576 by proliferation) demonstrated significant cytokine release (IFN-gamma) and proliferation in response to the L552S peptide pools, but not to the control peptide pool. Twenty five of the T cell lines which exhibited
30 specific activity were restimulated on the appropriate pool of L552S peptides and reassayed on autologous DCs pulsed with the individual peptides or recombinant

WO 02/04514

PCT/US01/22058

203

protein made in *E. coli*. Approximately 13 immunogenic peptides were recognized by the T cells from the entire set of peptide antigens tested. These 13 peptides are (*) in Table 25.

- In some cases the peptide reactivity of the T cell line could be mapped to a single peptide but some could be mapped to more than one peptide in each pool. This result indicates that all 13 peptides may be naturally processed epitopes of the L552S protein.

Table 25

10

L552S CD4 Overlapping Peptides			
Peptide	Sequence	Peptide SEQ ID NO:	DNA SEQ ID NO:
1	MRCHAHGPSCLVTAITREEG	1945	1974
2	HGPSCLVTAITREEGGPRSG	1946	1975
3	LVTAITREEGGPRSGGAQAK	1947	1976
4*	TRFEGGPRSGGAQAKLGCCW	1948	1977
5*	GPRSGGAQAKLGCCWGYVSP	1949	1978
6*	GAQAKLGCCWGYSPRSTWN	1950	1979
7*	LGCCWGYSPRSTWNPDRRF	1951	1980
8	GYSPRSTWNPDRRFWTPQT	1952	1981
9	RSTWNPDRRFWTPQTGPGEQ	1953	1982
10	PDRRFWTPQTGPGEGRHERH	1954	1983
11	WTPQTGPGEGRHERHTQTQN	1955	1984
12	GPGEGRHERHTQTQNHASP	1956	1985
13	RHERHTQTQNHASPRSPVM	1957	1986
14	TQTQNHASPRSPVMESPKK	1958	1987
15*	HTASPRSPVMESPKKKNQQL	1959	1988
16	RSPVMESPKKKNQQLKVGL	1960	1989
17	ESPKKKNQQLKVGLHLGSR	1961	1990
18*	KNQQLKVGLHLGSRQKKJR	1962	1991

WO 02/04514

PCT/US01/22058

204

19*	KVGILILGSRQKKIRIQLRS	1963	1992
20*	HLGSRQKKIRIQLRSQCATW	1964	1993
21*	RQKKIRIQLRSQCATWKVICK	1965	1994
22*	IQLRSQCATWKVICKSCISO	1966	1995
23*	SQCATWKVICKSCISQTPGIN	1967	1996
24*	KVICKSCISQTPGINLDLGS	1968	1997
25	SCISQTPGINLDLGSQVVK	1969	1998
26	TPGINLDLGSQVVKIIPKE	1970	1999
27*	LDLGSQVVKIIPKEEHCKM	1971	2000
28	GVKVKIIPKEEHCKMPEAGE	1972	2001
29	IIPKEEHCKMPEAGEEQPV	1973	2002

EXAMPLE 28

HLA RESTRICTION OF L552S

5

To determine the HLA restriction, a panel of fibroblasts matched at one or two HLA alleles were transduced with pBIB L552 or infected with adenovirus L552 at an MOI of 50:1. These targets were tested against a D77 L552 CD8 clone in an ELISPOT assay with 10000 fibroblasts, 10000 T cells and 5 U/mL IL-2 per well. The results indicate that the restriction is with either B*4402 or Cw*0501.

To decide between the two, COS-7 cells were transfected with a combination of pcDNA3 L552S and pcDNA3 B*4402 or pBIB Cw*0501. These targets were again tested against the same D77 L552 CD8 clone as above in an ELISPOT assay with 10000 COS-7, 10000 clone 14 and 5 U/mL IL-2 per well. The clone recognized the L552 and B*4402 transfected COS-7 cells.

In further studies, three different tumors, 659-22, 3-90T and HTB 183, that had been previously analyzed by real time RT-PCR for L552 message level were selected and transduced with B*4402 or Cw*0501 as a control. 659-22 and 3-90T contain L552 message, while HTB 183 was the negative control. After two rounds of selection, the tumors were analyzed by FACs for their B44 expression level. 659-22

WO 02/04514

PCT/US01/22058

205

were found to endogenously express high levels of B44, but it was not determined which of the B44s, B*4402, B*4404, etc, were being expressed. Approximately 20% of the 3-90T express B44. The HTB 183 express B44 quite well. When the above clone was tested against these tumors, 659-22 transduced with B*4402 was recognized.

5 Thus, the results indicate that L552 is restriction by IILA-B*4402.

EXAMPLE 29

EXPRESSION OF L552S AND XAGE-1

10

In this Example, real-time RT-PCR analysis was performed in order to delineate the expression of L552S from XAGE-1 in lung cancers. The real-time RT-PCR analysis was performed using specific primers localized in the 5' unique region of L552S, the 5' unique region of XAGE-1 and the 3' common region. Specific messages

15 for L552S and XAGE-1 were detected in two lung tumor samples but not in the normal lung samples.

Identical expression profiles were observed between 5' unique region of L552S and the common 3' sequences. The message level for XAGE-1 detected in lung tumors using the 5' unique primers of XAGE-1 was much lower compared with L552S.

20 However, the extreme secondary structure posed by XAGE-1 could hamper the cDNA synthesis of the 5' sequence unique to XAGE-1. Thus, it appears from the Northern analysis that XAGE-1 may be a more abundant isoform.

25

EXAMPLE 30

EST EXPRESSION PROFILE OF L552S AND XAGE-1

To further evaluate the expression profile of L552S and XAGE-1, an electronic express profiling was performed for each antigen. This was done by

30 searching with the same specific primers as disclosed in Example 29 against a public EST database. Results of this profiling confirm that there are two isoforms of the gene.

WO 02/04514

PCT/US01/22058

206

The ratio of expression between L552S and XAGE-1 is about 1:5. In addition, L552S and XAGE-1 seem to be expressed in Hepatoma, CML, germ cell tumor, Ewing's Sarcoma, and Alveolar Rhabdomyosarcoma.

5

EXAMPLE 31

L985P LUNG TUMOR ANTIGEN SURFACE EXPRESSION

Small cell lung carcinoma (SCLC) cell lines, NCI-H69, NCI-H128, HTB-171, HTB-173, HTB-175, and DMS 79, were grown in DMEM (Gibco) containing 10% FBS (Hyclone) at 37° C with 7% CO₂. These growing SCLC cell lines were then subjected to FACS analysis to determine whether L985P is expressed on the surface of these SCLC cell lines using an anti L985P peptide polyclonal sera raised against the predicted extracellular region of L985P.

For FACS analysis, cells were collected and washed with ice cold staining buffer (PBS+1%BSA +Azide + 10 µg/ml human IgG). The cells were next incubated for one hour on ice either with no primary antibody, with irrelevant rabbit IgG, whole molecule at a final concentration of 20 µg/ml, or with affinity purified rabbit polyclonal sera raised against L985P peptide #3482 (SEQ ID NO:1930) at a final concentration of 20 µg/ml. The sequence of the L985P peptide #3482 (SEQ ID NO:1930) represents the predicted extracellular region of the L985P protein. The cells were washed 2 times with staining buffer and then incubated with a 1:100 dilution of a goat anti-rabbit Ig(H+L)-FITC reagent (Southern Biotechnology) for 30 minutes on ice. Following two washes, the cells were resuspended in staining buffer containing propidium iodide (PI), a vital stain that allows for identification of permeable cells, and analyzed by FACS.

In addition, Real-time PCR was performed to determine if L985P mRNA was expressed in these SCLC cell lines. The results of the FACS analysis and the Real time PCR of mRNA expression are presented in Table 26:

30

Table 26

SCLC Cell Line	mRNA Expression by Real Time PCR	Surface Expression by FACS
NCI-H69	+	+
HTB-173	+	+
HTB-175	-	-
HTB-171	-	-
NCI-H128	-	-
DMS 79	-	-

5

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

10

CLAIMS

1. An isolated polynucleotide comprising a sequence selected from the group consisting of:

(a) a sequence provided in SEQ ID NOs:878-1239, 1240, 1243, 1247, 1269, 1272, 1280, 1283, 1285, 1286, 1289, 1300, 1309, 1318, 1319, 1327, 1335, 1339, 1346, 1359, 1369, 1370, 1371, 1393, 1398, 1405, 1408, 1413, 1414, 1417, 1422, 1429, 1432, 1435, 1436, 1438-1442, 1447, 1450, 1453, 1463, 1467, 1470, 1473, 1475, 1482, 1486, 1491-1494, 1501, 1505, 1506, 1514-1517, 1520, 1522, 1524, 1535, 1538, 1542, 1543, 1547, 1554, 1557, 1559, 1561, 1563-1664, 1668, 1669, 1680-1706, 1708-1732, 1734, 1736-1757, 1759-1765, 1767-1770, 1772-1774, 1776-1789, 1824, 1826-1829, 1865-1868, 1873, 1875-1877, 1879, 1881, 1883, 1918, 1922, 1931, 1938, 1941 and 1974-2002;

(b) complement of the sequence provided in SEQ ID NOs:878-1239, 1240, 1243, 1247, 1269, 1272, 1280, 1283, 1285, 1286, 1289, 1300, 1309, 1318, 1319, 1327, 1335, 1339, 1346, 1359, 1369, 1370, 1371, 1393, 1398, 1405, 1408, 1413, 1414, 1417, 1422, 1429, 1432, 1435, 1436, 1438-1442, 1447, 1450, 1453, 1463, 1467, 1470, 1473, 1475, 1482, 1486, 1491-1494, 1501, 1505, 1506, 1514-1517, 1520, 1522, 1524, 1535, 1538, 1542, 1543, 1547, 1554, 1557, 1559, 1561, 1563-1664, 1668, 1669, 1680-1706, 1708-1732, 1734, 1736-1757, 1759-1765, 1767-1770, 1772-1774, 1776-1789, 1824, 1826-1829, 1865-1868, 1873, 1875-1877, 1879, 1881, 1883, 1918, 1922, 1931, 1938, 1941 and 1974-2002;

(c) sequences consisting of at least 20 contiguous residues of a sequence provided in SEQ ID NOs:878-1239, 1240, 1243, 1247, 1269, 1272, 1280, 1283, 1285, 1286, 1289, 1300, 1309, 1318, 1319, 1327, 1335, 1339, 1346, 1359, 1369, 1370, 1371, 1393, 1398, 1405, 1408, 1413, 1414, 1417, 1422, 1429, 1432, 1435, 1436, 1438-1442, 1447, 1450, 1453, 1463, 1467, 1470, 1473, 1475, 1482, 1486, 1491-1494, 1501, 1505, 1506, 1514-1517, 1520, 1522, 1524, 1535, 1538, 1542, 1543, 1547, 1554, 1557, 1559, 1561, 1563-1664, 1668, 1669, 1680-1706, 1708-1732, 1734, 1736-1757,

1759-1765, 1767-1770, 1772-1774, 1776-1789, 1824, 1826-1829, 1865-1868, 1873, 1875-1877, 1879, 1881, 1883, 1918, 1922, 1931, 1938, 1941 and 1974-2002;

(d) sequences that hybridize to a sequence provided in SEQ ID NOs: 878-1239, 1240, 1243, 1247, 1269, 1272, 1280, 1283, 1285, 1286, 1289, 1300, 1309, 1318, 1319, 1327, 1335, 1339, 1346, 1359, 1369, 1370, 1371, 1393, 1398, 1405, 1408, 1413, 1414, 1417, 1422, 1429, 1432, 1435, 1436, 1438-1442, 1447, 1450, 1453, 1463, 1467, 1470, 1473, 1475, 1482, 1486, 1491-1494, 1501, 1505, 1506, 1514-1517, 1520, 1522, 1524, 1535, 1538, 1542, 1543, 1547, 1554, 1557, 1559, 1561, 1563-1664, 1668, 1669, 1680-1706, 1708-1732, 1734, 1736-1757, 1759-1765, 1767-1770, 1772-1774, 1776-1789, 1824, 1826-1829, 1865-1868, 1873, 1875-1877, 1879, 1881, 1883, 1918, 1922, 1931, 1938, 1941 and 1974-2002, under highly stringent conditions;

(e) sequences having at least 75% identity to a sequence of SEQ ID NOs: 878-1239, 1240, 1243, 1247, 1269, 1272, 1280, 1283, 1285, 1286, 1289, 1300, 1309, 1318, 1319, 1327, 1335, 1339, 1346, 1359, 1369, 1370, 1371, 1393, 1398, 1405, 1408, 1413, 1414, 1417, 1422, 1429, 1432, 1435, 1436, 1438-1442, 1447, 1450, 1453, 1463, 1467, 1470, 1473, 1475, 1482, 1486, 1491-1494, 1501, 1505, 1506, 1514-1517, 1520, 1522, 1524, 1535, 1538, 1542, 1543, 1547, 1554, 1557, 1559, 1561, 1563-1664, 1668, 1669, 1680-1706, 1708-1732, 1734, 1736-1757, 1759-1765, 1767-1770, 1772-1774, 1776-1789, 1824, 1826-1829, 1865-1868, 1873, 1875-1877, 1879, 1881, 1883, 1918, 1922, 1931, 1938, 1941 and 1974-2002;

(f) sequences having at least 90% identity to a sequence of SEQ ID NOs: 878-1239, 1240, 1243, 1247, 1269, 1272, 1280, 1283, 1285, 1286, 1289, 1300, 1309, 1318, 1319, 1327, 1335, 1339, 1346, 1359, 1369, 1370, 1371, 1393, 1398, 1405, 1408, 1413, 1414, 1417, 1422, 1429, 1432, 1435, 1436, 1438-1442, 1447, 1450, 1453, 1463, 1467, 1470, 1473, 1475, 1482, 1486, 1491-1494, 1501, 1505, 1506, 1514-1517, 1520, 1522, 1524, 1535, 1538, 1542, 1543, 1547, 1554, 1557, 1559, 1561, 1563-1664, 1668, 1669, 1680-1706, 1708-1732, 1734, 1736-1757, 1759-1765, 1767-1770, 1772-1774, 1776-1789, 1824, 1826-1829, 1865-1868, 1873, 1875-1877, 1879, 1881, 1883, 1918, 1922, 1931, 1938, 1941 and 1974-2002; and

(g) degenerate variants of a sequence provided in SEQ ID NOs: 878-1239, 1240, 1243, 1247, 1269, 1272, 1280, 1283, 1285, 1286, 1289, 1300, 1309, 1318,

WO 02/04514

PCT/US01/22058

210

1319, 1327, 1335, 1339, 1346, 1359, 1369, 1370, 1371, 1393, 1398, 1405, 1408, 1413, 1414, 1417, 1422, 1429, 1432, 1435, 1436, 1438-1442, 1447, 1450, 1453, 1463, 1467, 1470, 1473, 1475, 1482, 1486, 1491-1494, 1501, 1505, 1506, 1514-1517, 1520, 1522, 1524, 1535, 1538, 1542, 1543, 1547, 1554, 1557, 1559, 1561, 1563-1664, 1668, 1669, 1680-1706, 1708-1732, 1734, 1736-1757, 1759-1765, 1767-1770, 1772-1774, 1776-1789, 1824, 1826-1829, 1865-1868, 1873, 1875-1877, 1879, 1881, 1883, 1918, 1922, 1931, 1938, 1941 and 1974-2002.

2. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) SEQ ID NOs: 1667, 1670-1675, 1678, 1806-1822, 1825, 1830-1833, 1834-1856, 1863, 1869-1872, 1874, 1876, 1878, 1880, 1884-1890, 1913, 1917, 1921, 1927-1930, 1932, 1934, 1937, 1940 and 1942-1973;

(b) sequences having at least 70% identity to the amino acid sequence as provided in SEQ ID NOs: 1667, 1670-1675, 1678, 1806-1822, 1825, 1830-1833, 1834-1856, 1863, 1869-1872, 1874, 1876, 1878, 1880, 1884-1890, 1913, 1917, 1921, 1927-1930, 1932, 1934, 1937, 1940 and 1942-1973;

(c) sequences having at least 90% identity to the amino acid sequence as provided in SEQ ID NOs: 1667, 1670-1675, 1678, 1806-1822, 1825, 1830-1833, 1834-1856, 1863, 1869-1872, 1874, 1876, 1878, 1880, 1884-1890, 1913, 1917, 1921, 1927-1930, 1932, 1934, 1937, 1940 and 1942-1973;

(d) sequences encoded by a polynucleotide of claim 1;

(e) sequences having at least 70% identity to a sequence encoded by a polynucleotide of claim 1; and

(f) sequences having at least 90% identity to a sequence encoded by a polynucleotide of claim 1.

3. An expression vector comprising a polynucleotide of claim 1 or a polynucleotide which encodes a polypeptide of claim 2 operably linked to an expression control sequence.

WO 02/04514

PCT/US01/22058

211

4. A host cell transformed or transfected with an expression vector according to claim 3.

5. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a polypeptide selected from the group consisting:

- (a) a polypeptide according to claim 2; and
- (b) a polypeptide having an amino acid sequence that is encoded by a polynucleotide sequence provided in SEQ ID NOs:1948-1951, 1959, 1962-1968 and 1971 or a complement thereof.

6. An isolated antibody or antigen-binding fragment thereof according to claim 5, wherein the polypeptide is provided in SEQ ID NOs:1927-1929 and 1942-1973.

7. A method for detecting the presence of a cancer in a patient, comprising the steps of:

- (a) obtaining a biological sample from the patient;
- (b) contacting the biological sample with a binding agent that binds to a polypeptide selected from the group consisting of:

- (i) a polypeptide of claim 2,
- (ii) a polypeptide having an amino acid sequence provided in SEQ ID NOs:1677, 1679, 1806-1822, 1925, 1926 and 1934,
- (iii) a polypeptide having at least 90% identity to an amino acid sequence provided in SEQ ID NOs:1677, 1679, 1806-1822, 1925, 1926 and 1934,
- (iv) a polypeptide having at least 95% identity to an amino acid sequence provided in SEQ ID NOs:1677, 1679, 1806-1822, 1925, 1926 and 1934,
- (v) a polypeptide encoded by a polynucleotide provided in SEQ ID NOs: 1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446,

WO 02/04514

PCT/US01/22058

212

1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933, and

(vi) a polypeptide encoded by a polynucleotide having at least 90% identity to a sequence provided in SEQ ID NOs: 1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-1299, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;

(c) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(d) comparing the amount of polypeptide to a predetermined cut-off value and therefrom determining the presence of a cancer in the patient.

8. A fusion protein comprising at least one polypeptide selected from the group consisting of:

- (i) a polypeptide of claim 2,
- (ii) a polypeptide having an amino acid sequence provided in SEQ ID NOs:1677, 1679, 1806-1822, 1925, 1926 and 1934,
- (iii) a polypeptide having at least 90% identity to an amino acid sequence provided in SEQ ID NOs:1677, 1679, 1806-1822, 1925, 1926 and 1934,
- (iv) a polypeptide having at least 95% identity to an amino acid sequence provided in SEQ ID NOs:1677, 1679, 1806-1822, 1925, 1926 and 1934,
- (v) a polypeptide encoded by a polynucleotide provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282,

WO 02/04514

PCT/US01/22058

213

1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933, and

(vi) a polypeptide encoded by a polynucleotide having at least 90% identity to a sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933.

9. A fusion protein according to claim 8, wherein the at least one polypeptide is provided in:

(a) SEQ ID NOs:1876, 1878, 1913, 1917, 1921, 1937 and 1940; and

(b) a polypeptide encoded by a polynucleotide provided in SEQ ID NOs:1875, 1877, 1914, 1918, 1922, 1938 and 1941.

10. An oligonucleotide that hybridizes to at least one sequence selected from the group consisting of:

(a) a sequence set forth in claim 1; and

WO 02/04514

PCT/US01/22058

214

(b) a sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933,

under highly stringent conditions.

11. A method for stimulating and/or expanding T cells specific for a tumor protein, comprising contacting T cells with at least one component selected from the group consisting of:

- (a) polypeptides according to claim 2;
- (b) a polypeptide having an amino acid sequence provided in SEQ ID NOs:1677, 1679, 1806-1822, 1901-1909, 1925, 1926 and 1934;
- (c) a polypeptide having an amino acid sequence at least 90% identity to SEQ ID NOs:1677, 1679, 1806-1822, 1901-1909, 1925, 1926 and 1934;
- (d) a polypeptide having an amino acid sequence at least 95% identity to SEQ ID NOs:1677, 1679, 1806-1822, 1901-1909, 1925, 1926 and 1934;
- (e) a polynucleotide according to claim 1;
- (f) a polynucleotide which encodes a polypeptide having an amino acid sequence as provided in (b), (c) or (d);
- (g) a polynucleotide having a sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481,

WO 02/04514

PCT/US01/22058

215

1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;

(h) complement of the sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;

(i) sequences that hybridize to a sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933, under highly stringent conditions;

(j) sequences having at least 90% identity to a sequence of SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523,

WO 02/04514

PCT/US01/22058

216

1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;

(k) sequences having at least 95% identity to a sequence of SEQ ID NOs: 1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;

(l) degenerate variants of a sequence provided in SEQ ID NOs: 1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933; and

(m) antigen-presenting cells that express a polypeptide according to (a), (b), (c) or (d),

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

12. An isolated T cell population, comprising T cells prepared according to the method of claim 11.

WO 02/04514

PCT/US01/22058

217

13. A composition comprising a first component selected from the group consisting of physiologically acceptable carriers and immunostimulants, and a second component selected from the group consisting of:

- (a) polypeptides according to claim 2;
- (b) a polypeptide having an amino acid sequence provided in SEQ ID NOs:1677, 1679, 1806-1822, 1901-1909, 1925, 1926 and 1934;
- (c) a polypeptide having an amino acid sequence with at least 90% identity to one of SEQ ID NOs:1677, 1679, 1806-1822, 1901-1909, 1925, 1926 and 1934;
- (d) a polypeptide having an amino acid sequence with at least 95% identity to one of SEQ ID NOs:1677, 1679, 1806-1822, 1901-1909, 1925, 1926 and 1934;
- (e) a polynucleotide according to claim 1;
- (f) a polynucleotide which encodes a polypeptide having an amino acid sequence as provided in (b), (c) or (d);
- (g) a polynucleotide having a sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;
- (h) complement of the sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-

WO 02/04514

PCT/US01/22058

218

1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;

(i) sequences that hybridize to a sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933, under highly stringent conditions;

(j) sequences having at least 90% identity to a sequence of SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;

(k) sequences having at least 95% identity to a sequence of SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560,

WO 02/04514

PCT/US01/22058

219

1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;

(l) degenerate variants of a sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933;

(m) antibodies according to claim 5;
 (n) fusion proteins according to claim 8;
 (o) T cell populations according to claim 12; and
 (p) antigen presenting cells that express a polypeptide according to (a), (b), (c) or (d).

14. A method for stimulating an immune response in a patient, comprising administering to the patient a composition of claim 13.

15. A method for the treatment of a lung cancer in a patient, comprising administering to the patient a composition of claim 13.

16. A method for determining the presence of a cancer in a patient, comprising the steps of:

(a) obtaining a biological sample from the patient;
 (b) contacting the biological sample with an oligonucleotide according to claim 10;
 (c) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

WO 02/04514

PCT/US01/22058

220

(d) compare the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence of the cancer in the patient.

17. A diagnostic kit comprising at least one oligonucleotide according to claim 10.

18. A diagnostic kit comprising at least one antibody according to claim 5 and a detection reagent, wherein the detection reagent comprises a reporter group.

19. A method for inhibiting the development of a lung cancer in a patient, comprising the steps of:

(a) incubating CD4+ and/or CD8+ T cells isolated from a patient with at least one component selected from the group consisting of:

(i) polypeptides according to claim 2;
 (ii) a polypeptide having an amino acid sequence provided in SEQ ID NOs:1677, 1679, 1806-1822, 1901-1909, 1925, 1926 and 1934,
 (iii) a polypeptide having an amino acid sequence with at least 90% identity to one of SEQ ID NOs:1677, 1679, 1806-1822, 1901-1909, 1925, 1926 and 1934,

(iv) a polypeptide having an amino acid sequence with at least 95% identity to one of SEQ ID NOs:1677, 1679, 1806-1822, 1901-1909, 1925, 1926 and 1934,

(v) polynucleotides according to claim 1,

(vi) a polynucleotide which encodes a polypeptide having an amino acid sequence as provided in (i), (ii), (iii) or (iv),

(vii) a polynucleotide having a sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412,

WO 02/04514

PCT/US01/22058

221

1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933,

(viii) complement of the sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933,

(ix) sequences that hybridize to a sequence provided in SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933, under highly stringent conditions,

(x) sequences having at least 90% identity to a sequence of SEQ ID NOs:1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-199, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446,

WO 02/04514

PCT/US01/22058

222

1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933,

(xi) sequences having at least 95% identity to a sequence of SEQ ID NOs: 1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-1299, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933,

(xii) degenerate variants of a sequence provided in SEQ ID NOs: 1241, 1242, 1244-1246, 1248-1268, 1270, 1271, 1273-1279, 1281, 1282, 1284, 1287, 1288, 1290-1299, 1301-1308, 1310-1317, 1320-1326, 1328-1334, 1336-1338, 1340-1345, 1347-1358, 1360-1368, 1372-1397, 1399-1404, 1406, 1407, 1409-1412, 1415, 1416, 1418-1421, 1423-1428, 1430, 1431, 1433, 1434, 1437, 1443-1446, 1448, 1449, 1451, 1452, 1454-1462, 1464-1466, 1468, 1469, 1471, 1472, 1474, 1476-1481, 1483-1485, 1487-1490, 1495-1500, 1502-1504, 1507-1513, 1518, 1519, 1521, 1523, 1525-1534, 1536, 1537, 1539-1541, 1544-1546, 1548-1553, 1555, 1556, 1558, 1560, 1562, 1676, 1707, 1733, 1735, 1758, 1766, 1771, 1775, 1790-1805, 1823, 1910, 1923, 1924 and 1933, and

(xiii) antigen presenting cells that express a polypeptide according to (i), (ii), (iii) or (iv),

such that T cell proliferate;

(b) administering to the patient an effective amount of the proliferated T cells; and thereby inhibiting the development of a cancer in the patient.

【国際公開パンフレット（コレクトバージョン）】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/004514 A2

(51) International Patent Classification: C07K 14/47

(21) International Application Number: PCT/US01/22058

(22) International Filing Date: 10 July 2001 (10.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/616,174 11 July 2000 (11.07.2000) US
09/651,563 29 August 2000 (29.08.2000) US
09/658,824 8 September 2000 (08.09.2000) US
09/671,325 26 September 2000 (26.09.2000) US
09/677,419 6 October 2000 (06.10.2000) US
09/702,705 30 October 2000 (30.10.2000) US
09/736,457 13 December 2000 (13.12.2000) US
09/849,626 3 May 2001 (03.05.2001) US

(71) Applicant (for all designated States except US): CORINA CORPORATION [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).

(72) Inventors: and

(75) Inventors/Applicants (for US only): WANG, Tongtao [US/US]; 8049 N.E. 28th Street, Medina, WA 98039 (US); WATANABE, Yoshihiro [JP/US]; 2366 78th Avenue S.E., Mercer Island, WA 98040 (US); HENDERSON, Robert A. [US/US]; 8004 192nd Street S.W., Edmonds, WA 98026 (US); JOHNSON, Jeffrey C. [US/US]; 25828 15th Place Saco, Des Moines, WA 98198 (US); RETTER, Marc W. [US/US]; 33702 N.E. 43rd Place, Carnation, WA 98014 (US); DURHAM, Margarita [US/US]; 8436 36th Avenue W., Seattle, WA 98199 (US); CARTER, Derrick [US/US]; 324 Summit Avenue E., Seattle, WA 98102 (US); FANGER, Gary, Richard [US/US]; 15906 29th Drive S.E., Mill Creek, WA 98032 (US); VEDVICK, Thomas S. [US/US]; 124 S. 300th Place, Federal Way, WA 98003 (US); BANGUR, Chaitanya S. [IN/US]; 2102 N. 105th Street, Apt. 1101, Seattle, WA 98133 (US); MCNABB, Andria [US/US]; 3602 N.E. 14th Court, Renton, WA 98055 (US); WANG, Aijun [CN/US]; 4106 273rd Place S.E., Issaquah, WA 98029 (US); FANGER, Neil [US/US];

3648 Whitman Avenue N., A100, Seattle, WA 98105 (US); SWITZER, Ann [US/US]; 3838 N.W. 85th Street, Seattle, WA 98115 (US); MCNEILL, Patricia, D. [US/US]; 1533 South 290th Place, Federal Way, WA 98003 (US); CLAPPER, Jonathan, D. [US/US]; 2149 Dexter Avenue N., #4, Seattle, WA 98109 (US).

(74) Agents: VERNIA, James, M.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7002 et al. (US).

(51) Designated States (nationally): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GR, GU, HK, HM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(54) Designated States (regionally): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LI, MC, NL, PT, SE, TR), OAPI patent (BF, BI, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
without international search report and to be republished upon receipt of that report
with sequence listing part of description published separately in electronic form and available upon request from the international bureau

(48) Date of publication of this corrected version: 27 December 2002

(15) Information about Correction: see PCT Gazette No. 52/2002 of 27 December 2002, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly lung cancer, are disclosed. Illustrative compositions comprise one or more lung tumor polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly lung cancer.



WO 02/004514 A2

【国際公開パンフレット（コレクトバージョン）】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/004514 A3(51) International Patent Classification: C12N 15/12
C07K 14/07, 14/705, 16/20, G01N 33/574, 33/58, C12Q
1/68, C12N 15/52, 5/10, A61K 38/17, 48/00, 39/395,
A51P 35/00

(21) International Application Number: PCT/US01/23058

(22) International Filing Date: 10 July 2001 (10.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/014,124 11 July 2000 (11.07.2000) US
09/054,353 29 August 2000 (29.08.2000) US
09/058,834 8 September 2000 (08.09.2000) US
09/071,375 26 September 2000 (26.09.2000) US
09/077,419 6 October 2000 (06.10.2000) US
09/702,705 30 October 2000 (30.10.2000) US
09/736,457 13 December 2000 (13.12.2000) US
09/849,676 3 May 2001 (03.05.2001) US(71) Applicant (for all designated States except U.S.): CORINA
CORPORATION [US/US]; 1124 Columbia Street, Suite
200, Seattle, WA 98104 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): WANG, Tongtong
[US/US]; 3049 N.E. 28th Street, Medina, WA 98029 (US);
WATANABE, Yoshihiro [JP/US]; 2366 78th Avenue S.E.,
Mercer Island, WA 98040 (US); HENDERSON, Robert,
A. [US/US]; 8901 192nd Street S.W., Edmonds, WA 98026
(US); JOHNSON, Jeffrey, C. [US/US]; 25828 5th Place
Street, Des Moines, WA 98198 (US); KETTER, Marc,
W. [US/US]; 73405 N.E. 43rd Place, Carnation, WA
98014 (US); DURHAM, Margarita [US/US]; 3404 36th
Avenue W., Seattle, WA 98119 (US); CARTER, Derrick
[US/US]; 321 Summit Avenue E., Seattle, WA 98102 (US);
FANGER, Gary, Richard [US/US]; 15906 - 29th Drive
S.E., Mill Creek, WA 98012 (US); VEDVICK, Thomas,
S. [US/US]; 124 S. 300th Place, Federal Way, WA 98003
(US); BANGUR, Chaitanya, S. [IN/US]; 2102 N. 105th
Street, Apt. 1-01, Seattle, WA 98103 (US); MCNABB,
Audria [US/US]; 3602 N.E. 11th Court, Renton, WA
98056 (US); WANG, Aijun [CN/US]; 3106 E. 9th PlaceS.E., Issaquah, WA 98029 (US); FANGER, Neil [US/US];
3648 Whitman Avenue N., A100, Seattle, WA 98103 (US);
SWITZER, Ann [US/US]; 3838 S.E. 85th Street, Seattle,
WA 98115 (US); MCNEILL, Pauline, D. [US/US];
1333 South 290th Place, Federal Way, WA 98003 (US);
CLAPPER, Jonathan, D. [US/US]; 2149 Dexter Avenue
N., 44, Seattle, WA 98109 (US).(74) Agents: VERNIA, James, M., Seed Intellectual Property
Law Group PLLC, Suite 6500, 701 Fifth Avenue, Seattle,
WA 98104 2092 et al. (US)(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EG, ES, FI, GB, GR, GU, HK,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZW), European
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LI, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

with international search report
with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau(88) Date of publication of the international search report:
19 June 2003

(15) Information about Correction:

Previous Correction:

see PCT Gazette No. 52/2002 of 27 December 2002, Sec-
tion IIFor two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette

(54) Title: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER

(57) Abstract: Compositions and methods for the therapy and diagnosis of cancer, particularly lung cancer, are disclosed. Illustrative compositions comprise one or more lung tumor polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compositions are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly lung cancer.



WO 02/004514 A3

【国際調査報告】

INTERNATIONAL SEARCH REPORT		Inter. Appl. No.
A. CLASSIFICATION OF SUBJECT MATTER		PCT/US 01/22058
IPC 7 C12N15/12 C07K14/47 C07K16/30 G01N33/574 G01N33/68 C12Q1/68 C12N15/62 C12N5/10 A61K38/17 A61K48/00 A61K39/395 A61P35/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FILES SEARCHED		
Magnetic disc, microfiche searched; information system followed by classification symbols		
IPC 7 C12N C07K		
Documentation searched which may refer to documents relevant to the state of the art that such documents are included in the files searched		
Documents have been examined during the international search (state of the base and, where practical, search terms used)		
EPO-Internal, SEQUENCE SEARCH, WPI Data, BIOSIS, EMBL		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Classification of document, with indication of the appropriate of the relevant passages	Relevance to claim No.
X	WO 99 38973 A (CORIXA CORP) 5 August 1999 (1999-08-05) page 1 -page 46; claims; examples	1-19
X	GURE ET AL: "Human lung cancer antigens recognized by autologous antibodies: definition of a novel cDNA derived from the tumor suppressor gene locus on chromosome 3p21.3" CANCER RESEARCH, AMERICAN ASSOCIATION FOR CANCER RESEARCH, BALTIMORE, MD, US, vol. 58, 1 March 1998 (1998-03-01), pages 1034-1041, XP002103188 ISSN: 0008-5472 the whole document	1-19
-/-		
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents: "A" document calling the general state of the art, which is not considered to be of particular relevance "E" earlier document, but published after the international filing date "L" document which may throw doubts on priority claim(s) on which is cited in the publication date of another citation (i.e. other special means (e.g. parallel) "O" document referring to removal of disclaimers, additions or other means "P" document published prior to the international filing date but later than the priority date claimed		
"T" later document published after the international filing date or priority date and in conflict with the application being filed by underlined the applicant or theory underlying the invention "X" mechanical of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being set forth in a claim or claim in the art "Z" document showing the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
4 June 2002		11.09.2002
Name and mailing address of the ISA European Patent Office, P.O. Box 1 Patentamt 2 AL 2280 (47) 1100 Tel (+41-70) 340-2040, Fax (+41-70) 340-2016		Author's name Madruga, J

Form PCT/ISA/210 (second sheet) July 1999

INTERNATIONAL SEARCH REPORT		Sheet 1 of Application No. PCT/US 01/22058
C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Character of document, with indication, where appropriate, of the relevant passages	Relevance to the prior art
X	CHEN S-L ET AL: "Isolation and characterization of a novel gene expressed in multiple cancers" ONCOGENE, BASINGSTOKE, HANTS, GB, vol. 12, no. 4, 15 February 1996 (1996-02-15), pages 741-751, XP002106555 ISSN: 0950-9232 the whole document	1-19
X	--- DATABASE EMBL [Online] ID: H51285106, 1 July 1997 (1997-07-01) "ab23f06.r1 Stratagene lung (#937210) Homo sapiens cDNA clone" Database accession no. AA488696 XP002199797	1-6, 8-10,17, 18
Y	abstract	7,11-16, 19
X	--- DATABASE EMBL [Online] 7 June 2000 (2000-06-07) "Homo sapiens cDNA" Database accession no. AW956090 XP002199798	1-6, 8-10,17, 18
Y	abstract	7,11-16, 19
X	--- DATABASE EMBL [Online] ID: H573234, 5 March 1995 (1995-03-05) Database accession no. T63/32.1 XP002199799	1-6, 8-10,17, 18
X	abstract	
X	-& HILLIER L ET AL: "GENERATION AND ANALYSIS OF 280,000 HUMAN EXPRESSED SEQUENCE TAGS" GENOME RESEARCH, COLD SPRING HARBOR LABORATORY PRESS, US, vol. 6, 1996, pages 807-828, XP002942261 ISSN: 1088-9051	1-6, 8-10,17, 18
Y	the whole document	7,11-16, 19
P,X	--- WD 00 55375 A (ALPHAGENE INC) 21 September 2000 (2000-09-21) SEQ ID NO: 87, SEQ ID NO: 88 page 443; claim 96 --- -/--	1-19

Form PCT/US 2002 (Distribution of Second Sheet) (A4) (9/1992)

page 3 of 3

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 01/22058
Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)		
This International Search Report has not been established in respect of certain claims under Article 17(2)(c) for the following reasons:		
1.	☒ Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210	
2.	☐ Claims Nos.: because they relate to parts of the information: Application that do not comply with the prescribed requirements to such an extent that no meaningful international Search can be carried out, specifically:	
3.	☐ Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).	
Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)		
This International Searching Authority found multiple inventions in this international application, as follows:		
see additional sheet		
1.	☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.	
2.	☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.	
3.	☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:	
4.	☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:	
1-19 (all partially)		
Remark on Protest: ☐ The additional search fees were accompanied by the applicant's protest. ☐ No protest accompanied the payment of additional search fees.		

International Application No. PCT/US 01/22058

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claim 11 (as far as in vivo methods are concerned) and claims 14, 15 and 19 are directed to a method of treatment of the human/ animal body, the search has been carried out and based on the alleged effects of the compound/ composition.

Although claims 7 and 16 are directed to a diagnostic method practised on the human/ animal body, the search has been carried out and based on the alleged effects of the compound/ composition.

International Application No. PCT/US 01/22358

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

Invention 1, Claims: 1-19 (all partly)

A polynucleotide comprising a sequence provided in SEQ ID NO:878, a sequence containing at least 20 contiguous residues of said sequence, a sequence having at least 75% identity to the polynucleotide, a degenerate variant of said sequence, a polypeptide encoded by said sequence, a vector and a host containing the polynucleotide, an antibody that specifically binds to the encoded polypeptide and a fusion protein comprising said polypeptide.

A pharmaceutical composition or vaccine comprising said polypeptide, and a method for inhibiting the development of a cancer in a patient, a method for stimulating and/or expanding T cells specific for a tumour protein, an isolated T cell population, a method for determining or monitoring a cancer in a patient and a diagnostic kit.

Inventions 2-1125, Claims: 1-19 (all partly)

As invention 1, the subject-matter of each invention being limited to a SEQ ID NO, wherein:

Invention 2 is limited to SEQ ID NO 879,
Invention 3 is limited to SEQ ID NO 890, etc.

...

Invention 1125 is limited to SEQ ID NO 2002

INTERNATIONAL SEARCH REPORT
 Information on patent family members

 International Application No.
 PCT/US 91/22658

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9938973 A	05-08-1999	AU 2344399 A	16-08-1999
		BR 9907259 A	22-01-2002
		CA 2319117 A1	05-08-1999
		CN 1292629 T	18-04-2001
		EP 1051489 A2	15-11-2000
		JP 2002516659 T	11-06-2002
		NO 20003853 A	27-09-2000
		WO 9938973 A2	05-08-1999
		ZA 9900693 A	21-10-1999
WO 0055375 A	21-09-2000	AU 3899700 A	04-10-2000
		WO 0055375 A1	21-09-2000
		AU 2883500 A	04-09-2000
		EP 1153121 A2	14-11-2001
		WO 0049134 A1	24-08-2000

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード (参考)
A 6 1 K 48/00	A 6 1 K 48/00	4 H 0 4 5
A 6 1 P 35/00	A 6 1 P 35/00	
A 6 1 P 37/04	A 6 1 P 37/04	
C 0 7 K 14/82	C 0 7 K 14/82	
C 0 7 K 16/32	C 0 7 K 16/32	
C 0 7 K 19/00	C 0 7 K 19/00	
C 1 2 N 1/15	C 1 2 N 1/15	
C 1 2 N 1/19	C 1 2 N 1/19	
C 1 2 N 1/21	C 1 2 N 1/21	
C 1 2 N 5/10	C 1 2 Q 1/68	A
C 1 2 Q 1/68	C 1 2 N 5/00	A
	A 6 1 K 37/02	

- (31)優先権主張番号 09/671,325
 (32)優先日 平成12年9月26日(2000.9.26)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 09/677,419
 (32)優先日 平成12年10月6日(2000.10.6)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 09/702,705
 (32)優先日 平成12年10月30日(2000.10.30)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 09/736,457
 (32)優先日 平成12年12月13日(2000.12.13)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 09/849,626
 (32)優先日 平成13年5月3日(2001.5.3)
 (33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

- (72)発明者 渡部 良広
 アメリカ合衆国 ワシントン 98040, マーサー アイランド, 78ティーエイチ アベニュー エス. イー. 2266
 (72)発明者 ヘンダーソン, ロバート エイ.
 アメリカ合衆国 ワシントン 98026, エドモンズ, 192エヌディー ストリート エス. ダブリュー. 8904
 (72)発明者 ジョンソン, ジェフリー シー.
 アメリカ合衆国 ワシントン 98198, デス モイネス, 15ティーエイチ プレイス ストリート 25828
 (72)発明者 レッター, マーク ダブリュー.
 アメリカ合衆国 ワシントン 98014, カーネイション, エヌ. イー. 43アールディー プレイス 33402

- (72)発明者 マーネラキス, マルガリータ
アメリカ合衆国 ワシントン 98199, シアトル, 36ティーエイチ アベニュー ダブ
リュウ. 3444
- (72)発明者 カーター, ダリック
アメリカ合衆国 ワシントン 98102, シアトル, サミット アベニュー イー. 32
1
- (72)発明者 ファンガー, ギャリー リチャード
アメリカ合衆国 ワシントン 98012, ミル クリーク, 29ティーエイチ ドライブ
エス. イー. 15906
- (72)発明者 ヴェドヴィック, トーマス エス.
アメリカ合衆国 ワシントン 98003, フェデラル ウェイ, エス. 300ティーエイ
チ プレイス 124
- (72)発明者 バンガー, チャイタニャ エス.
アメリカ合衆国 ワシントン 98133, シアトル, エヌ. 105ティーエイチ ストリ
ート 2102, アpartment ジェイ101
- (72)発明者 マクナブ, アンドリア
エヌ. イー. 11ティーエイチ シー. 3602
- (72)発明者 ワン, アイジュン
アメリカ合衆国 ワシントン 98029, イサクア, 213ティーエイチ プレイス エス
. イー. 3106
- (72)発明者 ファンガー, ネイル
アメリカ合衆国 ワシントン 98103, シアトル, ウィットマン アベニュー エヌ.
3648, エイ100
- (72)発明者 スウィッツァー, アン
アメリカ合衆国 ワシントン 98115, シアトル, エヌ. イー. 85ティーエイチ ス
トリート 3838
- (72)発明者 マクネイル, パトリシア ディー.
アメリカ合衆国 ワシントン 98003, フェデラル ウェイ, サウス 290ティーエイ
チ プレイス 1333
- (72)発明者 クラッパー, ジョナサン ディー.
アメリカ合衆国 ワシントン 98109, シアトル, デクスター アベニュー エヌ. 2
149, ナンバー4

F ターム(参考) 4B024 AA01 AA12 BA80 CA01 CA07 GA11 HA12 HA20
4B063 QA19 QQ02 QQ08 QQ42 QQ53 QQ79 QQ96 QR32 QR48 QR55
QS33 QS34 QX01
4B065 AA93Y AB01 AC14 BA02 CA24 CA25 CA44 CA46
4C084 AA02 AA03 AA07 AA13 AA19 BA01 BA02 BA41 NA14 ZB092
ZB262
4C085 AA13 AA14
4H045 AA10 AA11 AA20 BA41 CA41 DA76 EA28 EA51 FA74